

Curso 2009/10
HUMANIDADES Y CIENCIAS SOCIALES/9
I.S.B.N.: 978-84-7756-948-0

JUAN JOSÉ COLOMINA ALMIÑANA

**Los problemas de las teorías
representacionales de la conciencia**

Director
ANTONIO MANUEL LIZ GUTIÉRREZ



SOPORTES AUDIOVISUALES E INFORMÁTICOS
Serie Tesis Doctorales

*Als meus pares,
per fer-ho possible...*

LOS PROBLEMAS DE LAS TEORÍAS REPRESENTACIONALES DE LA CONCIENCIA

ÍNDICE DE CONTENIDOS

| | |
|--|-----------|
| Agradecimientos | 13 |
| Prólogo | 17 |
| Capítulo 1. Introducción: una perspectiva sobre los estudios actuales en filosofía de la conciencia | 35 |
| 1.1. La peculiaridad del contenido fenoménico | 39 |
| 1.2. La distinción entre contenido conceptual y contenido no-conceptual | 41 |
| 1.3. La conciencia y el problema mente-cuerpo | 44 |
| 1.4. Los problemas de la conciencia | 60 |
| 1.4.1. El problema de la propiocepción y la perspectiva en primera persona | 61 |
| 1.4.2. El argumento del conocimiento | 62 |
| 1.4.3. El hiato explicativo | 63 |
| 1.4.4. El problema de la causalidad de lo fenoménico | 64 |
| 1.4.5. El problema de los duplicados físicos | 65 |
| 1.4.5.1. El gemelo de Chalmers | 66 |
| 1.4.5.2. Robots | 67 |
| 1.4.5.3. Científicos marcianos | 68 |
| 1.4.5.4. El hombre del pantano | 69 |
| 1.4.6. El espectro invertido | 70 |
| 1.4.7. La variación estándar | 71 |
| 1.4.8. La visión borrosa | 72 |
| 1.4.9. El problema de la transparencia | 73 |
| 1.4.10. El problema de la luminosidad | 74 |
| 1.4.11. El problema de la localización del contenido fenoménico | 75 |
| 1.4.12. El problema de los miembros-fantasma | 76 |
| 1.4.13. El problema de la percepción sin percepciones: la visión ciega | 76 |
| 1.4.14. El problema de la conciencia dividida | 77 |
| 1.4.15. El problema de la unidad de lo consciente | 78 |
| 1.5. Recapitulación | 79 |
| Capítulo 2. Representación y conciencia | 81 |
| 2.1. ¿Qué es una representación? | 84 |
| 2.1.1. Tipos de conciencia | 84 |
| 2.1.1.1. El carácter cualitativo de la experiencia | 87 |
| 2.1.1.2. El carácter subjetivo de la experiencia | 87 |
| 2.1.2. Aproximaciones al carácter subjetivo de la conciencia | 89 |
| 2.1.2.1. Teorías conceptualistas | 90 |
| 2.1.2.2. Teorías funcionalistas | 91 |
| 2.1.2.3. Teorías cognitivas | 93 |

| | |
|---|------------|
| 2.1.2.4. Teorías de nivel superior | 96 |
| 2.1.3. El contenido representacional de la conciencia | 101 |
| 2.1.3.1. Fodor sobre conceptos | 101 |
| 2.1.3.2. Contenido conceptual y no-conceptual en la experiencia fenoménica | 103 |
| 2.1.3.3. ¿Son independientes el contenido conceptual y el no-conceptual? | 105 |
| 2.2. Aproximaciones representacionales de la conciencia | 108 |
| 2.2.1. El carácter subjetivo desde una perspectiva representacionista | 108 |
| 2.2.1.1. Teorías informacionales | 111 |
| 2.2.1.2. Teorías teleológicas | 113 |
| 2.2.1.3. Primitivismo | 116 |
| 2.2.1.4. Representacionismo intencional | 117 |
| 2.2.2. Tipos de representacionismo | 120 |
| 2.2.2.1. Representacionismo puro, fuerte y débil | 121 |
| 2.2.2.2. Modos de caracterizar el contenido de las experiencias | 122 |
| 2.2.2.3. Representacionismo amplio y restringido | 122 |
| 2.2.2.4. Representacionismo reductivo y no-reductivo | 122 |
| 2.2.3. Cognición, representación y <i>qualia</i> | 123 |
| 2.2.3.1. El carácter cualitativo (representacional) de la experiencia | 123 |
| 2.2.3.2. Diferentes usos del término <i>qualia</i> | 123 |
| 2.2.3.3. ¿Qué estados mentales poseen carácter fenoménico? | 124 |
| 2.2.4. El representacionismo como escuela dominante | 125 |
| 2.2.4.1. Argumentos a favor | 126 |
| 2.2.4.2. Argumentos en contra | 126 |
| 2.3. Recapitulación | 128 |
| | |
| Capítulo 3. El dualismo representacional de Searle | 131 |
| 3.1. El misterio de la conciencia | 134 |
| 3.1.1. Francis Crick y la búsqueda científica del alma | 136 |
| 3.1.2. Gerald Edelman y la topobiología del cerebro | 138 |
| 3.1.3. Roger Penrose: microtúbulos y citoesqueletos | 139 |
| 3.1.4. Daniel Dennett y la negación de la conciencia | 140 |
| 3.1.5. David Chalmers y la mente consciente | 141 |
| 3.1.6. Conclusión: cómo convertir un misterio en un problema (para todos) | 141 |
| 3.2. El naturalismo biológico de Searle | 143 |
| 3.2.1. La intencionalidad de la mente | 145 |
| 3.2.1.1. Diferencias entre la intencionalidad de la mente y del lenguaje | 145 |
| 3.2.1.2. Acción, intencionalidad y percepción | 147 |
| 3.2.1.3. Redes conceptuales y trasfondo en los estados mentales | 150 |
| 3.2.1.4. Conclusión: implicaciones del trasfondo y las redes conceptuales en la vida mental de los individuos conscientes | 154 |
| 3.2.2. La conciencia como propiedad emergente del cerebro | 155 |
| 3.2.2.1. El compromiso aristotélico | 155 |
| 3.2.2.2. Las características estructurales de la conciencia | 156 |
| 3.2.2.3. El principio de conexión y la ciencia del inconsciente | 158 |

| | |
|---|------------|
| 3.2.2.4. El papel causal de la mente: ¿el dualismo de propiedades de Searle? | 163 |
| 3.2.2.5. Conclusión: el precio a pagar por el dualismo | 170 |
| 3.3. Recapitulación | 174 |
| Capítulo 4. Michael Tye y la teoría representacional fuerte de la conciencia | 177 |
| 4.1. La teoría intencional de Tye | 181 |
| 4.1.1. Clases de representacionalismo | 181 |
| 4.1.2. El intencionalismo de Michael Tye | 182 |
| 4.2. La teoría PANIC y su aplicación sobre el contenido fenoménico | 183 |
| 4.2.1. La teoría PANIC de Tye | 184 |
| 4.2.2. El contenido de la conciencia y los problemas respecto de la percepción | 185 |
| 4.2.3. El contenido de la conciencia en el estudio de las sensaciones corporales, las emociones y los estados de ánimo | 190 |
| 4.3. Los obstáculos para una teoría representacional fuerte respecto de la conciencia | 197 |
| 4.3.1. Revisando la teoría PANIC | 198 |
| 4.3.1.1. La condición de preparación del contenido fenoménico | 199 |
| 4.3.1.2. Objetos abstractos y contenido fenoménico | 201 |
| 4.3.1.3. Contenido conceptual frente a contenido no-conceptual en Tye | 202 |
| 4.3.1.4. La necesidad de la intencionalidad del contenido fenoménico | 205 |
| 4.3.2. El análisis de Tye del problema de la Variación Estándar | 206 |
| 4.3.2.1. Presentando el problema: el rompecabezas del “verdadero” azul | 206 |
| 4.3.2.2. Una primera solución: adaptándose a las circunstancias | 207 |
| 4.3.2.3. Una segunda solución: la alternativa de Tye | 209 |
| 4.3.2.4. Conclusión: diferencia sin dualidad de propiedades | 210 |
| 4.3.3. Por una nueva versión del contenido de la experiencia fenoménica | 211 |
| 4.3.3.1. Conciencia y atención | 212 |
| 4.3.3.2. Desplumando la gallina | 219 |
| 4.3.3.3. El contenido de la experiencia fenoménica como un demostrativo | 221 |
| 4.3.4. Espectro invertido y <i>qualia</i> ausentes en Tye | 225 |
| 4.3.4.1. El análisis de Tye del problema del espectro invertido | 225 |
| 4.3.4.2. El problema de los <i>qualia</i> ausentes en Tye | 227 |
| 4.3.5. Mary, <i>zombies</i> y el hiato explicativo: la solución de Tye | 228 |
| 4.3.5.1. “¡Ay Mary, Mary! ¿Realmente aumenta tu conocimiento?” | 229 |
| 4.3.5.2. El problema de los <i>zombies</i> en Tye | 231 |
| 4.3.5.3. ¿Existe eso que podríamos llamar hiato explicativo? | 233 |
| 4.4. Recapitulación | 236 |
| Capítulo 5. El representacionalismo débil de Dennett | 239 |
| 5.1. El desarrollo filogenético de la conciencia: evolución e intencionalidad en la filosofía naturalista de Dennett | 243 |
| 5.1.1. El papel de la evolución en la intencionalidad de la mente | 243 |
| 5.1.1.1. El relojero ciego | 244 |
| 5.1.1.2. Organismos biológicos e Inteligencia Artificial | 246 |
| 5.1.1.3. ¿Por qué plantea un problema la intencionalidad derivada? | 247 |

| | |
|--|------------|
| 5.1.1.4. El verdadero problema de la representación errónea | 252 |
| 5.1.2. Tipos de mentes: criaturas creadas | 255 |
| 5.1.2.1. Las criaturas darwinianas | 256 |
| 5.1.2.2. Las criaturas skinnerianas | 256 |
| 5.1.2.3. Las criaturas popperianas | 257 |
| 5.1.2.4. Las criaturas gregorianas | 257 |
| 5.1.3. La aparición del pensamiento | 258 |
| 5.1.3.1. Cuando las palabras hacen cosas con nosotros: la aparición de la representación | 259 |
| 5.1.3.2. Cooperación, imitación y egoísmo | 260 |
| 5.1.3.3. Fabricando nuevas herramientas: la irrupción de la semántica | 262 |
| 5.1.4. Conclusiones: las dificultades de la teoría evolutiva dennettiana | 264 |
| 5.2. El desarrollo ontogenético de la conciencia: el modelo de los borradores múltiples | 266 |
| 5.2.1. La implausibilidad de la existencia de un teatro cartesiano en la conciencia | 267 |
| 5.2.2. La heterofenomenología y el modelo de los borradores múltiples | 270 |
| 5.2.2.1. Dos tipos de revisiones de la experiencia: Orwell vs. Stalin | 271 |
| 5.2.2.2. La heterofenomenología como método para la conciencia | 273 |
| 5.2.3. Quineando los <i>qualia</i> | 274 |
| 5.2.4. El papel normativo de los estados mentales (incluidos los fenoménicos) | 276 |
| 5.3. La teoría dennettiana como una teoría representacional de Orden Superior | 277 |
| 5.3.1. La teoría descriptiva de Dennett frente a otras teorías de Orden Superior | 278 |
| 5.3.1.1. Teorías de Primer Orden (FOT) y Teorías de Orden Superior (HOT) | 279 |
| 5.3.1.2. Las dificultades de las Teorías de Orden Superior | 280 |
| 5.3.1.3. La Teoría Descripciónista de Orden Superior de Dennett | 284 |
| 5.3.2. Críticas a la teoría de Orden Superior de Dennett a partir de consideraciones biológicas y neurofisiológicas | 286 |
| 5.3.2.1. “Pero, la evolución no funciona de ese modo” | 286 |
| 5.3.2.2. Algunos datos desde las neurociencias | 289 |
| 5.3.2.3. Revisando la metáfora dennettiana | 292 |
| 5.4. Addenda: la acción moral como paradigma en Dennett | 293 |
| 5.4.1. La noción de individuo en el Argumento Básico de Galen Strawson | 294 |
| 5.4.2. Las respuestas al Argumento Básico de Strawson | 296 |
| 5.4.3. Dennett sobre libertad y responsabilidad | 299 |
| 5.5. Recapitulación | 305 |
| 6. Hacia una teoría no sólo representacional de la conciencia | 307 |
| 6.1. Conciencia e intención: el contenido de los estados fenoménicos | 309 |
| 6.1.1. Intención, representación y el carácter fenoménico | 310 |
| 6.1.2. Los problemas del fisicalismo causal | 312 |
| 6.1.3. ¿Tienen eficacia causal los estados mentales con carácter fenoménico? | 315 |
| 6.2. ¿Por qué deberíamos abandonar la teoría de la representación en favor de una teoría intencional de la conciencia? | 320 |
| 6.2.1. Los problemas fáciles, revisados | 320 |
| 6.2.1.1. Ni aun viviendo en un mundo <i>zombie</i> | 321 |
| 6.2.1.2. El fisicalismo es verdadero aunque Mary aprenda algo nuevo | 326 |

| | |
|---|-----|
| 6.2.1.3. El curioso caso de la heladería mágica | 331 |
| 6.2.1.4. ¿Podemos plantear una solución materialista al problema de la unidad de la experiencia? | 332 |
| 6.2.2. El problema difícil, revisado | 333 |
| 6.2.2.1. ¿Existe realmente el hiato explicativo? | 333 |
| 6.2.2.2. Entonces, ¿por qué los <i>qualia</i> parecen no ser físicos? | 340 |
| Conclusiones generales | 343 |
| Referencias bibliográficas | 383 |

AGRADECIMIENTOS

Desde que en el verano de 2007 decidiera cambiar completamente de tema de trabajo (había estado dedicado hasta entonces por completo a la Filosofía del Lenguaje) y decidiera centrarme en el desarrollo de una teoría de la conciencia, numerosas son las deudas que (en uno u otro sentido) he ido adquiriendo.

El mérito principal de que este trabajo haya podido ser llevado a término se lo debo a Antonio Manuel ('Manolo') Liz Gutiérrez, quién, sin conocerme de nada, accedió amablemente a dirigir mi investigación. De Manolo he recibido mucho en los últimos meses, y no sólo desde el punto de vista académico, sino también desde el plano personal. El tiempo que ha dedicado a leer por completo esta tesis y a realizar las correcciones necesarias y ayudarme en las partes más conflictivas dice mucho a favor de su profesionalidad. Pero, todavía más importante, la cálida acogida que él y Margarita Vázquez Campos me proporcionaron dentro del grupo de investigación LEMA (perteneciente a la Universidad de La Laguna) y en las ocasiones que me recibieron en Tenerife todavía dice mucho más a favor de su humanidad. En estos últimos meses, realmente desde que los conocí en septiembre de 2007, he recibido sus constantes muestras de apoyo académico y sus constantes muestras de apoyo personal. A ambos, gracias.

Muchos también son aquellos de los cuales me he enriquecido en los últimos años. Quiero agradecer las largas conversaciones y los valiosos comentarios que sobre varios de los temas desarrollados en esta tesis han tenido a bien compartir conmigo Vicente Raga Rosaleny (incansable compañero de fatigas y, sobre todo, amigo), David Pérez Chico (con el que además pude compartir unos fantásticos días de estímulo intelectual en Granada en septiembre de 2008), Andrés L. Jaume (tal vez, el mejor ejemplo que dispongo de lo que sería el trabajo bien hecho) y Quico Ribes (quién desde el comienzo de mi andadura me ha acompañado sin esperar recibir nunca nada a cambio). Pero también son muchos los que pacientemente han soportado mis constantes comentarios y desarrollos improvisados (algunos de ellos a deshora) pero que, muy pacientemente han sabido entender mi estado de ánimo: Josep Vidal, Miguel Ángel Martínez Arnau y Borja Colón de Carvajal.

Por último, pero no menos importante, quiero recordar a aquellas personas que desde siempre han estado a mi lado sin cuestionar nunca nada de lo que he hecho y que, sin embargo siempre han estado dispuestas a cederme su tiempo, su compañía, su apoyo y su cariño: mis abuelos Pepe, Amelia, Rigoberto y María, mis hermanas Yolanda y Cristina, mi sobrina Edurne y, especialmente, a mis padres Juan José Colomina Iváñez y María Teresa Almiñana Faus (a quienes dedico este trabajo). Sin todos ellos, ninguno de los momentos de mi vida *consciente* hubiese tenido sentido.

* * *

Algunos fragmentos de esta tesis han sido redactados empleando algunos textos ya publicados. Quiero agradecer a todas estas publicaciones su confianza.

El apartado 4.1. y fragmentos de la sección 4.2.3. apareció parcialmente ilustrada en "Tomando conciencia de nuestro cuerpo. Las explicaciones de Tye sobre las sensaciones

corporales”, en la revista *Dilema* de la Universidad de Valencia, Vol. XII, No.1, 2008 (pp. 103-123). El resto de la sección 4.2.3. se reformuló para la ponencia “On Tye’s intentionalist view about experience of emotions”, impartida dentro del marco del *1st Workshop on Language and Emotions*, celebrado en Granada en septiembre de 2008.

La sección 4.3.2. apareció como capítulo en el libro, editado por M. Vázquez y A. Alonso, *Periferias: el extremo como término medio* (Madrid, Verbum, 2008, pp. 93-102) como “¿Es subjetiva la experiencia de color? Sobre el argumento representacionista de Tye acerca de la Variación Estándar”.

El grueso del punto 5.1. se compone de reformulaciones de varios artículos publicados en diferentes revistas de divulgación. La revista *Arbor* del CSIC publicó en el Vol. 179, No. 705, de septiembre de 2004 (pp. 247-263), mi texto “El papel de la evolución en la teoría de Dennett”. El Instituto Tecnológico de Monterrey (México), publicó en *Revista de Humanidades del Tecnológico de Monterrey*, No. 19, otoño 2005 (pp. 169-192), mi trabajo “Diseñando la mente”. La revista *Argos*, de la Universidad Simón Bolívar de Venezuela, publicó en su Vol. 23, No. 44, 1r semestre 2006 (pp. 6-20) mi artículo “Intencionalidad y representación errónea en la teoría filosófica de Daniel C. Dennett”. La revista *Contrastes* de la Universidad de Málaga, en su Vol. XI de 2006 (pp. 97-115) publicó mi texto “Criaturas y *creaturas*: conciencia, evolucionismo e intencionalidad en la filosofía naturalista de Dennett”. En el libro colectivo editado por A. Alonso Martos, J. D. Mateu Alonso y V. Raga Rosaleny, *Surcar la cultura* (València, Pre-textos, 2006, pp. 79-86) apareció mi capítulo titulado “El fundamento biológico de la cultura en Daniel C. Dennett”.

Una versión previa de la sección 5.3.2.1. fue defendida como comunicación en junio de 2009 en el Congreso Internacional “La idea de evolución: 150 años después de Darwin”, celebrado en Salamanca por el Instituto de Estudios de la Ciencia y la Tecnología con el título “La visión evolutiva de Dennett respecto de la intencionalidad y la conciencia”.

Una versión reducida del apartado 5.4. apareció con el título “Algunas notas acerca de la filosofía de la acción moral en Daniel C. Dennett” en el volumen colectivo *TecnoÉtica: temas éticos relacionados con la ciencia, la tecnología y sus implicaciones sociales* (Barcelona, Publicacions de la Universitat de Barcelona/Instituto de Tecnoética, 2008, pp. 180-188). Una versión revisada y ampliada (co-escrita con Vicente Raga Rosaleny) apareció con el título de “Determinismo, autoconfiguración y Principio de Posibilidades Alternativas en la filosofía de la mente y de la acción de Daniel C. Dennett” en A. Fernández-Caballero, M. Manzano, E. Alonso y S. Miguel (eds.); *Una perspectiva de la Inteligencia Artificial en su 50 aniversario* (Albacete, Publicaciones de la Universidad de Castilla-La Mancha, 2006, pp. 161-172) y en la revista *Límite* de la Universidad de Tarapacá de Chile, Vol. II, No. 15, de 2007 (pp. 61-80) con el título “Algunas observaciones sobre la responsabilidad (moral) de nuestras acciones”.

La sección 6.1.2. es una reformulación de una comunicación presentada en el Congreso de la SLMFCE celebrado en Granada en el año 2006 y que apareció en sus actas, editadas por F. Martínez y L. M. Peris (Granada, Publicaciones de la Universidad de Granada, 2006, pp. 467-470) con el título “Sobre libertad y necesidad: una crítica al fisicalismo”.

Una versión previa de la sección 6.2.1.1. aparecerá durante 2009 en *Revista de Filosofía* de la Universidad Complutense, Vol. 33, No. 2, con el nombre “Una solución materialista a la corazonada *zombie*”.

* * *

La traducción de los textos citados a lo largo de esta tesis y que originalmente están escritos en inglés es mía.

PRÓLOGO

*¡Oh! ¿Quién puede sostener fuego en la mano
Pensando en el helado Cáucaso?
¿O embotar el agudo filo del hambre
Por la simple imaginación de un festín?
¿O revolcarse desnudo en la nieve de diciembre
Pensando en el calor de un fantástico verano?
(Shakespeare, Ricardo III)*

*Dos almas, ¡ay!, moran en mi pecho.
Una quiere separarse de la otra.
Una, con órganos tenaces,
Se aferra al mundo en intenso deleite amoroso;
La otra se eleva con fuerza desde el polvo
Hacia las regiones de los excelsos antepasados.
(Goethe, Fausto)*

PRÓLOGO

1. Contextualización

La noción de conciencia siempre ha estado vinculada, de alguna u otra forma, a la problemática general de la relación existente entre el cuerpo (como sustancia material) y el alma o espíritu (en tanto que sustancia incorpórea y pensante). Dicha contraposición se encuentra presente en todas las culturas, e inunda sus mitos y sus religiones, germen de toda su vida cotidiana. Dichas creencias religiosas, en su sentido más amplio, tienen que ver con ciertas concepciones animistas, en el seno de las cuales se forja la concepción de la existencia de un alma inmaterial e imperecedera que establece el dualismo entre ella y el cuerpo como la teoría de mayor calado y antigüedad de la que se tiene constancia, una concepción del mundo que acaba por impregnar las creencias cotidianas de las gentes que animan dichas culturas y que acaba por conformar, teniendo en cuenta siempre dicha dicotomía, su sentido común y su lenguaje ordinario. En tanto que la constitución de la separación entre cuerpo y alma afecta tanto a la idea de la posible autoconfiguración e identidad de los sujetos (mediante la interacción de sus partes para configurar la capacidad cognoscitiva necesaria para realizar toda actividad, como a la posibilidad de su permanencia más allá de la muerte del cuerpo (lo que daría sentido a la vida y a la muerte más allá de toda sujeción material), se convierte tal vez en el problema de mayor envergadura al que podemos hacer frente y al que, por supuesto, toda reflexión filosófica, tarde o temprano, debe acudir.

Por el término alma (posteriormente reconvertido, en la tradición moderna a partir de Descartes, en la noción de *res cogitans* o mente), al menos en esta tradición antigua, debemos entender el principio vital que anima a todo ser viviente. En tanto principio biológico, la noción de alma puede ser entendida como una realidad material y corpórea, una *psique*. Pero también puede ser caracterizada como un elemento inmaterial, distinto del cuerpo, que insufla vida a todo organismo, desde los vegetales a los animales superiores y al hombre.

Ya en los escritos homéricos podemos encontrar esta dualidad a la hora de definir la concepción del alma. Por un lado, encontramos el alma como aliento vital, como soplo de vida, que se consume con la muerte del cuerpo. Por otro lado, encontramos a la *psique*, al alma imperecedera, que será liberada tras la muerte del cuerpo y acudirá a ocupar su lugar en el Tártaro.

El desarrollo de este dualismo místico y religioso inunda los primeros siglos de la vida griega que nos relatan los pocos escritos que se conservan. Pero con la aparición de las primeras interpretaciones ontológicas de la naturaleza, con la aparición de la escuela pre-socrática como paradigma revolucionario que rompe con la concepción mística de la separación entre alma y cuerpo, y con la concepción religiosa y mítica de la explicación de los hechos que acontecen en el mundo, surgen nuevas interpretaciones de esta dicotomía. Por un lado, encontramos a aquellos que consideran que el alma debe ser un principio inmaterial e imperecedero, una tradición que se remonta hasta Pitágoras de Samos y Empédocles y que recogerán las enseñanzas de Sócrates y Platón. Por otro, podemos diferenciar distintas posiciones naturalistas

que desde Demócrito a Epicuro, pretenden explicar el alma a partir de la composición de pequeñas partículas elementales (los átomos).

El dualismo mente-cuerpo platónico, como es bien sabido, se encuadra dentro de un dualismo todavía más amplio que afirma la existencia de un mundo aparente de experiencia sensible (al que pertenecería el cuerpo humano, sujeto a limitación y cambio) como contrapuesto a un mundo inteligible (al que pertenece el alma, como inmortal e imperecedera, a la que le es propio el verdadero conocimiento intuitivamente accesible). El alma es propia del mundo inteligible, donde se encuentra en su naturaleza divina contemplando las Ideas, aunque de manera contingente haya caído al mundo sensible y haya sido encarcelada en un cuerpo que limita su conocimiento a aquello que pueden captar sus sentidos y limita su acción a lo mundano.

En su *Fedón*, donde recoge y renueva elementos tenidos presentes en su *República*, Platón muestra la separación tajante que existiría entre el cuerpo y el alma. Dicha contraposición tiene por base un profundo carácter moral, en tanto que se manifiesta en la actitud misma de Sócrates frente a la aceptación de su muerte. La serenidad y preparación con la que Sócrates afronta su condena a muerte, que supone su liberación del mundo sensible, muestra la preparación del que espera la llegada del momento en el que abandone las limitaciones sensible y vuelva a poder contemplar las Ideas.

Pero que el cuerpo esté limitado frente al alma racional no quiere decir que la concepción del cuerpo platónica sea sólo negativa. Acorde con la mentalidad griega del '*mens sana in corpore sano*', como deja constancia en la *República*, Platón concibe una necesaria armonía entre el cuerpo y el alma para poder alcanzar la virtud en el mundo sensible, una virtud que sólo puede alcanzar aquel cuyo cuerpo está equilibrado (mediante la práctica de la gimnasia) y cuya mente está ejercitada para la sabiduría (a partir de la música, las matemáticas y la filosofía). De nuevo, vemos aquí un elevado grado de moralidad en la teoría platónica. No sólo nos insinúa Platón la necesidad de una correcta relación entre lo mental y lo corpóreo para una vida saludable, sino también para poder desempeñar una vida recta en la que la correcta medida en el justo momento permita guiar nuestras acciones hacia una vida buena, entendida en el sentido de hacer en todo momento aquello que nos corresponde hacer por naturaleza.

El planteamiento de Aristóteles, a diferencia del platónico, es naturalista y biológico. El alma pertenece al campo físico, donde se sitúan los seres naturales dotados de vida, en tanto principio vital que anima los cuerpos orgánicos. El alma y el cuerpo son una y la misma sustancia, una sustancia que constituye los seres vivos. El alma sería la forma, el acto vital, que anima el cuerpo, y que permite a los seres materiales llevar a cabo sus propias funciones vitales. Al igual que la anterior tradición que considera el alma como principio de vida, asociada a la respiración y la sangre como soplo vital, Aristóteles concibe un alma como principio vital del cuerpo. Pero esta visión se constituye a partir de una nueva perspectiva biológica que persigue la máxima consistencia científica posible. Por dicho motivo, se prioriza el valor de todos aquellos datos observados con la intención de poder alcanzar el mayor grado de descripción

posible de las propiedades de los seres vivos, a la vez unión de las características somáticas y psíquicas que lo componen.

Desde este planteamiento materialista, la vida se caracteriza por las actividades básicas de nutrición, crecimiento, reproducción y muerte. Dichas funciones vitales se dan en la naturaleza de modo cada vez más complejo y desarrollado, lo que permite diferenciar una serie sucesiva de grados de complejidad en las mismas funciones básicas. Visto desde este punto de vista, las funciones superiores se sustentan sobre la base de las funciones básicas y primarias, un conjunto de funciones que le son propias a todo organismo por aquello que es. El hombre, en tanto que ser superior, representará el esquema de toda la naturaleza y, por su complejidad funcional, la relación existente entre sus funciones corporales y sus funciones mentales puede ser extremadamente problemática. Vemos, entonces, que en el hombre se da la conjunción de funciones vegetativas, nutritivas y sensitivas con funciones como la memoria, la imaginación y las emociones (de naturaleza claramente mental).

Pero, a diferencia de la teoría platónica, la cuestión cambia radicalmente cuando nos preguntamos acerca de la posibilidad de la existencia de un intelecto puro, de un pensamiento como actividad intrínseca y exclusiva del alma. Aristóteles sostendrá, entonces, una teoría doble. Dicha teoría del doble entendimiento, una actividad tanto activa como pasiva, donde Aristóteles abraza un oscuro eclecticismo. La única parte separable de lo físico, y por tanto inmortal, nos dice el estagirita, es el entendimiento activo, el entendimiento agente, en tanto que propio de acción, mientras que el entendimiento pasivo, aquel solamente capacitado para captar lo sensible, parece con el cese del cuerpo. Así, aunque en un principio parece que Aristóteles sostiene un monismo sustancial, también Aristóteles defiende un cierto dualismo que será debate de las generaciones posteriores.

En este sentido, en la tradición cristiana que se inicia a principios del siglo II, Alejandro de Afrodisia afronta la separación aristotélica como crucial a la hora de poder atender a la confrontación entre lo mundano y lo divino. Propio de las almas es lo divino, guiado por la actividad intelectual capaz de acerca los hombres a Dios, mientras que la mundanidad está sujeta al cuerpo. Pero la actividad del alma debe poder imponerse frente a la actividad mundana del cuerpo. La teoría del doble entendimiento se radicaliza con los pensadores medievales como Averroes, que ve en la posibilidad de la existencia de un intelecto con dos vertientes la puerta de salida al problema que, desde siempre, ha supuesto para un planteamiento determinista como el cristiano la libertad que, supuestamente, Dios prioriza en los humanos. Frente a la tradición platónica, que afirma la dualidad alma-cuerpo como contingente y que asegura la inmortalidad del alma y su llegada al mundo inteligible, la tradición aristotélica prefiere apelar a la bondad y voluntad divina para explicar la presencia de dos entendimientos capaces de dar cuenta de los hechos del mundo. Por un lado, porque el hombre tiene un entendimiento pasivo, está sujeto a los avatares propios de la naturaleza, a la determinación que su pasividad le impone, al padecimiento del acontecer. Pero, por otro lado, en tanto que poseedor de un intelecto activo, el hombre está obligado a actuar. Precisamente, sus acciones son las que determinarán sus actos futuros. La posibilidad de conciliar libertad y necesidad permite que, en su justa medida, el

entendimiento agente pueda interactuar con el mundo para proporcionarse las herramientas necesarias para forjar su propia salvación.

En una línea también voluntarista, pero negadora de la posibilidad del doble entendimiento, San Agustín nos aporta una visión conciliadora del problema mente-cuerpo. El cuerpo tiende a la dispersión, por su debilidad, en tanto que se ve atraído por la mundanidad. El alma, en cambio, debe buscar el amparo divino y, guiado por su voluntad infinita, debe tender hacia la salvación a partir de la asimilación de la enseñanza y el ejemplo de Dios. Dicha concepción del alma, entendida como única, como voluntad eterna y agente, es la que trasciende a la historia y la que se tendrá como la única posibilidad de entendimiento, forjando el sentido de alma de la cristiandad, desde Pedro Abelardo a Tomás de Aquino.

Si en el mundo antiguo la imagen que predomina sobre la realidad es la de un organismo completo, compuesto por todas sus partes anímicas, en tanto que modelo dinámico y vitalista, con su propia intencionalidad, con sus fines y propósitos, en el mundo moderno dichas explicaciones teleológicas y causales, la explicación de por qué las cosas se comportan del modo en que lo hacen, pasa a un segundo plano. La pregunta esencial ya no es por qué suceden las cosas, sino que lo esencial es la pregunta por el cómo actúan los objetos del mundo. Ya no se trata de averiguar las fuerzas internas, vivas y animadas, de las cosas del mundo real, sino que el objetivo es preguntarse acerca de las fuerzas mecánicas que mueven las cosas hacia la acción. A partir del siglo XV, caracterizado por un considerable aumento del número y variedad de mecanismos y máquinas, el modelo de la mente y del cuerpo como una máquina se incrementa, hasta el punto de dominar todo estudio de la realidad.

En este sentido, el mundo actuaría sincrónicamente como el mecanismo de un reloj, dotado desde el inicio por la acción que le otorga Dios como relojero único. Todos los aspectos de la realidad, entonces, estarán claramente mediatizados por una serie de leyes que, en conjunto, permiten dotar de armonía al universo y que, una vez descubiertas, permiten entender la realidad como una máquina perfectamente engranada en la que todo tiene un sentido de conjunto sincrónico. En esta imagen mecanicista, el hombre aparece como un autómatas cuyo cuerpo está sujeto a las leyes mecánicas de la física, establecidas por Galileo y Newton, y que es dirigido por un alma que lo habita, como el marinero dirige su navío. Visto así, como será la visión cartesiana, la metáfora del mecanismo sólo puede aplicarse a la sustancia extensa, al cuerpo, pero no a la sustancia pensante, la mente. A partir de esta idea esencial, las sucesivas generaciones tendrán por objetivo dilucidar las estrechas relaciones existentes entre estas dos sustancias, de cómo existe entre el alma y el cuerpo una relación no-causal que permite explicar las acciones del cuerpo como planteadas por el alma que lo habita (Malabranche y Leibniz) o a partir de la asimilación de las leyes mecánicas del mundo como necesarias (Spinoza).

Aunque la relación entre la mente y el cuerpo es un tema amplio, los estudios empíricos de Locke y Hume tienen en común un interés inusitado por cuestiones puramente psicológicas, por cuestiones que afectan directamente al funcionamiento de la propia mente humana. Frente al planteamiento innatista y esencialista cartesiano y centrándose en la capacidad cognitiva humana, en el origen de las ideas como mecanismos de conocimiento, están más interesados en

estudiar la relación y la asociación de ideas a partir de la noción de dato de los sentidos, en tanto que mediación psíquica de la relación entre los elementos de la mente y los elementos del mundo externo. Y esta relación tiene claras implicaciones a la hora de actuar, en tanto que el sujeto se dejará asesorar por el resultado de dicha asociación, hasta el punto de vincular directamente su actitud moral a dicho entendimiento.

Berkeley, por su parte, reducirá el mundo de las cosas materiales al conjunto de sensaciones que la mente humana posee de ellas, siendo la sustancia espiritual una mente capaz de captar las sensaciones del cuerpo, que quedan a un nivel de pura materialidad.

Precisamente Kant radicaliza el punto de vista idealista, pero teniendo en cuenta las conclusiones materialistas, cuando concibe la razón humana como la guía de toda acción. A partir de la asimilación de las categorías básicas que permiten conformar la realidad (y cuya presencia en la mente tiene sentido sólo a partir de la unificación entre concepto y contenido), Kant afirma la existencia de una apercepción trascendental que permite al sujeto humano diferenciarse como sujeto mental y de acción del resto de elementos puramente materiales y que, además, permite forjar su propia imagen como sujeto libre, permite configurar su propia identidad personal a partir de los elementos que le son propios como mente trascendental.

A partir de la orientación científica que se comienza a imponer a mediados del siglo XIX, dominada por el positivismo y el evolucionismo, comienza a forjarse la actual imagen de la mente como encarnada en un cuerpo, una imagen que rehuye de todo dualismo. Desde esta concepción, la mente y la conciencia es el producto de los procesos fisiológicos propios de un cuerpo y, aunque no se niega la existencia de elementos mentales, éstos quedan como puros epifenómenos al poder ser explicados a partir de las presencia de estados físicos. Con la aparición de la psicología científica de Wundt y, sobre todo, de W. James se opta por el establecimiento de un paralelismo total entre la mente y el cuerpo, una relación que aconseja a los psicólogos la defensa de una correspondencia total entre los estados de conciencia y los procesos neurofisiológicos del cerebro.

Con la publicación en 1949 de *El concepto de lo mental* de Ryle, la anterior concepción dualista comienza a abandonarse y la relación existente entre mente y cuerpo adquiere precisamente este tono identificador: el problema mente-cuerpo se reformula y adquiere el habitual sentido contemporáneo. Ahora, el problema de la relación entre cuerpo y mente es ver en qué sentido los estados físicos producen los estados mentales.

2. Antecedentes de la cuestión

Aunque el tema general de conciencia ha sido tratado de muy variadas maneras desde antiguo, debemos ser capaces de centrar nuestro trabajo en un área mucho más restringida y modesta.

Ya los antiguos filósofos griegos, como hemos indicado al principio, mostraron una especial predilección por temas relativos a la filosofía de la mente (aunque inicialmente entendida como la contraposición de un alma racional habitando en un cuerpo contingente material, que actúa a modo de prisión). En la *República*, Platón ya se preguntaba acerca de

temas limítrofes (como la felicidad o la fortaleza de carácter), temas ampliados en sus textos *Fredo* y *Fedón*, y mostró un interés sobre la ontología del alma humana, estratificándola en tres partes bien definidas. Aristóteles, interesado también por estos temas, realizó estudios acerca de estas cuestiones, vinculándolas directamente con un estudio materialista del mismo, dado que comenzó a considerar la posibilidad de que, al menos ciertas de nuestras capacidades, pudieran ser compartidas con otras especies animales. Ambos filósofos crearon escuela, y los temas sobre los que trataron fueron ampliados por sus sucesores a partir de desarrollos vinculados de algún u otro modo a una teoría ontológica relativa al alma y a un desarrollo epistemológico relativo al tema concreto de las virtudes (los seguidores de Platón) o bien a un estudio antropológico vinculado estrechamente con la biología animal y desarrollos de corte materialista relativos al tema de la razón como generadora de categorías.

Con la aparición del cristianismo, el tema de la conciencia se convirtió en un lugar más o menos común y central, en tanto que la conciencia (humana, por supuesto) permitía el punto necesario de inflexión para salvaguardar el libre albedrío, vinculada directamente con una ontología en donde la categoría de alma se convierte en prioritaria, en tanto que permitiría la posibilidad de que el hombre pudiera actuar de manera libre a partir de la asimilación de las ordenanzas divinas. La posibilidad de la existencia de un alma inmortal, unida a un cuerpo perecedero, permitía la benevolencia de Dios y permitía salvaguardar la posibilidad de una ontología en donde toda la realidad estaba causada directamente por las Categorías presentes en la Mente Divina.

Pero, como hemos comentado, no es precisamente este tipo de conciencia el que suscita nuestro interés. Nuestro tema nace más o menos a finales del siglo XVII, cuando Descartes transforma el viejo problema ontológico de la posibilidad de la existencia de un alma unida a un cuerpo en un problema epistemológico. Es en este momento cuando, directamente dirigido sobre el individuo particular, toma relevancia la tesis de una realidad externa (de la cual podríamos dudar, por ser inferida tan sólo a partir de los datos indirectos que nos proporcionan nuestros órganos perceptores) como contrapuesta a una realidad interna (clara y distinta, por ser dada directamente a nuestra conciencia). La transformación aquí de la contraposición cuerpo-alma y su reformulación en el clásico problema Mente-Cuerpo, supone el punto de inflexión desde el cual parte nuestra problemática.

Pasando por las críticas materialistas que los filósofos ingleses Locke y Hume realizan de las tesis cartesianas y la ampliación de las mismas realizadas por filósofos posteriores (como Stuart Mill o, más recientemente, Russell) y teniendo en cuenta los desarrollos dualistas (vinculados directamente a un neo-idealismo naciente) realizados por Berkeley y otros filósofos embelesados por la idea de una mente unida a un cuerpo (como si existiera un fantasma habitando nuestra máquina, como indica la tan manida metáfora de Ryle), llegamos a comienzos del siglo XX, donde nuestro tema se amplía a partir de los estudios realizados por el positivismo lógico del Círculo de Viena y las teorías de los conductistas (tanto los metodológicos como los lógicos).

Dentro de esta reformulación contemporánea, la mente humana deja de ser entendida como un alma para pasar a ser considerada como una sustancia mental, directamente vinculada con la posesión de un instrumento racional capaz de, por un lado, crear una imagen del exterior y del interior de la misma que, en dos momentos diferentes, permitiera al sujeto, por una parte, generar una imagen del mundo bajo una serie de interpretaciones racionales englobadas bajo categorías universales de catalogación y, por otra parte, a partir de ella, le permitieran forjar aquellas estrategias comportamentales que le permitirían racionalmente hablando conseguir sus objetivos. Por otro lado, esa misma imagen del mundo generada a partir de las propias intuiciones racionales debe permitir al sujeto su conformación como una personalidad establecida y su desarrollo pro-activo que desemboque en una correcta interacción con su entorno, en una corriente amplia de acciones.

Con la llegada del conductismo, y un creciente interés por determinar de una manera científica sólo aquello que pudiera ser demostrado bajo la aplicación de los métodos experimentales de la ciencia física, tiende a abandonarse la posibilidad de una explicación dual de la naturaleza humana, cada vez más entendida a partir de aquello que puede ser aprehendido bajo el método de las ciencias. En este sentido, sólo se atiende a aquello que puede ser directamente observado por experimentación (la conducta) y se tiende a confeccionar explicaciones que sólo tengan en cuenta elementos explicables desde el punto de vista externo al sujeto, porque aquello interno (entendido como mental, si realmente existe) no puede ser captado por dicho método.

A la luz de estas nuevas perspectivas, no exentas de dificultades, nacen nuevos intentos de explicación. El materialismo de la identidad postula, precisamente, una identificación de todo estado tenido por mental con estados cerebrales del sujeto, ya sea en términos de casos (esto es, momentos concretos de una misma clase de estado) o en términos de tipo (esto es, estados que pueden ser entendidos como casos concretos de un mismo estado). A partir de aquí, surgen nuevos intentos de eliminar la mente: la mente puede ser explicada en términos puramente físicos, por lo que deberíamos concluir que todo aquello que existe son procesos físicos y no estados mentales, pudiéndose prescindir de ellos.

Con la aparición en la década de 1970 de la teoría funcionalista, y aunque los anteriores intentos no desaparecen, se pretende una explicación de la mente en términos funcionales. La mente, nos dicen, es un conjunto de procesos funcionales implementados en el cerebro (o en cualquier otro sistema) capaz de generar salidas adecuadas a las entradas sensoriales pertinentes. En este sentido, se reformula el inicial problema en términos representacionales, en tanto que el sistema en cuestión representa sus entradas sensoriales de modo tal que es capaz de conseguir la salida pertinente en cada caso. Precisamente, en la actualidad, esta teoría representacional se ha potenciado de muy diferentes maneras, hasta el punto de ser la escuela dominante en Filosofía de la Mente.

En las últimas décadas, el estudio de la mente ha tenido presente tres problemas básicos. En primer lugar, la intencionalidad presente en todo estado mental (esto es, la capacidad de todo estado mental de ser respecto de algo, de estar dirigido hacia algo externo). La explicación que

actualmente es habitual encontrar de esta peculiar característica suele darse en términos naturalistas: la explicación más plausible de esta intencionalidad puede darse dentro de los estudios científicos, en tanto que esta intencionalidad puede explicarse como un hecho natural (como lo es la gravedad). En segundo lugar, encontramos el tema del contenido mental. Suele entenderse que el contenido de la mente se da en un nivel representacional que, de modo directo o indirecto, está relacionado con la presencia de algún tipo de realidad externa al propio sujeto. Dicho nivel representacional permite la aparición de una imagen general del mundo entorno que, de un modo u otro, permite al sujeto su correcto desenvolvimiento y desarrollo a partir de la manipulación de dichas representaciones para obtener la mejor respuesta en cada momento. En tercer lugar, y no menos importante, encontramos la problemática relativa a la conciencia fenoménica (o del carácter puramente cualitativo de ciertos estados mentales). Ésta se refiere directamente a la peculiar característica de cierto tipo de estado mental que, parece (al menos en principio), no puede dirigirse directamente sobre nada externo al propio sujeto y que, a diferencia de otro tipo de contenido, no tiene un carácter objetivo, sino que presenta un marcado carácter de subjetividad, en tanto que se refiere al particular modo en que los sujetos experimentamos (o sentimos) determinados estados.

Aunque dentro del estudio de la mente, en este sentido analítico expuesto, habitualmente se ha dejado de lado esta tercera problemática, por ser excesivamente difícil, varios autores han centrado sus intereses en ella en los últimos años. Es el caso de Ned Block o de David Chalmers, que identifican una serie de problemas fáciles (de una manera relativa, claro) relativos a la explicación de ciertas dificultades que podemos encontrar en el estudio de la conciencia, como el peculiar problema de tener que explicar dicho carácter subjetivo a partir de métodos objetivos, y que están directamente vinculados a un problema difícil (el clásico problema mente-cuerpo), reformulado ahora en dos vertientes: una primera vertiente ontológica (cuál es la naturaleza de dichas sustancias y cuál es la relación establecida entre ellas) y una segunda vertiente epistemológica (cómo explicar la subjetividad presente en determinados estados a partir de nuestro método descriptivo científico, puramente objetivo). Ha llegado a tal extremo dicha pertinencia que se ha establecido un socorrido lema para designar a esa peculiar contraposición entre el carácter subjetivo de los contenidos cualitativos de la mente y la posibilidad de una explicación científica de los mismos: parece que existe un ‘hiato explicativo’ que impediría dicha descripción científica, en tanto que una visión reductiva aportada por el método objetivo de la ciencia siempre dejaría fuera la explicación del carácter marcadamente subjetivo que dicho contenido fenoménico parece tener.

Diferentes estudios de corte materialista, como los propuestos por John Searle, Daniel C. Dennett o Michael Tye, deben ser tenidos en cuenta. Por dos motivos principales: uno, por su particular interés en dilucidar el modo en que dicha subjetividad puede ser explicada de un modo objetivo y, otro, por el empleo de determinados resultados aportados por las ciencias neurofisiológicas que permitirían (al menos en principio) rendir cuentas de dicha subjetividad desde el punto de vista científico, sin necesidad de llegar a ser reductivos (como sería el caso de Searle o Tye) o, incluso, eliminando dicha subjetividad (el caso de Dennett).

En nuestro país, en los últimos años, se han sucedido varias publicaciones relativas al problema general planteado por el contenido mental. Pero, al menos que yo conozca, no ha aparecido ninguna (al menos en castellano) específicamente destinada al esclarecimiento concreto de la problemática relativa a la explicación del carácter fenoménico (subjettiva y cualitativamente dependiente) que muchos de nuestros estados mentales poseen.

Dentro de los más relevantes de estos estudios podríamos mencionar dos volúmenes colectivos.

1. *Pensando la mente. Perspectiva en filosofía y psicología*, editado por P. Chacón y M. Rodríguez para Biblioteca Nueva en 2001, y que recoge artículos de varios especialistas respecto del contenido mental (García-Carpintero, Liz, Moya, Rivière, etc.).

2. *La mente humana*, editado por F. Broncano para Trotta en 1995, y que recoge estudios generales acerca de varios aspectos relacionados directamente con la Filosofía de la Mente preparados por especialistas en el tema (García-Carpintero, Liz, Acero, Gomila, García Suárez, Villanueva, etc.).

También recientemente han aparecido varios manuales relativos a la temática general de la Filosofía de la Mente, pero todos ellos dejan de lado (o lo tratan muy poco, dado que habitualmente se considera como un problema mayor, al que debería dedicársele una atención especial, pero extremadamente dificultosa) la cuestión del carácter cualitativo de nuestros estados mentales. Cabría destacar, por su pertinencia para una introducción general a nuestro tema, los volúmenes:

1. *Perspectivas actuales en filosofía de la mente*, confeccionado por Manuel Liz y editado en 2001 por la Consejería de Educación del Gobierno de Canarias. El volumen trata temas actuales directamente relacionados con aspectos del contenido mental y realiza un repaso general sobre las aportaciones particulares realizadas en los últimos años por filósofos internacionales de renombre, como John Searle, Daniel C. Dennett o Patricia y Paul Churchland.

2. *Filosofía de la mente*, de Carlos Moya fue editado en 2004 por el Servei de Publicacions de la Universitat de València (original en catalán, pero también disponible en castellano). Recoge en sus varios capítulos las diferentes respuestas que desde las escuelas más relevantes del siglo XX se han venido aportando al estudio de la mente desde la filosofía analítica (desde el materialismo de la identidad al materialismo eliminador, pasando por el funcionalismo), prestando un interés especial al problema del vínculo entre mente y acción.

3. *Filosofía de la Mente y de la Ciencia Cognitiva*, de José Hierro-Pescador, publicado en 2005 por Akal, repasa también las diferentes posiciones que la Filosofía Analítica ha tenido a lo largo del siglo XX respecto de la problemática de la mente.

A pesar de ser escasos los trabajos relativos a la problemática específica sobre el contenido cualitativo y subjettivo de la conciencia, el carácter fenoménico de nuestros estados mentales, en los últimos años han aparecido algunos estudios que, de algún u otro modo, tienen una relación directa con algunos aspectos del tema de la conciencia. Será el caso de los siguientes volúmenes:

1. *Qué es una emoción*, de David Casacuberta, editado en 2000 por Crítica. En este trabajo el autor nos presenta una introducción analítica al problema concreto del contenido de las emociones (uno de los estados fenoménicos que podemos encontrar asociados a la filosofía de la mente). Presenta una revisión, partiendo de las propuestas antiguas hasta desembocar en algunas propuestas contemporáneas, de las diferentes teorías que sobre la emoción han ido sucediéndose a lo largo de la historia de la filosofía, y concluye con una visión propia (de corte tele-funcionalista) que pretende una respuesta a la hora de definir qué sea una emoción.

2. *Mente y mundo. Aproximación neurológica*, de Juan Vázquez, editado en 2007 por Akal, pretende una presentación del carácter subjetivo de los estados perceptivos a partir del análisis de varios resultados obtenidos por las neurociencias. El problema de este volumen es que, haciendo una mala interpretación de dichos resultados, tomados de manera parcial y totalmente partidista, presenta una defensa de un cierto dualismo que relega nuestros estados fenoménicos de percepción a la posesión de una sustancia mental (a modo de alma) difícil de defender actualmente. Una ampliación de esta crítica puede verse en Colomina (2008d).

3. *La conciencia fenoménica y el naturalismo contemporáneo* es el nombre del monográfico que presenta la revista *Teorema* en su volumen XXVII, Número 3, de 2008. Es el único volumen del que tengo constancia que, en lengua castellana, recoge artículos directamente relacionados con el tema que pretendo defender en mi tesis. Recopila trabajos de especialistas internacionales respecto de los temas más relevantes sobre la filosofía de la conciencia.

También, en los últimos años, apareció una tesis doctoral en la Universidad de La Laguna (también bajo la dirección de Manuel Liz) que trataba, al menos de una manera indirecta, un tema crucial a la hora de enfrentarnos con el contenido de nuestros estados fenoménicos: la distinción entre contenido conceptual y no-conceptual. La tesis lleva por título *Psicología Popular: simulación vs. Teoría de la teoría* y fue presentada en 2005 por Tamara Ojeda.

3. Objetivos generales

Como hemos comentado en el anterior punto, la idea general de la tesis es presentar una crítica a las posiciones actualmente dominantes en el estudio de la filosofía analítica de la conciencia, entendiendo por conciencia precisamente aquellos estados de la mente caracterizados de un modo particular por su carácter fenoménico. Esto es, entendiendo por conciencia aquel conjunto de estados que, de uno u otro modo, cumplen dos condiciones clave a la hora de aparecer a nuestra mente: en primer lugar, disponen de un marcado carácter subjetivo (es decir, a cada individuo concreto se le presenta de una particular manera las sensaciones que tiene, permitiendo así establecer un idiosincrásico modo según el cual experimenta dichas sensaciones concretas) y, en segundo lugar, dichas sensaciones son entendidas por el sujeto concreto de un modo cualitativamente relevante según el caso (es decir, las sensaciones percibidas contienen un cierto carácter cualitativo en tanto que resaltan características concretas en cada momento según la perspectiva subjetiva de cada sujeto).

Perseguimos precisamente una revisión crítica de dichas teorías representacionales porque adolecen de varios defectos que, de un modo claro, imposibilitan una aproximación certera y acertada a la noción de conciencia y que, por tanto, imposibilitan una explicación correcta y coherente de la misma. El error más habitual que encontramos en los análisis anteriormente expuestos es, precisamente, la asimilación (e, incluso, la confusión entre dos tipos diferenciables de conciencia). Una primera conciencia entendida como puramente fenomenológica (claramente fenoménica y cualitativamente dependiente que, sin embargo, dispondría de poderes causales sobre el sujeto consciente) y una conciencia por acceso (claramente representacional y funcional que, sin embargo, influiría claramente en el papel comportamental del sujeto). Esta asimilación haría incurrir demasiado habitualmente al representacionalismo en la ilusión de que todo estado de conciencia es causal (en tanto que mezclaría causalidad fenomenológica y funcionalismo representacional), haciendo entrar en una explicación puramente comportamental todo dato descriptivo de la conciencia misma. Pero conocemos estados que, a pesar de ser claramente conscientes, no parecen tener un papel marcadamente causal en la vida mental del sujeto. Tan sólo debemos pensar, por ejemplo, en el estado de nerviosismo y ansiedad que despierta una primera cita con la persona amada o el estado de desasosiego y desánimo que puede producir un imprevisto en nuestro trabajo. Dichos estados, a pesar de que pueden ser conscientes, no tienen una influencia decididamente causal en nuestra vida cotidiana pero, pretendemos sostener, sin embargo, si pueden tener un papel puramente fenoménico en nuestra vida mental, en tanto que es inevitable que inunden (en la literatura actual suele emplear el término ‘coloreen’) el resto de nuestros estados mentales, ya sean representacionales (como el resto de estados de conciencia) o proposicionales (como las creencias o los deseos).

Teniendo presente precisamente esta diferenciación, pretendemos defender una alternativa a las teorías representacionales de la conciencia en tanto que éstas tan sólo resaltan el papel funcional que sobre el comportamiento tienen nuestros estados mentales, intentando evitar toda explicación de los estados fenoménicos como puramente fenomenológicos en tanto que irrelevantes a la hora de explicar nuestra conducta, como si tan sólo fuéramos receptáculos pasivos y pasivos de aquellos estados conscientes (percepciones, emociones o estados de ánimos) que nos inundan y acontecen en nuestra mente. Pretendemos reivindicar una contrapartida activa en dicha posesión de estados mentales con carácter fenoménico. Defenderemos la posibilidad de que nuestros estados fenoménicos de conciencia, aunque relegados a un papel causal menor, realmente tengan una capacidad efectiva de influir en nuestras decisiones cotidianas y de influir marcadamente en nuestras acciones.

La idea principal, sobre la que pivotará todo nuestro trabajo, es el reconocimiento de un componente activo en la conciencia, y no sólo de un componente pasivo. Es decir, estamos dispuestos a atribuir a la conciencia un papel causal, aunque a menudo es difícil, como veremos, que podamos identificarlo con precisión. Ello nos permitiría poder salvaguardar un papel activo en la conciencia a la hora de explicar las estrategias comportamentales de los sujetos conscientes, un papel activo que el representacionalismo no suele reconocer.

Por supuesto, ello nos obligará a tener en cuenta las amplias discusiones que sobre nuestro tema actualmente están llevándose a cabo por excelentes especialistas de talla internacional y atender concienzudamente a las disputas en los foros especializados, incluso atendiendo a desarrollos interdisciplinarios que se ofrecen desde disciplinas distintas, aunque afines a la Filosofía, como la Inteligencia Artificial o la Neurofisiología.

Estamos obligados a comenzar por una contextualización de nuestra problemática noción, haciendo especial hincapié en el constante cambio de sentido que, a lo largo de la historia del pensamiento, ha ido sufriendo hasta llegar, de la mano de la reformulación cartesiana, al calado que actualmente ha adquirido. Nos centraremos especialmente en la revisión que desde mediados del siglo XX viene haciéndose de ella, precisando su lugar dentro del estudio general sobre la mente y, especialmente, diferenciando los tratamientos distintos que ha recibido según la escuela concreta que de la misma se ha ocupado, desde el funcionalismo al representacionalismo. A ello dedicaremos un primer capítulo introductorio, en el que además atenderemos a la problemática del llamado ‘problema difícil’ de la conciencia (la vinculación de la conciencia a una serie de elementos físicos, como derivado del clásico problema mente-cuerpo) y a la identificación que actualmente se realiza de varios ‘problemas fáciles’ vinculados de algún u otro modo con cuestiones de investigación neurológica.

Posteriormente, nos veremos obligados a matizar los diferentes sentidos que actualmente se tienen de conciencia para, precisamente, centrar el tipo concreto de definición que una teoría representacional suele tener de la misma. Pero también estamos obligados a diferenciar entre distintos tipos de representación, con el objeto de identificar aquellos acertados de aquellos no tan acertados (aunque todos sean relevantes, sobre todo por los ríos de tinta que han sido capaces de derramar). Dentro de esta diferenciación, también deberemos atender a las diferentes posiciones que, tanto dentro como fuera del movimiento representacional, se han sucedido en las últimas décadas. Todo ello con el fin último de precisar en la medida de lo posible la noción concreta de conciencia que cada planteamiento (y cada autor, incluso) está manejando y qué tipo de contenido representacional está dispuesto a otorgar a la misma.

Ello nos obligará a detenernos en el estudio de tres autores representacionistas que, a mi entender, representan cada una de las tres diferentes posiciones extremas que pueden adoptarse dentro de esta corriente. John Searle es relevante por ser el ferviente defensor del naturalismo biológico, un epígrafe que esconde tras de sí un representacionalismo de corte dualista que pretende explicar la conciencia como un producto biológico excretado por el cerebro (en un sentido epistemológico) pero que, sin embargo, considera irreducible (desde el plano metafísico) la subjetividad de la conciencia. Michael Tye se erige como el mayor defensor del intencionalismo representacional, una teoría que defiende la explicación objetiva de la conciencia a partir del análisis concreto de dificultades particulares del contenido fenoménico que, sin embargo, salvaguarda su carácter subjetivo a partir de la apelación a contenidos representacionales que todo sujeto es capaz de captar del exterior. En el caso de Daniel C. Dennett, que pretende una eliminación de todo contenido fenoménico, sin embargo, encontramos un débil carácter representacional en todo estado de conciencia, en tanto que lo

identificaría con una ilusión o ficción útil que el sujeto es capaz de fabricar a partir de inferencias realizadas sobre los datos captados por sus sentidos.

Todas estas críticas puntuales nos llevarán a un conjunto homogéneo de datos que nos permitirán defender una noción de conciencia no sólo representacional que, como hemos indicado antes, permitan indicar un carácter causal en nuestros estados fenoménicos de conciencia que sean capaces no sólo de (de)mostrar nuestra pasividad ante las sensaciones que nos acontecen sino, también, inferir un papel totalmente activo de los mismos en la formación de estrategias de acción, en tanto que consiguen colorear por completo todo nuestro mundo mental.

4. Plan general de trabajo

La tesis se compone de un prólogo, seis capítulos centrales y una conclusión.

En el prólogo hemos intentado contextualizar el tema concreto a tratar, la conciencia, desde la aparición en los escritos antiguos como vinculada directamente con la noción de alma hasta su actual reformulación en términos funcionales y representacionales.

En el primer capítulo, se presenta el tema de la conciencia como un tema interno al estudio global de la mente. Tras un recorrido de la actual situación dentro de la Filosofía Analítica de la Mente, se atiende a las diferentes interpretaciones que (aunque suelen tratarla de modo secundario) las diferentes escuelas del siglo XX han realizado de la noción de conciencia. Posteriormente, se pasa a presentar una distinción entre contenido conceptual y contenido no-conceptual dentro de los estados de conciencia, en tanto que muestra la posibilidad de que todo estado consciente pueda ser entendido como resultado de una mediación o como una propiedad capaz de ser captada sin necesidad de apelar a un contenido proposicional o vinculado semánticamente. Por último, atenderemos a los diferentes problemas que podemos identificar dentro de la filosofía de la conciencia: un problema difícil (derivado del clásico problema Mente-Cuerpo) y una lista de 15 problemas secundarios (o fáciles) que serán analizados en detalle a lo largo de la tesis, que van desde la dificultad de atender al carácter subjetivo de la conciencia desde un método objetivo como es el científico hasta la dificultad que nos muestran ciertas lesiones cerebrales a la hora de centrar el carácter cualitativo de las percepciones concretas.

En el segundo capítulo, intentamos un doble objetivo. Por un lado, centrar la noción de conciencia que actualmente es relevante en los numerosos estudios que al respecto encontramos en la literatura científica y, por otro, centrar los diferentes tipos de representacionalismo (y de representación) que actualmente están en boga en el panorama intelectual. Para ello realizamos una contraposición entre teorías representacionales y no-representacionales con la intención de mostrar sus diferencias y con la intención de salvaguardar de cada una de ellas aquellos aspectos que pueden ser utilizados y, mediante una feroz crítica, intentar deslegitimar los intentos de una explicación puramente representacional del estudio de la conciencia. Por supuesto, también presentamos, a partir de una contraposición entre ellas, las diferentes posiciones representacionistas actualmente relevantes, con el mismo fin: salvar aquello que puede ser

provechoso para nuestro estudio y deshacernos de aquellos equívocos que sólo pueden empañar nuestras soluciones.

Los capítulos 3, 4 y 5 presentan pormenorizadamente las posiciones de los representantes de las tres posiciones representacionistas que, a mi juicio, son más relevantes. El motivo principal es que los tres pretenden asumir un punto de partida no dualista y sensible a la ciencia física (incluyendo la biología evolucionista), aunque no siempre consiguen mantener ese objetivo inicial (como será el caso de Searle). Los tres pretenden también ser realistas respecto a la conciencia, aunque (como veremos) dentro de un espectro en el que Dennett estaría muy cercano al eliminativismo y Searle muy cercano al misterianismo.

El capítulo 3 analiza el naturalismo biológico de John Searle. A partir de la presentación, en un primer momento, de los puntos fuertes de su teoría, intentamos desmontar los pilares sobre los que se edifica este naturalismo biológico de Searle en tanto que, creemos, comete dos errores fundamentales. En un primer momento, su intento de naturalizar la conciencia es ficticio, en tanto que su verdadero interés está supeditado a salvaguardar la subjetividad presente en los estados mentales. En un segundo momento, bajo la gran cantidad de argumentos retóricos que incorpora a su discurso (del que destacamos el Trasfondo y el Principio de Conexión entre lo consciente y lo inconsciente) realmente podemos concluir que su interés principal es defender un dualismo conceptual que, aunque intente reducir epistemológicamente lo mental a lo físico, salvaguarde metafísicamente la irreductibilidad de la conciencia.

En el capítulo 4 analizamos el representacionalismo fuerte de Michael Tye. Dicho representacionalismo estaría interesado en presentar todo contenido de las experiencias fenoménicas (en general) como teniendo estrechas relaciones con el aparato nervioso y neuronal (según muestran los recientes estudios neurológicos), conteniendo tanto componentes sensitivos como afectivos (en un sentido evaluativo, pero carente de toda carga moral), permitiendo representar aquellas características relevantes de los objetos (o los objetos enteros) involucrados en las sensaciones y en las experiencias fenoménicas, y como no-conceptual. Pero la característica principal que Tye resalta del contenido de una experiencia fenoménica es que ésta siempre es transparente: esto es, que toda experiencia fenoménica está presente de modo inmediato en nuestras vivencias aunque no existiría un lugar determinado hacia el que se dirija nuestra atención para sentir dicha experiencia (dado que se daría al unísono con todas las demás experiencias conscientes), lo que permitiría evaluar los objetos del mundo de modo introspectivo y no cabría el error a la hora de individualizar e identificar una determinada sensación como una sensación de algo en particular. Desde nuestro punto de vista, la exigencia de transparencia es demasiado estricta porque, aunque es cierto que a veces permitiría la explicación respecto de los errores de juicio, impide la posibilidad de explicar la noción empírica de error tanto en los estados proposicionales como en los fenoménicos por convertir a la vivencia de la experiencia en virtualmente infalible. Es decir, a diferencia de Tye (que aceptará un cierto millianismo a la hora de definir el contenido fenoménico, consideramos pertinente, al menos, tener en cuenta un cierto descriptivismo de carácter fregeano a la hora de

caracterizar e individualizar el contenido de una experiencia consciente. Y todo ello porque, además de parecerse una ilusión (por ser demasiado restrictivo) el requisito de transparencia, que hechos como los mencionados ocurran nos hace dudar de la posibilidad de que todo contenido fenoménico consciente, sea no-conceptual. Porque parece necesario que una explicación adecuada de las experiencias fenoménicas debe dar correcta cuenta del modo en que los seres conscientes nos comportamos en determinadas ocasiones, que esas reacciones naturales responden a nuestros intereses, y estos preceden a toda representación y que una interpretación estrictamente no-conceptual no parece satisfacer estos requisitos (al dejar fuera ciertos elementos conceptuales presentes en nuestra experiencia, siendo necesario apelar de tanto en tanto a ellos, como veremos en el capítulo final), por lo que se hace necesario un estudio de la experiencia fenoménica que respete la intencionalidad presente en la conciencia. Precisamente con la intención de mostrar esas carencias en la teoría de Tye, pretendemos realizar un análisis pormenorizado de la concepción representacional fuerte que nuestro autor actualmente defiende, comenzando por presentar una caracterización del modo en que debe ser entendido su representacionalismo. Seguidamente, una vez considerado el carácter que en la teoría de Tye adquieren las experiencias con contenido consciente, analizaremos su teoría PANIC, una teoría que persigue presentar todo Contenido de la conciencia como Preparado, Abstracto, No-conceptual e Intencional. Por último, nos detendremos en el análisis de los argumentos que Tye defiende para caracterizar el contenido de las experiencias corporales, las emociones y los estados de ánimo como contenidos PANIC, presentando así de manera completa su teoría de la conciencia y sus soluciones a determinados problemas relevantes en el estudio de la conciencia.

En el capítulo 5 analizaremos el representacionalismo ficcionalista de Daniel C. Dennett. Los estudios de Dennett, a pesar de sus fines puramente filosóficos, pueden encuadrarse dentro del campo de las ciencias cognitivas. Por su naturaleza multiforme, esta disciplina combina metodologías, teorías y conceptos pertenecientes a diversas esferas científicas, unos ámbitos no siempre fácilmente compatibles metodológicamente pero, nunca, excluyentes. En pocas palabras, el estudio de las tesis de Dennett obliga a que nuestra investigación sea interdisciplinar: nos involucra con terminología filosófica y científica y nos obliga a emplear diversas metodologías (biología, neurología, psicología, etc.) para hacer justicia a la profundidad de los pensamientos transmitidos en los escritos de nuestro autor, profundidad que se irá abriendo paso a medida que avancemos en nuestra investigación y vayamos enfrentándonos a los problemas y los conceptos. El problema principal con la teoría de Dennett lo encontramos precisamente en sus conclusiones. Aunque se empeña en demostrar desde la teoría evolutiva que la conciencia no es más que un producto biológico determinado a partir de las leyes del proceso de la selección natural y apuesta por una eliminación en términos neurobiológicos de los aspectos fenoménicos de nuestra vida mental, pretende salvaguardar también un aspecto representacional a partir de una noción de conciencia como ficción útil que lo hará recaer en los inconvenientes que identificaremos en todo planteamiento representacionalista en los anteriores capítulos.

Concluimos en el capítulo 6 presentando una alternativa a las teorías representacionales. A partir de nuestro objetivo general, intentaremos mostrar los aspectos de pasividad y de actividad que el contenido fenoménico tiene en los sujetos conscientes, defendiendo una teoría no sólo representacional de la conciencia basa en dos aspectos relevantes. En primer lugar, una defensa apoyada en los últimos descubrimientos neurobiológicos de los mecanismos cerebrales que permiten una concepción materialista de la conciencia, como es la nuestra. En segundo lugar, un intento de solución desde dicha posición a los problemas iniciales (indicados en el capítulo 1) identificados en el estudio de la conciencia.

El apartado de conclusiones generales recoge sumariamente los objetivos principales de los capítulos anteriores, mostrando los resultados obtenidos.

CAPÍTULO 1. INTRODUCCIÓN: UNA PERSPECTIVA SOBRE LOS ESTUDIOS ACTUALES EN FILOSOFÍA DE LA CONCIENCIA

Todo se refugia en la mente. El 'mundo' para ser y reconocerse siempre tan poco; el Ser para conocerse, comunicarse y complicarse. El cerebro humano es un lugar donde el mundo se pincha y se pellizca para asegurarse de que existe. El hombre piensa, luego yo existo, dice el Universo.
(Paul Valéry, *Estudios filosóficos*)

Nadie tiene ni la más remota idea de qué es la conciencia, o de por qué existe, o qué hace posible su existencia (por no decir de qué está hecha).
(Jerry Fodor, "You can't argue with a Novel")

I. INTRODUCCIÓN: UNA PERSPECTIVA SOBRE LOS ESTUDIOS ACTUALES EN FILOSOFÍA DE LA CONCIENCIA

Imaginemos por un momento que tenemos un fuerte dolor en la espalda. Quién nos mandaría a nuestra edad habernos ‘pegado esa paliza’ con la bicicleta esta mañana. El dolor es insoportable. Al llegar a casa tenemos que tumbarnos en el suelo para intentar mitigarlo, pero a pesar de ello, todavía sentimos ese dolor. Después de la ducha, nos aplicamos una pomada calmante. Los efectos anti-inflamatorios de dicho producto poco a poco comienzan a demostrar su eficacia y los productos analgésicos que contiene comienzan a hacer más tolerable el dolor. ¿Qué ha ocurrido? ¿Hemos dejado de tener el dolor de espalda? La respuesta más plausible es que el daño todavía persiste, sólo que los productos químicos que componen el ungüento que acabamos de aplicarnos en la zona afectada han reducido nuestra sensación consciente respecto del dolor de nuestra espalda.

Pensemos en un ejemplo análogo. Hemos ido al dentista para que nos extraiga una muela del juicio que nos está dando unas más que considerables molestias. El dolor que sentimos, incluso al masticar, dificultan nuestras labores cotidianas. Cuando nos encontramos sentados en la silla del dentista, con la boca abierta, nos aplican anestesia local. Poco a poco el dolor de muela se mitiga. ¿Ha desaparecido nuestro dolor? Todo nos indica que el daño está todavía presente, que debe persistir, dado que la muela sigue en su sitio y la situación no ha cambiado tan profundamente como para que el daño haya desaparecido. Pero ya no sentimos el dolor. La anestesia local, como los productos analgésicos de nuestra pomada anti-inflamatoria, han removido momentáneamente nuestra experiencia consciente respecto del dolor.

El mundo cambia radicalmente si dejamos de percibir de manera consciente. Imaginemos por un momento cómo cambia nuestro mundo si cerramos los ojos. Con los ojos abiertos, las experiencias conscientes nos inundan. No podemos dejar de percibir colores, formas, tamaños, etc. Pero cuando cerramos los ojos, esas experiencias también desaparecen. Dejamos de tener experiencias visuales conscientes. ¿Dónde se han ido esas experiencias? ¿Han desaparecido? Como en los dos ejemplos anteriores, la respuesta más intuitivamente acertada parece ser que ha desaparecido nuestra experiencia consciente. En el caso de tener los ojos cerrados, dejamos de tener experiencia visual consciente porque, obviamente, nuestro sentido de la vista se ha encontrado con un impedimento físico: tener los ojos cerrados. Tan sólo debemos volver a abrir los ojos para recuperar nuestra visión normal consciente de las cosas que pueblan el mundo. En los ejemplos del dolor, como hemos dicho, el daño no desaparece. Lo que desaparece es nuestro acceso consciente a la experiencia de dolor porque, cuando aplicamos la pomada a la lesión ocasionada en nuestro cuerpo, desaparece la posibilidad de *sentir* dolor. Como bien nos ha demostrado la experiencia, cuando los efectos de la anestesia pasan, el dolor de muelas se mitiga porque ya nos han extraído esa dichosa muela que tanto fastidio nos estaba provocando (aunque ahora sentimos el punzante dolor que nos está produciendo ese vacío rellenado ahora por una pequeña herida en nuestra boca taponada por un par de puntos quirúrgicos). Cuando los efectos analgésicos de la pomada desaparecen, el dolor de nuestra

espalda aflora de nuevo y es que ¡quién nos mandaría a nosotros haber cogido con tanta fuerza la bicicleta esta mañana!

Ejemplos de este tipo pueblan todo nuestro mundo consciente. El principal objetivo de este trabajo es identificar una serie de dificultades presentes en los análisis que respecto de la conciencia fenoménica podríamos considerar en las teorías representacionales que, actualmente, dominan los estudios en filosofía de la mente. Ello nos permitiría, al menos, encontrar los problemas que toda teoría debería poder superar a la hora de intentar explicar la conciencia, lo que nos guiará a la hora de ofrecer nuestra propia respuesta a qué es la conciencia y qué permite que nuestra subjetividad esté poblada de experiencias conscientes similares. Para ello creemos necesario comenzar ofreciendo una breve visión de los anteriores intentos de aproximación a dicho concepto. Para poder llegar a definirla, deberemos primero evaluar los intentos de explicación que en los últimos años han ido apareciendo en la literatura especializada acerca de la conciencia, pero también deberemos intentar evaluar los diferentes problemas que han ido apareciendo a partir de dichos intentos.

En este capítulo, comenzaremos ofreciendo un primer apartado que incluya una presentación de las características que presenta el contenido que podemos encontrar en los estados mentales con carácter fenoménico, un contenido diferente de aquel que podemos encontrar en las actitudes proposicionales (y otros estados mentales afines), caracterizable como puramente conceptual. Precisamente por esa peculiaridad del contenido proposicional, en la actualidad ha aparecido una dicotomía entre el contenido que podemos denominar conceptual y aquel contenido definible como no-conceptual. Dedicaremos el segundo apartado a establecer una primera diferenciación entre ambos, intentando en todo momento definir ambas nociones y viendo en qué sentido se emplean en los actuales estudios dentro del campo de la filosofía de la mente. En el tercer apartado pretendemos mostrar una retrospectiva que analice, de modo general, las diferentes aproximaciones al clásico problema mente-cuerpo. De dicho problema desprenderemos, de modo más particularizado, un problema concreto que se centra en exclusiva a la diferenciación existente (y la dificultad de equiparar) entre los estados fenoménicos con estados físicos: el conocido como problema *difícil* de la conciencia. Desde este punto de vista, existiría una imposibilidad, determinada por un ‘hiato explicativo’, de reducir toda experiencia (subjetiva y cualitativamente sentida por un sujeto) a estados neurofisiológicos, dada la diferencia clara entre las características objetivas de los segundos, que entrarían en clara contradicción con las propiedades subjetivas que presenta toda sensación. Al atender precisamente a los intentos de explicación de dicho problema fuerte, surgen una gran variedad de problemas sencillos, *fáciles*, en tanto que se centran de un modo particular en cuestiones más o menos concretas dentro de la multiplicidad de la experiencia consciente. Intentamos presentar un elenco más o menos exhaustivo de dichos problemas actuales de la conciencia en el apartado cuarto, no con la intención de agotar el tema, sino más bien con la de presentar una variada perspectiva que incluya, por un lado, la mayor cantidad posible de ejemplos y contraejemplos explicativos que nos permita dotarnos del material suficiente para poder afrontarlos y, por otro, para intentar presentar lo más fielmente posible la amplitud del tema que nos atañe.

Comencemos, pues, con una primera presentación del contenido fenoménico que podemos encontrar en determinados estados mentales.

1.1. La peculiaridad del contenido fenoménico

Es comúnmente aceptada la idea de la naturaleza relacional de los estados mentales. Es decir, los estados mentales suelen ser considerados como las relaciones existentes entre agentes psicológicamente dotados y sus contenidos. Esta relación es claramente patente si atendemos a actitudes proposicionales como las creencias (podemos explicar la creencia de David de tener un buen trabajo como la relación existente entre David como sujeto mental consciente y el contenido proposicional ‘tener un buen trabajo’), los deseos (podemos explicar el deseo de Jordi de tomar una cerveza como la relación existente entre Jordi como sujeto mental consciente y el contenido proposicional ‘tomar una cerveza’) o las percepciones (podemos explicar la percepción de Manolo de un tomate rojo como la relación existente entre Manolo como sujeto mental consciente y el contenido proposicional ‘(en un determinado lugar) hay un tomate rojo’).

Podemos caracterizar los contenidos de dichos estados mentales como elementos marcadamente particulares porque:

1. el hecho de que alguien posea un estado mental característico con un determinado contenido tan sólo demuestra la existencia de ese alguien y no indica nada respecto de la existencia de lo que es expresado por dicho contenido (es decir, un individuo puede tener perfectamente una percepción respecto de algo que es irreal, que es ilusorio),

2. el hecho de que alguien posea un determinado estado mental con un contenido característico no nos dice nada respecto de la verdad o falsedad de dicho contenido (es decir, alguien puede poseer una creencia y demostrarse que es falsa),

3. el hecho de que alguien posea un determinado estado mental con un cierto contenido parece depender por entero de las condiciones externas que acompañarían a la perspectiva subjetiva del agente (es decir, alguien puede querer tomar una cerveza y no poder hacerlo porque su nevera está vacía, dejando insatisfecho su deseo) y, además,

4. el hecho de que alguien posea un determinado estado mental con un contenido determinado ejerce un poder causal sobre la acción del sujeto (es decir, tener un determinado deseo D incita al sujeto a cumplirlo y, si se tiene la creencia de que haciendo A se puede cumplir D, es muy probable que el sujeto acabe haciendo A).

Pero también es cierto que, además de ocupar un lugar dentro de la explicación de la conducta, los contenidos mentales tienen también un carácter marcadamente normativo. Si los contenidos mentales tienen un papel causal respecto de la acción consciente (o inconsciente) deben poseer también un carácter de explicación general. Pero si, como hemos dicho, los contenidos mentales poseen unas condiciones de verdad que permiten indicar cuándo se satisface o no se satisface un determinado contenido mental según su ajuste a las circunstancias, entonces nos encontramos con un importante problema a la hora de establecer leyes generales de explicación de cómo lo mental tiene una relación causal con lo físico.

Es cuando surge esta necesidad de establecer leyes naturales que expliquen esta relación entre estado mental y contenido cuando nos encontramos con el programa de naturalización del contenido, empeñado en demostrar la existencia de una serie de estados físico-corporales que permitirían determinar el estado mental en el que se encuentra el sujeto en cada momento. Para esta concepción, la presencia de determinados estados mentales es idéntica a la posesión de determinados estados físicos (cerebrales y funcionales). El contenido de dichos estados físicos permitiría determinar de un modo suficiente el papel causal de dicho estado mental. Porque dicho programa de naturalización afirma que el significado de dichos estados se estipula a partir de una serie de representaciones de ciertas características físicas, demostrando así la vinculación natural entre la mente y la intencionalidad. El problema con el naturalismo de lo mental lo encontramos cuando atendemos a la posibilidad de tener que rendir cuentas de un estado mental con un contenido erróneo (como en el caso de la creencia falsa, el deseo insatisfecho o la percepción ilusoria): ¿qué ocurre entonces? ¿Cómo explicar la posibilidad de la existencia de actitudes proposicionales erróneas si vinculamos la posesión de un estado mental con la representación de un determinado estado físico?

Habría dos modos diferentes de afrontar este problema. Por un lado, encontramos a aquellos que consideran que todo estado mental es esencialmente diferente de un estado físico y que el elemento relacional de la mente con su contenido depende de propiedades intrínsecas a la mente que constitutivamente se dan de un modo totalmente independiente de las condiciones externas, aunque estas condiciones podrían causarlo. Por otro lado, encontramos a aquellos que consideran que las condiciones de verdad del contenido mental no vienen dadas por la restrictividad de las propiedades intrínsecas de los sujetos sino que vienen establecidas por una serie de condiciones externas al mismo, por lo que optan por explicar dicho contenido como contextualmente dependiente de los hechos externos a la mente.

Pero todavía quedaría un ulterior punto a tratar respecto de la naturaleza del contenido mental. Cuando atendemos a actitudes proposicionales, como mi creencia en que mañana a las 22:00 horas actúa *After Forever*, mi grupo musical favorito, o mi deseo de beber una cerveza, parece que el contenido de dichos estados mentales puede ser expresado a través de elementos conceptuales. Esto es, entrarían dentro de las categorías culturalmente dependientes que vehiculan nuestros significados y que, a menudo, requieren un conocimiento previo de aquello a lo que se refieren. Pero parece que cuando atendemos a ciertos contenidos mentales fenoménicos (esto es, contenidos ya no tan claramente definidos como conceptualmente dependientes sino vinculados directamente con nuestra peculiar distribución subjetiva, con cómo nos aparecen a nosotros las cosas) como mi sensación de dolor o mi orgasmo (y otros tipos de estados conscientes caracterizados por la *vivencia* o toma de conciencia [*awareness*] y otros casos-límite, como mi ver el color rojo), el modo en que éstos se relacionan con sus respectivos estados mentales parece cambiar con respecto a las actitudes proposicionales. El motivo principal es que, en las segundas, la relación con un elemento objetivo parece estar garantizada (aunque sabemos que, a menudo, dicha relación puede estar equivocada, como hemos comentado) mientras que aquel elemento proposicional con el cual se relacionarían los

estados mentales fenoménicos parece ser más difícil de precisar o, incluso, como algunos pensadores actualmente sostienen (llegando a ser considerada la posición dominante), puede llegar a ser totalmente inexistente. De este modo, ni tan siquiera podríamos estar seguros del todo de que existiera una relación proposicional. ¿Cómo debemos, entonces, tratar esa relación?¹

1.2. La distinción entre contenido conceptual y contenido no-conceptual

Como hemos comentado en el apartado anterior, el contenido fenoménico tiene un carácter marcadamente particular. Precisamente en este apartado pretendemos ofrecer una presentación de las dos clases de contenido que, habitualmente, se considera cuando se habla de estados mentales pero, por supuesto, en tanto que aplicado particularmente a aquellos estados con carácter fenoménico.

Parece que, en última instancia, el contenido fenoménico no puede ser reducido a contenido proposicional porque, si así fuera, como sostienen sus detractores, se perdería la riqueza experiencial de los sujetos. Es decir, se perdería esa gran cantidad de experiencias individuales e individualizadas que todo sujeto tiene cuando se relaciona con su mundo circundante. En este sentido, si consideramos que todo contenido mental, incluido el experiencial-fenoménico, puede ser tratado como conceptual (es decir, dependiente de la categorización objetiva y objetivante vinculada lingüísticamente a una cultura o a una serie de categorías universales identificadas de modo previo), parece que la gran cantidad de contenido experiencial subjetivamente dependiente que todo individuo es capaz de percibir se ve menoscabada y limitada a su exposición bajo categorías externas.

Con vistas a salvaguardar precisamente esta riqueza experiencial presente en la vida mental de todo sujeto consciente nacieron las teorías del contenido no-conceptual (Evans 1982; Cussins 1990; Peacocke 1992; Bermúdez 1998). Desde esta perspectiva, el contenido no-conceptual (CNC) se define, de modo general, como el contenido mental que puede ser adscrito a todo individuo consciente sin necesidad de tener que adscribirle, también, todo el aparato conceptual requerido para especificar de modo determinado dicho contenido (Cussins 1990: 370). Esto es, en este sentido, el CNC puede establecerse como toda aquella organización que a nivel sub-doxástico se requiere para poseer un determinado estado mental sin necesidad de presuponer la determinación doxástica de dicho estado.²

¹ En las líneas anteriores hemos defendido el carácter peculiar que el contenido fenoménico tiene respecto de las actitudes proposicionales. Pero no queremos dejar de constatar que las actitudes proposicionales también tienen sus propias características y problemas definitorios. De todos modos, si una cosa queda clara a partir del análisis pormenorizado de dichos problemas y de dichas características, un análisis que no vamos a realizar aquí por desviarse sustancialmente de nuestros propósitos, es que el contenido semántico de las actitudes proposicionales nos aporta una perspectiva fiel de los avances y logros de los actuales desarrollos de la filosofía del lenguaje. Para una panorámica de la relación entre las actitudes proposicionales y su contenido puede verse con provecho Acero, Frápolli y Romero 1997 y 1998.

² La distinción entre estados doxásticos y estados subdoxásticos es establecida para dar cuenta, por una parte, de todos aquellos estados que, de uno u otro modo, tienen algún elemento intencional (como las creencias, los deseos, etc.) a un nivel personal y, por otra parte, todos aquellos estados que desempeñan un papel causal en la historia cognitiva del sujeto sin llegar a tener un elemento intencional como las actitudes proposicionales (Stich 1979: 499).

Aunque en un primer momento la teoría del CNC nació con la intención de servir como base de fundamentación del contenido conceptual (Evans 1982), en los últimos años ha evolucionado para convertirse en una herramienta explicativa capaz de dar cuenta de, por un lado, cuándo puede decirse que un individuo consciente posee un determinado concepto y, por otro lado, cuándo una criatura consciente es capaz de adquirir determinadas habilidades conceptuales. Aunque a primera vista pueda parecer que esta teoría del CNC es similar a aquella defendida por los teóricos de la ‘teoría de los datos de los sentidos’, a diferencia de aquellos (que defendían el carácter fenoménico de los mismos), éstos consideran el CNC como una construcción de la mente (como un producto de la actividad mental individual), capaz de reconstruir para sí misma la realidad a partir de la representación particular del mundo. Es decir, lo que defienden es que la mente es capaz de reconstruir por sí misma, a partir de una labor de interpretación representacional, una imagen de la realidad desde los datos que es capaz de recoger y procesar del mundo externo.

Vistas así las cosas, para esta teoría, el contenido conceptual siempre se instanciará sobre el CNC. Esto es, siempre que en el mundo personal aflore algún tipo de conceptualización será debido a la presencia a un nivel sub-personal de un estado representacional que respalde dicho contenido conceptual. La pregunta que, entonces, queda abierta ahora es, si puede decirse con estas palabras, ¿en qué sentido podemos decir que el CNC es prioritario con respecto al contenido conceptual? En la literatura actual es habitual encontrar tres tipos de explicación de dicha prioridad (Kim 1998: cap. 8; Braddon-Mitchell y Jackson 1996: cap. 11; Davies 1998): ontológica, epistémica y analítica. El problema principal lo encontramos cuando analizamos detenidamente las últimas dos posibilidades. Respecto de la supuesta prioridad epistémica, parece que en realidad ocurre lo contrario, es decir, que es el contenido conceptual el que es prioritario respecto del CNC, ya que el único modo en que un individuo puede tener conocimiento proposicional de su CNC procede tan sólo a partir del conocimiento que posee a un nivel conceptual. Del mismo modo, con respecto a la presunta prioridad analítica, parece ocurrir lo mismo, ya que sólo a partir de un análisis plenamente conceptual pueden ser tratadas aquellas nociones clave que permitirían explicar la existencia del CNC. En estos dos sentidos, considerar que el CNC es prioritario sobre el contenido conceptual supondría negar la comprensión que un individuo posee de sus propios conceptos, conceptos bajo los cuales puede tener un acceso a la explicación de su CNC. En última instancia, lo que le quedaría al teórico del

Los estados subdoxásticos, así entendidos, serían estados psicológicos que permitirían al sujeto cognitivamente dotado la posibilidad de almacenar información acerca de los rasgos del sistema cognitivo, en general, que permiten la existencia de juicios perceptivos o sensoriales a partir de la posesión, a un nivel biológico, de mecanismos de procesamiento de dicha información que mediarían entre los estímulos externos y la formación de los mencionados juicios. Dicha dicotomía estaría avalada por la imposibilidad de que los estados subdoxásticos pudieran ser accesibles a la conciencia y por la imposibilidad de que éstos puedan ser combinados con estados intencionales que permitan aportar razones para la acción (contraviniendo así la tesis de la integración inferencial presente en las actitudes proposicionales que postula el necesario carácter relacional de las creencias que permite la confección de una red de inferencias potenciales. Cf. Stich 1979), pero también sostendría esta dicotomía la posibilidad de la existencia de estados psicológicos, como los estados subdoxásticos, carentes de contenido conceptual (Davies 1986 y 1989). Para una crítica de dicha distinción y de los criterios que parecen sustentarla, véase Skidelsky 2007.

CNC es, como ocurre por ejemplo con Peacocke (1992 y 1994) o Prinz (2002), presuponer que tanto el contenido conceptual como el CNC se complementan y se explican simultáneamente.

Una mención especial requiere la posibilidad de que el CNC sea ontológicamente prioritario respecto del contenido conceptual. En última instancia, aquellos que consideran la prioridad ontológica del CNC se basan, normalmente, en la supuesta independencia del CNC con respecto del contenido conceptual. Lo que suelen aducir para sostener esta independencia es la posibilidad de la existencia de CNC sin necesidad de la existencia de un contenido conceptual que lo determine. Un claro ejemplo de esta independencia viene marcado por el denominado Principio de Autonomía:

(PA) Es posible que una criatura esté en un estado mental con contenido no-conceptual aunque dicha criatura no posea conceptos (Bermúdez 1998: 61).

De todos modos, aunque la existencia de cierto tipo de criaturas a las que podemos atribuir estados mentales fenoménicos con contenido no-conceptual (como en animales sin lenguaje articulado) parece necesaria para poder suscribir la independencia del CNC, parece también que no es suficiente. Además sería necesario demostrar que nunca sería posible encontrar una criatura que posea conceptos sin la presencia de CNC.

De todos modos, tampoco existe un claro consenso entre los teóricos del CNC respecto a la posibilidad de la total independencia del mismo del contenido conceptual. Peacocke (1994), sin ir más lejos, rechaza el (PA). Mientras que para Bermúdez, como hemos visto, es posible la presencia de CNC sin la presencia de contenido conceptual (aunque ambos pueden ser definidos como representacionales (Bermúdez 1998: 51), lo que los distinguiría es el tipo de representación),³ para Peacocke (1994) sería más adecuado sostener una posición débil en la que el CNC siempre requeriría de la presencia de algún tipo de concepto, lo que no supone que todo CNC sea reducible a contenido conceptual (para ello sería necesario que ambos contenidos pudieran agotar la explicación de la experiencia de la que dan cuenta).

A pesar de los grandes esfuerzos que los teóricos del CNC dedican a argumentar acerca de la representacionalidad del mismo, consideramos que no parece del todo adecuado sostener una noción representacional del CNC, por dos motivos principales. En primer lugar, si como Bermúdez adoptamos una posición fuerte y consideramos que el CNC es independiente del contenido conceptual, entonces resultaría imposible fundamentar el contenido conceptual sobre el CNC porque, en última instancia, nos faltarían las bases adecuadas para saber en todo momento si ante una experiencia determinada el contenido conceptual que pueda poseer un individuo está respaldado por el CNC que dicho individuo posee. Pero del mismo modo, y en segundo lugar, si como Peacocke consideramos la posibilidad de defender una posición tan débil que permitiera asimilar CNC y contenido conceptual, entonces obtendríamos una tercera

³ Dedicaré la sección 2.1. a la cuestión de qué es una representación y a las diferentes definiciones que habitualmente podemos encontrar en la literatura al uso de la noción de representación.

instancia metafísicamente indistinguible de ‘los datos de los sentidos’, teniendo que pagar el alto precio cognitivo que ello supondría.⁴

En este sentido, quedarían dos caminos a seguir. O bien, por un lado, intentamos defender una posición representacional intermedia que permita evitar los inconvenientes de las posiciones anteriormente mencionadas o bien, por otro lado, debemos abandonar la representacionalidad del contenido mental fenoménico en beneficio de otro tipo de explicación.⁵

1.3. La conciencia y el problema mente-cuerpo

Un hecho que debería llamarnos poderosamente la atención al leer la literatura filosófica actual sobre la conciencia es que parece una idea general y comúnmente aceptada que tan sólo planea un problema respecto de su estudio: el llamado *hard problem*. Aunque este problema no es nuevo. Se refiere al viejo problema filosófico, anclado habitualmente a los estudios generales sobre filosofía de la mente, del problema mente-cuerpo. En pocas palabras, este problema se refiere a la dificultad de conectar dos instancias aparentemente tan (metafísicamente) separadas como son nuestra base física y nuestro aparato mental.

Dedicaremos este apartado, como hemos comentado al comienzo del capítulo, a mostrar una perspectiva general de los estudios más relevantes sobre filosofía de la mente que, de algún modo u otro, en los últimos años han sido influyentes a la hora de analizar el problema mente-cuerpo. Esta retrospectiva nos permitirá, primero, ofrecer una visión de conjunto de la problemática tal y como actualmente se entiende y, segundo, permitirá una contextualización de nuestro tema principal: el estudio contemporáneo dentro de la filosofía analítica de la mente de la conciencia fenoménica.

Se suele dar por supuesto, al menos en la Psicología Popular o del sentido común, de marcado carácter cartesiano, la existencia de dos instancias (o sustancias): por una parte, tendríamos la sustancia extensa, el cuerpo material, como base física que actúa dentro de un mundo marcado mecánicamente, que sirve de base comportamental a, por otra parte, una sustancia pensante, la mente, como base agentiva de la conducta. Si consideramos, como en realidad es común hacer, la existencia de estas dos sustancias separadas, entonces, aparece la delicada cuestión de tener que dar cuenta de su interconexión. ¿Cómo pueden conectarse entre

⁴ He criticado pormenorizadamente la teoría de los datos de los sentidos en Colomina 2007 y en Colomina (por aparecer). En estos textos dejo constancia de la dificultad subyacente a todo intento de basar las percepciones en una entidad metafísicamente separada como es la noción de dato de los sentidos porque, por un lado, apelar a la configuración interna de entidades mentales separadas de los objetos para dar cuenta de su contenido de experiencia multiplicaría innecesariamente los objetos metafísicamente relevantes y, por otro lado, nos obligaría a tener que establecer nuevas relaciones causales que permitieran vincularlas de algún modo a la auto-configuración material de las bases neurológicas que explican la percepción de los objetos externos.

⁵ Este intento de explicación aportado por el representacionalismo se muestra en las secciones 2.1.3.2. y 2.1.3.3. Una respuesta final, apostando más por un abandono del representacionalismo, la intentaremos ofrecer en el capítulo VI. En este último capítulo ofreceremos una defensa de una noción de intención en tanto que separada de la noción de representación que intente, por un lado, salvar las dificultades que las teorías representacionales presentan a la hora de explicar los estados mentales conscientes con carácter fenoménico (y que iremos analizando a lo largo de este trabajo) y que, por otro lado, permitan ofrecer una respuesta adecuada a los diferentes problemas que identificamos en el estudio de la conciencia sin necesidad de apelar a nociones internas al propio sujeto perceptor o metafísicamente dependientes de entidades separadas de las condiciones materiales que los objetos físicos cumplen a la hora de ser percibidos.

sí estas dos sustancias? ¿Cómo puede una ejercer un poder causal sobre la otra? ¿En qué sentido podemos decir que el causante de nuestra conducta no tiene un soporte físico? ¿O sí que lo tiene? En pocas palabras, ¿cómo podemos explicar la necesaria conexión entre una sustancia física (y extensa) con una sustancia mental (e inextensa)? Muchas, y muy variadas, han sido las diferentes respuestas aportadas a lo largo de los años.

Para el filósofo francés René Descartes, la mente existe como una sustancia independiente perteneciente a un nivel metafísicamente separado del mundo material. En su teoría de la mente, por tanto, Descartes defendió la existencia de dos sustancias, la pensante y la material, cuya existencia pertenecía a dos reinos totalmente separados, el mental y el físico, aunque interconectados. En el nivel físico, encontramos tan sólo materia en movimiento continuo. Las leyes de dicho movimiento estarían marcadas por las leyes mecánicas establecidas por la física newtoniana. Las partículas materiales que configuran nuestros órganos estarían en movimiento continuo y serían capaces de captar la realidad material, los objetos del mundo externo, permitiendo así obtener gran información respecto del entorno. A partir de dicha información, la mente humana (caracterizada por producir pensamiento sin disponer de ninguna característica material) construiría toda una serie de impresiones (olores, sabores, sonidos, etc.) que definirían aquella información respecto de los objetos del mundo exterior. Esta sería una proto-formulación de la distinción entre cualidades primarias y secundarias: las cualidades primarias son aquellos datos que directamente captarían nuestros sentidos materiales y que serían pertenecientes a los objetos del mundo, mientras que las cualidades secundarias serían aquellas construcciones que la mente confeccionaría a partir de dicha información, permitiendo así establecer cómo se le aparecerían a un sujeto particular dichos elementos pertenecientes al mundo material. De este modo, el mundo material produce efectos mentales, ya que los datos captados por los sentidos posibilitan la actividad mental. Pero también la mente tiene un poder causal respecto de lo material, ya que determinadas impresiones permiten a la mente controlar de manera activa al cuerpo.

El principal problema para Descartes, llegados a este punto, es explicar en qué sentido podemos decir que ambos reinos separados pueden interactuar uno sobre el otro. Descartes encontró plausible establecer un punto físico de unión entre lo material y lo mental: la glándula pineal. En este punto del cerebro humano, característico por su situación (se encuentra en el centro del cuerpo calloso) y por su asimetría (es la única parte del cerebro que no dispone de dos partes o de un correlato en los dos hemisferios), Descartes situó la interacción entre las fuerzas físicas y el poder causal mental.

Descartes estableció una teoría marcadamente dualista. Separar en dos reinos interconectados la naturaleza humana dificulta sobremanera la posibilidad de su explicación racional. No sólo porque tendríamos que explicar la actividad por separado de cada una de ellos, sino (sobre todo) porque debemos dar correcta cuenta de la conexión existente entre los mismos.

En este sentido, una respuesta obvia a la dificultad cartesiana es evitar la necesidad de tener que explicar dicha separación. Un ejemplo característico lo encontramos en la teoría idealista de George Berkeley. Partiendo del supuesto de la inexistencia del mundo material,

Berkeley afirmó que la experiencia tan sólo es un constructo de nuestra naturaleza mental. Tan sólo habría para Berkeley impresiones producidas por nuestra actividad mental, unas impresiones que no se correlacionarían con el mundo material, porque nada en nuestra experiencia mental estaría conectado con él. En este sentido se perdería el reino de lo físico, dado que los sentidos no captarían las características materiales de las cosas, sino que éstas serían ya un producto de la misma realidad mental. En este sentido, todo lo que existe es aquello que nuestra realidad mental es capaz de construir, sólo existe aquello que nuestra mente percibe.

El problema principal con la teoría de Berkeley, por supuesto, es la eliminación del mundo material. ¿Cómo es posible negar la existencia de algo que, al menos en apariencia, es tan obvio como la existencia de árboles, montañas y estrellas? ¿Cómo es posible negar la existencia del mundo material que nos circunda? Siempre podríamos rechazar la teoría berkeleyana con el famoso (y divertido) experimento de Samuel Johnson quien, pateando una piedra, le digo a Berkeley: ‘Yo refuto su teoría así’.

A pesar de esta crítica mordaz, la corazonada idealista ha sido aceptada tradicionalmente. En este sentido, John Stuart Mill defendió que los objetos del mundo físico no tienen una naturaleza separada del mundo mental. Es decir, el mundo físico puede ser reducido al mundo de la *vivencia* particular de cada individuo. En pocas palabras, el mundo material se reduce al modo en que a cada cual se le aparece, ya que los objetos materiales no serían más que posibilidades permanentes y continuas de sensaciones. Siguiendo su estela, Bertrand Russell (en una de sus etapas más famosas) defendió la idea de que el mundo físico no es más que una construcción lógica a partir de los datos de los sentidos que construye nuestra percepción. En este mismo sentido, el positivismo del Círculo de Viena también defendió una teoría del mundo físico como construcción mental a partir de la percepción.

De todos modos, gran parte de los esfuerzos de los filósofos y psicólogos desde el inicio del siglo XX ha sido, precisamente, rechazar este idealismo. Con este objetivo nació el conductismo. Para esta escuela, el mundo psicológico no puede ser construido a partir de la introspección sobre los estados subjetivos de los individuos, sino que debe ser reconstruido a partir de los estados objetivos de éstos, es decir, debe ser reconstruido sobre aquellos movimientos corporales públicamente observables: su comportamiento. A partir de sus experimentos de laboratorio basados, por un lado, en el ensayo-error y, por otro, en los refuerzos y castigos, estipularon que aunque la base de toda conducta es genética, ésta puede ser modificada por el entorno en el cual se desarrolla un sujeto. Esto es, que lejos de ser producto de nuestros estados internos, nuestro aparato mental es producto de la interacción del mundo externo en el cual nos encontramos situados.

Siguiendo las doctrinas del conductismo pero a un nivel más abstracto (y sin emplear el apoyo de los experimentos de laboratorio), la escuela del conductismo lógico, con Gilbert Ryle a la cabeza, defendió la incoherencia de la noción de experiencia individual subjetiva y, por tanto, de introspección. En este sentido, un estado mental tan sólo es aquella disposición públicamente observable a comportarse de cierto modo. Siguiendo esta filosofía, Ryle ridiculizó la tradición cartesiana con su famosa crítica al ‘fantasma en la máquina’: si consideramos como

hace el dualismo cartesiano la posibilidad de la existencia de dos reinos totalmente separados, el mental y el físico, entonces la mente no puede dejar de ser considerada más que como un fantasma (un alma) que maneja los hilos de la enorme marioneta que sería nuestro cuerpo. El principal inconveniente para esta imagen sería, por supuesto, su inconsistencia con la realidad comportamental humana, ya que postularía la existencia de una instancia diferente (un duplicado nuestro) que, a modo de homúnculo, guiaría nuestra conducta desde el interior de nuestro cuerpo, dejando en el aire quién sería realmente el responsable de nuestra conducta, ya que parecería que la acción de nuestro cuerpo no depende realmente de nosotros mismos.

Aunque hemos comenzado este apartado insistiendo en que el problema central de la conciencia básicamente se refiere al viejo problema mente-cuerpo, la formulación actual de este problema respecto de la conciencia es relativamente reciente. El desarrollo de la visión científica del mundo realizada a partir de la segunda mitad del siglo XX ha mostrado una construcción del mundo marcadamente diferente de aquella que tenían nuestros antepasados. Y, aunque a grandes rasgos, la dificultad sigue siendo la misma (¿cómo encajar en la realidad física la realidad de nuestra vida mental?), la nueva división del mundo físico en moléculas, átomos y elementos cuánticos que interactúan entre sí para conformar nuestra realidad material parecen desplazar la posibilidad de la existencia separada de estados mentales conscientes. Si históricamente, como hemos visto, la materia era considerada como el elemento a situar fuera de la realidad en beneficio de la supremacía de lo mental (una supremacía que, por cierto, todavía sigue vigente en nuestros días a través de la imagen popular del sentido común), actualmente es la materia y su explicación a la hora de constituir la realidad la que ocupa el lugar dominante, siendo la mente como tal la que desaparece bajo la explicación física.

En este sentido, el filósofo David Chalmers (1996) diferenció entre los problemas ‘fáciles’ de la conciencia y el problema ‘difícil’ (*hard problem*). Los primeros tan sólo se referirían a problemas objetivos respecto del estudio de la base cerebral de la conciencia, problemas que por supuesto tan sólo son relativamente fáciles, dado que se refieren a problemas experimentales respecto, por ejemplo, del modo en que el cerebro es capaz de procesar los estímulos ambientales que dan lugar a nuestras sensaciones o respecto del modo en que el cerebro es capaz de integrar la gran cantidad de información externa para producir estados internos (Chalmers 1996: 16). Todos estos problemas, por supuesto, proceden de la necesidad de *tomarse en serio* la ciencia física actual y de considerar que la conciencia tiene, en última instancia, una base material que la convierte en un fenómeno totalmente natural sometido a las mismas leyes que cualquier otra propiedad física del mundo. El segundo, el problema realmente importante, sería dilucidar en qué sentido podemos decir que todo este enorme procesamiento de información del cerebro está acompañado de una vida mental experimentada (Chalmers 1996: 16), una vida mental que sería interna y que permitiría a cada individuo, de modo subjetivo e individualizado, sentir de modo diferente las mismas fuerzas físicas que interactúan con nosotros. Este problema, según Chalmers, surgiría de la necesidad de *tomarse en serio* la conciencia.

Cuando nosotros hablamos de la existencia de estados mentales conscientes, como por ejemplo estados de dolor, experiencias perceptivas o, incluso, sueños, aportamos conjuntamente tanto explicaciones subjetivas como objetivas de los mismos. En un sentido laxo, nosotros estamos dando cuenta de propiedades y características objetivas de nuestra actividad psicológica pero, también, estamos aportando una visión subjetiva de todas aquellas sensaciones que se producen en nosotros ante la presencia de dichos estados. Esto es, por una parte hablamos acerca de lo que percibimos y por otra parte hablamos de lo que sentimos. No sólo decimos que estamos percibiendo, por ejemplo, un dolor en nuestra espalda, sino que también informamos de lo que es para nosotros tener un dolor en la espalda.

En este sentido, el filósofo Thomas Nagel defendió la importancia de estas últimas características subjetivas. En su trabajo clásico Nagel (1974) sugirió que deberíamos diferenciar entre la posibilidad de llegar a saber qué mecanismos le permiten a un murciélago percibir de la posibilidad de llegar realmente a saber lo que es (*what is it like*) percibir como un murciélago. Es decir, la ciencia actual puede llegar a determinar las características físicas que permiten al murciélago percibir; esto es, es capaz de llegar a definir de manera exhaustiva las características físicas y objetivas que permiten a un murciélago percibir a través de un sistema de eco-localización tan particular. Pero de ningún modo podríamos llegar nunca a saber qué se siente realmente al percibir la realidad de esta manera, precisamente porque no somos murciélagos y porque carecemos de dicho sistema de eco-localización. Para poder llegar siquiera a imaginar qué es sentir como lo hace un murciélago deberíamos poseer la capacidad experiencial de los murciélagos y, aún así, todavía nos faltaría el elemento subjetivo que permite individualizar para cada sujeto lo que sería su particular modo de sentir (o percibir) la realidad. En otras palabras, la ciencia actual está capacitada para poder llegar a ofrecer una explicación del modo en que los murciélagos pueden *percibir* la realidad, pero nunca estará capacitada para ofrecer una explicación de lo que es para un murciélago *sentir* la realidad.

En este sentido, parece que se abre una brecha (un hiato, como suele traducirse el término inglés *gap* para referirse a este contexto fenoménico) entre la descripción científica de la realidad y el mundo de la experiencia (Levine 1983). Vistas así las cosas, el problema central de la conciencia no se referirá a la posibilidad de explicar de manera objetiva la conciencia, sino de dar correcta cuenta de la conciencia fenoménica de la experiencia. Las explicaciones que han perseguido dar una respuesta adecuada a este problema han sido de lo más variopintas a lo largo de los últimos años.

En un primer momento encontramos una posición negativa. Desde este punto de vista, la conciencia no sería explicable, al menos en principio, porque estamos, nosotros y nuestra ciencia, limitados conceptualmente para poder dar correcta cuenta de todos los procesos de los cuales es capaz nuestra conciencia. Esta posición *misterianista* aboga por dejar a un lado la pretensión de encontrar una explicación total al fenómeno de la conciencia, no porque no exista, sino porqué la posibilidad de su explicación completa nos supera. Dentro de esta tradición filosófica, encontramos a naturalistas respecto de la mente como McGinn (1989) que, a pesar de establecer una explicación naturalizada respecto de parte de nuestro contenido mental, acaba por

advertir la imposibilidad epistémica de ofrecer una teoría completa que dé cuenta de la conciencia fenoménica (Cf. McGinn 1983, 1991 y 2004. También Kim 2005).

Siguiendo también esta estela, pero desde un marcado carácter anti-naturalista, Wittgenstein (1953) aporta razones en la misma dirección para negar la posibilidad de reducir la conciencia a una mera explicación naturalista y naturalizadora, estableciendo una categoría diferente para los estados conscientes. También filósofos subjetivistas, como Chisholm (1989), Nagel (1974) o Searle (1983, 1992, 1997 y 2004)⁶ consideran acertada esta concepción neo-dualista que, aunque no contempla la existencia separada de dos sustancias metafísicamente distinguibles (al estilo cartesiano) sí consideran aceptable la existencia separada de dos instancias explicativas a nivel causal epistémicamente diferenciables: por un lado, un elemento natural con un soporte físico y material y, por el otro lado, un elemento también natural pero no reducible a su soporte físico y material.

Explicaciones causales como éstas acaban por afirmar la naturaleza epifenoménica de la conciencia, como la aportada por Jackson (1982), o por defender una posición interaccionista neo-cartesiana que asocia un yo (trascendental) con un cuerpo (Popper y Eccles 1978).

De todos modos, los actuales avances en las ciencias físicas y en las ciencias cognitivas también han permitido nuevas argumentaciones a favor de una irreductibilidad de la conciencia. Este sería, por ejemplo, el caso del actual desarrollo de la física cuántica. Desde el punto de vista microfísico, las partículas que componen las neuronas parecen estar sujetas a un cierto grado de indeterminación que permitiría que en un momento determinado el cerebro sea consciente pero que, en último término, no podríamos explicar, porque en un momento determinado nunca seríamos capaces de establecer al mismo tiempo ni la velocidad de dichas partículas ni la dirección del espín de las mismas, como bien nos indica el principio de incertidumbre de Heisenberg (Penrose 1989; Lapiedra 2005; Chalmers 1996: cap. 10). Pero también sería el caso de, por ejemplo, la psicología cognitiva. En este sentido, y empleando los argumentos semánticos de la referencia directa (defendidos por Kripke (1971 y 1972), especialmente), Chalmers (1996: cap. 4) ha conseguido establecer una lógica bi-dimensional capaz de separar en dos tipos diferentes de intenciones los marcos globales de clases, dejando por un lado cabida a los elementos de clase natural y permitiendo la existencia de elementos no-reducibles físicamente, por otro.⁷

Vinculada directamente a esta concepción anti-reductiva de la conciencia, también aparece la tesis que defiende la intencionalidad como una propiedad básica de los seres conscientes. En este sentido, la intencionalidad sería una propiedad con un carácter diferente al de los elementos naturales, quedando su estatuto en un ámbito trascendental. Pero quienes apuestan por un marco naturalista de explicación de la intencionalidad (es decir, quienes consideran que la intencionalidad es una clase natural explicable desde las ciencias naturales)

⁶ La teoría subjetivista de Searle, que podríamos denominar ‘dualismo epistemológico’, o más convencionalmente ‘naturalismo biológico’, será analizada pormenorizadamente en el capítulo III.

⁷ Un análisis pormenorizado, que incluye propuestas tanto a favor como en contra de esta lógica bi-dimensional, puede verse en García-Carpintero and Macià (2006).

parecen discrepar a la hora de establecer la pertinencia de una explicación natural respecto de las propiedades que nuestro vocabulario intencional expresa.

En un primer momento podemos encontrar a aquellos que tajantemente niegan la posibilidad de que las actitudes proposicionales sean significativas, es decir, que sean portadoras de las cualidades semánticas que intuitivamente les aplicamos, por lo que optan, de uno u otro modo, por una eliminación del vocabulario intencional a favor de un vocabulario físico, cayendo en un anti-realismo intencional que parece negar el fenómeno del significado (Boghossian 1992). Dentro de esta posición eliminativista podemos distinguir dos estrategias básicas. Por un lado, encontramos a aquellos que directamente eliminan cualquier tipo de apelación a intenciones, como Feyerabend (1963) y Rorty (1970), cayendo en un mundo sin significado, o la posición más científicamente realista de Paul y Patricia Churchland, que defienden la sustitución del vocabulario mental por un vocabulario neurofisiológico, también en un sentido eliminativista (Paul Churchland 1979, 1989 y Patricia Churchland 1986). Por otro lado, encontramos a aquellos que a pesar de considerar irrelevante el vocabulario intencional prefieren mantenerlo en base a su practicidad, a su conveniencia pragmática, acabando por establecer una especie de perspectivismo pragmático o de instrumentalismo. Siguiendo la estela de Quine (1960), que a partir de sus famosos argumentos sobre la indeterminación de la traducción y de la inescrutabilidad de la referencia afirmó una indeterminación semántica a la hora de atribuir actitudes proposicionales que lleva inevitablemente a asociar su conveniencia a la compatibilidad y dependencia que la vincularían directamente con las demás atribuciones (su holismo), más que a hacerla depender de su verdad o falsedad, encontramos a Dennett (1978 y 1987),⁸ que apuesta por mantener una actitud intencional frente a toda conducta que permita considerar a un organismo *como si* en realidad fuera consciente y atendiera a razones, una actitud pertinente porque permite a los organismos coordinarse y predecir conductas futuras, tanto propias como ajenas. Con este mismo espíritu instrumentalista encontramos la posición de Stich (1983 y 1990). Para este autor, la atribución de actitudes proposicionales, esto es, de intenciones, no dependería ni de ningún tipo de ejercicio interpretativo ni de ningún tipo de equilibrio reflexivo (una actitud más propia de perspectivismos subjetivistas como el de Dennett), pero tampoco dependería de ningún tipo de pragmatismo filosófico, como el de Rorty (1982), que acaba por aceptar la utilidad necesaria de la atribución de actitudes proposicionales, sino que dependería de solucionar un problema tecnológico (como lo denomina Liz 2003: 98). Para Stich, ningún proceso cognitivo tiene un valor epistémico, porque no depende ni de razones ni de su relación con otros procesos. Un proceso cognitivo simplemente se tiene o no se tiene, y su valor no va más allá de las posibilidades que permite. Un ejemplo claro lo encontramos en la inteligencia: los seres inteligentes son inteligentes no porque les sea útil o porque podamos interpretar sus conductas como si fueran inteligentes o respondieran a razones inteligentes, sino porque son inteligentes. Del mismo modo, el ser intencional, o consciente, es intencional o es consciente porque tiene intencionalidad, o conciencia, no porque actúe según razones o porque su actitud pueda interpretarse como si fuera intencional o consciente. Si se

⁸ A analizar la perspectiva de Dennett dedicaremos el capítulo 5.

comprobara que éste es el caso, entonces lo que quedaría sería una eliminación del vocabulario estandarizado que permite referirse a estas propiedades (un vocabulario dependiente del sentido común, que tanto Stich como los Churchland elevan al estatuto de ciencia, denominándola *Folk Psychology*, por su presunta fuerza explicativa) en favor de un vocabulario técnico dependiente de las Ciencias Cognitivas que permitiera rendir cuentas de los procesos desde el punto de vista meramente mecánico del organismo poseedor de los mismos, evitando así cualquier posibilidad de apelación a elementos subjetivos y/o interpretativos.

Siguiendo la línea aportada desde la filosofía eliminativista, aunque es temporalmente previa a la misma, encontramos la posición defendida por los teóricos de la identidad. Aunque habitualmente se suele asociar la Teoría de la Identidad con un materialismo de corte reduccionista, y a menudo se la ha identificado directamente con él, lo cierto es que ésta y el Eliminativismo (aunque en última instancia persiguen el mismo objetivo) difieren en el tratamiento que realizan de los términos mentales. A pesar de que ambas escuelas nacen con la pretensión de aportar soluciones a las dificultades que provocan los análisis disposicionales de los estados cognitivos, volitivos e intencionales por parte del conductismo, inevitablemente entendidos como residuos de una metafísica que todavía conserva el vocabulario mentalista del cartesianismo, ambos planteamientos parten de tesis muy diferentes.

La Teoría de la Identidad (TI), a diferencia del Eliminativismo, todavía conservará un resquicio conceptual de dualismo al considerar la existencia de una relación de identidad entre lo mental y lo físico, entre la mente y el cerebro, al afirmar que las propiedades o tipos de estado mental son propiedades o tipos de estados del sistema nervioso central. Como podemos observar, esta teoría es de carácter materialista y se dirige a determinar la naturaleza física de las propiedades de los estados mentales, una naturaleza física en tanto que estas propiedades son propiedades del cerebro, asumiendo una total identidad entre ellas. De este modo, una propiedad mental, como la experiencia de dolor, no está tan sólo determinada causalmente por un estado neurológico (por ejemplo, la excitación de las células C del cerebro), sino que realmente existe una identidad entre la propiedad mental 'dolor' y la propiedad física 'excitación de las células C del cerebro', del mismo modo que existe una identidad entre la propiedad macrofísica 'calor' y la propiedad microfísica 'energía cinética de las moléculas'.

El Eliminativismo, por su parte, no es una teoría respecto de las propiedades mentales, sino que por el contrario es una teoría respecto de los términos cotidianos con los cuales el lenguaje ordinario se refiere a determinados hechos físicos y cómo, en última instancia, los primeros carecen de referente (esto es, carecen de una realidad a la que representar) por ser puramente intensionales, con lo que acaba por concluir que mejor sería abandonarlos en beneficio de un lenguaje puramente extensional (el científico) capaz de referirse a lo realmente existente. Desde este punto de vista, no existiría una propiedad mental a la que podamos denominar 'dolor', sino que lo único existente sería la propiedad física 'excitación de las células C del cerebro', permitiendo así sustituir toda referencia a lo mental por términos físicos; del mismo modo que no podemos decir que existen dos sustancias tales que una es 'agua' y la otra es 'H₂O', sino que en realidad todo aquello que llamamos 'agua' es todo aquello que

extensionalmente es 'H₂O', pudiendo eliminarse el primer término (cotidiano) por su referencia física.⁹

Vistas así las cosas, la relación entre lo físico y lo mental, para la TI, sería una relación reductiva, en tanto que consideraría que los estados mentales son estados cerebrales. Por su parte, para el eliminativismo, esta relación sería eliminativa: no habría ninguna propiedad mental, sino que todo lo que existe son propiedades físicas. La relación sería tan sólo conceptual y no metafísica. Esta diferenciación es crucial para nosotros porque, por un lado, nos permitirá aportar nueva luz a la contraposición entre reducción y eliminación, que demasiado a menudo ha sido confundida (Bernstein 1968: 203ss.) y también, por otro lado, nos permitirá matizar las diferencias entre los argumentos de Rorty y los argumentos de algunos seguidores de la TI.

En un artículo seminal, Place articula las primeras nociones sobre la posible identidad entre lo mental y lo físico. Aquí, Place parte de la hipótesis de que los enunciados introspectivos no tienen por qué describir objetos no-físicos y que toda teoría que asevere lo contrario recae en la denominada 'falacia fenomenalista' al considerar posible la existencia de propiedades imposibles de explicar por las ciencias físicas. En palabras del propio Place,

“si suponemos, por ejemplo, que cuando un sujeto informa acerca de una post-imagen de color verde está afirmando la existencia en su mente de un objeto literalmente de color verde, está claro que estamos en presencia de una entidad que no tiene cabida en el mundo. Este error, al que me referiré como la 'falacia fenomenalista', es el error de suponer que, cuando el sujeto está describiendo su experiencia... está describiendo las propiedades de objetos en un tipo peculiar de pantalla interna” (Place 1956: 35).

La única posibilidad alternativa a esta falacia es, afirma, sostener que todo proceso fenomenológico es reducible, en última instancia, a procesos neurológicos. El modo en que un ser consciente siente su introspección, el modo en que describe su 'cómo qué es' (*what it is like*), depende de la capacidad de describir las propiedades reales, físicas, de los objetos del mundo. De este modo, la experiencia subjetiva que alguien siente respecto de las cosas puede ser descrita por medio de propiedades físicas de objetos físicos determinados, sin tener necesidad de recurrir a propiedades fenoménicas de objetos no-físicos. De este modo, nos dice,

“cuando describimos una post-imagen como siendo de color verde, no decimos que exista algo que sea verde, sino que decimos que tenemos una experiencia tal como la que habitualmente tenemos cuando observamos una mancha de luz verde a la que hemos aprendido a describir así” (*Ibid.*).

⁹ Es común entre los filósofos eliminativistas negar la realidad del contenido fenoménico de los estados mentales o *qualia*. Como hemos indicado, esta negación viene establecida por la irrelevancia explicativa de los mismos. De lo que se trataría es, entonces, y como irónicamente establece el título de un famoso texto de Dennett, de *quinear* los *qualia*, de negar firmemente su existencia o importancia como algo real o significativo (Dennett 1988: 519). En última instancia, esta estrategia supondría un ataque a una posición cartesiana de la mente (que defendiera la posibilidad de la existencia de un teatro donde se representarían los estados fenoménicos) porque, si así fuera, podría darse el caso de la existencia de propiedades cualitativas hacia las cuales el sujeto no estaría dirigido ni tendría ningún tipo de acceso, perdiendo así dichas cualidades su carácter normativo. Una defensa actual de varias posiciones reivindicativas de la existencia de los *qualia*, entre las que se encuadran argumentos dualistas y a favor de la teoría de los datos de los sentidos, puede encontrarse en Wright (ed.) 2008.

Vistas así las cosas, para Place es plausible que las experiencias fenoménicas sean realmente idénticas a procesos que tienen lugar en nuestro sistema nervioso central; es decir, en última instancia, Place sostiene que toda experiencia fenoménica introspectiva es reducible a un estado cerebral determinado explicable desde las ciencias neurofisiológicas.

Aunque la TI nace con la intención de ser una teoría alternativa capaz de mejorar la explicación de los estados mentales aportada por el conductismo, al no considerarse una tesis semántica (esto es, una tesis respecto del significado del contenido psicológico introspectivo, al estilo del conductismo), sino una mera hipótesis científica (esto es, un conjunto de enunciados contingentes *a posteriori*), conseguirá escapar de las objeciones habituales realizadas a las teorías que buscan establecer equivalencias lógicas entre términos. Así, por ejemplo, al igual que enunciados de identidad del tipo “La temperatura es la energía cinética de las moléculas de los cuerpos” o “Júpiter es el planeta más grande del sistema solar” o “El número de protones en el núcleo de una molécula de uranio es 92”, el enunciado “Los estados mentales son estados del sistema nervioso central” es un enunciado que se confirmará por medio de la investigación y la observación científicas y no a partir de ningún análisis conceptual *a priori*. Como Place afirma,

“la proposición ‘La conciencia es un proceso cerebral’ no es, en mi opinión, ni auto-contradictoria ni evidente, sino que conforma una hipótesis científica razonable, del mismo modo que el enunciado ‘Un relámpago es movimiento de cargas eléctricas’ es una hipótesis científica razonable” (Place 1956: 30).

Los críticos de la teoría de la identidad no han estado nunca del todo satisfechos con esta explicación. Muchos de ellos mantienen la imposibilidad de establecer equivalencias lógicas (y menos aún, semánticas) entre enunciados psicológicos y enunciados de carácter mental y conductual porque siempre es posible imaginar situaciones donde difieran en sus condiciones de verdad. La única posibilidad de realizar dichas equivalencias es encontrar definiciones en un lenguaje físico de esos estados psicológicos. Pero esta práctica es implausible porque, cada vez que realicemos la definición de un estado psicológico, ésta estará compuesta por términos no-mentales, físicos. Sin embargo, al desmarcarse de la identificación lógica o semántica, los teóricos de la identidad consiguen desembarazarse de esta crítica. Gustosamente aceptan que podemos imaginar situaciones en las que, a pesar de la equivalencia entre enunciados psicológicos y físicos, difieran en sus condiciones de verdad. Pero esto sólo demuestra que la teoría de la identidad no es necesariamente verdadera, sino hipotética, algo que desde el principio postulan sus defensores. Los teóricos de la identidad parten del supuesto de que TI es una hipótesis empírica y que, en último término, enunciados físicos definen contingentemente enunciados psicológicos, teniendo acceso a ellos a partir de la investigación empírica.¹⁰

¹⁰ Contra esta tesis propuesta por la TI, Kripke propone su tesis acerca de la necesidad de todo enunciado de identidad (Kripke 1971 y 1972). A grandes rasgos, lo que vendría a sostener Kripke es la permanencia ontológica de los objetos mencionados en todo enunciado de identidad. Podemos entender el sentido en que “Hesperus es Phosphorus” o “Cicerón es Tulio” son verdades necesarias *a posteriori* porque estos enunciados requieren de la investigación empírica para poder asentir su verdad, pero también en tanto que están constituidos por dos designadores rígidos y, en tanto que su referencia está fijada por las entidades o propiedades esenciales que refieren, pueden considerarse como

Pero, tal vez, la objeción más radical a TI sea aquella que niega la identidad entre enunciados mentales y enunciados neurofisiológicos a partir del hecho que muestra que, a diferencia de los segundos, los estados mentales no tienen una localización espacial. Pero, en casos como, por ejemplo, el dolor, esta crítica es salvable porque, en última instancia, somos nosotros mismos los que localizamos el dolor (en la cabeza, en la espalda, en la rodilla...) pero, además, aunque sepamos que no existen células C en la rodilla, sabemos que sí existen terminaciones nerviosas en ella que la conectan con el cerebro. La crítica renueva sus fuerzas cuando se aplica a otro tipo de estados mentales distintos del dolor, dado que el caso del dolor (como también el del color) son casos-límite, pero ¿qué ocurre con otros casos como, por ejemplo, las emociones o los estados de ánimo? Para muchos, en especial aquellos que como Malcolm (1984) siguen las sendas del lenguaje ordinario, parece extraño hablar de la localización de los estados mentales, pero para TI no lo será si, como defienden, éstos son equivalentes a estados neurofisiológicos.

Un ejemplo claro lo encontramos en la defensa que de TI realiza Armstrong, que extiende la identidad a todos los estados mentales, incluidas las actitudes proposicionales, con lo que hace casar la TI con nuestro uso ordinario del concepto 'mente' y la muestra como una alternativa razonable al conductismo. No sin otro propósito escribe:

“La identificación de los estados mentales con estados físicos cerebrales es perfectamente inteligible... si logramos una forma correcta del análisis de los conceptos mentales. ¿Cómo es posible que los estados mentales sean estados físicos cerebrales? La respuesta consiste en proponer un análisis del concepto de 'estado mental' que sea independientemente plausible y permita dicha identificación. Así, el filósofo acondiciona el camino para una doctrina de primer nivel que, ya sea verdadera o falsa, es importante: una concepción puramente fiscalista del ser humano” (Armstrong 1981: 40).

Así, la filosofía no debe tan sólo remitirse a la ciencia para comprobar la TI, sino que también puede incluir argumentos que demuestren dicha hipótesis a partir de la caracterización causal de nuestra concepción ordinaria de lo mental, abriendo el camino a una nueva teoría respecto de los estados mentales, el funcionalismo (en el que no me voy a detener aquí).

Como hemos visto, la TI es la teoría que postula la equivalencia entre propiedades mentales y propiedades físicas. A este respecto, Feigl escribe:

“La tesis de identidad que quiero clarificar y defender afirma que los estados de experiencia directa que los seres humanos conscientes 'experimentamos' y aquellos que confidencialmente adscribimos a los animales superiores son idénticos a... aspectos de procesos neuronales... que se tienen-en-la-experiencia y que (en el caso de los seres humanos) se conocen *por familiaridad* como idénticos a los objetos de conocimiento *por descripción* que nos aporta el conductismo y éstos, a su vez, son idénticos a los que describen las ciencias neurológicas” (Feigl 1959: 446).

Desde su punto de vista, existirían dos tipos diferenciados de conocimiento respecto de las experiencias conscientes. Por un lado, tendríamos aquellas experiencias que pueden ser conocidas por familiaridad, que requieren de predicados fenoménicos, y aquellas experiencias

auténticos enunciados de identidad. Identidad que, por otra parte, es necesaria a partir de las tesis esencialistas de la metafísica kripkeana.

conocidas por descripción, que son perfectamente descriptibles en términos físicos. Por lo tanto, Feigl establece una diferenciación entre aquellas propiedades cognitivas y fenoménicamente experimentadas que no son verificables de manera objetiva, sino que su acceso sólo es posible en primera persona, porque son subjetivamente dependientes, y aquellas propiedades experimentadas cuya verificación atiende a un aspecto objetivo y en tercera persona. Pero Feigl no quiere significar con ello que aquello que podemos conocer por familiaridad sea ilusorio o meramente aparente, sino que también tiene una realidad similar a aquella que podemos obtener por medio de la descripción. Ambos tipos de conocimiento serían legítimos y válidos, ya que ambos refieren a cosas existentes.

Pero también es verdad que este tipo de distinción epistémica ha sido criticada muy duramente por admitir aquello que Sellars llama ‘el mito de lo dado’ (Sellars 1963) porque, en última instancia, no está del todo claro en qué sentido podemos verificar la relación que los individuos disponen con su conocimiento en primera persona, con su conocimiento por familiaridad. Feigl es consciente de ello, pero aún así considera que debemos mantener esta distinción, porque de lo contrario deberíamos presuponer la inexistencia de lo que denomina ‘sentimientos brutos’. Pero éstos existen, aunque sean idénticos a procesos neurofisiológicos.

El problema principal con esta interpretación de la TI es que asume algunos de los supuestos dualistas que pretende refutar. Por una parte, al diferenciar entre dos clases de conocimiento respecto del mismo hecho físico, Feigl también mantiene una diferenciación entre dos modos de verificación diferentes: uno subjetivo y otro objetivo. Por otra parte, al considerar la presencia de dos tipos de conocimiento, uno en primera persona y otro en tercera persona, los convierte en irreconciliables e irreducibles porque, si consideramos (como hace Feigl) que una ciencia neurológica futura podría ser capaz de describir y explicar adecuadamente los procesos cerebrales, la explicación fenoménica estaría de más y haría recaer su teoría en el epifenomenalismo. En pocas palabras, al diferenciar entre dos tipos diferentes de explicación, Feigl necesariamente recae en un dualismo epistémico entre dos tipos irreconciliables de conocimiento. En último término, lo que Feigl no es capaz de justificar es en qué sentido debemos entender la identidad entre la explicación mental y la física porque, si consideramos que esta cuestión se refiere tan sólo a afirmar la identidad de la referencia de ciertas expresiones en dos tipos independientes de conocimiento, entonces recaemos en una cuestión terminológica y no metafísica, convirtiéndose nuestro problema inicial en un pseudo-problema sin solución (Bernstein 1968: 207). Pero, como defiende Feigl, que la pretensión de la TI no persiga una revisión metafísica sino una adecuación del uso cotidiano del concepto de lo mental (como hemos visto que ocurría con la teoría de Place), no convierte la problemática entre el cuerpo y la mente, el llamado problema mente-cuerpo, en un pseudo-problema, sino que permite eliminarlo, también, mediante la identificación de lo mental con lo neurofisiológico (Feigl 1960).

A diferencia de Feigl, Smart no persigue encontrar de modo unívoco aquellas propiedades físicas a las que son reducibles las propiedades mentales, sino que considera que “debe ser posible afirmar la existencia de procesos físicos irreducibles, pero no afirmar la existencia de propiedades físicas irreducibles”. Esto es, la identidad propuesta presupone la

existencia de tipos y no de casos.¹¹ Pero, ¿por qué lo contrario supondría un problema para TI? En el caso de Feigl, suponer que un enunciado de propiedades mentales pueda diferir en el valor de verdad en su reducción a propiedades físicas imposibilita la TI. Pero si, como sugiere Smart, consideramos que sólo existen procesos y entidades físicas (destituyendo la legitimidad del uso cotidiano del vocabulario mental), entonces desaparece esta dificultad.

Rorty localiza aquí un error sustancial de la TI. La insistencia en la existencia de dos tipos de conocimiento acerca de las entidades conscientes, un conocimiento que las caracteriza como físicas y otro que las caracteriza como mentales, no hace más que mantener el viejo y metafísico vocabulario dualista en lugar de disolverlo, cuando lo que de verdad deberíamos hacer es conseguir que nuestras descripciones en primera persona, no inferenciales, de sensaciones puedan ser traducidas a un lenguaje de tipo fisicalista, el científico, capaz de deshacerse de los inconvenientes de la metafísica dualista. Al intentar legitimar ambos tipos de conocimiento, pero a la vez afirmar la reductibilidad del mental al físico, la TI tan sólo consigue crear una tensión porque parece ir contra el modo habitual en el que expresamos nuestras sensaciones. Lo que deberíamos buscar es el modo en que podamos perfectamente hablar de sensaciones como estados neurofisiológicos, sin promover la ambigüedad y la asimetría existente entre el conocimiento en primera y en tercera persona y apostando siempre por la constitución de un nuevo vocabulario más cercano, más científico, que sustituya el vocabulario ordinario y permita disolver el malentendido creado por el vocabulario residual cartesiano que todavía inunda nuestro lenguaje ordinario.¹²

Siguiendo la corriente anti-conductista iniciada por la TI, el funcionalismo (metodológico, al menos) proporcionará una interpretación naturalizada de lo mental en general, y de la conciencia en particular, que por un lado postula un nivel teórico vinculado a la psicología (pero independiente de su reducción neurofisiológica, subdoxástica, o intencional, doxástica) y por el otro lado un nivel práctico que postula todo un vocabulario conceptual capaz de proporcionar el marco que permita el desarrollo de dicha actividad psicológica. En este sentido, para un funcionalista los estados mentales conscientes se caracterizan por su peculiar papel funcional dentro de la entera dinámica mental, asegurando así la posibilidad de la múltiple realizabilidad de los estados conscientes y justificando el nivel teórico general de explicación a partir de una estratificación de las diversas funciones de las que son capaces los diferentes procesos cognitivos y los diferentes estados que los implementan. Al considerarse que dichos

¹¹ Existe aquí una diferencia marcada entre lo que se podría denominar un tipo (*type*) y un caso (*token*). Los segundos hacen referencia a un numéricamente único ejemplar concreto de un estado mental determinado, mientras que los primeros hacen referencia al conjunto de todos los ejemplares que pertenecen a la misma categoría.

¹² En la actualidad, la TI ha sido prácticamente abandonada por las contundentes críticas recibidas, por un lado, por los defensores de la teoría directa de la referencia que, como Kripke (1971 y 1972), ven inconsistente la identidad entre cerebro y mente a partir de argumentos semánticos basados en la lógica modal (véase nota 10) y, por otro lado, por los filósofos funcionalistas defensores de la múltiple realizabilidad. Un reciente intento por hacer renacer este espíritu de identidad es el ofrecido por Polger (2004). En esta obra, nuestro autor se posiciona en una teoría de la identidad (de tipos) entre la mente y el cerebro que la hace consistente con los hechos empíricos, afirmando por un lado que la experiencia consciente es el fenómeno mental fundamental (a diferencia de la visión dominante en la ciencia actual, que considera que la conciencia es sólo una pequeña parte de los procesos mentales humanos), aunque sin aportar ningún avance científico ni ningún correlato neuronal que permita tal identificación, y, por otro lado, que es posible explicar el contenido fenoménico a partir de la múltiple realizabilidad.

estados conscientes son semánticamente relevantes, se entienden como intercambios y procesamientos de información, de modo similar a la dinámica estratificada de estados cognitivos ejercida por un computador, por lo que (además de la pertinente metáfora de la mente como procesador) se establece una jerarquía que es capaz de distribuir adecuadamente todos los mecanismos elementales que permiten el funcionamiento de dicho proceso, hasta llegar a un nivel neurofisiológico en el que la estructura y la función cognitivas coinciden.¹³ Sin embargo, surgen diferencias pronunciadas a la hora de establecer dicha jerarquización. Por un lado, encontramos a los defensores del lenguaje del pensamiento (Cf. Fodor 1975; Harman 1973; Prinz 2003 y 2004), que postulan la existencia a nivel mental de un sistema simbólico y representacional que es manipulado en la actividad cognitiva para generar estados intencionales, con propiedades cuasi-lingüísticas, que permitan dar cuenta de los procesos cognitivos. Por otro lado, encontramos a aquellos que consideran que no existe tal simetría entre estados funcionales y estados intencionales, por lo que no sería posible reducir los segundos a mera función, debiendo precisamente por ello concebir los estados intencionales a un nivel más general (Stalnaker 1984 y, especialmente, Stalnaker 2008).

En la misma línea funcionalista, encontramos la posición ortodoxa de Loar y Block. Para dicha posición, la mente no es más que una estructura dinamizada que contiene estados que pueden expresarse según el papel funcional ejecutado, un papel que determinará su rol inferencial, permitiendo establecer así, su valor intencional. El problema principal al que debe enfrentarse esta concepción es su tratamiento monádico del contenido semántico, ya que ello imposibilita la composicionalidad del significado y, lo que sería peor, dificulta la posibilidad de establecer una relación entre individuo y mundo (recayendo en un solipsismo, como bien indica Fodor 1980). La salida a esta encrucijada debe buscarse en la separación entre la explicación funcionalista de las relaciones intencionales posibles y la explicación de las características semánticas de los contenidos de dichas relaciones intencionales, permitiendo así establecer una relación causal entre ambos tipos de explicación. Ahora la dificultad recae sobre el tipo de explicación adecuada que dé cuenta de dicha relación causal entre las propiedades representacionales de los estados intencionales y el mundo (Cf. Block and Stalnaker 1999). Aunque existe un consenso más o menos estable sobre la necesidad de aportar una explicación naturalista a dicha relación que permita, por un lado, evitar la circularidad que supondría tener que explicar dicha relación en términos semánticos y, por el otro, explicar tanto el modo en que podemos diferenciar entre representaciones correctas e incorrectas (su normatividad) como determinar adecuadamente la corrección de la atribución de representaciones, existe una disonancia a la hora de explicar la naturaleza de dichas representaciones.¹⁴

¹³ Un ejemplo claro lo encontramos, por ejemplo, en Lycan (1986), donde postula diferentes niveles en la naturaleza que permiten sostener conceptos diferentes de reducción que permitirían la equiparación entre lo físico y lo mental.

¹⁴ No todos los intentos de subsanación de la teoría funcionalista se basan en dicha explicación. Algunos han preferido negar la presunta simetría entre estados intencionales y contenido semántico, apostando por una relación de dependencia asimétrica que, por un lado, permita explicar la posibilidad de la existencia de representaciones erróneas y, por el otro, permita aportar una explicación del contenido en términos diferentes a los conceptos mentales tradicionales. Cf. Fodor 1990. Una aproximación al tratamiento que Fodor realiza del contenido conceptual se presenta en la sección 2.1.3.1.

Quienes como Devitt (1988) o Papineau apuestan por una explicación histórico-causal, establecen una explicación de la relación entre representación y mundo a partir de la interacción entre el sujeto y el entorno, tal y como viene especificada en la teoría de la referencia directa de Putnam (1975). Pero, como dicha teoría da por supuesta la existencia de dichas intenciones y lo que trata es de precisar su contenido, se ve obligada a tener que explicar la relación referencial de las intenciones sin tener que contar con la carga semántica misma (para evitar la temida circularidad explicativa), lo que hace depender dicha relación de una aleatoriedad ocasional que permita aportar los mecanismos causales necesarios, lo que tampoco es especialmente deseable. (Un desarrollo de las tesis de la teleología en general, y de la posición de Papineau, en particular, se presenta en la sección 2.2.1.2.).

Con vistas a superar esta dificultad, Dretske (1981) elabora una teoría de la representación dependiente de la noción de información. En este sentido, para Dretske, el contenido semántico de todo estado consciente es contenido informacional transmisible. La relación entre representación y mundo se daría de modo causal, siendo el mundo el que permitiría una descarga de información que ocasionaría la implementación de un estado determinado, tan sólo dependiente de la situación externa que se haya producido. El problema principal para esta teoría sería la imposibilidad de cometer un error en la representación.¹⁵ Con vistas a superar esta dificultad, Dretske (1988) aporta una visión mucho más funcional. Las representaciones serían aquellos estados que tienen una relación nomológica con cierto tipo de estados del mundo y que, además, tienen para el organismo una función indicadora intrínseca. Ahora la dificultad se traslada a una identificación indeseable, dado que no es posible definir la noción de función de forma independiente de la noción de representación. (Una ampliación de la posición de Dretske y de su teoría informacional se presenta en la sección 2.2.1.1.).

Millikan (1984 y 1993) adopta una noción biológica y evolutivamente selectiva de función. Lejos de querer determinar el contenido de un estado en base a lo que indica o representa (como lo haría una noción general de función), Millikan pretende determinar dicho contenido a partir de las condiciones necesarias (tanto externas como internas) que permitirían que dicho estado funcionara adecuadamente dado un determinado contexto. Dichas condiciones son las que constituirían el contenido representacional del estado en cuestión, permitiendo así una explicación teleofuncional del contenido que, por un lado, aportaría una explicación a la posibilidad de error en la representación y, por otro, permitiría una explicación a partir de razones y predicciones de la relación entre intención y contenido.¹⁶

Siguiendo con nuestro recorrido por los actuales planteamientos respecto de la explicación de la conciencia fenoménica, podemos encontrar también en la actualidad teorías materialistas que aportarían respuestas basadas en los recientes desarrollos de las ciencias neurofisiológicas a los problemas planteados por el carácter fenoménico de determinados estados mentales. El estudio pionero de este tipo de propuestas lo encontramos en la hipótesis

¹⁵ El análisis pormenorizado de esta imposibilidad se ofrece en el apartado 5.1.

¹⁶ Una aproximación teleofuncional a la problemática de las emociones, además de una buena introducción al planteamiento de Millikan, puede encontrarse en Casacuberta (2000). Por nuestra parte, hemos defendido una posición teleosemántica respecto del lenguaje en Colomina (2008).

neuronal lanzada por Crick (1989). En pocas palabras, lo que estaría defendiendo Crick es que la conciencia humana no es más que una banal fusión de neuronas cerebrales, que todo nuestro comportamiento, incluida la ilusión a la que llamamos mente (y su principal característica: la conciencia) no es más que el producto de las reacciones bioquímicas que tienen lugar en las células neuronales de nuestro cerebro. En este sentido, por ejemplo, el hecho regular de que exista una oscilación neuronal de 40 Hz. podría explicar la existencia de la conciencia (Cf. Crick and Koch 1990). Siguiendo la estela neurocientífica, Ramachandran and Blakeslee (1998) analiza aquellas áreas del córtex cerebral involucradas en la construcción de representaciones mentales respecto de estados corporales y respecto el campo visual. Defiende la idea de la existencia de circuitos cerebrales especializados localizados en los lóbulos temporales (por ejemplo, la amígdala, el hipotálamo o el córtex insular) que permiten la representación de estados fenoménicos. Por su parte, Baars (1988 y 1997) defiende la posibilidad de que los recientes descubrimientos neurocientíficos sostengan la metáfora dualista de que la mente consciente puede ser interpretada como un teatro (en sentido cartesiano), donde la memoria, el lenguaje del pensamiento y la estructura biológica conformarían el escenario, los contenidos fenoménicos de conciencia serían los actores, la atención sería la luz escénica, los contextos de conciencia serían los aparatos de escena, el yo (como agente de acción y observador) sería el director y la audiencia sería todos los demás procesos inconscientes que automáticamente actuarían a nivel bioquímico en el sistema nervioso central y el cerebro.

La mayoría de las teorías de la conciencia actuales, como hemos visto, tienden a naturalizar su realidad bajo la reducción de su existencia a procesos cerebro-nerviosos o, al menos, redefinen su realidad en términos físicos. Aunque estas explicaciones tienen su parte de razón, para Chalmers el problema central no desaparece: la realidad de la conciencia subjetiva sigue persistiendo, porque todos tenemos la sensación de vivir de modo interno la realidad, a todos nos parece que vivimos *nuestra* realidad. En este sentido, y en palabras de Chalmers (1996: 17), lo que requerimos es una explicación, y no una reducción, de qué sea aquello que llamamos experiencia fenoménica.

Pero también, por otro lado, han tomado fuerza recientemente aquellas teorías que defienden una explicación, aunque también de corte materialista, no-reductiva de la conciencia. Estas teorías ofrecen una explicación de nuestra fenomenología consciente a partir de representaciones del mundo circundante formadas por esa misma base física, pero (a diferencia de las teorías reductivas) aceptan que estas representaciones no pueden identificarse con las bases físico-químicas del cerebro, aunque la mayoría de ellas presuponen que sobrevienen a las mismas.¹⁷

¹⁷ La noción de superveniencia o sobreveniencia fue introducida por Kim (1978) y desarrollada en Kim (1984a) y, especialmente, en los textos recogidos en Kim (1993) para referirse de un modo no-reductivo a la intuición de que la mente, aunque no puede ser identificada *sensu stricto* con las propiedades físicas que la componen y permiten, puede ser explicada a partir de las mismas teorías fisicalistas que posibilitan una explicación de las propiedades físicas del mundo. Para una revisión general de las diferentes nociones de sobreveniencia, véase Stalnaker (1996) y especialmente Liz (2001). Para una revisión de la noción de superveniencia aplicada directamente a la problemática de la concepción física de la conciencia, véase Díaz-León 2008.

1.4. Los problemas de la conciencia

Hemos dedicado el anterior apartado a mostrar una breve perspectiva de las explicaciones que a lo largo de los últimos años las escuelas filosóficas y psicológicas más relevantes han aportado para dar cuenta del problema central de la filosofía de la mente en general (el problema de la interacción entre el cuerpo físico y la mente) y, en particular, del problema central de la conciencia. Pero como también dijimos anteriormente, la conciencia no se reduce tan sólo a este problema ‘difícil’: existen una gran cantidad de problemas vinculados de una u otra manera al problema central y que, para poder aportar una verdadera respuesta a qué sea la conciencia, deben ser afrontados tarde o temprano.¹⁸ Dedicaremos este punto a mostrar los problemas más relevantes a los que se enfrenta toda teoría de la conciencia y a los que, inevitablemente, debería poder aportar una respuesta en algún sentido. Pero antes, queremos ofrecer una breve contextualización que permita ver la profundidad que, desde la antigüedad, ha marcado a nuestro tema.

Desde la antigüedad, el tema relativo a la conciencia ha estado directamente vinculado con el alma. Ya en los estudios clásicos de Platón o Aristóteles vemos la presencia directa de elementos conscientes directamente vinculados a la posesión de un alma racional. A partir de los estudios de Descartes, como hemos indicado en el anterior apartado, el problema se reformula en términos de la relación existente entre un cuerpo (material) y una mente (racional).

Históricamente hablando, el problema de la conciencia tiene una larga tradición como un problema relativo a los aspectos puramente característicos de la sustancia mental. La cuestión principal era llegar a poder demostrar cómo, de algún modo, la propia mente era capaz de estar informada de las propiedades físicas externas al propio cuerpo (en tanto que se suponía una relación entre los elementos mentales del alma racional y los elementos físicos presentes en el mundo externo) y de las propiedades externas (relativas al cuerpo) e internas (relativas a la misma mente). La dilucidación de esta relación entre lo mental y lo físico permitía llegar a establecer un concepto de persona más o menos estable para, por un lado, permitir atribuirle racionalidad y, por otro, poder atribuirle libertad de acción. Desde este punto de vista, los actuales problemas de la conciencia (que en breve analizaremos) apuntan precisamente a intentar aportar una explicación más o menos plausible a dicha relación entre un cuerpo material y la (presunta) existencia de una mente, en tanto que metafísicamente separada de lo físico. Aunque no de un modo directo y explícito, como era más habitual antaño, la contemporánea reformulación de los problemas clásicos de la filosofía de la conciencia permiten, por un lado, ofrecer una imagen más realista de la realidad mental (en tanto tiene presente los desarrollos y avances de las ciencias físicas destinadas a la investigación y descubrimiento de los procesos

¹⁸ Por supuesto, no todos los filósofos están de acuerdo a la hora de aceptar la existencia ni del problema ‘difícil’ ni de los problemas ‘fáciles’. La mayoría de filósofos que sostienen de alguna manera una teoría fisicalista de la mente y de la conciencia afirman la necesidad de explicar ciertas dificultades a la hora de describir la conciencia, pero consideran que éstas son tan sólo eso, dificultades a explicar, y no inexplicables (empleando un lenguaje chomskiano, que dichas dificultades son tan sólo problemas y no misterios). Tarde o temprano, sostienen, será posible ofrecer una aproximación científicamente situada de todas esas dificultades y que, en definitiva, puede ofrecerse una explicación fisicalista del (supuesto) misterio de la conciencia. Un ejemplo relevante de la negativa a aceptar la existencia de dichos problemas ‘fáciles’ la encontramos en Lowe (1995).

cerebrales) y, por otro, evitar una serie de errores categóricos que, según parece, impiden (tanto metafísica como epistémicamente hablando) una explicación fiscalista de la conciencia y de sus problemas afines. Atendamos, pues, a esos problemas.

1.4.1. El problema de la propiocepción y la perspectiva en primera persona

Este problema surge en el seno de las teorías fiscalistas de la conciencia a la hora de tener que aportar argumentos explicativos a un hecho intuitivamente cierto: cómo podemos explicar que el contenido de las experiencias fenoménicas que percibimos sea físico si consideramos que los estados fenoménicamente conscientes son necesariamente privados y percibidos de manera interna y subjetiva por los sujetos perceptores. Parece más razonable concluir que, a pesar de que los objetos que permiten la existencia de estados mentales con contenido fenoménico son de naturaleza física, los mismos estados fenoménicos, sin embargo, no pueden ser físicos. Este problema referido a la subjetividad de las experiencias fenoménicas es un producto derivado de la distinción epistémica entre el acceso privilegiado en primera persona y el acceso objetivo en tercera persona a los estados mentales en general.¹⁹

Este problema parece dirigirse hacia la diferenciación ontológica entre dos clases de objetos metafísicamente diferenciables. Mientras que, por un lado, la realidad de las cosas mismas es física y el modo en que son percibidas es objetivo, por otro lado, la realidad de las experiencias conscientes con contenido fenoménico no puede ser física porque el modo en que son percibidas es diferente del modo objetivo en que percibimos las cosas físicas. Las experiencias fenoménicamente conscientes son percibidas subjetivamente, mientras que los objetos que las provocan son objetivamente describibles.

En este sentido, Nagel (1974) estableció una diferenciación entre el saber cómo siente un murciélago y realmente sentir como un murciélago. En última instancia, viene a concluir Nagel, nuestra ciencia actual puede llegar a descubrir y desarrollar de un modo eficiente y exhaustivo el mecanismo cerebro-nervioso que permite al sistema eco-localizador de los murciélagos captar su mundo circundante y permitir, así, explicar detalladamente el modo en cómo un murciélago percibe el mundo. Lo que jamás podrá hacer la ciencia, precisamente porque el acceso que tiene a esta información respecto del sistema perceptor de los murciélagos es descriptivo, objetivo y en tercera persona, es explicar cómo sería percibir a la manera de los murciélagos, algo sólo reservado para el acceso en primera persona (propiocepción) a las experiencias conscientes. En este sentido, lo que Nagel está estableciendo es una diferenciación entre *saber cómo* es percibir el mundo a través del mecanismo perceptor de un murciélago (un

¹⁹ Son muchos y variados los argumentos que desde una posición internista del contenido mental se han articulado respecto de la dificultad que las teorías externistas sobre el contenido mental tienen a la hora de rendir cuentas a la intuición cotidiana del acceso privilegiado al conocimiento de los propios estados mentales que proporciona la perspectiva de la primera persona. De uno u otro modo, todos acaban apelando a la necesidad de la introspección para afirmar la superioridad de dicho acceso en primera persona. Pero también muchos filósofos externistas han contraargumentado a esta objeción intentado conciliar externismo respecto del contenido mental y acceso privilegiado, como vimos en la anterior sección al analizar las dificultades de la TI, intentado mostrar que no existe ninguna incompatibilidad entre ellos. Un buen ejemplo de intento de conciliación entre ambos puede encontrarse, también, en McLaughlin and Tye (1998).

conocimiento puramente informacional y descriptivo) y realmente *sentir cómo* es el mundo a través del sistema perceptivo de los murciélagos (un conocimiento no-descriptivo, sino sensitivo, cualitativo).

Vistas así las cosas, parece querer decirnos Nagel, todos aquellos teóricos que sostengan que las experiencias conscientes con carácter fenoménico son físicas deberán explicar cómo el color, el dolor y el estado depresivo de ánimo, por ejemplo, y otros objetos mentales de experiencia pueden tener características tan sólo interna, privada y subjetivamente cognoscibles y experienciables. Pero todos aquellos teóricos que consideran que existe una diferencia ontológica entre los objetos cognoscibles de modo objetivo (es decir, los objetos con propiedades físicas) y aquellos cognoscibles tan sólo de modo subjetivo (es decir, los objetos de experiencia anteriormente mencionados) también estarán obligados a rendir cuentas de la manera en que ambos modos se interrelacionan entre sí y a explicar la base que sustenta la existencia de objetos independientes de la realidad física.

1.4.2. El argumento del conocimiento

En Leibniz (1840), nuestro autor defiende la posibilidad de construir una máquina cuya estructura permita producir pensamientos, percepciones y sentimientos. Ese engendro mecánico, nos dice Leibniz, parece funcionalmente idéntico a nosotros (casi como si fuera una réplica física de nuestro sistema mental consciente),²⁰ pero realmente carece de aquello que a los humanos nos hace conscientes. Es decir, aunque los engranajes que componen los mecanismos de nuestra máquina permiten la aparición de percepciones y sentimientos, realmente ésta carece de aquello verdaderamente importante: la posibilidad de sentirlos como tales. Nuestra máquina más bien parece poder simular estados mentales conscientes, pero realmente, nos dice Leibniz, carece del conocimiento necesario para poder darle importancia y significado a dichos estados: realmente carece de la posibilidad de ‘sentir en sus carnes’ dichos estados. La conclusión que Leibniz perseguía era diferenciar las causas físicas que permitían a los mecanismos físicos simular la posesión de estados mentales de los estados mentales conscientes que realmente poseen los seres humanos. Y dado que parece que la máquina es capaz de simular estados mentales conscientes sin llegar a sentirlos realmente, entonces Leibniz concluye que los estados conscientes deben tener una naturaleza diferente de aquella física. Nuestra máquina puede implementar diversos estados mentales y comportarse cómo si realmente tuviera estados conscientes; pero realmente no los tiene porque no sabe cómo sería tener un estado consciente: le falta la introspección que acompaña a toda experimentación de cualquier estado mental consciente.

Una reformulación actualizada del argumento leibniziano del conocimiento la encontramos en el experimento mental que nos presenta Jackson (1982 y 1986). Imaginemos a Mary, una muchacha que ha sido criada desde su más tierna infancia en un entorno carente de toda impresión cromática; sólo ha tenido acceso a contemplar el color blanco, el negro y la

²⁰ Este argumento del conocimiento originario de Leibniz también muestra una proto-formulación de otro de los problemas que veremos en el apartado 1.4.5., el problema de los duplicados físicos o ‘zombies’.

escala de grises que va entre uno y otro. (Podemos imaginar todo aquello que creamos necesario para conseguir hacernos una idea de cómo sería la vida cotidiana para Mary: siempre iría completamente vestida y con guantes, portaría permanentemente unas oscuras gafas de sol, todo aquello que pudiera leer o ver por televisión debería ser a través de fotocopias o de un televisor en blanco y negro, etc.). Un buen día, tras finalizar sus largos años de estudio sobre la ciencia del color, Mary se convierte en la mayor experta a nivel mundial sobre el color. Mary ha llegado a saber *todo* aquello que, desde el aspecto físico, puede saberse sobre el color. Pues bien, nos dice Jackson, Mary puede saber todo aquello que desde el punto de vista físico puede llegar a saberse sobre el color, pero siempre le faltará algo esencial para llegar a saberlo todo sobre el color. Lo que le faltaría a Mary es, precisamente, saber qué se siente al percibir el color. Es decir, aunque Mary sabe todo aquello que es necesario desde el aspecto físico para que se dé el color, a Mary realmente le falta el percibir el color para llegar realmente a saber qué es el color. Para defender su conclusión, Jackson aporta algún ejemplo: imaginemos que un buen día sus captores le regalan una rosa roja. ¿Qué sucede entonces? Para Jackson lo que sucede es algo tan sencillo como que Mary conoce algo nuevo cuando consigue ver por primera vez el rojo de la rosa: la visión del color rojo le permite percibir el rojo por primera vez y le permite conocer qué se siente al percibir el color rojo, algo que totalmente desconocía, según el argumento de Jackson, hasta el momento de percibir la rosa.

Muchos son los seguidores del argumento de Jackson, que acaban por aceptar la presencia de algo esencial en la percepción del color que permite diferenciarlo del conocimiento de los elementos físicos que consiguen causar la percepción del color. Pero muchos son también aquellos que consideran que argumentos como los de Jackson recaen en un internismo innecesario al otorgarle un papel excesivo y primordial a la introspección, por lo que prefieren optar por una explicación física y material de la percepción del color que les permite concluir que Mary realmente no llega a conocer nada nuevo cuando percibe por primera vez el color rojo de la rosa, ¡porque ya sabía todo aquello que debía conocer sobre el color! Para estos filósofos materialistas, básicamente lo que diría Mary en éste y en otros casos similares sería algo parecido a esto: ‘Vaya, mira. El color de esta rosa es rojo porque estimula la parte de mi cerebro que me permite percibir las cosas de color rojo, tal y como indican mis manuales y mis investigaciones corroboran’.²¹

1.4.3. El hiato explicativo

¿Cómo podemos explicar el hecho de que nuestro cerebro sea responsable de la producción de estados mentales con carácter fenoménico? ¿Cómo podemos explicar que los cambios en nuestro sistema físico permitan variaciones en nuestro sistema fenoménico (que intuitivamente parece tan diferente del sistema físico)? Como bien hizo notar T. H. Huxley,²²

²¹ Aunque en secciones posteriores analizaremos pormenorizadamente tanto el argumento de Jackson y algunas de sus variantes como contra-ejemplos al mismo, véase para una visión actualizada de esta temática el excelente volumen colectivo Ludlow, Nagasawa y Stoljar (2004).

²² “¿Cómo es que algo tan notable como un estado de conciencia pueda surgir de la irritación del tejido nervioso? Parece tan sorprendente como la aparición del Genio cuando Aladino frota su lámpara” (Huxley 1866).

parece existir una brecha explicativa (*explanatory gap*), un hiato, entre el modo en que podemos describir los cambios en el sistema perceptivo físico y los cambios en el sistema fenoménico que permite sentir esos cambios. Intuitivamente parece que existe algo inexplicable en la generación de estados mentales conscientes a partir de estados puramente físicos (Levine 1983 y 2007).

Pero, ¿realmente existe dicho hiato? Muchos filósofos consideran que no. Veamos un ejemplo. Imaginemos que conseguimos fabricar un dispositivo (habitualmente denominado ‘cerebroscopio’) que, conectado adecuadamente a la cabeza de un paciente, nos permite saber qué cambios se están produciendo en su cerebro. Imaginemos también que descubrimos que cuando nuestro paciente tiene un estado mental con carácter fenoménico, por ejemplo la percepción de algo rojo o la presencia de dolor, las neuronas identificadas para tal fin se excitan mientras que, en los momentos en los que no existe tal estado mental, permanecen en letargo. Parece plausible afirmar la existencia de una especie de mecanismo que permitiría explicar cómo la excitación del sistema cerebro-nervioso de nuestro paciente implementa un determinado estado con carácter fenoménico, anulando así la posibilidad de una brecha abierta entre la explicación física y la generación de contenido fenoménico consciente. La explicación científica basada en la ciencia neurofisiológica bastaría para rendir cuentas a la explicación de cómo se generan los estados fenoménicos: tan sólo deberíamos estudiar el modo en que la activación de ciertas neuronas implementan estados mentales con carácter fenoménico diferente según el caso.

Según estos filósofos materialistas, en el mundo natural la generación de estados, procesos o propiedades de nivel superior (esto es, aquello que a un nivel personal denominamos estados mentales con contenido fenoménico) puede ser explicado a partir de la implementación de estados, procesos o propiedades de nivel inferior, o primario (esto es, aquello que a un nivel sub-personal denominamos estados químico y micro-físicos del cerebro y el sistema nervioso central), demostrándose así que también la conciencia fenoménica es un fenómeno natural que forma parte del mundo físico, lo que permitiría una unificación en sentido explicativo entre los puntos de vista subjetivo y objetivo respecto de la naturaleza del contenido fenoménico.

Por supuesto, hay todo un cúmulo de filósofos que no aceptan esta explicación, para ellos, reductiva. En este sentido, consideran, si fuera verdad que los estados fenoménicos pueden reducirse a estados físicos, todavía quedaría por explicar cómo y por qué esto es así. En pocas palabras, todavía quedaría por explicar cómo los cambios objetivos y físicos del cerebro y del sistema nervioso central son capaces de generar las experiencias y los sentimientos subjetivos que todos parecemos percibir y a los cuales tan sólo parece que tenemos un acceso directo a través de la introspección.

1.4.4. El problema de la causalidad de lo fenoménico

Como hemos dicho anteriormente, parece que la explicación reductiva aportada por el fisicalismo estándar tiende a identificar estados mentales con estados físicos, por lo que todo cambio en nuestros estados mentales y/o todo mecanismo capaz de causar un estado mental

consciente deben ser entendidos como físicos. Pero también existen teorías que, evitando todo eliminativismo respecto de lo mental, no identifican los estados mentales con estados físicos, pero sí consideran que todo estado mental superviene a estados físicos (en el sentido de Kim, anteriormente mencionado, ver nota 17). Vistas así las cosas, esto es, si consideramos plausible que la explicación fiscalista es incompleta y no consigue rendir cuentas del modo en que ocurren los cambios en nuestro aparato mental (esto es, no consigue explicar de modo adecuado cómo y por qué cierto tipo de estímulos externos y/o internos provocan la aparición en determinado tipo de individuos de cierto tipo de estados mentales con contenido fenoménico), ¿cómo explicar la aparición de dichos estados conscientes?

Un intento de explicación no-fiscalista lo aportan tanto Nagel (1974) como Jackson (1982). La conclusión de ambos trabajos persigue un objetivo claro: no existe una identificación posible (ni tan siquiera una sobrevenida posible) entre los estados físicos de un individuo y sus estados mentales. Esto es, la posibilidad de un cambio en sus estados físicos no permite explicar un cambio en sus estados mentales conscientes; los estados físicos no causan los estados mentales conscientes, y viceversa. Al sostener esta tesis, tanto Jackson como Nagel recaen en el epifenomenalismo respecto de lo mental. Si, como ambos sostienen, las experiencias sensitivas tienen un contenido fenoménico que es epifenoménico, esto es, que no tiene ningún poder causal sobre la realidad física, entonces su mera existencia no puede explicar cómo ocurren los cambios en los estados conscientes ni, mucho menos, puede explicar por qué los seres conscientes tienen determinadas sensaciones ante la presencia de determinados estímulos, por lo que podemos prescindir de ellos para aportar una explicación de la conciencia. Una solución propuesta sería la posibilidad de eliminarlos (Cf. Dennett 1988a y 1991: 412-414).

1.4.5. El problema de los duplicados físicos

Otro de estos escollos aparentemente insalvables que se han expuesto como traba para una explicación fiscalista de la conciencia consiste en una impresión intuitiva dependiente de la idea de sentido común que nos dice que no es posible establecer una teoría física total de la conciencia porque siempre quedaría algo que dicha teoría sería incapaz de explicar: la esencia de lo consciente, su aspecto cualitativo. Para reforzar dicha tesis, algunos de los filósofos *misterianistas* (a los que hemos apelado anteriormente) lanzaron una bomba intuitiva que muestra en qué consistiría este elemento a partir de la comparación entre un ser que tiene experiencias conscientes y un ser que, aunque aparentemente se comporta de modo tal que parece actuar conscientemente, realmente carece de ese contenido consciente del cual sí dispone un ser con conciencia. A este experimento mental se lo ha denominado *la corazonada 'zombie'*.²³

²³ Aunque el problema de los 'zombies' es antiguo, debemos su actual formulación a las discusiones nacidas en el seno de las Conferencias Tucson sobre conciencia del año 1996, donde se tuvieron en cuenta los análisis que Dennett (1991) y Chalmers (1996) habían llevado a cabo al respecto. Es Dennett (2001) quien rebautiza el problema como la corazonada 'zombie'. Pero no deberíamos olvidar que la primera presentación de la problemática de los *zombies* en vistas a criticar la imposibilidad de un dualismo de lo mental en el seno de una discusión respecto de la conciencia se la debemos a Kirk (1974a y 1974b). Véase también Campbell (1970). De todos modos, la problemática respecto de

Comenzaremos nuestra discusión presentando a los ‘zombies’. Según la literatura al uso, un ‘zombie’ es una réplica exacta molécula por molécula de cualquier individuo consciente (esto es, físicamente indistinguible de ese mismo ser consciente) que, aunque su conducta y comportamiento son socialmente aceptables (e incluso parece responder causalmente a los estímulos externos e internos), no posee experiencias verdaderamente conscientes.

Los argumentos acerca de la posibilidad de la existencia de seres carentes de experiencias conscientes se remontan a los escritos de Descartes y Leibniz, autores que postularon la existencia de engendros capaces de confundirnos a través de su conducta simuladamente consciente pero que, realmente, respondían a principios mecánicos carentes de toda cualidad consciente. Tal vez la descripción actual más conocida de qué sea un ‘zombie’ la ofrece David Chalmers:

“Esta criatura es idéntica a mí molécula por molécula e idéntica en todas las propiedades de bajo nivel postuladas por una física terminada, pero carece por completo de experiencia consciente” (Chalmers 1996: 133).

Un ‘zombie’, por tanto, por definición, no sería más que un conjunto funcional de materia (orgánica o inorgánica) que tan sólo parece actuar del modo en que cualquiera de nosotros actuaría cuando tenemos alguna experiencia consciente pero que, sin embargo, no dispone realmente de aquello que, parece, caracterizaría nuestra conciencia: carecería de *qualia*, de contenido cualitativamente consciente.

Lo que pretenden hacernos ver todos aquellos filósofos (habitualmente recibirían el nombre de chauvinistas, por considerar que tan sólo los seres humanos estarían dotados de dicho carácter cualitativo) que apuestan por la corazonada ‘zombie’ es que la mente humana parece poseer un elemento esencial que la hace aparecer de modo especial, tanto desde un punto de vista metafísico como epistemológico: la mente humana dispone de experiencia consciente que la diferenciaría de todo aquello que carece de dicha cualidad (aporte metafísico) y que dicha cualidad experiencial consciente estaría vedada a un acceso más allá de la autoridad en primera persona (aporte epistémico), y dicha cualidad es lo que permitiría diferenciar un estado consciente de uno que no lo es. Pero éste, como sostendremos a lo largo de todo este trabajo, no es el caso.

Aunque en un primer momento parezca que un ‘zombie’ sería cualquier ser que se comporte de un modo normal pero que carece de toda cualidad experiencial consciente, debemos diferenciar (como hacemos a continuación) distintas clases de individuos zómbicos (o duplicados físicos) según el alcance del argumento que sus autores han pretendido explotar.

1.4.5.1. El gemelo de Chalmers

El ejemplo de ‘zombie’ que actualmente se tiene más presente en las discusiones es el esbozado en Chalmers (1996), que lo presenta del siguiente modo. Nos dice que imaginemos un

los duplicados físicos de seres conscientes no se reduce tan sólo a la existencia de los *zombies*, como mostramos más abajo.

gemelo suyo, idéntico físicamente hasta en su más mínimo detalle, que se comporta como si tuviera experiencia consciente pero que, realmente, carece de todo contenido de conciencia. Como es físicamente idéntico a Chalmers, porque la selección natural ha hecho evolucionar del mismo modo a ambos individuos (Chalmers 1996: 163-164), ello se reflejará de dos modos: tendrá la misma actividad funcional que éste cuando tiene experiencias conscientes, procesando la información adecuadamente, respondiendo acertadamente a los *inputs* sensoriales con *outputs* conductuales, etc., pero también será psicológicamente indistinguible de Chalmers en tanto que es capaz de percibir, de degustar, de oler, etc., aunque carecerá de la experiencia consciente que, normalmente, acompaña a dichas acciones. (Ciertamente, los anteriores dos puntos están relacionados, pero son distintos).

Como el mismo Chalmers afirma, la existencia real de un ser de estas características es un tanto imposible, pero lo que interesa no es su realidad física sino que lo que aquí verdaderamente nos interesa es mostrar, cuanto menos, que sea lógicamente coherente la posibilidad de la existencia de seres idénticos físicamente a nosotros que, sin embargo, carecen de un contenido fenoménico.²⁴

1.4.5.2. Robots

Aunque durante décadas ha estado presente en la literatura y en el cine, y aunque desde Turing y el surgimiento de la Inteligencia Artificial se ha tenido muy presente, también Chalmers comienza a conceder cierta relevancia filosófica a la posible existencia de robots que funcional y psicológicamente sean idénticos a un individuo consciente (pero que ya no serían productos evolutivamente idénticos) y que, a pesar de que sus estados internos son capaces de producir salidas electro-químicas en respuesta a entradas sensoriales, carecen de experiencia consciente de todo aquello que son capaces de hacer (Chalmers, 1996: 323ss.).

Un argumento que refuerza esta intuición es la habitación china (Searle 1980, 1984: 33-48 y 1992: 77-94). El escenario de este experimento mental es el siguiente. Imaginemos que tenemos una habitación cerrada en la que se encuentra un individuo que tan sólo dispone de

²⁴ Es habitual encontrar argumentos materialistas contra los ‘zombies’ que apelan a la imposibilidad metafísica de los mismos a partir de la crítica a la noción de necesidad presente en la teoría kripkeana. Chalmers responde a estas críticas a partir de un análisis bidimensional (que divide en dos tipos de intensiones los mundos posibles) de la teoría de la necesidad de la identidad que, aunque le imposibilita salvar la existencia metafísicamente posible de ‘zombies’, le permite mantener su existencia lógica (su posibilidad, al menos, conceptual), estableciendo lo que conocemos como Argumento de la Concebibilidad (Chalmers 1996: 175ss.):

(AC) 1. Los *zombies* son concebibles.

2. Todo aquello que es concebible es posible, por tanto

3. Los *zombies* son posibles.

Muchos de los actuales intentos materialistas de argumentar contra la corazonada ‘zombie’ atacan a (AC) porque consideran ilegítimo el paso de la concebibilidad a la existencia (más próximos a mecanismos *Deus ex machina*) e, incluso, algunos de ellos consideran filosóficamente irrelevante (o poco interesante) la posibilidad de la existencia de ‘zombies’. Desde mi punto de vista, el (AC) no es del todo plausible, pero considero importante hacer frente a la posibilidad conceptual de la existencia de los ‘zombies’ porque, como dicen sus propios defensores, si es cierta la posibilidad (aunque tan sólo sea como contra-ejemplo) de la existencia de los ‘zombies’, entonces el materialismo se (de)mostraría falso. Pero, como intentaremos demostrar más tarde (véase sección 6.2.1.1.), éste no creemos que sea, necesariamente, el caso. Uno de los mejores argumentos actuales contra (AC) puede encontrarse en la primera parte de Kirk (2005) y en Pineda (2008). Un buen volumen colectivo que recoge textos a favor y en contra de (AC) puede encontrarse en Carpintero and Macià (2006).

acceso al mundo exterior a través de dos rendijas: una de entrada (por la cual se le introducen láminas de cartulina en la que aparecen impresos símbolos en un idioma que desconoce (pongamos por caso, el chino) y una de salida (por la cual debe deslizar también láminas que lleven impresos símbolos también en chino que, sin embargo, respondan de manera adecuada a las órdenes que aparecían impresas en las láminas impresas de entrada). Para poder dar correcta cuenta de dicha tarea, nuestro prisionero tan sólo dispone de un enorme manual de instrucciones de manejo del chino que indica qué símbolos serían las salidas adecuadas a las entradas escritas en chino, aunque él desconoce por completo aquello que está haciendo. Según muestra este experimento, y según concluye Searle, en ningún caso podemos afirmar que el individuo encerrado en la habitación es conocedor del chino, porque aunque es capaz de aportar las respuestas de salida correctas a las entradas escritas en chino que se le introducen por la primera rendija, en ningún caso comprende aquello que dichas láminas en chino le están diciendo. Del mismo modo, nos dice Searle, actúa un instrumento mecánico como un computador.

Aunque la pretensión de Searle (1980) no es estrictamente hablando un argumento acerca del privilegio de lo consciente, sino un argumento en contra de la que denomina Inteligencia Artificial Fuerte que impide reducir semántica a sintaxis en beneficio de la mente humana, que la salvaría de una explicación funcionalista computacional, en última instancia, lo que el experimento de la habitación china nos aportaría es una concepción de la mente humana como privilegiada respecto de la (posible) mente de otros seres (mecánicos). El de Searle es un argumento que pretende diferenciar a los seres realmente pensantes (o conscientes de lo que piensan, en el sentido en que tan sólo aquellos seres cuyos pensamientos estén dotados de semántica tienen realmente experiencia consciente sobre ellos) de todos aquellos seres que sólo parece que piensan. Porque tendrían un acceso sintáctico sobre su contenido mental, siendo un argumento que presenta una versión de silicio de los ‘zombies’ (Searle 1992).

Siguiendo con el espíritu del argumento de Searle, Chalmers asimila los resultados del experimento de Block (1978), la nación china: imaginemos que se conecta a todos los habitantes de China mediante equipos de radio de tal modo que simulen la organización funcional que presenta un cerebro, asignándole incluso a cada individuo una función concreta dentro del sistema representado. Imaginemos que se les proporciona una entrada sensorial a la que responden funcionalmente con una salida a partir de la computación pertinente de los datos manejados. Siguiendo lo que ya había establecido inicialmente Block, Chalmers concluye que, aunque la salida responda perfectamente a la entrada, ello no significaría que el conjunto de individuos chinos que conforman el sistema entraron en un determinado estado mental del mismo modo que decimos que el cerebro se encuentra en un determinado estado mental, por lo que sería conveniente concluir que el cerebro proporciona genuinos estados mentales frente a otros sistemas que tan sólo pueden simularlos.

1.4.5.3. Científicos marcianos

Un argumento a favor de una teoría materialista de la conciencia podemos encontrarlo en Lewis (1978). A pesar de que afirma que aquello que caracteriza a la conciencia es su

carácter fenoménico (Lewis 1990: 473ss.), Lewis lo reduce a un conjunto funcional informacional cuya misión sería activar aquellos estados físicos con los que se identificaría cada estado mental consciente. Éste sería el caso, por ejemplo, de la vida mental de un ser extraterrestre. Aquello que caracterizaría el dolor de un marciano (aunque la base de su estructura fuera de, por ejemplo, silicón o azufre y no, como en los humanos, el carbono) sería la activación de aquellas zonas de su sistema cerebral y nervioso (biológica y funcionalmente hablando) que ocasionan en el individuo la sensación de dolor, del mismo modo que podemos decir que (según nuestra particular distribución subdoxástica) tenemos un dolor cuando tenemos las fibras C estimuladas. Contra este argumento reduccionista se suele aportar la siguiente evidencia de sentido común: aquello que cuenta como dolor no es la simple activación de un estado cerebro-nervioso, sino que realmente el individuo esté sintiendo dolor.²⁵

1.4.5.4. El hombre del pantano

Es hora de atender al último de los ejemplos de duplicado físico que me gustaría resaltar, el hombre del pantano:

“Supongamos que un rayo cae sobre un árbol muerto en una zona pantanosa; yo estoy de pie junto a él. Mi cuerpo es reducido a sus elementos, mientras que, por pura coincidencia (y a partir de moléculas diferentes), el árbol se convierte en una réplica física de mí mismo. Mi réplica, El Hombre de los Pantanos, se mueve exactamente como yo lo hacía; de acuerdo con su naturaleza, abandona los pantanos, se encuentra con mis amigos y parece reconocerlos, y en apariencia responde a sus saludos en inglés. Se traslada a vivir a mi casa y parece escribir artículos sobre la interpretación radical. Nadie puede notar la diferencia.

Sin embargo, *hay* una diferencia. Mi réplica no puede reconocer a mis amigos; no puede reconocer cosa alguna, puesto que nunca la conoció anteriormente. No puede saber los nombres de mis amigos (aunque desde luego parece saberlos), no puede recordar mi casa. No puede querer decir lo mismo que yo con la palabra ‘casa’, por ejemplo, puesto que el sonido ‘casa’ emitido por él no fue aprendido en un contexto que le diese su significado correcto, o algún significado siquiera. En realidad, no veo cómo se podría decir que mi réplica quiere decir algo con los sonidos que emite o que tiene pensamientos” (Davidson 1987: 46ss).

Aunque la pretensión de Davidson en este trabajo era cuestionarse, desde una perspectiva externista, la presunta incorregibilidad de los pensamientos en primera persona, su argumento permite aportar un elemento más a nuestra discusión. Aquello que diferenciaría al Davidson original de su réplica sería que, aunque Davidson del Pantano (así llamaremos a su réplica) se comporta de modo tal que parece tener los mismos pensamientos conscientes, no tendría realmente el mismo contenido experiencial consciente que el primero porque no comparten la misma historia causal cognitiva (al modo en que lo entendía Putnam (1975) en sus experimentos de la Tierra Gemela, lo que parecería permitirle aportar argumentos a favor de un acceso externista al contenido semántico), sino sólo la distribución físico-funcional, de lo que se concluye la imposibilidad de reducir lo mental a lo físico, entendido localmente como estados neurofisiológicos (abriendo de nuevo el hiato explicativa y cerrando las puertas a la revisión externista del contenido mental).

²⁵ La problemática respecto del dolor ha cobrado fuerza en los últimos años como un derivado de los estudios respecto de las sensaciones corporales. Atenderemos a esta noción más tarde. Véase sección 4.2.3.

1.4.6. El espectro invertido

Como nos dice Dennett (1991: 401), la posibilidad de que existan experiencias fenoménicas invertidas es una de las ideas filosóficas con más tradición y con más fuerza de toda la historia de la filosofía de la mente.

La primera discusión de este problema se la debemos a Locke:

“Nadie podría acusar de falsedad a nuestras ideas simples si, por una diferencia en la estructura de sus órganos, el mismo objeto pudiera producir en la mente de hombres diferentes ideas distintas. Esto es, si un mismo objeto pudiera producir que la mente de un hombre lo viera como ‘violeta’ y la mente de otro hombre lo viera como ‘carmesí’, y viceversa. Si esto sucediera, nunca lo sabríamos porque nunca podríamos introducirnos en la mente de dichos hombres para observar el modo en que sus órganos perciben” (Locke 1689: II, xxxii, 15).

En pocas palabras, a lo que Locke está apuntando es a la dificultad existente a la hora de determinar si dos sujetos conscientes normales en condiciones normales perciben la misma experiencia subjetiva ante la misma sensación.

Tradicionalmente, el escenario del problema del espectro invertido es el siguiente. Imaginemos una persona, digamos Miguel Ángel, con un sistema visual peculiar: su experiencia del color es sistemáticamente invertida con respecto a los demás. Esto es, a Miguel Ángel se le aparecen de color rojo los objetos que para todos los demás son de color verde (y viceversa), por ejemplo; es decir, ante el mismo objeto intencional, Miguel Ángel tiene un *qualia* diferente al nuestro (Shoemaker 1991). Pero esta peculiaridad ha pasado desapercibida porque no ha trascendido a su comportamiento lingüístico: Miguel Ángel ha aprendido a llamar ‘objetos de color verde’ a todos aquellos objetos que a él se le aparecen como de color rojo, ya que en cierto sentido ha aprendido a emplear los términos de color del mismo modo que nosotros.

Alguien podría decir que lo que ocurre es que el sistema visual de Miguel Ángel está conectado de manera diferente al nuestro, es decir, podríamos presuponer que su sistema visual es funcionalmente diferente del nuestro (Block 1978). Pero para que el experimento mental del espectro invertido tenga éxito, lo que debemos suponer es que el sistema visual de Miguel Ángel es funcionalmente idéntico al nuestro y que lo que varía sustancialmente es su fenomenología, esto es, el modo en que a él se le aparecen las cosas. Si esto fuera así, si existiera una diferencia subjetiva en el modo en que los sujetos podemos sentir las experiencias del color pero no existiera una diferencia funcional en nuestro sistema visual, entonces parece plausible la existencia de experiencias invertidas a partir de la presencia de sistemas visuales funcionalmente idénticos (Cf. Lycan 1973; Shoemaker 1982).

Existe otro escenario habitual, el ofrecido por Block (1990), que emplea las intuiciones de Putnam sobre la Tierra Gemela. Imaginemos a Josep, que ha sido transportado sin que él se diera cuenta a una Tierra Gemela, idéntica en todo a nuestra Tierra menos en la experiencia del color verde. Aunque también allí hablan castellano, cuando ellos allí dicen ‘verde’ lo que realmente quieren decir es ‘rojo’, es decir, tienen invertido su sistema visual respecto del de Josep (y viceversa), dado que cuando él dice ‘verde’ quiere decir ‘verde’ y no ‘rojo’. ¿Qué ocurre en este caso? Para Block, los habitantes de la Tierra Gemela deben tener una distribución

visual funcionalmente diferente de la de Josep, porque no se refieren a la misma experiencia ante la presencia del mismo color. Para él, el espectro invertido no es posible: una diferencia en la experiencia subjetiva del color presupone una diferencia funcional en el sistema visual. Una conclusión similar se extrae de Rey (1998).

Filósofos externistas, como Dretske, consideran que el argumento de Block es falso porque se excede a la hora de formular sus argumentos. Para Block, la teoría funcionalista permite explicar la adquisición de *qualia* a partir de la determinación estrecha del contenido mental, pero el contenido cualitativo no tiene porqué ser estrecho, sino que puede ser entendido en sentido amplio (en el sentido en que lo defiende Dretske 1995), por lo que puede darse el caso que dos individuos molecularmente indistinguibles tengan experiencias diferentes ante la misma percepción (Dretske 1996).

La que hemos presentado en las últimas líneas es la versión intersubjetiva del espectro invertido. Esto es, la versión en la cual se supone la diferencia en la experiencia fenoménica subjetiva entre un sujeto y todos los demás a la hora de percibir el color del mismo objeto externo. Pero existe también una versión intrasubjetiva. En este segundo caso, lo que debemos imaginar es que un mismo sujeto, pongamos en este caso que dicho sujeto es Borja, tiene experiencias subjetivas diferentes ante el mismo objeto intencional en dos momentos diferentes. Pero del mismo modo, y aunque es posible afirmar que en el caso intrasubjetivo el sujeto sería subjetivamente sensible al cambio que se produce en su campo cualitativo, el problema radica en la misma cuestión, ¿pueden existir dos personas (o una misma) que a pesar de ser funcionalmente idénticas tengan experiencias subjetiva y fenomenológicamente invertidas respecto del mismo objeto intencional?

1.4.7. La variación estándar

Aunque este problema está emparentado con la anterior problemática (incluso podría considerarse como una variante inter-subjetiva del espectro invertido), presentamos ahora un problema relativamente diferente. A diferencia del anterior, que priorizaba la posibilidad de la diferencia físico-funcional capaz de causar una determinada experiencia con contenido fenoménico, este problema se refiere directamente a la posibilidad de que una misma sensación o experiencia pueda parecer diferente a dos individuos *a priori* perceptivamente *normales*, se refiere a una diferencia puramente fenoménica en la manera en la que tienen acceso al conocimiento de una vivencia concreta y no ya a su implementación funcionalmente dependiente. Si fuera cierta esta afirmación (es decir, si se demostrara la existencia de diferente contenido de conciencia en los dos individuos dada una misma experiencia), entonces podría llegar a afirmarse la existencia de un aspecto puramente subjetivo en la conciencia: el contenido fenoménico de una experiencia. Este problema es el denominado problema de la Variación Estándar.

Pongamos un ejemplo recurrente dentro de la literatura filosófica: el problema del verdadero azul. Parece ser que es comúnmente aceptado que la gente varía en su modo de tener experiencias según su edad, su sexo e, incluso, su estatus social. Por ejemplo, parece ser que

cuando a alguien, digamos Vicente, se le presenta una cierta muestra de color, digamos la muestra de azul 527 de cualquier catálogo, a él le puede parecer como azul *puro* o *verdadero*; mientras que a otra persona, digamos Andrés, le puede parecer que la misma muestra de azul 527 pertenece a un azul ligeramente teñido de verde. Intuitivamente, suponemos que al menos uno de los dos debe estar equivocado, dado que no podemos esperar que la muestra de azul 527 sea a la vez “verdadero” azul y azul ligeramente verdoso. Pero si tenemos en cuenta, como nos dice el clásico trabajo Block (1999), que no existen clases privilegiadas de perceptores, entonces, ¿cómo podemos afirmar que uno de los sujetos perceptores anteriormente mencionados se equivoca en su experiencia de la muestra de azul 527 mientras que el otro está en lo cierto?

Este problema, que habitualmente es tenido como un problema referente a la objetividad de los tipos fisicalistas relativos al color o al modo en que nos representamos el carácter fenoménico de la experiencia de color, aparece como más general en el momento en que atendemos a la problemática de tener que adoptar un punto de vista sobre el diferente papel causal que una determinada experiencia de color tiene sobre dos sujetos distintos pero perceptivamente normales.

Podríamos decir, como haría un fisicalista, que tanto Vicente como Andrés están equivocados, que la muestra de azul 527 realmente es otro tipo de azul del que ellos perciben, dado que lo que tiene color es el objeto exterior y físico, y éste implementaría el mismo estado mental en los dos sujetos, por lo que ambos deberían ver el mismo color; si no es así, se debe a que ambos están en un error a la hora de percibir el color de la muestra. Pero, como hemos visto, sí que es posible que dos sujetos tengan experiencias diferentes ante la misma sensación.

También podríamos decir, como haría un internista, que nada de lo que podemos observar tiene color, ya que el color es algo que nosotros proyectamos sobre el mundo (que son cualidades secundarias en el sentido disposicional de Locke), por lo que es posible que ambos estén en lo cierto. Pero si aceptamos que nada en mi cabeza tiene color (Byrne and Tye 2006; Dennett 1991: 382), entonces esta posibilidad debe ser desechada.

Entonces, si ambos no pueden estar equivocados ni ambos pueden llevar razón, ¿en que sentido podemos decir que un sujeto experimenta una sensación de color y que ésta puede ser diferente dependiendo del sujeto que la percibe?

1.4.8. La visión borrosa

El problema de la visión borrosa se refiere a la posibilidad de que un sujeto normal, con un sistema perceptivo normal, en un momento determinado, perciba de manera borrosa aquello que se presenta ante su campo visual. Dicha visión puede aparecer como distorsionada por efecto de algún utensilio (pensamos en unas gafas que producen el emborronamiento de la visión cuando se llevan puestas, por ejemplo) o bien por el efecto momentáneo de algún tipo de anomalía en su sistema visual o cerebral. Este tipo de situaciones plantea un problema porque, parece ser, en un momento determinado la visión del sujeto queda relegada a contemplar el mundo a partir de esta anomalía. La explicación de dicha dificultad, de la posibilidad de que en

un momento determinado al sujeto le pueda aparecer como borrosa su visión, nos vincula directamente con la posibilidad de tener que explicar en qué sentido su captación del mundo cambia radicalmente frente a otros sujetos que no tienen dicha dificultad impuesta por la peculiaridad de emborronamiento de su visión, siendo posible contemplarlo como una variante del espectro invertido, sólo que aquí la visión cambia por completo tras el emborronamiento (aunque sea por un instante) del campo visual.

1.4.9. El problema de la transparencia

Supongamos que nos encontramos frente al famoso cuadro *Paseo a la orilla del mar* de Sorolla. Experimentamos una rica variedad de formas y colores. Para experimentar adecuadamente nuestra percepción del cuadro, prestamos atención a nuestra experiencia visual y a su fenomenología. ¿Qué sucede? ¿Qué es lo que estamos haciendo? Según una gran tradición filosófica (destacaríamos, por ejemplo, el intencionalismo brentiano), lo que sucede es que estamos prestando atención al cuadro y a su gran variedad de detalles, que nosotros estamos prestando atención a algo que está fuera de nosotros (al cuadro) y a varias cualidades externas pertenecientes a dicho objeto (las características físicas del cuadro, sus colores, sus formas, etc.), y éstos permiten que nosotros experimentemos dicha sensación visual de manera fenoménicamente subjetiva. Por otro lado, existen autores (especialmente aquellos englobados dentro de corrientes representacionales) que consideran que el modo en que subjetivamente sentimos la experiencia fenoménica de la visión del cuadro, el modo en que se nos aparece dicha experiencia fenoménica, no depende de la experiencia del objeto externo o de sus cualidades, sino que depende de que estamos teniendo una determinada experiencia visual con un cierto tipo (representacional) de carácter fenoménico (Moore 1922: 22; McGinn 1982: 13; Harman 1990; Shoemaker 1990; Tye 1991: 119; Tye 1995: 30-31; Tye 2007: 30-31). Pero, ¿realmente es así?

Si, como los intencionalistas, consideramos que directamente sentimos la experiencia a partir del objeto externo, entonces parece que perdemos la riqueza subjetiva que intuitivamente observamos a la hora de tener experiencias con contenido fenoménico. Pero si, con los representacionistas, consideramos que no somos conscientes de la experiencia misma, sino que somos conscientes de estar teniendo una determinada experiencia con un contenido fenoménico concreto, entonces parece que perdemos la posibilidad de que sea el mundo el que proporcione dichas experiencias, recayendo de nuevo en la introspección el peso de la prueba y, en este sentido, el problema recaería en determinar qué cantidad de experiencia interna sería necesaria para considerar consciente una experiencia (convirtiendo el problema de la transparencia en un problema clásico de sorites).²⁶

²⁶ La paradoja del sorites se refiere a la antigua dificultad que se encuentra a la hora de afirmar cuántos granos de arena hacen falta para decir que tenemos un montón de arena (de ahí su nombre, del griego *sorites*: montón). Del mismo modo, también existe una dificultad aplicada concretamente al problema de la fenomenología de la mente, en tanto que aparece un problema similar aplicado a los seres conscientes. Desde este punto de vista, el problema radicaría en determinar qué cantidad de estados fenoménicos debe poseer un individuo para ser considerado como un ser consciente o, al menos, fenoménicamente relevante. Cf. Tye 1995 para una formulación de este problema.

Si tenemos presente que, a menudo, empleamos descripciones diferentes para describir los contenidos cualitativos que tenemos de determinadas experiencias, en este sentido deberemos admitir que nuestras experiencias son transparentes porque siempre tendremos la posibilidad de describir de manera diferente una determinada sensación. La dificultad reside, creemos, en el hecho de que no siempre podemos afirmar la existencia de un acceso consciente a dicho contenido, por lo que parecería entonces difícil admitir que toda experiencia sea transparente al impedirse de este modo la posibilidad de una descripción adecuada de la experiencia concreta que se está teniendo en ese momento.

1.4.10. El problema de la luminosidad

Este problema es relativamente reciente (además de ser una versión epistémica fuerte del argumento ontológico de la transparencia que hemos visto anteriormente). Según Williamson (1998) y, especialmente, Williamson (2000: cap. 2), debemos reconocer que es difícilmente aceptable suponer que todo estado mental sea luminoso; es decir, que sobre todo estado mental el sujeto, de manera consciente, sostenga una posición de conocimiento y que, además, dicho conocimiento sea él mismo un estado mental que el sujeto consciente posea. Esta conclusión se extrae a partir de la diferenciación que Williamson realiza entre el creer y el conocer. La creencia es un estado mental interno que relaciona al sujeto con un objeto del mundo, mientras que el conocimiento es una estrategia explicativa que nada tiene que ver con la creencia ni con la certeza subjetiva de la posesión de estados mentales. Parece ser que requerimos algo diferente de la creencia para conocer. Aunque la teoría de Williamson es una crítica a la definición de conocimiento como creencia verdadera justificada, tiene implicaciones en el tratamiento del contenido mental, especialmente el fenoménico. Del mismo modo, parece que requerimos de algo más que de la propia vivencia, de la introspección, para saber aquello que fenoménicamente estamos experimentando.

En este sentido, consideremos el siguiente ejemplo. Abrimos los ojos después de haber dormido toda la noche. La madrugada ha sido especialmente fría y, ahora, cuando nos disponemos a darnos la pertinente ducha matutina, observamos que, como es costumbre, el ventanal del baño se ha quedado abierto durante toda la noche. Por supuesto, el frío de la calle también está presente en nuestro baño. Sentimos frío. Antes que nada, cerramos la ventana y conectamos una pequeña estufita de aire caliente para evitar congelarnos en la ducha. Paulatinamente, el ambiente dentro del baño se va caldeando. Es curioso, porque ahora mismo, aunque notamos todavía el frío que durante la noche ha inundado el baño, también sentimos el calor que proporciona nuestra estufita. ¿Qué ocurre realmente? ¿Sentimos frío o sentimos calor? ¿Podemos diferenciarlos? ¿Podemos saber realmente cuando sentimos uno y cuando sentimos el otro? ¿Qué cantidad de vivencias requerimos para decir que lo que sentimos es calor y qué

cantidad se requiere para decir que lo que sentimos es frío? (De nuevo, la formulación del problema parece adquirir la forma del viejo problema sorites).²⁷

1.4.11. El problema de la localización del contenido fenoménico

Según la psicología popular, parece que las experiencias con contenido fenoménico son necesaria y esencialmente privadas y subjetivas, que nosotros no podemos sentir vuestras experiencias y vosotros no podéis sentir las nuestras; parece que la explicación más plausible sería considerar estas experiencias como sensaciones en tanto que objetos mentales. Pero esta interpretación se encuentra con serias dificultades cuando tiene que aplicarse a determinadas partes del cuerpo. Por ejemplo, tiene muchos problemas cuando tiene que localizarse en algún lugar concreto porque, ¿dónde podemos encontrar dichas experiencias? Si, por ejemplo, tenemos un dolor en la espalda, parece plausible considerar que el dolor está en nuestra espalda. Pero si hemos considerado que el dolor es un objeto mental no podemos a la vez decir que se encuentra en nuestra espalda, porque deberá pertenecer a nuestra mente y no a un objeto material como lo es nuestra espalda, perfectamente localizable en la parte trasera de nuestro cuerpo.

El sentido común todavía podría insinuar que, en casos de experiencia fenoménica, como es el caso del dolor, que éste se encuentre ‘en’, por ejemplo, nuestra espalda debe ser entendido en un sentido especial, en un sentido causal; que el dolor y la sensación de dolor de nuestra espalda son una y la misma cosa porque una causa la otra, independientemente de que no podamos localizar espacialmente el dolor en nuestra espalda. Pero esto tampoco parece muy plausible por dos motivos: el primero, porque de ser así, sólo podríamos conocerlo de modo *a posteriori* a partir de la investigación empírica (algo que no sería posible porque tan sólo podríamos realizarla sobre el cuerpo y no sobre la mente), y el segundo, porque el sentido de ‘en’ que emplean es un tanto rebuscado, totalmente alejado del uso ordinario de ‘en’.

Pero no les funcionan las cosas mejor a otros que consideran plausible la posibilidad de que el dolor se encuentre ‘en’ el cuerpo porque es una representación que se proyecta sobre él (Maund 2006; Aydede 2006) porque esta respuesta implicaría que es la representación misma lo que realmente se está sintiendo y no el dolor que tengo. Además, esta concepción también implicaría situaciones divertidas (aunque imposibles) como la siguiente: imaginemos que tenemos un dolor en nuestro pulgar (nos hemos golpeado el dedo con un martillo al intentar clavar un clavo) y que, como acto reflejo, introducimos nuestro dedo en nuestra boca con la intención de calmar el dolor que sentimos. Según la interpretación representacionista anteriormente citada, el dolor en tanto representación se proyectaría ‘en’ nuestro dedo, pero como nuestro dedo está ‘en’ nuestra boca, donde realmente tendríamos el dolor es ‘en’ nuestra boca y no ‘en’ nuestro dedo.

²⁷ Un problema añadido al del sorites fenoménico podría ser, también, la posibilidad de la existencia de conceptos fenoménicos vagos. Cf. Papineau 1993 para una tesis a favor de la existencia de dichos conceptos y Evans 1982 para el caso contrario.

Un problema añadido con el que se encontraría aquel que defendiera esta concepción es con la posibilidad de equivocarnos en la atribución del propio dolor: supongamos que tenemos un dolor en nuestro brazo izquierdo; según la interpretación previa, el dolor sería una representación que proyectamos sobre nuestro brazo. Pero imaginemos que nuestro dolor en el brazo no se debe a ninguna herida o anomalía en nuestro brazo, sino a un problema derivado de una disfunción cardíaca (como en el caso de los infartos). Entonces, nuestro dolor en el brazo izquierdo no estaría localizado en nuestro brazo, sino realmente debería localizarse en nuestro corazón, por lo que también deberíamos descartar esta posibilidad explicativa. Entonces, ¿dónde podemos localizar el contenido fenoménico de nuestras experiencias?

1.4.12. El problema de los miembros-fantasma

Este problema, relacionado con el anterior, se refiere a determinados casos de desórdenes psíquicos que afectan a personas que o bien tienen completamente paralizado algunos de sus miembros (como un brazo o una pierna) ya sea de manera permanente o momentánea (como en caso relatado en Sacks 1984) o bien a personas que de un modo u otro han perdido o les ha sido amputado alguno de dichos miembros (como en los casos relatados en Sacks 1987: caps. 3-6). Estas personas insisten a menudo en que sienten dolor (el caso del picor sería similar) en ese miembro que les falta. ¿Cómo decir que sienten dolor o picor en dicho miembro si lo que realmente les ocurre es que carecen de dicho miembro?

Si suponemos que el dolor o el picor es proyectado sobre la representación realizada de nuestro cuerpo, entonces volvemos a encontrarnos con la dificultad añadida de tener que explicar la relación de causalidad existente entre una sensación física como es el dolor o el picor y el objeto mental que es nuestra representación del propio cuerpo. Pero si suponemos que el dolor o el picor es proyectado por el miembro mismo, entonces la tesis fallaría al faltar también la parte del cuerpo 'en' la que, se supone, estaría localizada la sensación de dolor. ¿Cómo, entonces, podemos explicar la presencia de una sensación cuando falta el miembro sobre el cual deberían sentirse?

1.4.13. El problema de la percepción sin percepciones: la visión ciega

Relativo también a cierto trastorno cerebral, el problema de la visión ciega es habitual en aquellas personas que tienen una lesión cerebral que afecta la región post-genicular (en especial, el córtex occipital) del cerebro encargada de mostrar el campo visual del sujeto. Estos sujetos, aparentemente ciegos por dicha lesión, bajo ciertas circunstancias son capaces de interactuar con los objetos de dicho campo visual sin llegar nunca a percibirlos (Weiskrantz 1986). Es decir, el sujeto afectado por esta visión ciega es capaz de ser consciente de todo aquello que ve pero sin tener un acceso epistémico a dicha experiencia, es capaz de ser afectado por el contenido fenoménico de su visión pero sin llegar a ser consciente de la propia experiencia del ver. ¿Cómo encuadrar la posibilidad de que una persona pueda tener un contenido fenoménico que influya en su comportamiento sin ni siquiera ser consciente de que está teniendo las experiencias que le permiten tener dicho contenido?

1.4.14. El problema de la conciencia dividida

Otro tipo de lesión cerebral que una teoría de la conciencia debería ser capaz de explicar es la relativa a los sujetos con el cerebro escindido (*split-brain subjects*). Un cerebro humano está separado en dos hemisferios simétricos. Imaginemos que, como en algunos casos experimentales aplicados a pacientes con epilepsia, una operación quirúrgica secciona el cuerpo calloso que une las dos mitades de nuestro cerebro pero que, en nuestro caso imaginario, esta operación no llega a provocar que alguno de estos dos hemisferios pierda el acceso a la conciencia. Una operación de dicha índole, sería de suponer, como muestra los experimentos de Sperry (1968), permitiría generar sujetos con alguna especie de bifurcación mental. Esto es, en un mismo cuerpo, parecerían co-existir dos sujetos diferentes con una conciencia independiente cada uno que se disputarían la posesión y control causal del cuerpo que ambos ocupan, incluso llegando al punto de no llegar a ser jamás consciente ninguno de ellos de la existencia del otro (como una especie de *Jeckyl and Hyde* desmitificado). ¿Cómo explicar esta bifurcación de la conciencia?

Varias han sido las propuestas que a lo largo de los años se han pretendido como la explicación correcta. En un sentido laxo, podría decirse que el caso del cerebro escindido provoca la existencia de dos sujetos con dos mentes separadas (Pucetti 1972). En otros casos, se podría optar o bien por una explicación donde ocupe un papel dominante la parte racional afirmando que todas las respuestas producidas por la parte derecha del cerebro escindido pasarían a ocupar un lugar en el inconsciente del sujeto único resultado de la operación o bien por afirmar que existe una única persona con dos flujos separados de conciencia, cada uno referente a dos tiempos diferentes desde el momento de la escisión (Parfit 1987). También podría sostenerse que el número de personas resultante de la operación de bisección es indeterminado, o al menos dependiente del concepto de persona que estemos definiendo en cada momento (Nagel 1971). Visto desde otro punto de vista, podría decirse que el sujeto con el cerebro escindido es un sujeto único cuya conciencia fenoménica está momentáneamente dividida en dos bajo cierto tipo de condiciones especiales experimentales, pero que tarde o temprano dicha escisión de conciencia volverá a unificarse bajo un único flujo consciente (Mark 1980). De un modo similar, también encontramos quienes consideran que los sujetos con el cerebro escindido tienen un flujo unificado de conciencia pero un acceso dividido a la misma (Bayne y Chalmers 2003).

Ya sea de un modo o ya sea de otro, podemos englobar las diferentes respuestas en dos grandes bloques. Por un lado, tendríamos a aquellos pensadores que consideran que los sujetos con el cerebro escindido son una única persona sin ninguna división en su conciencia fenoménica. Por el otro lado, encontramos a aquellos que defienden que los sujetos con el cerebro escindido realmente padecen una separación en su flujo de conciencia, hasta el punto de generarse varios sujetos diferentes. ¿Cómo explicar estas posibilidades?

1.4.15. El problema de la unidad de lo consciente

Un problema especialmente importante actualmente trabajado dentro de las recientes investigaciones en psicología cognitiva y en neurofisiología es el problema relativo a cómo nuestro sistema visual es capaz de captar al unísono tanto el tamaño como la forma y el color de los objetos. Nosotros podemos ver un objeto como mediano, redondo y verde (podemos identificar un objeto de estas características como un globo), pero dichas características son experimentadas por partes diferentes de mi sistema visual (Marr 1982) aunque sea capaz de captar el objeto de dichas características como un todo. En nuestra experiencia, por tanto, el tamaño, la forma y el color, aunque se implementan por separado, son experimentados como unificados en un mismo objeto. Experimentamos un objeto mediano, redondo y verde. ¿Cómo puede ser este el caso? Cualquier teoría de la conciencia que pretenda ser completa debe poder contestar adecuadamente a esta pregunta, debe poder dar cuenta del problema de la unidad del objeto de conciencia. (Ello no quiere decir que nosotros no podamos experimentar por separado, por ejemplo, el color o la forma de un objeto determinado. Podemos quedarnos tan sólo con el color verde del globo, o con su forma redondeada, si fuera el caso, si sirviera a nuestras pretensiones, pero de modo general, la experiencia de un objeto aparece como unificando todas sus características relevantes porque, nos dicen sus estudios, existe una oscilación neuronal a 40 Hz. que engloba todos los eventos neurológicos relevantes en la visión de un objeto. Cf. Crick y Koch 1990).

Una segunda clase de unidad sobre la que debería llamarse la atención es la unidad neurofisiológica. Todos los estados conscientes parecen estar unificados neurológicamente porque parece que todos ellos se realizan a través de un mecanismo neuronal único o se implementan en la misma región cerebral. Una teoría total de la conciencia debería también poder ofrecer una explicación de este tipo de unidad.

Pero la conciencia también aparece como subjetivamente unificada. Es decir, los estados conscientes de un mismo sujeto aparecen entrelazados. Algunos estados mentales conscientes presuponen la existencia de otros. Por ejemplo, la existencia de estados conscientes de nivel superior presupone la existencia de estados de nivel primario, dado que tan sólo la existencia de estados a nivel sub-personal permite la aparición de estados conscientes a los que el mismo sujeto puede tener acceso. Pero no sólo eso, además, existen estados conscientes de nivel primario que presuponen la existencia de estados del nivel superior. Es decir, la existencia de estados prioritarios y/o especialmente relevantes presupone la necesaria presencia de estados superiores que los provoquen o los sustenten. Además, toda esta unidad de la conciencia debe mantenerse a través del espacio (es decir, debe poder englobar de manera adecuada el entorno que es experimentado) y a través del tiempo (es decir, debe perdurar el acceso a dichos estados conscientes por parte del sujeto). ¿Cómo puede darse toda esta unidad fenoménica en el mismo sujeto? (Esta sería una reformulación del clásico problema de Hume, donde la identidad personal se subordinaría, a grandes rasgos, a la continuidad personal. Cf. Hume 1748).

1.5. Recapitulación

Llegados a este punto no nos queda más que echar la mirada atrás y recapitular lo dicho hasta el momento.

La idea con la que comenzamos este capítulo era mostrar cómo existe una diferencia crucial entre los estados mentales con contenido proposicional y aquellos estados mentales que se caracterizan por ser fenoménicos. A diferencia de los primeros, los estados fenoménicos, podemos decir, tienen una peculiar naturaleza en tanto que aparecen de un modo subjetivamente dependiente al individuo que los posee. Visto de este modo, el contenido de estos estados fenoménicos no puede determinarse como proposicional, en tanto que los elementos subjetivos que aparecen tienen tanto una dimensión cualitativa como un acceso privilegiado desde la primera persona que los hacen aparecer como intrínsecamente dependientes.

Siguiendo el hilo argumental propuesto al inicio, continuamos presentando la diferenciación entre contenido conceptual y contenido no-conceptual. A diferencia de los estados proposicionales (y aunque ya comentamos la existencia de autores que tienen un punto de vista diferente), los estados fenoménicos no parecen tener un contenido conceptual, en tanto que parecen depender de la distribución sub-doxástica de los propios individuos más que de aquella distribución que, a un nivel personal, inunda su vida mental.

También presentamos una perspectiva histórica respecto de la problemática más general del problema mente-cuerpo y del problema más particular respecto del presunto hiato explicativo entre los estados mentales y los estados físicos. En dicha retrospectiva atendimos a los planteamientos clásicos que, desde Descartes y el materialismo británico, han permitido formular en los actuales términos nuestra problemática y realizamos una revisión de las posiciones más representativas, desde la teoría de los datos de los sentidos hasta el funcionalismo (pasando por el conductismo lógico, el eliminativismo y la teoría de la identidad) que a lo largo del siglo XX han permitido la actual evolución de los estudios sobre la conciencia.

En el apartado final, presentamos una revisión de los problemas planteados por Tye (1995). Tye ofreció un análisis de tan sólo diez problemas relativos al estudio de la conciencia. Aunque todos ellos están presentes en nuestra compilación, creemos que su lista es insuficiente, en tanto que deja fuera elementos cuyo análisis es actualmente imprescindible para la conciencia, como la luminosidad de lo mental. Aunque es cierto que los posteriores trabajos de Tye (2000 y 2003) recogen algunos de los problemas que yo presento aquí, lo hacen de una manera separada, en tanto que pretenden llamar la atención de manera particularizada sobre ciertos aspectos que, en su opinión, son particularmente relevantes. Presentándolos de este modo, creo, se permite tener una mejor visión global respecto de los problemas que podemos encontrar a la hora de realizar un análisis de la conciencia y que, en definitiva, nos servirán de guía durante el resto de este trabajo.

CAPÍTULO 2. REPRESENTACIÓN Y CONCIENCIA

*Parece absurdo para nosotros que un dolor, un estado
de ánimo o un deseo pueda ser acerca del mundo
sin ningún soporte, de modo independiente.
Una experiencia es imposible sin alguien que experimente.
El mundo interno presupone la existencia de la persona
de quién es dicho mundo interno.*
(Gottlob Frege, "El pensamiento: una investigación lógica")

*La conciencia tan sólo es una red de comunicación
entre los seres humanos; sólo es aquello que permite
su desarrollo. Un ser humano, en solitario,
que viviera en estado salvaje,
no tendría necesidad de ella.*
(Friedrich Nietzsche, *La gaya ciencia*)

II. REPRESENTACIÓN Y CONCIENCIA

En los últimos años, la visión representacionista de la mente y la conciencia ha ido adquiriendo cada vez mayor protagonismo en el panorama filosófico. A pesar de que existen variadas alternativas que tienen por base la noción de representación, existen diferentes posiciones dentro la misma, dependiendo del modo en que entienden el concepto de representación, dependiendo del tipo de concepto de *qualia* que posean o dependiendo del tipo de propiedades con las que caractericen a los estados fenoménicos.

A lo largo de este capítulo pretendemos varios objetivos. El primero de ellos es mostrar una perspectiva general respecto a la conciencia y las diferentes nociones que de la misma pueden manejarse. A partir del análisis de diferentes posiciones no-representacionales, perseguimos encontrar opiniones divergentes sobre el modo en que puede entenderse el carácter fenoménico de determinados estados mentales sin necesidad de recurrir a ningún tipo de estado o contenido representacional implícito. A ellas enfrentaremos las posiciones representacionales clásicas (como la de Fodor o Dretske) con el objeto de mostrar, como ya indicamos en nuestro planteamiento inicial, por un lado, el modo en que existe una divergencia cada vez más excluyente entre aquellos que consideran que el contenido de las experiencias fenoménicas debe ser entendido de modo conceptual y aquellos que consideran, por el contrario, que debe entenderse en sentido no-conceptual. Por otro lado, intentaremos también mostrar que, a pesar de los diferentes tipos de aproximación representacionista que podemos encontrar actualmente en la literatura, todos ellos comparten el mismo entusiasmo respecto de una noción clara de representación, en tanto que consideran que, de algún u otro modo, la experiencia fenoménica representa el mundo de un modo concreto a un sujeto perceptor consciente determinado. Para ello, analizamos los diferentes criterios representacionales que las teorías actualmente al uso defienden a la hora de definir sus posiciones y a la hora de explicar la posesión (por partes de los sujetos) de estados mentales con carácter fenoménico. Ello nos obligará, también, a definir brevemente los diferentes tipos de *qualia* que dicha teorías conciben, en tanto que existen diferencias sustanciales a la hora de definir aquello que sería el carácter cualitativo y subjetivo de dichos estados. La idea, al final del capítulo, es mostrar qué argumentos pueden esgrimirse a favor de sostener una teoría representacional de la conciencia actualmente, cuando existen tantos contraejemplos en contra. Todo ello nos dotará de las herramientas necesarias para definir el representacionismo presente en algunas teorías actuales respecto de la conciencia fenoménica que nos permitirá, como es el objetivo final de esta tesis, tomar distancia de ellas y preparar el camino para presentar todo un aparato crítico que nos guiará hasta una alternativa (creemos) viable a sus planteamientos respecto de los problemas fáciles y difíciles de la conciencia (tal y como han sido definidos en el capítulo anterior. Véase especialmente apartados 1.3. y 1.4.).

2.1. ¿Qué es una representación?

Según hemos presentado la noción de conciencia en el capítulo anterior, a lo largo de la historia de la filosofía han existido diferentes planteamientos a la hora de tratar la cuestión del modo en que nosotros vivimos nuestro mundo interior consciente. Lo que sí parece ser una realidad es que tradicionalmente se ha mantenido y acrecentado una diferenciación entre dos tipos distintos de conciencia. En un primer lugar, encontraríamos la conciencia fenomenológica; en un segundo lugar, tendríamos la conciencia representacional, proposicional, de contenido conceptual, de acceso o psicológica (según las diferentes y variadas caracterizaciones de las que ha sido objeto en los últimos años de estudio e investigación). Muchas veces, se ha dejado de lado algunos aspectos de esta separación en tanto que, por razones explicativas, se ha potenciado las características de una de ellas más que las de la otra con el objetivo de presentar una explicación de la conciencia. El problema principal que pretendemos analizar en este apartado es cómo el olvido (o directamente el haber pasado por alto) de algunas de dichas características dificulta la presentación de una teoría de la conciencia capaz de rendir cuentas de la intuición que todos tenemos de que nuestras propias sensaciones son subjetiva y cualitativamente experimentadas pero que, de algún u otro modo, no podemos olvidar que tienen una base física que permite su implementación.

Teniendo siempre presente este objetivo, analizamos por separado los dos tipos de conciencia identificados, haciendo especial hincapié en que, aunque la dicotomía no es (en un sentido estricto) tan profunda, ciertos aspectos de cada una de ellas son necesarios en un intento de explicación de la conciencia. A partir de ahí, presentamos una serie de críticas a aquellas posiciones que simplemente han considerado, por separado, uno de los dos tipos de conciencia (o, a lo sumo, ciertas características desconectadas), obteniendo una definición deficiente de lo que sea la conciencia. Acabaremos por concluir que, aunque dicha separación es necesaria porque permite diferenciar determinadas características importantes en el análisis de la noción de conciencia, dicha separación no es excluyente. Sino que, por el contrario, en el estudio de la conciencia se deben tener presentes ambos tipos de conciencia (con sus respectivas características), lo que permitiría un acercamiento plausible a la explicación de nuestra noción habitual de conciencia.

2.1.1. Tipos de conciencia

La conciencia fenomenológica es el estado mediante el que nosotros, seres conscientes, nos damos cuenta de modo subjetivo de nuestras propias vivencias, esto es, es indistinguible de las experiencias mismas, es el carácter marcadamente cualitativo de las propias vivencias. Según Block (1995: 228):

“La conciencia fenomenológica es la experiencia misma. Las propiedades de la experiencia fenomenológica son las propiedades experienciales, esto es, un estado es fenomenológicamente consciente si tienen propiedades experienciales. La totalidad de dichas propiedades experimentales de un estado fenomenológicamente consciente son los aspectos de lo que sería para alguien sentir dicho estado”.

Estados característicos de conciencia fenomenológica serían el ver, el oler, el oír o el estar deprimido. Es esencial a la naturaleza de estos estados tener diferente contenido fenoménico, porque precisamente es la diferencia en el contenido intencional de estos estados la que nos indica en qué estado concreto nos encontramos. En este sentido, la conciencia fenomenológica tiene a menudo contenido y es representacional, en tanto que es imagen de las cosas, pero es diferente de cualquier propiedad cognitiva (dado que presupone la existencia de un modo esencial de experiencia), intencional (dado que presupone referencia a algo concreto) o funcional (dado que tan sólo podría definirse en términos causales). Vistas así las cosas, la experiencia de oír un ruido procedente de nuestra izquierda es diferente de la experiencia de oír un sonido procedente de nuestra derecha, en tanto que (a pesar de ser representacionalmente similares, por ser ambos dos ejemplares concretos de la experiencia de oír) intencionalmente tienen un contenido distinto (Block 1995: 230). A pesar de que existe una cierta coincidencia a la hora de ofrecerle un papel central en la vida mental, muchos autores (como ya hemos mostrado en la introducción), son reacios a considerar que se tenga un acceso inmediato a la conciencia fenomenológica, estableciendo la existencia del temido hiato explicativo.

Por su parte, la conciencia representacional (entendida como proposicional o de contenido conceptual) se encargaría de proporcionarnos un acceso a los contenidos informacionales de los estados mentales. En este sentido, la conciencia es esencialmente funcional, en tanto que su contenido representacional es empleado en el razonamiento y en el control racional de las acciones y del lenguaje (Block 1995: 231).

“Un estado es consciente por acceso si, en virtud de tenerlo, una representación de su contenido es 1) inferencialmente promiscua, esto es, está preparada para ser empleada como premisa de razonamiento y 2) está preparada para el control (racional) de la acción y 3) está preparada para el control racional del lenguaje”.

Las diferencias entre ambos conceptos de conciencia son, según Block (1995: 231), marcadas. Por un lado, el contenido de la conciencia fenomenológica es fenoménico y causal, mientras que el contenido de la conciencia por acceso es representacional y funcional. Así, parece que tan sólo la conciencia representacional puede tener un papel funcional y racional a la hora de dirigir nuestro comportamiento, porque sólo éste tipo de conciencia es capaz de mostrar cierta normatividad al reflejar cierta imagen de los objetos externos permitiendo, así, guiar nuestra acción. Por otro lado, los estados de la conciencia fenomenológica son ejemplares de un determinado tipo de estado. Es decir, la sensación de dolor es fenoménicamente consciente en tanto que cualquier dolor, aunque cambie en contenido, será sobre ese tipo de experiencia. Mientras que, por el contrario, la conciencia de acceso es una noción totalmente funcional. En este caso, todo contenido sería dependiente de la estructura en la que se implemente. Es decir, lo que hace que un estado sea consciente de acceso es lo que una representación de su contenido puede hacer dentro del sistema consciente (las conductas que provoca, por ejemplo). Visto de este modo, entonces, ante determinados estados conscientes, pongamos por caso mi experiencia visual de observar un tomate maduro, experimento una sensación de color rojo (conciencia fenomenológica) que me aporta toda la información relevante para que sepa que está listo para

ser recolectado y que puedo emplearlo, por ejemplo, en una ensalada (conciencia representacional).

Una dificultad que actualmente encontramos para poder dar correcta cuenta de la conciencia es, principalmente, el olvido de la existencia de estas dos nociones separadas de conciencia. Pero que existan dos conceptos separados y distintos no quiere necesariamente decir que sean contradictorios o antagónicos. Los dos conceptos son complementarios entre sí en tanto que se refieren a fenómenos distintos de los estados mentales, fenómenos que son explicables al mismo tiempo. En este sentido, la tendencia segregacionista actual, como la ha denominado Güzeldere (1997: 11), aquella intuición que afirma que si creemos que la conciencia es fenomenológica entonces no puede ser funcional y que si creemos que la conciencia es funcional entonces no puede ser causal, es la tendencia dominante dentro de los estudios actuales sobre conciencia. Pero existen alternativas viables. Como bien dice Güzeldere, deberíamos instaurar una tendencia integracionista, donde la conciencia sea caracterizada teniendo presente el modo en cómo siente y experimenta y que dicho modo de experimentar no puede conceptualizarse si se deja de lado lo que hace. De todos modos, avanzo ahora que la diferenciación empleada por Block no es del todo correcta. Como mostraremos a lo largo de las siguientes páginas, pero especialmente en el capítulo final, aunque diferenciamos entre dos tipos de conciencia, ello no quiere decir que exista esa separación realmente, ni que existan límites marcadamente definidos entre ambas nociones. Más bien podríamos apostar por una noción de conciencia que permita una explicación del comportamiento de los seres conscientes como guiados de manera racional según, por una parte, sus habituales modos de acción (como se comportaría habitualmente un ser consciente en determinadas situaciones) y, por otra parte, permita explicar su comportamiento en base a sus propios intereses. Además, no está del todo claro en qué sentido podemos decir que la conciencia fenomenológica no es funcional, en tanto es más que obvio que la percepción concreta de determinadas características cualitativas mueve al sujeto perceptor a actuar de determinado modo, por lo que puede tener un marcado interés para su control racional.

Según hemos visto, entonces, dentro de la misma experiencia consciente debemos diferenciar dos aspectos relevantes. Volvamos a nuestro ejemplo de observar un tomate maduro. Cuando tenemos una percepción de un tomate maduro debemos distinguir, por un lado, el carácter cualitativo de la experiencia (su aspecto de rojez) de, por otro lado, el carácter subjetivo de ver un tomate (lo que es para mí percibirlo). Parece razonable afirmar que ambos componentes pueden ser presentados por separado, con independencia uno del otro. Pero esto no es del todo correcto. Parece plausible presuponer que el componente cualitativo de una experiencia perceptiva no puede variar de modo independiente de su componente subjetivo, en tanto que el carácter representacional de una experiencia estaría determinado por completo por su contenido proposicional. Es decir, aquello que la experiencia representa es lo que permite decir de qué experiencia estamos hablando.

2.1.1.1. El carácter cualitativo de la experiencia

El carácter fenoménico de una experiencia es una propiedad que nos muestra sus elementos cualitativos (*qualia*, podríamos llamarlo, según la tradición). En este sentido, existen varias nociones distintas de *qualia* que pueden ser empleadas, cada una, para fines diferentes (Lycan 2000. Véase la sección 2.2.3.2.). Pero todas ellas tienen en común una idea general. El carácter cualitativo de una experiencia, su vivencia, es el modo concreto en que la experiencia perceptiva representa el mundo para un sujeto (Evans 1982: 226; McGinn 1989: 58; Harman 1990: 34).

En este sentido, el contenido de una experiencia especifica el modo en que el mundo se aparece a un sujeto concreto de percepción, por lo que parece claro que ambas nociones no pueden darse por separado. Cuando hablamos de la vivencia de una experiencia, se muestran al unísono tanto su carácter cualitativo como su componente subjetivo.

2.1.1.2. El carácter subjetivo de la experiencia

Si hay algo que haga consciente una experiencia para mí, ese algo es el tener dicha experiencia. Debo darme cuenta de tener dicha experiencia. Nosotros, habitualmente, nos damos cuenta de nuestras experiencias. Así, las experiencias conscientes no son estados que, simplemente, acontecen en nosotros a un nivel sub-personal, sino que son estados de los cuales nos damos cuenta. Normalmente esto ocurre porque focalizamos nuestra atención sobre ellos. Si nouviésemos un modo de focalizar nuestra atención sobre ellas, entonces no seríamos capaces de darnos cuenta de nuestras experiencias.

En este sentido, entonces, sería conveniente establecer una diferenciación entre nuestra atención consciente concreta y nuestra atención consciente periférica.¹ Porque cuando prestamos atención a una experiencia, a menudo, dicha experiencia no se nos aparece de modo claro. Sino que dicha experiencia puede ir entremezclada con otras experiencias. Visto así, entonces, parece plausible entender nuestra vida consciente como holista, en tanto que es complicado, a veces, separar nítidamente las varias experiencias que se están teniendo en un mismo momento. Pongamos un ejemplo. Estoy conduciendo de camino a casa. A los lados de la calzada hay muchos objetos que, sin duda, veo pero a los que no presto atención porque mi atención está puesta por completo en la carretera y en los otros vehículos para evitar en la medida de lo posible una colisión. Sé que entre los objetos que se cruzan en mi trayectoria existen toda una serie de vallas publicitarias. El mensaje de dichas vallas raramente podría enunciarlo de modo consciente, pero sin dudar un instante podríamos decir que los veo (si no fuera así, ni siquiera los publicistas aconsejarían hacer el desembolso que debe suponer al año

¹ Esta diferencia se verá especialmente clara cuando presentemos la diferencia entre posiciones representacionales de nivel superior claramente disposicionales (como la de Carruthers) en tanto que enfrentada a posiciones auto-representacionales de corte internista (como la de Kriegel). La diferencia crucial entre ellas es que, como posteriormente analizaremos con más detalle, frente a posiciones disposicionales que tan sólo consideran la existencia de una tendencia a la sensibilidad que un elemento externo provoca en el sujeto en un momento concreto, la existencia de un sentido periférico interno defendida por el auto-representacionalismo de Kriegel permite explicar el modo en que un sujeto puede atender al unísono a todos los elementos fenoménicos presentes en su conciencia. Véase la sección 2.1.2.4.

para un empresario colocar uno de dichos anuncios). Nuestra conciencia, por tanto, no tiene unos límites claros en este sentido. Podemos insistir, aquí, en que nuestra vivencia de las experiencias conscientes exhibe un cierto carácter subjetivo en tanto que al tener una experiencia consciente también tenemos una percepción periférica de ella. Vistas así las cosas, parece plausible presuponer que nuestra vivencia de nuestras experiencias conscientes es una cierta forma de darnos cuenta de dicha periferia de modo no-perceptivo. En este sentido, como ya en su momento estableció el filósofo Thomas Reid, deberíamos realizar una diferenciación entre lo que es tener una percepción concreta y lo que es tener una sensación concreta.

“El agradable olor que sentimos [cuando olemos una rosa], en sí mismo, es tan sólo una sensación. Afecta a nuestra mente de cierto modo, y dicha afección de la mente debe ser considerada sin la intervención de ningún pensamiento o de cualquier otro objeto. Esta sensación no puede ser nada más que el modo en que es sentida. Su esencia consiste en su ser sentida; y cuando no es sentida, deja de ser... La percepción, en cambio, siempre tiene un objeto externo, y el objeto de nuestra percepción es, en este caso, las propiedades que la rosa presenta cuando la capto a través de mi sentido del olfato” (Reid 1785: 242).

La distinción entre sensación y percepción es crucial en este caso, dado que como nos muestra la cita parece que la única diferencia perceptible entre ellas sería, precisamente, la posibilidad de que la percepción esté siempre dirigida a un objeto externo. Esto es, según Brentano (1874), que muestre intencionalidad.² A lo largo de las próximas páginas, y de los sucesivos capítulos, recuperaremos esta distinción para afrontar críticamente la posibilidad de que dicha diferenciación no sea tan clara, después de todo. Anticipando hechos, en el representacionalismo fuerte (por ejemplo, el de Dretske o el de Tye) a menudo suele considerarse que las sensaciones también tienen un carácter intencional, aunque ello no quiera decir que se asimile percepción con sensación ni que la intencionalidad sea el marco de lo mental. Véase, por ejemplo, Dretske (1981: 135-153).

Ante esta afirmación de la existencia de un carácter no-perceptivo en las sensaciones, podría objetarse que todavía involucra una concepción cartesiana de la mente, en tanto que la noción de un darse cuenta subjetivo interno presupondría la posibilidad de la existencia de una incorregibilidad del conocimiento sobre los propios estados. Dentro de esta imagen, de herencia cartesiana, la mente tendría perfecto conocimiento de todos los estados mentales, en tanto que sería transparente. Un determinado individuo siempre tendría acceso privilegiado a sus propios estados, tendría un conocimiento infalible sobre sus propios estados mentales, en tanto que todo contenido siempre es accesible al sujeto. Pero, como sabemos, la posibilidad de error existe, por lo que no parece del todo plausible sostener por mucho tiempo esta idea.

Además, este darse cuenta subjetivo también es reivindicado como especialmente relevante en tanto que parece implausible suponerlo como separado de la misma vivencia de la experiencia. Porque si mi experiencia concreta de ver un tomate maduro tuviera que deberse a varias representaciones (por un lado, la vivencia misma en tanto que cualitativamente roja y, por

² “Todo fenómeno mental se caracteriza por lo que los escolásticos de la Edad Media llamaron la inexistencia intencional (y también mental) de un objeto y nosotros podríamos llamar, aunque en términos no totalmente carentes de ambigüedad, la referencia a un contenido, una dirección hacia un objeto (por el que no hemos de entender una realidad en este caso) o una objetividad inmanente” (Brentano 1874: 124).

otro, el darse cuenta subjetivo de la experiencia, el modo en cómo se me da la experiencia), entonces dichas representaciones podrían ser inconscientes. Pero, por definición, lo inconsciente no se puede catalogar como algo de lo que podemos darnos cuenta, por lo que éste no puede ser el caso. Así, el carácter cualitativo y el carácter subjetivo parecen necesitarse mutuamente, por lo que no es demasiado recomendable anticipar la posibilidad de una separación radical entre percepción y sensación.

De todos modos, como bien sabemos, existe también otro aspecto importante que resaltan aquellos que pretenden realizar esta separación entre la percepción y la sensación. Éstos tienden a remarcar que tan sólo la percepción parece estar directamente relacionada con otros estados mentales a partir de su contenido intencional, en tanto que plenamente cognitivos, como por ejemplo las creencias y los deseos. Porque sólo la percepción permitiría la aparición de este tipo de estados proposicionales, en tanto que posibilitaría preparar todo el contenido cognitivo aportado para su acceso al nivel de las actitudes proposicional. Avanzándonos un poco a nuestras conclusiones, diremos aquí que parece no ser del todo cierto que un estado puramente fenoménico no tenga posibilidad de conectarse directamente con otros estados. Pensemos tan sólo en el hecho de alguien que está deprimido. Tal vez, impulsado por su estado de abatimiento, piensa que todo el mundo está en su contra, cree que nadie le hace caso y desea que la tierra se lo trague.

También podemos avanzar aquí que el problema del carácter subjetivo de una experiencia se refiere a cómo puede y debe incluirse tal carácter subjetivo en la vivencia misma de la experiencia. En este sentido, muchos han sido los intentos para conseguir unificar en la misma experiencia su componente cualitativo y su componente subjetivo sin dejar de lado ningún aspecto relevante de su vivencia concreta. En este sentido, ¿cómo puede un estado mental dar cuenta, a la vez, de su rojez y de su darse un objeto concreto de modo determinado a un sujeto (para el caso de observar un tomate maduro)?

2.1.2. Aproximaciones al carácter subjetivo de la conciencia

Como avanzamos anteriormente, el principal problema que presenta el carácter subjetivo de la conciencia es la dificultad que supone comprender cómo un cierto estado mental puede ser incluido en su propia vivencia. Es decir, en qué sentido podemos decir que un mismo estado mental tiene dos aspectos (aparentemente) diferenciados, su carácter cualitativo (su rojez, en el caso del tomate maduro) y su carácter subjetivo (el ser para mí de un cierto modo). Por supuesto, si concebimos la mente como una entidad sustancial separada de un cuerpo físico, la dificultad desaparece, en tanto que podemos vincular necesariamente ambas propiedades a las diferentes sustancias. El problema aparece cuando intentamos afrontar esta dificultad desde un punto de vista fisicalista. En este caso, si consideramos que sólo existe una entidad posible, la material, ¿cómo podemos rendir cuentas de esta separación? Aparece, entonces, el problema de la naturalización del carácter subjetivo de la conciencia.

Muy diferentes y variadas son las respuestas que se han sucedido intentando solucionar esta problemática. Veamos algunas de las más relevantes.

2.1.2.1. Teorías conceptualistas

Como nos dice uno de los máximos defensores del dualismo de la mente en la actualidad, la naturalización del carácter subjetivo requiere mostrar que la relación entre el sujeto y sus vivencias (y más específicamente, entre las vivencias y la intencionalidad) es una relación reflexiva, es decir, una relación en la que tener un estado mental es en sí mismo suficiente para darse cuenta de él (Levine 2001: 69). Los intentos para explicar esta noción de subjetividad vivencial, siempre según Levine, han pasado por mantener una de las dos siguientes estrategias (erróneas, desde su punto de vista, como veremos).

Por un lado, estarían aquellos que, como Gennaro, han defendido una relación conceptual entre sujeto y vivencia.

“Podemos comprender la conciencia como involucrando la propiedad de estar acompañada por un PMP [Pensamiento Meta-Psicológico (MET en inglés, Meta-psychological Thought)], pero deberíamos individualizar los estados consciente de manera amplia, esto es, de modo que tratásemos la conciencia como una propiedad intrínseca de dichos estados. Según esto, el PMP formaría parte del estado consciente. En este sentido, estaríamos hablando de ‘un punto de vista intrínsecamente amplio (PvIA)’ (WIV en inglés, Wide Intrisicality View)” (Gennaro 1996: 23).

Desde este punto de vista, Gennaro no parece estar negando que la experiencia y la vivencia de la experiencia se den de modo separado. Lo que parece estar negando es que sean estados separados. Estaría reivindicando que vivencia y experiencia son partes separadas de un único y mismo estado, permitiendo así hacer justicia a la dualidad que aparentemente presenta la experiencia consciente, pero sin dejar de ser fisicalista.

“El PvIA también da cuenta de la tesis histórica de que los estados mentales conscientes están, en algún sentido, dirigidos hacia sí mismos. El punto de vista de Brentano nos dice que un estado mental consciente es, en parte, sobre sí mismo. Estrictamente hablando, sobre PvIA, un estado no se dirigirá sobre sí mismo. Habría una direccionalidad interna o una relacionalidad interna en el estado mental consciente... pero el PMP es dirigido a partes del estado del que forma parte” (Gennaro 1996: 27-28).

El punto de vista de Gennaro incurriría, según Levine (2001: 169), en tres dificultades básicas. Primero, que el carácter subjetivo de la experiencia sea un fenómeno real no presupone que pueda ser sustituido *salva veritate* por un correlato puramente conceptual. Ello implicaría poder mostrar el lugar donde acontece dicha sustitución y donde se localiza dicho correlato, porque necesariamente debería poderse mostrar la base física concreta que permite su existencia (si consideramos, como hace Gennaro, que podemos identificar sensación con estado físico). Segundo, la estrategia conceptual de Gennaro no lograría dar cuenta de la dualidad (aparente) de la experiencia consciente, en tanto que no consigue explicar adecuadamente cómo experiencia y vivencia son dos aspectos separados de un mismo estado mental. Tercero, que la experiencia y la vivencia llegaran a demostrarse como aspectos separados de un mismo estado no quiere decir que un sujeto concreto fuera capaz de saberlo, llevándonos de nuevo al hiato explicativo (Levine 1983), en tanto que es posible que un sujeto sepa p y que sepa q y que no sepa que p&q. Vistas así, las cosas, siempre según Levine, deberíamos dejar de lado la estrategia conceptual avanzada por Gennaro. Desde nuestro punto de vista, la estrategia conceptual es demasiado

arriesgada en tanto que presupondría la posibilidad de la existencia necesaria de un marco conceptual previo inscrito de modo innato en nuestros genes que sería implementado en nuestros estados cerebrales. Respetaremos, por tanto, las críticas de Levine a Gennaro (pero sólo estas críticas), lo que no quiere decir que no consideremos problemática, cuando menos, la acepción de la tercera conclusión (Diremos alguna cosa más respecto a la posibilidad de la existencia de un marco conceptual innato en la sección 2.1.3.3., y volveremos a ella en la sección 6.2.2., cuando analicemos la inviabilidad de que todo estado fenoménico pueda ser entendido como teniendo siempre un contenido conceptual).

Por otro lado, según Levine, encontraríamos a aquellos que, como funcionalistas y representacionistas, han sostenido la existencia de una relación real no-conceptual entre los sujetos (y su intencionalidad mental) y las experiencias mismas. Attendamos, ahora, a las aproximaciones funcionalistas (de las aproximaciones representacionales nos encargaremos en los apartados siguientes y en los capítulos 3, 4 y 5).

2.1.2.2. Teorías funcionalistas

El funcionalismo individualiza los estados mentales en términos de su diferente papel funcional en la vida mental del sujeto. En este sentido, mantiene una diferenciación entre papel funcional y contenido del estado mental. El papel funcional sería una abstracción de todas aquellas causas y efectos que provoca, o que es capaz de provocar, un determinado estado, mientras que el contenido de dicho estado es el estado neurofisiológico del cerebro que tiene dicho papel causal. Si esto es así, el estado neurofisiológico realiza el estado mental pertinente en virtud de ocupar el papel funcional adecuado para implementar el estado mental concreto. Vistas así las cosas, para el funcionalismo, hay un cierto papel funcional desempeñado por un determinado estado neurofisiológico que determina el carácter cualitativo de una experiencia de rojo, esto es, la rojez (para la experiencia de observar un tomate maduro) y hay un cierto papel funcional desempeñado por un determinado estado neurofisiológico que determina el carácter subjetivo de la vivencia de una cierta experiencia.

Según estas aproximaciones funcionalistas, como ya hemos dicho, el papel funcional de un estado mental sería una abstracción del conjunto total de sus relaciones de causa-efecto. Teniendo presente esta afirmación, entonces, si dos estados mentales, digamos M1 y M2, tienen el mismo efecto E, pero M1 tarda tres milésimas de segundo más en implementarse que M2, entonces diremos que esta insignificante diferencia es irrelevante a la hora de establecer su papel funcional, dado que lo que nos interesa es abstraer y mantener el papel funcional que tanto M1 como M2 son capaces de desempeñar, y éste sería el mismo (Block 1986). Aunque ello no quiere decir que en un momento determinado la diferencia temporal sí sea relevante, por ejemplo cuando dicha diferencia temporal tuviera consecuencias causales cruciales para explicar las acciones de un agente (debemos tener en cuenta que, muchas veces, ésta sería la única evidencia experimental disponible, como nos muestran los estudios sobre tiempos de reacción que nos ofrece la psicología cognitiva). Visto de este modo, el papel funcional sería un papel inferencial (tal y como lo describe Field 1977). Dicho papel inferencial de un estado mental

puede caracterizarse como un conjunto ordenado de todos aquellos estados de los cuales puede inferirse dicho estado mental concreto. Si el conjunto de los estados mentales de los cuales puede ser inferido un estado mental M1 coincide con el conjunto de los estados mentales de los cuales puede ser inferido un estado mental M2, entonces M1 y M2 compartirían el mismo papel funcional.

Esta es la estrategia que persigue Carruthers. Para este autor, deberíamos apelar a la semántica del papel inferencial en nuestro intento de acercarnos al carácter subjetivo de la experiencia consciente, en tanto que cuando tenemos una determinada sensación, por ejemplo la rojez al contemplar un tomate maduro, estaríamos dispuestos a inferir (no-deductivamente) no sólo pensamientos primitivos, sino también pensamientos de nivel superior. En este sentido, un estado mental actualmente exhibiría un nivel periférico, en tanto que potencialmente podríamos llegar a darnos cuenta de dichos pensamientos si focalizáramos nuestra atención sobre ellos (Carruthers 2000: 242-252), algo que es dudoso.

Por su parte, evitando esta dificultad, Shoemaker considera que un estado mental podría calificarse como una experiencia de rojo y como un darse cuenta de la experiencia de rojo, al unísono, si aquel elemento que cumple la función de ambos puede caracterizarse funcionalmente hablando como la característica que permite la experiencia de rojo y el darse cuenta de la experiencia de rojo (Shoemaker 1994: 288-289). En este sentido, sería tan sólo una mera cuestión empírica determinar el estado neurofisiológico de nuestro cerebro que ocuparía este papel funcional. Si dicho estado fuese descubierto, entonces exhibiría el carácter subjetivo determinado que permitiría indicar la experiencia concreta de que se trata. Pero, en todo caso, estaríamos hablando de un elemento no-conceptual, un elemento natural difícil de asociar con el carácter subjetivo de las vivencias fenoménicas. Ello nos llevaría a suponer la existencia de dos dificultades aparentemente insalvables dentro del planteamiento funcionalista (al menos, dentro del planteamiento realizado por Shoemaker).

En primer lugar, para que el papel funcional inferencial fuera plausible debería implementar un contenido intencional pero no, como pretendería Shoemaker, cuando constituye el contenido intencional. Esto es, sólo cuando el papel funcional determine el contenido intencional, tendría sentido hablar de una relación inferencial entre estado mental y contenido. Pero parece plausible presuponer que la naturaleza dual de la experiencia, la existencia de una experiencia y de una vivencia de una experiencia, debería ser explicable en términos intencionales pero no en términos funcionales dado que no es posible reconocer la existencia de un término medio que permita establecer una relación entre ambos, por ser inconsistente la suposición de una entidad tal (Newman 2004).

En segundo lugar, la existencia de un papel funcional se desarrollaría en términos disposicionales y no en términos intencionales. El papel funcional de un estado mental tan sólo sería una propiedad disposicional, mientras que la vivencia interna presente en una experiencia sería una propiedad categórica. Desde este punto de vista, cuando un estado mental se convirtiera en consciente, el sujeto adquiriría una característica que anteriormente no poseía: adquiriría un nuevo concepto, en tanto que se presupondría que su *esse est percipi*, esto es, que

en el caso de las experiencias fenoménicas podríamos identificar o reducirlas a meras cualidades subjetivas, independientes de toda estructura objetiva (carentes de cualquier posibilidad de constituir una clase natural); que, aunque predicadas de los objetos externos, la posibilidad de su sensación dependería de que el sujeto poseyera una cierta familiaridad con la experiencia sensorial concreta, una familiaridad sólo disponible desde la perspectiva de la primera persona (McGinn 1983: 5-9). Pero esta réplica tendría una dificultad. Si fuera realmente posible que cuando un individuo tuviera una experiencia éste adquiriera un nuevo concepto, entonces estaría adquiriendo nuevos conceptos durante toda su existencia. En este sentido, deberíamos presuponer la existencia de elementos no-conceptuales (lo que no quiere decir que sean funcionales, como hemos desmentido anteriormente vistas las dificultades de Shoemaker y Carruthers) que permitieran a los sujetos conscientes tener experiencias nuevas sin necesidad de aprender nada nuevo. En este sentido, si lográramos encontrar dichos elementos, podríamos establecer una solución fisicalista al problema difícil de la conciencia al evitar el inconveniente de tener que explicar la presencia de conocimientos previos a partir de argumentaciones epifenomenalistas (además de posibilitarnos una solución al problema concreto del moderno argumento del conocimiento formulado por Jackson en su clásico ejemplo de Mary, expuesto en la sección 1.4.2., como intentamos demostrar en la sección 6.2.1.2.).

2.1.2.3. Teorías cognitivas

Precisamente con la intención de encontrar dichos elementos no-conceptuales que permitieran explicar el carácter intencional de las experiencias y de las vivencias particulares de las experiencias fenoménicas, evitando en la medida de los posibles las aproximaciones funcionalistas, encontramos la teoría cognitiva de Baars.

El enfoque de Baars parte de considerar la conciencia como un fenómeno natural más, como uno más de los fenómenos de tercera persona públicamente observables, por lo que considera necesario adoptar el mismo método de estudio que se emplea con los restantes fenómenos naturales. En concreto, la teoría de Baars pretende realizar una modelización cognitiva, es decir, persigue presentar

“un modelo de la dinámica causal implicada en los procesos cognitivos que explique las causas de la conducta de un agente cognitivo” (Chalmers 1996: 153).

Para Baars, la experiencia consciente debe considerarse como una construcción teórica que debe inferirse de la evidencia (Baars 1988: 9). Para alcanzar dicha evidencia, en primer lugar, deberíamos poder contrastar los estados y procesos conscientes y los estados y procesos no-conscientes, en tanto que ambos se diferenciarían del acceso que podría tenerse a ellos en base a una observación pública. Dicho contraste sería posible en tanto que ambos tipos de estado tienen el mismo estatuto psicológico que otras construcciones teóricas propuestas por la ciencia. En este sentido, lo que se requiere para el estudio de la conciencia es un análisis comparativo que permita proporcionar la base empírica para explicar el desarrollo teórico de

dicha abstracción, partiendo de casos simples como los ejemplos de visión ciega³ (Baars 1988: xvi-xvii; Baars 1997: 187). Este planteamiento debe ser global, en tanto permitiría conseguir una teoría integradora basada en datos concretos. El modelo propuesto sería expresable en un lenguaje computacional, en un lenguaje de procesamiento de la información, en tanto que sería un lenguaje neutro que se mantendría al margen de la experiencia consciente, permitiéndonos así hablar tanto de estados mentales conscientes como inconscientes, siendo también más adaptable a los sucesivos avances y descubrimientos científicos (Baars 1988: 13). En este sentido, Baars partirá de una noción operativa de conciencia.

“Consideramos que alguien es consciente si puede decir de manera inmediata que es consciente de un evento y podemos verificarlo de modo independiente a su información” (Baars 1988: 15).

Visto así, formarían también parte de la conciencia tanto las experiencias conscientes de los preceptos como la experiencia consciente que un individuo tuviera de los conceptos abstractos que pueblan nuestra actividad mental, separando así entre la conciencia fenomenológica y la conciencia por acceso (anteriormente precisadas).

El primer paso para poder explicar la conciencia, según nuestro autor, sería (de acuerdo con su método de análisis comparativo) tener una clara concepción de aquellos fenómenos que podríamos considerar inconscientes. En su jerga, un estado sería inconsciente si es el resultado del funcionamiento de un sistema especializado, es decir, en tanto que cada vez es realizado de modo más mecánico. Estos sistemas mecánicos trabajarían con representaciones, en tanto que una representación “es un objeto teórico que tiene una semejanza abstracta con algo fuera de ella” (Baars 1988: 44). Dichas representaciones sufrirían cambios a partir de procesos y el conjunto de procesos que permite confeccionarlas y modificarlas, a su vez, constituirían de continuo una serie de numerosos procesadores de información representacional (Baars 1988: 50). Todos estos procesadores actuarían de manera automática, ciega, en tanto que se podrían concebir como un conjunto de destrezas y habilidades que operarían en el sujeto de modo inconsciente. De este modo, las capacidades cognitivas, en general, podrían verse como el conjunto elaborado de una gran cantidad de operaciones realizadas de modo inconsciente por estos procesadores altamente especializados, que acabarían por conformar módulos que interactuarían unos con otros para crear y generar experiencias conscientes integradas. En este sentido,

“el procesamiento detallado a lo largo del sistema estaría altamente descentralizado. Cada módulo se compondría y descompondría de modo variable, dependiendo del objetivo a perseguir y del contexto involucrado. Los procesadores especializados son capaces de adaptarse a nuevas entradas, pero tan sólo de un modo restringido” (Baars 1988: 64).

Frente a los estados inconscientes encontramos los fenómenos conscientes, que están asociados, según Baars, a lo que denomina un Espacio Global de Trabajo. Este espacio de trabajo

³ Recordemos que el problema de la visión ciega ya fue presentado en el apartado 1.4.12.

“es semejante a una pizarra en una clase o a una estación de radio en una comunidad humana porque en él se produce un intercambio de información que permite a los procesadores especializados interactuar entre sí” (Baars 1988: 74).

En este sistema de procesadores que es el cerebro, no encontraríamos

“ningún ejecutivo central [en tanto que es un sistema distribuido], esto es, un sistema único que asigna los problemas a los especialistas específicos o los dirige para realizar la tarea. Para diferentes trabajos, procesadores distintos pueden comportarse como ejecutivos, en tanto ejercerían el control de modo flexible. El control sería, así, esencialmente descentralizado” (Baars 1988: 87).⁴

La existencia de un espacio global de trabajo se volvería necesaria porque la información entre los diferentes procesadores debe ser internamente consistente, dado que, de lo contrario, se degradaría el sistema debido a la competencia entre los sistemas de procesamiento. En este sentido, el intercambio computacional debe ser informativo, en tanto debe ser transmitido de modo consciente.

La noción clave de información debe entenderse en un sentido clásico.

“Somos conscientes de un suceso sólo cuando existe en un contexto estable, pero no cuando es tan predecible que no hay alternativas concebibles” (Baars 1988: 178).

Vista de este modo, la experiencia consciente no es una función directa de la estimulación física, sino que dependería de la información real que el mundo aporta en cada momento (en un sentido clásico, según la cantidad de incertidumbre que la información del exterior sea capaz de reducir). En este sentido, la probabilidad de que cualquier suceso sea consciente es dependiente de su valor informativo, puesto que la información sería una condición necesaria de toda experiencia consciente.

Toda esta teoría de la conciencia, para Baars, estaría apoyada por la gran cantidad de evidencias empíricas que actualmente nos proporciona la ciencia neurológica (Baars 1988: cap. 3). Por un lado, el sistema nervioso central, con el cerebro al frente, sería entendido como un sistema distribuido en paralelo, compuesto por múltiples procesadores especializados, y muchas de las estructuras cerebrales importantes, como el córtex o la amígdala, podrían considerarse como una colección de módulos distribuidos. Por otro lado, algunos de estos módulos podrían competir o colaborar entre ellos para acceder al Sistema de Activación Reticular-Talámico Extendido, que sería, según Baars, la parte del sistema nervioso que realizaría las funciones del espacio global de trabajo. Este sistema de activación se compondría por un sistema capaz de recibir la información de todas las estructuras importantes del cerebro (la Formación Reticular) y se extendería hasta alcanzar e involucrar a los núcleos inespecificados del tálamo cerebral y un sistema de proyección difuso que enviaría gran cantidad de fibras a toda la corteza con el fin de englobar la mayor cantidad posible de conexiones corticales en el intercambio de

⁴ Otro autor que considera la posibilidad de que un sistema de procesamiento permita conformar una conciencia (o mejor, un nivel consciente) carente de centralización es Dennett, como veremos en el capítulo 5.

información. Pero además de lo anterior, también parecería científicamente demostrado que la información transmitida por los diferentes sistemas de procesamiento cerebral puede ser transmitida de modo global a otras partes del sistema nervioso, especialmente al manto cortical del cerebro.

A pesar de la plausibilidad de este tipo de enfoque cognitivo de la conciencia, parece que no conseguiría dar cuenta de modo correcto de un aspecto sumamente importante de la misma. La teoría de Baars parece dejar fuera del análisis a la conciencia fenomenológica. El modelo de Baars permite explicar a la perfección el acceso del sujeto consciente a la información, a la atención, a las numerosas capacidades introspectivas conscientes, etc., pero no es capaz de proporcionar una explicación satisfactoria de por qué dichos procesos deben estar siempre acompañados por una experiencia consciente. Un ejemplo claro podríamos encontrarlo en el hecho de que la teoría de Baars no es capaz de decirnos por qué el sujeto concreto es capaz de experimentar la información recibida en el espacio global de trabajo. “¿Por qué la accesibilidad global debería dar origen a la experiencia consciente?” (Chalmers 1996: 154). Volvemos a encontrarnos con el mismo problema que anteriormente habíamos identificado en el funcionalismo: si el precio a pagar por naturalizar la conciencia es perder, precisamente, aquello que la hace interesante (esto es, la vivencia concreta de la experiencia consciente, los *qualia*), entonces parece que algo está fallando a la hora de ofrecer la explicación de los fenómenos conscientes.

2.1.2.4. Teorías de nivel superior

Rosenthal parte de la idea de la existencia en la tradición filosófica de dos corrientes dentro del estudio de la mente: una caracterización cartesiana, que identificaría lo mental con lo consciente (con el inconveniente de que impediría cualquier explicación de la conciencia desde lo mental, porque al identificar mente con conciencia cualquier apelación a lo mental presupondría ya la conciencia), y una caracterización naturalizada (directamente vinculada con Aristóteles) que negaría la identificación de lo mental con lo consciente y que afirmararía que lo mental (y la conciencia) puede ser perfectamente explicado como uno más de los numerosos fenómenos naturales que pueblan el mundo. Según este segundo planteamiento, aquello definitorio de los estados mentales es la posesión de toda una serie de propiedades intencionales o de propiedades fenomenológicas (o ambas), siendo la conciencia una característica de lo mental externa a dichas propiedades. Aunque el mismo Rosenthal se ubica dentro de esta segunda corriente, considera posible la explicación de la conciencia desde la intuición del sentido común que subyace a la concepción cartesiana. Así, para Rosenthal,

“los estados conscientes son estados mentales de los que somos conscientes de encontrarnos en ellos. Y, de modo general, ser consciente de algo consiste en tener un pensamiento de alguna clase particular sobre ello. De acuerdo con esto, sería natural identificar el ser consciente de un estado mental con la posesión de un pensamiento simultáneo de que uno se encuentra en dicho estado. Cuando un estado mental es consciente, la conciencia que tenemos de dicho estado es, de alguna manera, inmediata. Por tanto, podemos considerar que dicho pensamiento no estará mediado por ninguna inferencia ni por ninguna entrada sensorial [algo que lo diferenciaría del planteamiento anteriormente atribuido a Carruthers], pero dado que un estado mental es consciente sólo si está acompañado del pensamiento de nivel superior adecuado, podemos explicar nuestro

ser consciente de dicho estado mental mediante la hipótesis de que dicho estado causa dicho pensamiento de nivel superior” (Rosenthal 1986: 335-336).

Vista de este modo,

“la conciencia es una propiedad relacional, la propiedad de estar acompañado por pensamientos de nivel superior y, por tanto, ciertos procesos causales deben mediar entre estos procesos y nuestra conciencia de los mismos. Puesto que dichos estados pueden estar conectados causalmente a diferentes pensamientos, entonces nosotros podemos ser conscientes de dichos estados de un modo distinto en diferentes momentos. La apariencia de los estados mentales, por lo tanto, no sería automáticamente coincidente con la realidad” (Rosenthal 1986: 354-355).

En este sentido, como afirma nuestro autor, no todo pensamiento de nivel superior será consciente. Para que un pensamiento sea consciente requiere de la existencia de un pensamiento relacionado con él que sea, todavía si cabe, superior a él. Así, para que un pensamiento de segundo nivel sea consciente, debe ir asociado y acompañado de un pensamiento de tercer nivel que dé cuenta de la posesión de dicho pensamiento de segundo nivel. En este hecho reside la diferencia importante del planteamiento de Rosenthal, en tanto permite diferenciar entre ser consciente de un estado mental y ser introspectivamente consciente de un estado mental. Sólo cuando somos introspectivamente conscientes de un estado mental somos conscientes también de nuestros estados mentales superiores. Como vemos, Rosenthal se enfrenta de modo claro al planteamiento cartesiano, que indicaba precisamente lo contrario: que todo estado mental es introspectivamente consciente. Suponer esto, nos dice Rosenthal, es descabellado, en tanto que esperar que todos los pensamientos de nivel superior sean conscientes supondría que deberíamos tener infinitos pensamientos en cualquier momento que fuésemos conscientes de un estado mental determinado.

Esta concepción, según Rosenthal,

“no implica necesariamente una teoría materialista o naturalista de la mente, ya que de hecho es totalmente compatible con un dualismo de sustancias, [pero] encaja a la perfección con los planteamientos materialistas en tanto que infiere que lo que hace que los estados mentales sean conscientes es su causar pensamientos de nivel superior que indiquen que nos encontramos en dichos estados. Además, los materialistas podrían sostener que dicha estructura causal se debe a conexiones neuronales” (Rosenthal 1986: 339).

Vista así, la concepción de Rosenthal es capaz de explicar adecuadamente los estados fenomenológicos, porque es capaz de rendir cuentas de la estrecha conexión existente entre estar en un estado consciente y ser consciente del estado. Porque para atribuir conciencia de un estado mental concreto a un sujeto, el pensamiento de nivel superior correspondiente debe referirse a dicho estado, y el único modo en que dicho pensamiento puede ser de un estado mental concreto es refiriéndose a alguien como estando en dicho estado mental concreto.

El problema principal, como en el anterior planteamiento, parece ser la confusión conceptual existente entre las dos nociones de conciencia y la falta de explicación de algunos aspectos de la conciencia, como el aspecto cualitativo de la experiencia consciente a partir de pensamientos de nivel superior. Pero una crítica mucho más concreta la podemos descubrir

ahondando un poco en el tema. Pensemos tan sólo en el siguiente ejemplo. Me acabo de levantar, todavía somnoliento. Me dirijo a la cocina y, como cada mañana, me dispongo a preparar unas tostadas mientras se calienta la leche. Pongo la mesa y me siento. Distráido, mientras escucho las noticias matutinas por la radio, noto un extraño olor. Las tostadas, ¡se me están quemando! El simple olor de las tostadas ha hecho que yo supiera de inmediato que las tostadas se estaban quemando (inspirado en Dretske 1993: 265-266). No podría saber que la tostada se está quemando si no supiera qué es una tostada y no supiera qué significa que algo se está quemando. Parece, pues, claro que ser consciente de un hecho incluye algún tipo de actitud que muestre la posesión (o al menos, el uso) de conceptos. En este sentido, parece razonable suponer la existencia de estados conscientes que no requieren de la presencia de pensamientos de nivel superior, lo que pondría en duda los planteamientos de Rosenthal.

Siguiendo con su intento de ofrecer una explicación naturalista reduccionista de la conciencia, Carruthers aceptará la diferenciación entre conciencia fenomenológica y conciencia por acceso, llegando a afirmar que la conciencia fenoménica es la experiencia misma (Carruthers 2001: 61).

“Hay algo distintivo en *cómo* sería sentir la experiencia en cuestión, y estos estados son fenomenicamente conscientes. Como Block (1995) dijo: la conciencia fenoménica es la experiencia misma”.

En este sentido, nuestro autor también apela a diferentes hechos y evidencias neurocientíficas. Pensemos tan sólo, propone, en nuestro sistema visual. Poseemos, parece ser, al menos dos sistemas diferentes. Por un lado tendríamos un sistema conceptualizador (localizado en los lóbulos temporales de nuestro cerebro) cuyo contenido es consciente. Por otro lado, tendríamos un sistema sensorio-motor (situado en los lóbulos parietales de nuestro cerebro) dotado de numerosos detectores de movimiento controlado *on-line*, de cuyo contenido no seríamos conscientes. Visto de este modo, tendríamos una explicación razonable a problemas concretos de la conciencia, como el problema de la visión ciega (véase 1.4.12.), en tanto que permitiría demostrar que el segundo sistema mencionado, el sistema con contenido inconsciente, también realizaría el mismo papel causal respecto del sujeto que el primer sistema visual, de cuyo contenido sí seríamos conscientes. Pero deberíamos poder también aclarar dicha diferencia, en tanto que Carruthers todavía nos debería una explicación de cómo puede llegar a ser consciente una experiencia determinada.

Frente a este requisito, Carruthers apela directamente a la distinción entre estados primitivos y estados de nivel superior. Sólo aquellos estados de los que somos capaces de darnos cuenta pueden ser considerados como fenomenológicamente conscientes. Los estados fenoménicos conscientes son aquellos cuyas propiedades son reconocidas por nuestras capacidades introspectivas. Vistos de este modo, los estados conscientes serían aquellos estados que podemos reconocer como realizados en nosotros de un modo no-inferencial en tanto que

están presentes en nosotros de un modo subjetivo (Carruthers 2000b).⁵ Dicho de otro modo, un estado consciente sería aquel estado de nivel superior que daría cuenta de las propiedades primarias que presentaría un estado mental y que permitiría darnos cuenta de la presencia subjetivamente dependiente de dicho estado en nosotros como sujeto de percepción y de experiencia.

En este sentido, el carácter cualitativo de una experiencia concreta sería un estado intrínseco y no-representacional, en tanto que aquello que realmente llegaría a ser consciente sería el pensamiento de nivel superior realizado por las capacidades cognitivas internas de los sujetos capaces de dar cuenta del mismo. Así, es posible ser un eliminativista respecto de los *qualia*, estableciendo su identidad con determinados estados cerebrales con determinadas funciones concretas (como hemos visto en la sección 2.1.2.3.) sin negar la existencia de la presencia de un carácter cualitativo que permita, primero, diferenciar los diferentes tipos de sensaciones y, segundo, indicarnos cómo sería tener una determinada sensación concreta, pongamos como ejemplo experimentar la rojez de un tomate maduro. Siendo así la situación, como hemos dicho, el mejor modo de explicar la conciencia sería negar la existencia de *qualia* (tal y como los presentaría la tradición) y establecer una relación entre el papel causal y el contenido intencional de los estados cerebrales y funcionales que se dan en nuestro sistema nervioso central, en vistas a defender un estatuto científico de la psicología intencional y reivindicando, así, la realidad del contenido intencional de nuestros propios estados (Cf. Botterill and Carruthers 1999: 227-271; Carruthers 2001: 65). Desde este planteamiento, la explicación reduccionista propuesta por Carruthers requiere establecer un nivel explicativo metafísicamente superveniente con implicaciones eliminacionistas. Por ello mismo, críticas como las expuestas por Chalmers podrían serle aplicadas, en tanto que, nos dice, es perfectamente concebible la existencia de estados funcionales e intencionalmente idénticos que, sin embargo sean fenomenológicamente distintos porque no podemos conceptualizar el aspecto sensitivo de las experiencias en términos funcionales o intencionales, estableciendo así el hiato explicativo entre funciones neurológicas y cognitivas y estados fenoménicamente conscientes. Pero, se defiende Carruthers, las críticas mencionadas podrían ser ciertas si concibiéramos los estados de nivel superior en un sentido funcional, pero éste no es el caso. Dichos estados están expuestos en términos de reconocimiento. Nuestros conceptos de experiencia consciente consistirían tan sólo en la posesión de capacidades inmediatas de reconocimiento de los diferentes tipos de estado fenoménico,⁶ siendo la capacidad concreta de reconocer una determinada experiencia como una experiencia fenoménica de un tipo determinado (digamos, nuestra capacidad de reconocer la rojez en la experiencia de un tomate maduro) lo que cuenta a

⁵ “Los estados fenoménicamente conscientes son aquellos que pueden ser reconocidos en sí mismos, de modo no-inferencial, en virtud del modo en que los sentimos o del modo en que se nos presentan de manera subjetiva” (Carruthers 2001: 62).

⁶ Directamente asociada con esta posibilidad, tendríamos también la existencia de conceptos cualitativos que, de un modo inmediato, podrían aflorar a la conciencia del sujeto perceptor, en tanto que respondería, de un modo general, a la experiencia concreta de la manera en que habitualmente lo haría todo ser consciente ante la presencia de las mismas características relevantes.

la hora de presentarlo como un estado fenoménico consciente y no el papel funcional concreto que sea capaz de realizar.

El principal problema que, a mi entender, ahora mismo se le presentaría a una teoría como la de Carruthers sería la de explicar exactamente en qué consisten dichas capacidades de reconocimiento de las cuales no somos conscientes pero que sí permiten, sin embargo, ser conscientes de sus contenidos de nivel superior, en tanto que darían cuenta de los estados fenoménicos primarios. Porque, para Carruthers, estos sistemas incluirían una serie de módulos capaces de crear creencias. La dificultad residiría principalmente en explicar el estatuto de dichos módulos, vistas las evidencias neurológicas que desmienten la existencia de los mismos.⁷

Un problema mucho más serio y mucho más concreto es el que plantearía la posibilidad de la existencia de estados mentales inconscientes que, sin embargo, tuvieran un papel causal. Desde el planteamiento de una teoría de nivel superior, sólo tendría sentido hablar de eficacia causal de los estados fenoménicos conscientes, esto es, de la eficacia causal de estados intencionales dependientes de los sistemas de reconocimiento descritos anteriormente. Pero cabe la posibilidad real de la existencia de estados intencionales inconscientes, al existir evidencias a la hora de aceptar la posibilidad de la existencia de una intencionalidad sin conciencia que nos obligarían a mostrarnos cautos a la hora de aceptar las conclusiones de Carruthers y Rosenthal. (También podrían aplicarse estas críticas a la teoría de Galen Strawson o Searle, en tanto que ambos aceptarían la idea de la existencia de estados inconscientes que, sin embargo, serían intencionales).⁸

La conciencia parece que no involucra tan sólo una serie de capacidades de reconocimiento, sino que presupone una mayor disponibilidad del contenido proporcionado por los diferentes procesos cognitivos que tienen lugar en el sistema nervioso y en el cerebro. Esto es, proporciona un nivel introspectivo dependiente más de un papel causal de las representaciones presentes en los sujetos que de los sistemas no-representacionales capaces de proporcionar determinados contenidos. La intencionalidad, por sí sola, no sería suficiente para la introspección; y como parece que la introspección es una condición necesaria de todo estado consciente, entonces la intencionalidad no parece ser suficiente para la conciencia. La intencionalidad, en sí, es tan sólo un recurso que permitiría indicar los objetos sobre los que se

⁷ Contra la modularidad de la mente ya me he pronunciado en Colomina 2008c. En este texto vengo a sostener que, contra ciertos autores como Pinker o Lightfoot que sostienen la posible existencia de un órgano de la mente específico para el lenguaje, los datos aportados en la actualidad por la teoría neurofisiológica permiten inferir, primero, una descentralización de las labores cognitivas que permitirían explicar la posesión de un lenguaje (directamente contra la posibilidad de la existencia de una gramática universal) y, segundo, la existencia de un nivel semántico externo al individuo que posibilita una explicación mucho más plausible de la originación del contenido semántico que la que son capaces de ofrecer las teorías de corte internista.

⁸ Galen Strawson reivindica que los estados inconscientes no son genuinos estados mentales: “estrictamente hablando, no existen fenómenos mentales disposicionales no-experienciales” (Strawson 1994: 167). Searle afirma que “nosotros no tenemos una noción de mente como separada de la conciencia” (Searle 1992: 18). Ello les llevará a presuponer un Principio de Explicación Intencional (Rey 2008: 60). Según el PEI, cuando el comportamiento de un sistema, o los procesamientos internos, sean tan sensibles que no puedan ser explicados mediante intencionalidad consciente, existe la posibilidad científica de que dicho comportamiento y dichos procesamientos estén causados por contenidos que soporten estados intencionales introspectivamente inaccesibles a los sujetos. Una crítica general contra la posibilidad de este tipo de teoría, donde se desechan por completo la posibilidad de estados inconscientes, la encontramos en Rey 2008. Para una amplia crítica al Principio de Conexión de Searle, análogo al PEI, véase el capítulo 2.

dirigiría la mente, pero no sería un proceso que hiciera conscientes a los estados mentales. Por tanto, parece necesario presuponer alguna clase de proceso perceptivo interno que permita, por un lado, indicarnos el tipo de sensación que estamos teniendo en cada momento y, por otro, mostrarnos el contenido representacional que toda experiencia parece transmitir (Rey 2008: 73-74).

2.1.3. El contenido representacional de la conciencia

Según hemos estado viendo en las anteriores secciones, un importante hecho a la hora de poder dar cuenta de la conciencia es el tipo de contenido que se presupone que deberían tener los estados mentales fenoménicos. A lo largo del apartado 1.2. dejamos constancia de una seria dificultad a la hora de presentar la diferencia fuerte entre el contenido no-conceptual (CNC) y el contenido conceptual de la conciencia. Como vimos allí, la discusión se realizó en términos de dependencia. Si consideráramos posible que el CNC fuera independiente del contenido conceptual, entonces deberíamos asumir que el contenido conceptual quedaría sin fundamento porque faltarían las bases sobre las que se asentaría el contenido proposicional de los estados mentales de los sujetos conscientes (Bermudez 1994). Pero, por otra parte, si defendiéramos la interdependencia del contenido conceptual y del contenido no-conceptual, entonces ello nos llevaría a presuponer la existencia de una tercera instancia que multiplicaría innecesariamente la realidad metafísica y psicológicamente subyacente (Peacocke 1994). En este sentido, y como defendimos al final del mencionado apartado, una opción viable era defender una noción neutra de representación que permitiera, por un lado, servir de base para el contenido conceptual de las experiencias fenoménicas del que los sujetos son conscientes y, por otra, evitar en la medida de lo posible el extender las entidades metafísicas. (La otra opción, la posibilidad de abandonar definitivamente una noción representacional de contenido, será considerada como la base sobre la que se desarrollará el capítulo final).

2.1.3.1. Fodor sobre conceptos

Dentro de la tradición, existen dos explicaciones predominantes a la hora de intentar explicar la naturaleza de los conceptos. La primera, que podríamos denominar el punto de vista psicológico, mantiene (como veremos) que los conceptos son representaciones mentales, esto es, son particulares mentales que tienen propiedades semánticas. Por el contrario, en segundo lugar, como contrapuesto, encontramos el punto de vista semántico. Según esta explicación, los conceptos no son representaciones mentales, porque los conceptos no pueden estar en la mente. Los conceptos, entonces, serían objetos abstractos de una cierta clase (tal y como defendió Frege). Desde su punto de vista, el significado de una oración declarativa es una proposición (o pensamiento, empleando el lenguaje fregeano). Dicha proposición debe entenderse como un objeto abstracto en tanto que es posible su existencia de modo separado de las mentes de los hablantes. Si oraciones en lenguas diferentes pueden transmitir el mismo significado es debido a que el significado de las mismas puede expresar la misma proposición, del mismo modo, objetos en 'mentes' diferentes pueden expresar el mismo pensamiento. Porque tanto en un caso

como en el otro lo que tenemos son modos diferentes de presentación (sentidos distintos) de dicho contenido proposicional. Peacocke sería uno de los máximos exponentes de esta posición en la actualidad. Hoy en día ha cobrado fuerza la posibilidad de una tercera vía: se considera plausible establecer una definición de concepto que combine tanto representaciones mentales como sentidos. Vistos de este modo, los conceptos serían representaciones mentales implementadas según el sentido concreto que expresan. No disponemos de espacio suficiente para desarrollar más esta idea. Para una exposición más detallada, véase Margolis and Laurence 2007. También Prinz 2002.

Hablar aquí del planteamiento de Fodor nos va a ser de suma utilidad porque nos permitirá avanzar ciertas explicaciones básicas necesarias a la hora de precisar la noción de contenido o representación conceptual (algo que reservamos para la sección 2.1.3.2.). El segundo capítulo de Fodor (1998) plantea cinco condiciones restrictivas respecto de lo que debería ser un concepto. Según la primera condición, un concepto es un particular mental. Visto de este modo, un concepto es un estado mental que es causa, y puede ser efecto, de otros estados mentales. Tiene una existencia efectiva en tanto que en el cerebro existe un determinado estado físico que permite instanciar dicho concepto, por lo que ante dos conceptos diferentes cabría suponer la existencia de dos estados físicos distintos que los implementaran.

La segunda condición es que un concepto es siempre una categoría, dado que permitiría delimitar el conjunto de objetos a los que referiría dicho concepto, agotando así su extensión.

En tercer lugar, los conceptos son composicionales, esto es, los conceptos pueden combinarse entre sí y el valor semántico de dicha composición será una función dependiente de los valores semánticos de los conceptos que son sus constituyentes. Esta característica composicional de los conceptos, además, permitiría explicar por qué el pensamiento es sistemático y tan productivo (Fodor and Pylyshyn 1988). Visto así, entonces, el pensamiento también debería ser composicional, por lo que también deberá tener una estructura repleta de elementos simples constituyentes.

Siguiendo con la lista de restricciones sobre los conceptos, Fodor plantea en cuarto lugar un requisito de adquisición: los conceptos (o, al menos, la gran mayoría de ellos) deben poder ser aprendidos. Cuando un individuo adquiere un nuevo concepto, nos dice Fodor, entonces el sujeto queda anclado nomológicamente a la propiedad que dicho concepto expresa. Es decir, una vez el sujeto conoce lo que engloba el concepto, se establece una dependencia asimétrica entre las instanciaciones correctas y las instanciaciones incorrectas (erróneas) de dicho concepto. Como Fodor advierte, al establecer la plausibilidad de esta condición, evitamos tener que rendir cuentas de hipótesis intermedias que permitan explicar el despliegue del concepto. Esto es, permite concebir los conceptos adquiridos de este modo como primitivos sin necesidad de tener que recurrir a nociones genéticas o innatas. Visto de este modo, entonces, la mayor parte de los conceptos que posee un individuo serán primitivos sin necesidad de tener

que presuponer, como haría por ejemplo Pinker (1997),⁹ que dichos conceptos están ya insertos de manera innata en la estructura conceptual inicial en los primeros estadios de la vida humana.

En último lugar, para Fodor, un requisito imprescindible es que los conceptos deben ser públicos y compartidos. Aplicando la distinción entre tipos y casos (o ejemplares, referido a la clásica contraposición entre *type and token*) y teniendo presente la tesis funcionalista de la múltiple realizabilidad, Fodor insiste en que existen particulares mentales (esto es, casos) que permitirían instanciar el mismo tipo de concepto. Visto así, entonces, dos individuos pueden ser poseedores del mismo concepto aunque implementen dos casos diferentes del mismo tipo mental. En este sentido, dos personas pueden diferir en sus experiencias particulares pero compartirán los mismos conceptos. Lo que subyace a esta explicación es una noción fuerte de identidad teórica aplicada a los conceptos que evitaría el relativismo conceptual.

2.1.3.2. Contenido conceptual y no-conceptual en la experiencia fenoménica

La definición de Fodor, como hemos visto, valdría para aquellos que defienden una concepción conceptual del contenido de las experiencias fenoménicas, pero no para aquellos que defienden una noción no-conceptual. ¿Cuál de las anteriores condiciones deberíamos abandonar para poder establecer una noción de contenido no-conceptual? El problema principal, como ya nos ha indicado Fodor, es que las cinco condiciones presentadas son necesarias para establecer una teoría representacional de la mente. Visto de este modo, advierte, eliminar una de las condiciones supondría establecer una teoría de la mente que sería no-representacional. O bien los seguidores del CNC cumplen las condiciones mencionadas o bien deben confeccionar una teoría no-representacional. Una alternativa viable sería, por supuesto, que el teórico del CNC confeccione una noción alternativa de representación.

Volvamos a la concepción del CNC de Bermudez (1994). La noción de contenido defendida allí tiene en cuenta las cuatro primeras condiciones de Fodor, porque, primero, la condición 1 establece una distinción entre una teoría realista y una instrumentalista del contenido mental, pero no nos dice nada respecto del estatuto del contenido mismo y, segundo, según la condición 2 permitiría también establecer un criterio de evaluación semántica para el CNC que permitiría diferenciar entre representaciones verídicas y falsas, un criterio que, según la condición 3, también permitiría la composicionalidad del contenido. Pero, además, según la condición 4, tampoco los teóricos del CNC estarían dispuestos a aceptar que el contenido es innato. Visto así, la condición que parecería demarcar la separación entre contenido conceptual y no-conceptual sería la condición 5 de Fodor, por lo que el defensor del CNC debería desestimarla.

El CNC es el contenido tal y como se presenta en la experiencia, es el modo según el cual experimentamos la realidad. En sentido estricto, entonces, si negamos dicha condición resultará que, presentados ante el mismo segmento de la realidad, dos individuos no implementarán el mismo CNC. Visto así, si las representaciones del CNC son idiosincrásicas, es

⁹ Para una crítica directa a la teoría expuesta en dicho volumen por Pinker, puede verse Fodor (2000).

decir, diferentes para cada uno de los sujetos conscientes del mismo segmento de la realidad, entonces cabe la duda razonable sobre el modo en que el CNC pueda proporcionar a dichos individuos los elementos necesarios para construir los conceptos comunes que comparten. Podríamos hacer caso a Stalnaker (1998a) y afirmar que no existe una identidad interindividual en el mundo conceptual, esto es, negar que dos individuos puedan compartir el mismo concepto instanciado por la misma experiencia fenoménica, que ambos individuos siempre tendrán contenidos diferentes ante una misma experiencia, por lo que la experimentarán de modo distinto. O bien podríamos considerar que McDowell (1994) lleva razón al insistir en que la esfera de lo conceptual abarca también el ámbito de las experiencias. Visto así, nuestros conceptos diferirían entre sí en la medida en que también difieran nuestras experiencias concretas, ampliándose cada vez más las diferencias entre los individuos cuanto más lejanos se encuentren los repertorios experienciales de los diferentes individuos. Aparece aquí un problema: cómo conseguir que los pensamientos y las intuiciones se conecten racionalmente (el viejo problema kantiano de la imposibilidad de la existencia de pensamientos vacíos y de intuiciones sin concepto).¹⁰ La solución que decide McDowell pasa por entender la realidad conceptual como algo que se extiende hacia la experiencia, siendo ésta pasiva pero siendo, también, reflejo de capacidades conceptuales. En este sentido, las impresiones en los sentidos tendrían, *per se*, contenido conceptual, pero ello no impediría que pudieran ser experiencias del mundo.

Una solución alternativa es la ofrecida por Cussins (1990). Para este autor, la posibilidad de representar objetos y propiedades emergería a partir de experiencia pre-objetivas, esto es, dicha posibilidad dependería de la estructura no-conceptual del contenido de las distintas capacidades de reconocimiento de los sujetos perceptores. Visto de este modo, los conceptos serían atómicos. Poseerían una estructura no-conceptual modelada según la experiencia y su estructura podría captarse a partir de la estructura representacional interna (no-semántica) de los mismos. Visto así, lo conceptual sería un constructo derivado del CNC, pero éste sería el causante de la experiencia concreta. El problema de esta perspectiva recaería en el papel irrelevante que los conceptos desempeñarían en la vida mental de los sujetos, dado que los

¹⁰ Frente a esta imagen tradicional kantiana, han reaccionado varios filósofos en la actualidad. Nöe (2004) propone una teoría que pretende mostrar en qué sentido podemos entender la percepción como un modo de acción. Nöe (2004: 11) la denomina perspectiva de entrada-salida, porque presupone que existe una conexión inmediata por medio de la intervención de mecanismos que conectarían intuiciones (datos entrantes como sensaciones) y pensamientos (datos salientes como respuestas a dichos estímulos de entrada). Esto es, propone la existencia en los organismos de determinadas capacidades sensorio-motoras corporalmente dependientes que permitirían dotarlo de un conocimiento concreto que le permitiría actuar en cada momento. Frente a esta teoría, Block (2001) ha argumentado que peca de conductismo porque presupondría que tener una experiencia sería como disponer de una determinada relación de entrada y salida. Pero esto no es del todo cierto. Lo que Nöe afirma estar reivindicando es la existencia de una especie de mecanismos que proporcionarían al organismo una serie de *know-how* que le permitirían realizar, de manera inmediata, toda una serie de razonamientos que le llevarían a la acción. Estaría proponiendo una especie de conocimiento práctico. Y ésta, por supuesto, no es una tesis conductista (Nöe 2004: 33). Otra defensora de una teoría similar es Hurley (1998). Véase Kobes (2000) y Kinsbourne (2000) para una serie de contundentes críticas, así como Hurley (2000) para una defensa de su actual posición. En la actualidad, Nöe defiende que la conciencia es algo que acontece en nosotros pero que no está localizada en el cerebro. Por supuesto, tendría una base neurológica, pero la conciencia no se encuentra en dicha base, sino que es algo que nosotros hacemos de manera activa mediante la interacción dinámica con nuestro entorno. Véase Nöe 2009.

primeros aparecerían como idealizaciones de las estructuras subyacentes de los segundos, lo que dificultaría la posibilidad de dotar de una naturaleza compartida a los propios conceptos.

Al apelar a mecanismos universales de interacción entre sujeto y entorno, podríamos conceder a la posición de Bermudez que las experiencias pre-objetivas sean idénticas entre individuos, siempre que consideremos plausible que éstas pudieran estar constituidas a partir de interacciones con elementos públicos. De este modo, el carácter compartido de las experiencias vendría garantizado por las restricciones establecidas entre el contexto y el sujeto perceptor. Desde este punto de vista, el contenido representacional basado en dichas propiedades relacionales poseería una naturaleza pública y compartida determinada por la regularidad de determinadas propiedades establecidas por el entorno y por el funcionamiento de los diversos sistemas cognitivos capaces de captar dichas propiedades.

Como hemos visto, las teorías del CNC que reivindican su carácter no-público tienen la dificultad añadida de tener que establecer un marco conceptual que no derive en relativismo. Pero las teorías que se empeñan en establecer criterios demasiado estrictos para evitar dicho relativismo, véase si no el caso de Bermudez, hacen dudar de la naturaleza pública del CNC. Si el CNC cumple el criterio de corrección, entonces parece ser el caso que dos ejemplares concretos del mismo tipo de CNC pueden poseer las mismas condiciones de corrección, lo que indicaría la implementación de la misma experiencia en ambos sujetos (aunque sólo sea así si se considera que dichas condiciones de corrección son suficientes para tener dicho concepto). Ello nos indicaría, precisamente, que la noción de publicidad no es tan crucial en la diferenciación entre CNC y contenido conceptual, por lo que parece difícil establecer dicha dicotomía.

2.1.3.3. ¿Son independientes el contenido conceptual y el no-conceptual?

Visto que no parece plausible dejar de lado la condición quinta de concepto establecida por Fodor, podríamos adoptar una estrategia mucho más arriesgada para procurar la correcta definición del papel del contenido conceptual y del contenido no-conceptual. Esta pasaría por intentar negar la condición segunda de Fodor, lo que supondría admitir que las entidades con CNC no poseerían ningún tipo de evaluación semántica, es decir, no podríamos decir de ellas que se refieren o denotan cosas o hechos, aunque sean irreales o abstractos (Manrique 1999). La mayor dificultad con esta propuesta la encontramos en que una entidad tal parece no poder dar cuenta de la posibilidad de representar erróneamente un estado de cosas, lo que supondría que no podría ser una representación en absoluto (Bermudez 1998: 84). Una dificultad mayor pasaría por tener que negar, también, la tesis de la composicionalidad de los conceptos, dado que el contenido de los compuestos no puede ser una función establecida por el valor semántico del contenido de sus componentes.

Una posibilidad alternativa pasaría por defender una posición representacionista como la defendida por McGinn (1989), que parte de la idea de la existencia de modelos mentales. Podríamos decir que en esta obra no queda claro del todo si McGinn considera que los modelos mentales serían representaciones con contenido no-conceptual. Pero parece claro que McGinn defiende precisamente que los modelos mentales serían de esta clase, dado que emplea la noción

de simulación para referirse a su implementación. Desde esta perspectiva, entonces, un modelo mental sería, en palabras de nuestro autor, una estructura mental capaz de simular un determinado aspecto de la realidad. Las representaciones basadas en un modelo adecuado se considerarían como verdaderas, mientras que las representaciones basadas en un modelo inadecuado deberán ser catalogadas como falsas. En este sentido, los modelos adecuados serán aquellos capaces de dotar al agente de las mejores opciones posibles para enfrentarse a la supervivencia y los problemas que le plantea su entorno. Visto de este modo, las representaciones erróneas pueden, a menudo, resultar adecuadas desde un punto de vista evolutivo (pensemos tan sólo en el joven lémur que emite el sonido de alarma cuando observa, a lo lejos, un águila culebrera cuando en realidad no supone una amenaza real para el grupo, pero que le permite ponerse a refugio y alertar a los demás miembros). Aunque, por el contrario, no parece necesario que una representación deba cumplir estrictamente una condición de adecuación teleológica. Siguiendo esta línea argumental, como nos dirán las teorías de Dretske o Millikan (como veremos un poco más tarde, sección 2.2.1.2.), los modelos mentales representacionales podrían cumplir perfectamente la función que les corresponde y ser puramente informativas, careciendo pese a ello de contenido conceptual.

Como vemos, lo realmente problemático aquí es establecer una teoría que pretenda ser representacional y, sin embargo, quiera prescindir del requisito de evaluación semántica. ¿Podríamos establecer una noción de CNC sin tener que acabar transformándolo en contenido conceptual? Manrique (1999: 641) opta por abandonar el Principio de Autonomía propuesto por Bermudez (1994) (y que ya analizamos en el primer capítulo, apartado 1.2.). Visto de este modo, el contenido conceptual y el contenido no-conceptual serían interdependientes: para que una entidad tuviera CNC requeriría la presencia de un concepto, pero para que una entidad tenga contenido conceptual se requiere la presencia de experiencias no-conceptuales. Entonces, un estado con contenido sería representacional y un estado representacional sería informativo, pero no a la inversa. Esto es, no todo estado informativo sería representacional, en tanto que el contenido de un estado informativo no podría constituirse como un objeto intencional a falta de un concepto. Es decir, sólo la presencia de un concepto permitiría indicar la experiencia concreta en la que se encuentra el sujeto en un momento determinado, en tanto que tan sólo se podría implementar si se da un determinado CNC. El CNC no podría categorizar si no es a través de un concepto concreto, pero el concepto sólo podría categorizar la realidad si está presentada de un modo concreto al sujeto perceptor por un determinado CNC, mostrándose necesaria, entonces, la interdependencia entre CNC y contenido conceptual.

Una consecuencia de establecer la interdependencia entre CNC y contenido conceptual sería tener que afirmar lo que Clark (1993) denomina innatismo conceptual mínimo (véase también Clark 2003). Esto es, la existencia de mecanismos mentales con los que todo sujeto cuenta desde su nacimiento y que le permitirán el desarrollo posterior de toda una serie de estructuras funcionales que actualizarán el mundo del modo adecuado según el caso. La peculiaridad de esta noción es que permitiría establecer toda una serie de mecanismos conceptualizadores que, por un lado, estarían orientados hacia la localización de aquellos

aspectos de la realidad más relevantes para permitir la subsistencia del organismo en su entorno (la supervivencia) y, por otro lado, permitirían establecer dichos conceptos sin necesidad de presuponerlos como entidades intencionales hasta que no son rellenados por el CNC a medida que tiene lugar el desarrollo de la historia cognitiva propia de cada uno de los individuos. Visto de este modo, el contenido de una experiencia puede incluir elementos conceptuales sin necesidad de ser representacional, lo que facilitaría las cosas, como veremos en el capítulo final, a la hora de tener que ofrecer una explicación satisfactoria del problema de la conciencia.

Si se tiene esto en cuenta, entonces, lo que propone este modelo es una explicación en la que los estados informacionales de la red inicial son de carácter no-conceptual para, según avanza la integración en el seno de la comunidad lingüística y se produce el desarrollo histórico y cognitivo particular de cada individuo, posteriormente constituirse como conceptos. Pero para que ello tenga éxito, debe negarse que los estados iniciales tengan contenido previo, en tanto que éste se establece en función de la historia de entrenamientos cognitivos particular de cada individuo, dependiente del nivel de organización determinado por el objetivo básico, de supervivencia, de cada organismo en concreto.

La crítica más feroz que una concepción tal de interdependencia podría sufrir es, precisamente, la posibilidad de que ambas nociones se tornen ininteligibles. Si consideramos que ambas nociones se necesitan mutuamente, esto puede llevar al error de presuponer que ambas nociones pueden identificarse como una única entidad. Pero existen casos concretos en los que parece razonable pensar que existe una disociación entre contenido conceptual y no-conceptual. Pensemos por ejemplo en los casos de agnosia visual. Según Bermudez (1998: 81), en determinados casos de agnosia asociativa, los errores de reconocimiento son debidos a cierto tipo de daño selectivo a nivel conceptual, por lo que el CNC quedaría más o menos a salvo. De modo contrario, en casos de agnosia aperceptiva, el daño se ha producido a nivel no-conceptual, siendo el nivel conceptual el que permanecería sano y salvo. Si esto fuera cierto, es decir, si pudiera demostrarse que en diferentes casos de agnosia el daño podría localizarse de manera concreta o bien en el nivel estructural primario (los casos de agnosia aperceptiva) o bien en el nivel de los sistemas de reconocimiento (los casos de agnosia asociativa), entonces parecería demostrarse, también, que el CNC se emplea en la construcción de conceptos pero no en el resto de implementaciones de conceptos.

Otro ejemplo podríamos encontrarlo en determinados casos de esquizofrenia alucinatoria (Manrique 1999: 645). Una alucinación esquizofrénica resulta incomprensible para una persona que no sufra la enfermedad porque supone una alteración en las estructuras más fundamentales, las de tiempo y espacio. El problema principal de este tipo de anomalía reside no en el hecho de que el sujeto construya entidades incorrectas en su imagen del mundo, sino en el estatuto ontológico alterado de dichas entidades. A diferencia de los otros dos casos de agnosia, en el caso de la esquizofrenia parece que el enfermo puede tener la mente literalmente dividida en un mundo de representaciones experienciales que no consigue articular con el mundo de sus representaciones conceptuales. Por ello mismo, la dificultad de explicar la disociación entre este tipo de representaciones en casos como los mencionados (también

ocurriría lo mismo en los casos de cerebro escindido que introdujimos en la sección 1.4.13.) parecerían presuponer una necesaria separación entre contenido no-conceptual y contenido conceptual. El problema principal con este punto de vista sería presuponer que tanto las experiencias como los conceptos pueden entenderse en un sentido representacional. Si en lugar de ello los entenderíamos desde un punto de vista diferente, por ejemplo, como dependientes de un sustrato computacional, podríamos llegar a mostrar cierta interconexión entre ambos, permitiendo incluso salvar problemas patológicos como los anteriormente mencionados.

2.2. Aproximaciones representacionales de la conciencia

Como hemos visto en el apartado anterior, el representacionalismo individualizaría los estados mentales en términos de su contenido representacional. Es normal distinguir entre aquello que es representado y el vehículo mediante el cual se da dicha representación. En este sentido, ambas nociones pueden ser interpretadas de modo diferente. Según una concepción materialista, ambas serían entidades físicas. El vehículo de transmisión sería un estado neurofisiológico del cerebro del sujeto consciente, mientras que el contenido de dicho estado es una determinada característica física del entorno en el que estaría situado el sujeto consciente. Visto así, la relación entre el representante y el representado sería una relación física.

En los próximos apartados atenderemos a los desarrollos representacionistas más relevantes en la actualidad.

2.2.1. El carácter subjetivo desde una perspectiva representacionista

A la hora de dar cuenta del carácter subjetivo de la experiencia consciente, un problema que sin duda debe asaltar a todo teórico es poder explicar la manera en la que un estado mental es capaz de ser intrínsecamente sobre sí mismo. Es decir, como hemos comentado anteriormente, una de las dificultades a la hora de construir una explicación de la conciencia recae en mostrar cómo podemos mantener que un estado mental fenoménico como lo es la experiencia puede tener un carácter subjetivo. En tanto que subjetivo, la experiencia se refiere a un objeto tal y como éste se me presenta a mí en mi experiencia. Por ello mismo, al intentar aportar respuestas concretas a esta problemática, podríamos distinguir entre dos aspectos diferentes de la intencionalidad de dicho estado experiencial. Por un lado, tendríamos el contenido intencional y, por otro, tendríamos el modo intencional.¹¹ Al igual que en las actitudes proposicionales, también en las experiencias podríamos aplicar la diferenciación entre contenido y actitud psicológica. La diferencia entre ver un fantasma y ver un perro es una diferencia de contenido, mientras que temer a un fantasma y ver un fantasma es una diferencia en la actitud psicológica de presentación de la experiencia.

¹¹ La diferenciación entre contenido y actitud es muy habitual cuando se habla de actitudes proposicionales. Pensemos tan sólo en la diferencia entre creer que está lloviendo y tener la esperanza de que llueva. Aunque ambas actitudes proposicionales tienen el mismo contenido (el que llueva), éstas se presentan de modo diferente. Igualmente, dos creencias diferentes pueden compartir la misma actitud psicológica, aunque difieran en contenido (piénsese tan sólo en la diferencia de contenido entre creer que el Teide es un volcán y creer en el Ratoncito Pérez).

Pero reconstruir el carácter subjetivo en términos de la actitud psicológica de la experiencia no es lo mismo que ofrecer una explicación fisicalista del mismo. Necesitaríamos para ello poder ver cómo los estados neurofisiológicos de nuestro cerebro son capaces de implementar las experiencias. La dificultad estriba en que parece imposible ver cómo puede realizarse esta explicación de una manera satisfactoria. Por ejemplo, Levine desarrolla un trilema que muestra de modo patente la dificultad de poder explicar neurofisiológicamente las experiencias subjetivas (Levine 2001: 171-173). Cuando alguien tiene una experiencia consciente de rojez (como nuestro amigo, aquel que experimenta un tomate maduro), el cual (supongamos) se implementa por un determinado estado neurofisiológico, hay tres diferentes descripciones bajo las cuales es capaz de representarse dicha experiencia. En primer lugar, puede representárselo en tanto que estado cerebral. En segundo lugar, puede representárselo en tanto que experiencia de rojez. O, en tercer lugar, puede representárselo de modo indéxico, esto es, en tanto algo capaz de realizar dicha representación.

Según Levine, deberíamos abandonar de inmediato la primera opción porque es implausible:

“La representación de la experiencia en tanto estado cerebral obviamente no funciona –no hay razones para pensar que una representación de primera persona ofrezca una descripción neurofisiológica” (Levine 2001: 172).

La dificultad con esta afirmación es, básicamente, que el único motivo que parece impulsar a Levine para sostenerlo es que la opción primera simplemente no casa de modo adecuado con su propia concepción de la fenomenología de la experiencia consciente en tanto que revelada de modo inmediato a nuestra introspección. Pero, admitámoslo, éste no es un argumento demasiado robusto. Pensemos tan sólo por un momento en el siguiente contraejemplo: pensemos en nuestro concepto de materia gris. Cuando pensamos en él, nuestro pensamiento tiene lugar, sin duda, de algún modo en estados neurofisiológicos del cerebro que, seguramente, incluirán también una cierta cantidad de materia gris en su composición. Por tanto, cuando pensamos en la materia gris, deberíamos presuponer que la materia gris se está auto-representando a sí misma. Pero esto no es lo que parece defender Levine cuando habla de introspección, por lo que deberíamos presuponer la existencia de estados a un nivel sub-personal que, sin embargo, permiten pensar dichos conceptos. (El ejemplo es una modificación de un ejemplo sobre el dolor aparecido en Hill 2005).

La segunda opción, para Levine, también es desechable. Porque la experiencia de rojez debe representarse a sí misma como una experiencia de rojez. Esto es, debe representarse a sí misma como empleando una descripción de experiencia roja. Pero, podríamos objetar, parece que cuando una experiencia se representa a sí misma de este modo, no sólo se representa como siendo de rojez, sino también se representa como siendo de rojez de un cierto modo concreto. Esto es, no se representa de un modo indeterminado, sino que se representa como una concreta descripción que determina cómo es dicha experiencia en concreto. Si esto es así, entonces la

réplica de Levine queda desautorizada, porque la experiencia cumple con su pertinente carácter subjetivo.

Contra la tercera opción, Levine insiste en que considerar la experiencia como un indéxico impide que pueda ser representada de modo introspectivo. Pero éste no parece ser el caso. Pensemos en una experiencia de rojez en tanto que representada de modo indéxico. Puede indicar perfectamente el estado que la representa. No es necesario pensar, como hace Levine, que la experiencia debe ser representada como algo (en este caso, la rojez), sino que podemos tan sólo pensar que la experiencia es representada de un modo no-conceptual.

Visto lo visto, no tenemos más remedio que considerar que el argumento de Levine no es conclusivo, en tanto que no es capaz de mostrar la necesidad de que todo estado consciente deba ser conceptual y no ha sido capaz de mostrar la irrelevancia de los hechos neurofisiológicos empleados. Pero, además, no es exhaustivo, dado que cualquier interpretación naturalizada de la representación, la conciencia o de la intencionalidad deberá, tarde o temprano, ofrecer una explicación reduccionista de la misma en términos de algún tipo de relación causal y, como sabemos, toda relación causal es anti-reflexiva por definición, por lo que no se respetaría el papel que el contenido fenomenológico tendría en la teoría de Levine.

Pero esto, también, podría ser un inconveniente para las teorías que persiguen la naturalización del carácter subjetivo de la conciencia, porque a la hora de establecer una interpretación para la auto-representación, ésta siempre incluirá una relación física entre el estado mental representante y el aspecto del entorno que es representado, siendo necesario establecer de modo determinado dicha relación. Si toda relación física es anti-reflexiva, entonces toda aproximación naturalista (aunque sea representacional) de la auto-representación de los estados mentales con carácter fenoménico será falsa porque deberíamos establecer de modo correcto que ningún estado mental es capaz de representarse a sí mismo (Kriegel 2005: 45).

Anteriormente hemos indicado el modo en que el funcionalismo intenta salvar este problema (véase sección 2.1.2.2.). Pero el funcionalismo no es el único intento de establecer una teoría naturalista de la conciencia. Podemos encontrar, también aproximaciones causales informacionales (como la de Fodor y la de Dretske), aproximaciones teleológicas (como las de Millikan o Papineau), aproximaciones de nivel primario (como las de Tye o Byrne), aproximaciones dualistas (como las de Chalmers o Searle) y aproximaciones de orden superior (como las de Carruthers, Rosenthal o Dennett). En las siguientes secciones atenderemos a varias de estas interpretaciones.¹²

¹² La concepción representacional de Fodor ha sido brevemente formulada en la sección 2.1.2.2. A analizar la teoría de Searle, y siguiendo nuestra planificación, dedicaremos el capítulo 3. Dedicamos el capítulo 4 a presentar críticamente la teoría de Tye. La teoría de Dennett se muestra críticamente en el capítulo 5. Las concepciones de Carruthers y Rosenthal han sido expuestas en la sección 2.1.2.4. Ninguna de ellas se mostrará en este apartado.

2.2.1.1. Teorías informacionales

Una teoría informacional considera la existencia de una relación de intercambio de información entre la experiencia y el objeto externo que la provoca. En este sentido, un estado mental representará una determinada experiencia sí y sólo si un caso concreto de dicho estado mental es nómicamente dependiente de la implementación de un determinado caso de experiencia concreta.

La teoría informacional más relevante es la ofrecida por Dretske. Para nuestro autor, la experiencia perceptiva (el carácter fenoménico de determinados estados mentales), las cualidades que determinan qué sería tener una experiencia concreta, están completamente determinadas por las propiedades de la experiencia que representan las propiedades de los hechos del mundo. Esto es, para Dretske, experimentar una serie de cualidades (experimentar el modo en que las cosas del mundo parecen ser) supone experimentar dichas propiedades como si realmente las tuvieran los objetos que están representados en nuestra experiencia. En pocas palabras, Dretske defiende un realismo perceptivo al considerar que las propiedades representadas en la percepción pertenecen a los objetos percibidos.

Las experiencias, por tanto, son representacionales. En este sentido, las experiencias son representaciones naturales (Haugeland 1981), en tanto que dichas representaciones no se derivan de los sujetos, sino que informan de cómo son los objetos del mundo que representan.

Si esto es cierto, debemos admitir la existencia de una serie de vehículos representacionales (esto es, objetos, hechos o condiciones de representación) que transportan contenidos representacionales (esto es, condiciones o situaciones objetuales que son representadas por los vehículos). Para Dretske, un vehículo estará situado en la cabeza, mientras que un contenido será siempre algo externo. Las representaciones tienen un objeto fijado según el modo en que representan (o mal-representan) cómo es el mundo. De este modo, el carácter fenoménico de nuestra experiencia es determinado por el modo en que la experiencia representa cómo son las cosas, las propiedades de los objetos representadas por ellas, pero no por los objetos en tanto experimentados. Si la experiencia es comprendida en términos representacionales, entonces parece claro que no puede ser considerada como conceptual. Para apoyar esta afirmación, Dretske (1981) realiza una diferenciación entre términos de codificación de la información analógica y digital, mientras que Dretske (1995) realiza una diferenciación entre formas de representación sistémica (de origen filogenético) y adquirida (de origen ontogenético). Visto esto, podríamos concluir que una teoría representacional de la experiencia debería distinguir entre la experiencia de las propiedades de un objeto (pongamos por ejemplo el caso de la visión, debemos diferenciar entre el tamaño, el color, la forma, etc., del objeto) y el juicio determinado que permite inferir que dicho objeto posee dichas características (por ejemplo, una creencia).

El hecho de que las propiedades que la experiencia permite individuar (esto es, aquellas características que permitirían determinar el tipo de experiencia que se tiene en cada caso) no sea de nada que el sujeto posea en la cabeza (a diferencia de su vehículo, que sí lo estará) permite indicar que dichas propiedades no pertenecen a la experiencia misma, sino que

dependen de los objetos que ésta representa. Visto así, los *qualia* no pueden ser más que una clase de propiedades físicas, propiedades que la experiencia representa como del objeto percibido. Cuando las experiencias son ciertas, entonces los objetos representados en la experiencia realmente poseen las propiedades representadas. En caso de ser falsa, los objetos no poseen dichas propiedades representadas. En este sentido, las propiedades que empleamos para individuar experiencias son las propiedades objetivas de los objetos experimentados, pero no las propiedades de la experiencia misma. Debemos diferenciar, entonces, entre las diferentes experiencias no en términos de sus propiedades, sino en términos de las propiedades que poseen los objetos de su representación. Pero las propiedades que empleamos para clasificar las diferentes experiencias no pueden ser las propiedades públicamente observables que poseen los diferentes objetos externos, sino que deben referirse a ellos. Pero sí deben serlo aquellas que permiten individuar una experiencia concreta. De este modo, como muestra en Dretske (1995), una teoría representacional plausible debe ser externista. Allí defendió que el contenido representacional no superviene sobre las bases biológicas de los sistemas perceptivos, dado que todo organismo biológicamente idéntico puede representar su entorno de modos diferentes dadas sus diferentes posibles historias cognitivas particulares. Dentro de una explicación representacional, entonces, individuos biológicamente indistinguibles pueden tener diferentes experiencias frente a los mismos estímulos.¹³ Visto así, el contenido representacional, por tanto, no superviene de modo interno a los estados mentales de los sujetos perceptores.

Ante esta tesis de Dretske, Bach (1997: 467) dice:

“La objeción obvia al externismo fenomenológico es que existen propiedades fenoménicas que, realmente, serían atribuibles a las experiencias mismas. Pensemos, por ejemplo, en las experiencias visuales borrosas... Los objetos [de estas experiencias] se ven diferentes, por supuesto, pero no se ve que hayan cambiado”.

La objeción de Bach intenta llamar la atención sobre el hecho de que, en casos concretos de visión borrosa (un problema que ya presentamos en la sección 1.4.8.), puede ser el caso que la experiencia no represente el objeto experimentado como borroso. Considerado de este modo, entonces, existirían propiedades fenoménicas que no serían propiedades reales de las

¹³ Dretske no considera que la intencionalidad sea el marco de lo mental, por lo tanto no puede considerar que los estados mentales con carácter fenoménico sean intencionales. Pero sí considera que dichos estados mentales son representacionales, esto es, presentan los objetos del mundo como siendo de una cierta manera, como teniendo determinadas propiedades. Y las representaciones, en este sentido, sí que serían intencionales y, por lo tanto, normativas. Vista de este modo, la normatividad implícita en la intencionalidad de los estados mentales, entonces, no puede depender de propiedades relacionales reales, dado que la existencia de dichas relaciones requeriría presuponer la existencia de sus *relata* (multiplicando, así, innecesariamente los objetos metafísicamente relevantes). Pero como se ha considerado Dretske a sí mismo como un reduccionista al afirmar la identificación entre vehículos de transmisión de contenido representacional y ciertas partes del cerebro (en tanto que sistemas de procesamiento cognitivo), entonces la intencionalidad debe ser considerada, primero, como algo real y, segundo, como una relación naturalmente explicable (como cualquier otro fenómeno natural). No le queda más remedio a Dretske que afirmar la existencia (necesaria) de una conexión histórica entre el desarrollo de los diferentes sistemas cognitivos de representación y los contenidos capaces de representar objetos del mundo (en tanto que representados de un cierto modo al sujeto). Cf. Dretske 2001.

cosas representadas en la experiencia. Dichas propiedades serían, como sugiere Bach, propiedades pertenecientes a la misma experiencia.

Una posible respuesta es dibujada en Dretske (2003: 76-79). Supongamos, nos dice Dretske, que alguien se pone, por ejemplo, unas gafas que le hacen percibir el mundo como borroso. En este caso, ¿deberíamos decir que el vehículo representacional (la parte de nuestro córtex visual donde se localizaría la experiencia visual concreta) se vuelve borroso? Seguramente, la respuesta correcta es que no. No Debemos suponer que nada en nuestras cabezas debe ser borroso para que los objetos externos nos aparezcan como borrosos. En este sentido, deberíamos decir que la propiedad de aparecer borroso un objeto es el modo en que la experiencia concreta nos presenta los objetos concretos, por lo que no necesitaríamos que la representación fuese borrosa para que las cosas fueran representadas como siendo borrosas. Así, cuando Bach describe una experiencia como borrosa, está confundiendo una propiedad intencional de una representación concreta (cómo la experiencia representa de modo concreto los objetos) con una propiedad de la misma representación. (Lo mismo le ocurriría a Peacocke (1983), cuando analiza casos concretos que incluyen errores de percepción (o malas representaciones) parecidas). Sin embargo, vemos que no ocurriría lo mismo (esto es, no es habitual encontrar a nadie que confunda entre propiedades intencionales y propiedades de una representación), por ejemplo, cuando hablamos de casos de representación verbal, porque que se describa el mundo como borroso no quiere decir que las palabras mismas sean borrosas. Visto así, por tanto, las propiedades indicadas deben ser propiedades relacionales que permiten a los objetos representados, pero no a las representaciones mismas, aparecer al sujeto perceptor de un cierto modo, en este caso, aparecer como borrosos.¹⁴

Concluyendo, la perspectiva representacional ofrecida por Dretske, por tanto, estaría defendiendo que el carácter cualitativo de la experiencia estaría dado no sólo por la información que es recibida, por ejemplo, por la retina (aunque sería crucial dicho proceso cognitivo), sino también por el modo en que dicha información se presenta al sistema perceptivo mismo (esto es, el modo en que es elaborado por los diferentes procesos cognitivos de los sujetos perceptores). La diferencia crucial, entonces, entre el carácter cualitativo que parecen tener las experiencias de dos sujetos biológicamente iguales, entonces, podría perfectamente ser explicada por las diferencias existentes en los procesos cognitivos particulares de elaboración de la información percibida, unos procesos que dependen por completo de la historia cognitiva particular que cada mecanismo de procesamiento ha sufrido a lo largo de su vida funcional.

2.2.1.2. Teorías teleológicas

No existe una gran diferencia entre la anterior teoría y la que aquí va a exponerse. Tal vez la diferencia crucial entre ambos tipos de interpretación recae en el hecho de que, mientras

¹⁴ Una nueva revisión de las teorías representacionales sobre la visión borrosa, así como las últimas aportaciones del punto de vista de Dretske al respecto, se proporciona en la sección 4.3.3.1., cuando analicemos la diferenciación de Tye entre estados mentales con carácter fenoménico y el resto de estados mentales. Respecto a la visión borrosa como un problema para la conciencia, véase sección 1.4.8.

para la primera la determinación de qué cuenta como contenido representacional depende por completo de la relación históricamente establecida entre la cantidad y el tipo de información involucrada y la elaboración particular que los sistemas cognitivos de cada sujeto realizan del contenido representado, para las segundas la relación existente entre un contenido y un estado representacional dependerá de los procesos evolutivos que han permitido seleccionarlos.

En este sentido, como dirá Millikan (1984, 1993 y 2003), los hechos biológicos que permitirán seleccionar una determinada función propia¹⁵ son, en sí mismos, normativos. (Una conclusión similar ofrece Neander 1995: 112, cuando indica que las normas biológicas suscriben normas semánticas). Millikan defenderá su posición a partir de una solución biológica al problema de la normatividad tratando el lenguaje y el pensamiento como categorías biológicas (Millikan 1984) y considerando la verdad y falsedad de los estados mentales en términos de éxito o fracaso basado en normas biológicamente establecidas (en términos funcionales asociados a explicaciones biológicamente dependientes. Cf. Millikan 1993).

Otra teoría teleológica la ofrece Papineau. Este autor se considera a sí mismo como un fisicalista. Parte de la idea de que el fisicalismo respecto de la conciencia debe ser entendido como una tesis empirista fuerte: debe ser entendido como una tesis de identidad entre las propiedades conscientes y las propiedades físicas. Esta concepción fisicalista, además, deberá considerarse como un argumento causal, en tanto que, considera, creemos en la completud de la ciencia física y concebimos que todo efecto físico tan sólo puede ser producido por causas físicas. Entonces, al concebir los estados mentales conscientes como capaces de producir efectos (comportamentales) físicos, estos estados mentales deben ser entendidos como causas físicas.

Una posible respuesta sería presuponer que una explicación de este tipo estaría enfrentada a la teoría del sentido común. Éste nos dice que, al pensar que todo estado mental además de proporcionar una causa física también supone reconocer una causa mental, nos llevaría a presuponer, entonces, la existencia de una sobredeterminación de los efectos físicos. Pero esto no sería del todo cierto si concebimos como posible la tesis de la identidad de los estados mentales con estados físicos. Pensemos en el caso del dolor. Si consideramos posible que el dolor tan sólo sea la estimulación de las fibras-C del cerebro, y nada más, entonces no parece posible decir que la existencia de efectos físicos del dolor (como el llevarnos la mano a la mejilla, por ejemplo, en el caso del dolor de muelas) pueda estar sobredeterminado por causas mentales, dado que la única explicación causal pertinente sería la de la existencia de causas físicas (la estimulación de las fibras-C) a un efecto físico (Papineau 1998a: 375-377).¹⁶

¹⁵ El concepto clave de la propuesta de Millikan es la noción de función propia. Para Millikan, el contenido correcto de un estado determinado se identifica con lo que dicho estado tiene la función propia de indicar. Esto es, la función propia de un determinado estado (u órgano) es la función que se supone que debe desempeñar, y no aquella función que de hecho acaba por realizar. En palabras de Millikan, “la definición de función propia... mira a la *historia* de un ítem con vistas a determinar su función más que a las propiedades o disposiciones presentes en dicho ítem” (Millikan 1993: 12).

¹⁶ “La identificación de las propiedades conscientes con las propiedades físicas requiere que las propiedades conscientes (de ser un dolor, por ejemplo) sean ellas mismas idénticas con las propiedades físicas (tener las fibras-C excitadas, para este caso), por lo que los hechos o estados relevantes no pueden ser idénticos si estas propiedades tampoco los son” (Papineau 1998a: 376).

Del mismo modo que Lycan (ver sección 2.2.1.4.), Papineau considerará implausibles, también, argumentos *a priori* como los de Chalmers o Jackson que vinculen esencialmente la presencia de conciencia a estados mentales irreducibles. Constatar la presencia de conciencia en dichos estados tan sólo es posible mediante la investigación empírica, en tanto es posible establecer un estudio científico de la conciencia que, por un lado, permita establecer la identidad entre propiedades conscientes con propiedades materiales y, por otro, permita establecer las propiedades atencionales de bajo nivel capaces de producir determinados estados mentales como conscientes (1998a). Esto es, la relación entre la existencia de determinadas características físicas y la presencia de determinados estados mentales que permitan la experiencia de determinadas sensaciones sólo puede ser establecida *a posteriori* por una ciencia física (Papineau 2003a).

Precisamente por esta razón Papineau critica la habitual comparación entre los conceptos fenoménicos y determinados nombres propios (nombres que denotan directamente a sus referentes). Papineau denomina a esta confusión la ‘falacia antipatética’ (Papineau 1993a y 1993b). La confusión de sus defensores estriba, principalmente, en suponer que la relación entre la mente y el cerebro es misteriosa, en tanto que presuponen la existencia de algún tipo de relación esencial (e inaprehensible) entre las experiencias fenomenológicas y las sensaciones concretas a las que se refieren. Por supuesto, la teoría de la identidad propuesta por Papineau pretendería acabar con dicha confusión, en tanto que el descubrimiento de la relación causal entre los efectos físicos (las sensaciones concretas) y las propiedades físicas que las provocan acabaría por demostrar la existencia de una identidad entre los conceptos fenomenológicos y determinadas propiedades físicas.

Visto de este modo, las experiencias fenomenológicas estarían provocadas por una serie de representaciones, entendidas en términos telesemánticos. Es decir, los contenidos representacionales de los estados cognitivos podrían ser analizados en términos de las condiciones requeridas por los mismos para servir a sus funciones biológicas, en tanto que permitirían una especialización en los estados que posibilitan las salidas comportamentales (Papineau 1998b) y la constitución de una serie de disposiciones conductuales (podríamos entenderlas como mejoras estructurales que permitirían un mejor desarrollo de los mecanismos de toma de decisiones y el razonamiento para la obtención de objetivos) que facilitarían las respuestas de los organismos (Papineau 2001). Por supuesto, toda esta estructura estaría mediada, al menos en los seres humanos, por la posesión del lenguaje (Papineau 2003b) y el desarrollo de las concretas adaptaciones cognitivas (entendidas, como hemos visto, como habilidades y capacidades) estaría garantizado a partir de una serie de elementos originarios culturalmente dependientes (Papineau 2005a), lo que haría que en última instancia todo concepto fenoménico se debiera entender como un objeto vago.¹⁷

¹⁷ “Cuando miramos dentro de nosotros mismos nos parece encontrar un límite definido. Los dolores, picores, las experiencias visuales y demás son conscientes mientras que el proceso que nos permite darles nombre o describirlos a través de estereogramas no lo es. Parece, entonces, que existen experiencias ‘medio-conscientes’... Reflexionando un poco, podemos ver que dichas experiencias son inequívocamente conscientes después de todo, en tanto que son como algo más que como nada” (Papineau 1993a: 125). El argumento más fuerte esgrimido contra la tesis de vaguedad de

2.2.1.3. Primitivismo

Algunos autores se han visto atraídos por el punto de vista realista que afirma que los objetos externos poseen realmente las propiedades que aparecen en nuestra experiencia. Pensemos en el caso concreto del color: para ellos, los objetos del mundo están coloreados. A diferencia de las anteriores posiciones presentadas en este apartado, a pesar de abrazar el realismo, el primitivismo rechaza el reduccionismo. Es este sentido, podríamos decir que este planteamiento es primitivista porque considera que la propiedad de ser rojo (por ejemplo, de un tomate maduro) es una característica extra, primitiva, de los tomates que no puede ser identificada con una disposición para producir experiencias de rojez o con la reflectancia de la luz que provocaría la propiedad experimentada de rojez. Al mismo tiempo, como hemos comentado, será realista porque considera que los tomates maduros tienen la propiedad de ser rojos. Esto es, que los colores (pero también las otras propiedades secundarias) de los objetos externos son propiedades del mundo, como lo son también la carga electromagnética de los polos de un imán o la dirección de un espín. Visto de este modo, pronto surge una analogía con el dualismo de propiedades. Si el primitivismo reconoce la existencia de propiedades físicas y de propiedades experienciales en los objetos, ¿no acaba por defender la existencia separada de dos entidades irreconciliables?

En defensa de esta posición, y contra esta crítica, McGinn (1996 y 1999) reivindica que la propiedad primitiva de ser rojo es una propiedad que superviene sobre algunas otras propiedades de, por ejemplo, los tomates maduros. Es decir, considera plausible que la propiedad de ser rojo de un tomate maduro emerge directamente de las propiedades del tomate que producen experiencias de rojo en los receptores de tomates. Visto así, por tanto, los tomates tienen la propiedad emergente de ser rojos, aunque no poseen ningún tipo de disposición (recordemos que una conclusión similar puede obtenerse, también, de los análisis reduccionistas que consideran que las propiedades secundarias se determinan a partir de las respuestas que provocan en los sujetos perceptores, pero sin ser reduccionista, aunque en esta perspectiva se afirma la existencia de disposiciones).¹⁸ Así, al evitar ser reduccionista, esta nueva visión de McGinn considera que la propiedad de ser rojo es una propiedad emergente de

Papineau defiende que, en última instancia, la vaguedad de los estados conscientes depende de la posibilidad de determinación del contenido conceptual que los definiría. Si dejamos de lado la posibilidad de que el contenido de los estados con carácter fenoménico sea conceptual, entonces parece desaparecer el conflicto que presenta la vaguedad (Antony 2006). Ello, por supuesto, también haría tambalearse la tesis fisicalista de Papineau. Pero ese es un precio que estaríamos dispuestos a pagar.

¹⁸ Aunque la posición clásica es la defendida por Locke, también McGinn (1983) defiende esta postura. Considera allí que la propiedad de ser rojo es una cualidad secundaria de los objetos externos, esto es, una propiedad que se puede definir en términos de la respuesta que despierta en los sujetos. De este modo, concibe la propiedad de un tomate de ser rojo como la propiedad de estar dispuesto para producir experiencias de rojo en los sujetos perceptores normales bajo condiciones normales. Los principales problemas puestos de relieve contra esta concepción parten de la idea de que es fenomenológicamente inconsistente, en tanto que las experiencias (como la experiencia de color) no pueden verse como disposiciones que despierten nada en los sujetos, sino que son tan sólo propiedades intrínsecas, no-relacionales, de los mismos objetos externos (Boghossian and Velleman 1989). Pero, todavía peor, la teoría de McGinn (como ya le ocurrió a la de Locke) parece ser circular al pretender identificar el color rojo, por ejemplo, con la disposición que normalmente produciría experiencias rojas, en tanto que se presupondría que la disposición para producir experiencias rojas sería aquello que normalmente produce experiencias concretas rojas cuando se ve el rojo, y así *ad infinitum*.

los tomates, y que depende de dicha propiedad tener (o no) causar en los perceptores la experiencia de rojo pertinente en virtud de sus propiedades físicas (aquí, ahora, la analogía podría establecerse con aquellas posiciones que consideran posible establecer un análisis reduccionista pero que creen que las propiedades secundarias son independientes de las respuestas que provocan en los sujetos perceptores, por lo que, además, se les podría aplicar las mismas críticas que hemos visto).¹⁹

A pesar de las ventajas mostradas, la posición de esta versión del primitivismo defendida por McGinn tiene varias dificultades. Por una parte, como la motivación principal del primitivismo es acomodar nuestra intuición habitual respecto de la experiencia fenoménica, no parece estar claro en qué sentido puede hacerse esto precisamente si lo que estamos defendiendo es una teoría que presupone la existencia de propiedades para un sujeto como dependientes de características objetivas (perfectamente aprehensibles desde métodos objetivos). Por otro lado, todavía más difícil de asimilar sería la idea de que los objetos físicos tengan propiedades emergentes que supervengan a otras propiedades físicas de los mismos.

2.2.1.4. Representacionalismo intencional

El intencionalismo defiende que “el carácter fenoménico es lo mismo que cierta clase de contenido intencional” (Tye 1995: 137). Visto de este modo,

“la mente no tiene ninguna propiedad especial que no sea agotada por sus propiedades representacionales, con o en combinación con la organización funcional de sus componentes” (Lycan 1996: 11).

En este sentido, podríamos definir el representacionalismo intencional, o intencionalismo, como la tesis que indica que el contenido representacional de las experiencias perceptivas determina su propio carácter fenoménico. Esto es, que no puede existir una

¹⁹ La posición que defiende un punto de vista reduccionista de las experiencias pero como independientes de las respuestas que provoca en los sujetos de percepción reivindica que las cualidades secundarias (como el color) son propiedades físicas relativas a los objetos (como sus propiedades de reflectancia). Visto así, el color rojo es la reflectancia de un determinado espectro de onda lumínica. Versiones diferentes de esta perspectiva reduccionista no-dependiente, las encontramos en Armstrong (1999), Byrne (2006), Byrne and Tye (2006), Dretske 1995), Marr (1982) y Tye (2000). Por supuesto, las críticas que han recibido también son contundentes. En primer lugar, para casos concretos de variación estándar (véase sección 1.4.7.), o de *mala representación asimétrica* (Byrne and Tye 2006: 242), dos sujetos tendrán diferentes experiencias de color porque, por supuesto, ambos han sufrido una evolución diferente de sus procesadores post-receptivos. Pero parece ser que esta concepción sería incompatible con la propuesta realizada, dado que si fuera posible que ante la misma propiedad física dos sujetos difirieran en su percepción, una de dos, o bien uno de ellos deberá estar equivocado o bien las cualidades secundarias no pueden ser reducidas a propiedades físicas. Una dificultad más importante atañe a la estructura del color. Según una teoría muy popular, existen colores primarios (rojo, azul y amarillo) y colores secundarios (que serían la combinación de dos colores primarios o secundarios). Visto así, parece posible afirmar que experimentamos los colores (ya sean primarios o secundarios) por las peculiares características de nuestro sistema visual (Hardin 1988 y 2003). Pero la idea de que los colores pueden dividirse en primarios y secundarios dependería tan sólo de nuestra experiencia peculiar de color y no de nuestro sistema neurológico capaz de captar los colores. Visto así, según esta teoría reduccionista dependiente, los colores como representados por nuestra experiencia tendrían una serie de características de las que realmente carece la estructura neurofisiológica capaz de captarlos y causar dicha experiencia de color, por lo que deberíamos concluir que los colores son diferentes de las propiedades físicas que las causan (lo mismo podría aplicarse a las restantes cualidades experienciales).

diferencia en el carácter fenoménico que determina de qué tipo de experiencia perceptiva concreta estamos hablando en cada momento si no existe una diferencia en el contenido representacional de las mismas. Por ejemplo, dos experiencias visuales concretas serán diferentes si ambas experiencias difieren en el contenido que las determina. En este sentido, el intencionalismo puede ser entendido como una teoría que o bien pretende reducir el carácter fenoménico de determinados estados mentales a intencionalidad o bien pretende explicar el carácter fenoménico de determinados estados mentales en términos de contenido intencional. Siendo esto así, por tanto, por cuestiones puramente metodológicas (en tanto que nos va a permitir diferenciar problemas concretos de cada una de las posiciones, al apelar a diferentes criterios para determinar los contenidos representacionales), vamos a contemplar la existencia de dos tipos diferentes de representacionalismo intencional, uno al que denominaremos reductivo, en tanto que persigue una asimilación de las representaciones con mecanismos funcionales (Lycan-Prinz), y otro que denominaremos fuerte, que pretende la asimilación del carácter fenoménico con contenido representacional no-conceptual (Tye-Byrne).²⁰

Lycan (1996) considera que los conceptos fenoménicos no son esencialmente intrínsecos, contra todos aquellos que, como Chalmers (1996) o Jackson (2003), consideran que sí, por establecerse como representaciones que podrían ser entendidas única y exclusivamente como pensamientos proposicionales *a priori*. Las experiencias fenoménicas deben ser entendidas en un sentido representacional en tanto que dichas experiencias son capaces de representar los objetos externos como poseyendo determinadas características físicas. En este sentido, la estipulación de dichas características no puede determinarse de modo intrínseco, esencial o *a priori*, sino que dependería por completo del determinado papel funcional de los mecanismos de procesamiento capaces de causar las experiencias. En este sentido, además de fallar todo intento de instaurar un argumento *a priori* respecto de la existencia de conceptos fenoménicos (al estilo de Kripke-Chalmers. Cf. Lycan 2003), debemos presuponer que existe toda una serie de elementos relacionales que permitirían establecer un vínculo causal entre representación (como contenido de la experiencia) y objeto, una relación perfectamente explicable (de modo empírico) a partir de elementos funcionales (Lycan 1996: 84) y psico-semánticos (Lycan 2002).

²⁰ Byrne (2001: 205) realiza una interesante distinción entre tipos de intencionalismo. Separa en dos categorías diferentes a los defensores del intencionalismo. Por un lado, encontraríamos a los defensores del intencionalismo intramodal. Éstos defenderían que no tiene porqué existir necesariamente diferencias fenoménicas dadas dos formas distintas de experiencia perceptiva (entre, por ejemplo, las experiencias visuales y las experiencias olfativas o auditivas). La (presunta) diferencia (presente tan sólo a un nivel personal) en el carácter fenoménico de dos experiencias perceptivas diferentes podría perfectamente ser explicada (de modo reduccionista) en términos del papel funcional que los sistemas cognitivos capaces de provocar dichas experiencias tienen en el sistema perceptor. Por otro lado, encontraríamos a los defensores del intencionalismo intermodal. Éstos defenderían la existencia de una diferencia en el carácter fenoménico de dos experiencias perceptivas distintas en base a la existencia de una diferencia en el contenido (representacional) de las mismas. A su vez, éstos últimos podrían ser subdivididos entre aquellos que consideran posible que esta diferencia pueda aplicarse también a las experiencias corporales (como el dolor, el picor o los estados emocionales o de ánimo), recibiendo el nombre de intencionalismo intermodal no-restringido, y aquellos que considerarían que la diferencia no podría ser aplicada a dichas experiencias corporales por exceder los límites de la experiencia fenoménica (intencionalismo intermodal restringido).

Partiendo de las tesis estipuladas en Jackendoff (1987), por su parte, Prinz considera que existen evidencias neurocientíficas que permiten afirmar la existencia de una jerarquía en la estructura neurofuncional de la percepción que permitiría establecer la existencia de un nivel intermedio de procesamiento que estaría situado de manera privilegiada frente a la conciencia (Prinz 2000). En este texto, Prinz aceptará la existencia de una visión de alto nivel que estaría asociada a las áreas concretas del córtex infero-temporal. Pero, como Jackendoff insiste, este nivel intermedio parecería darse en todas las localizaciones de la conciencia y no sólo en la visión. En la conciencia, entonces, parecería existir también una jerarquía que se extendería por todas las partes del cerebro. Dicha jerarquía sensitiva no sólo se produciría para la conciencia externa, sino también para la percepción interna de los estados corporales, conformando así las bases tanto de la experiencia introspectiva como las bases de la experiencia emocional (Prinz 2004). Pero también es cierto que este nivel intermedio no es suficiente para la conciencia, dado que parece claro que puede existir una activación de dicho nivel jerárquico sin producir conciencia.

Esto nos debería llevar a aceptar dos conclusiones. Por una parte, deberíamos considerar que la existencia de casos anómalos de percepción (como la visión ciega, tal y como fue presentada en la sección 1.4.13.) no puede interpretarse siempre como consciente, en tanto que no serían interpretables tan sólo como efectos de la memoria perceptiva. Mostrar la existencia de actividad en el sistema perceptivo no parece ser suficiente para aceptar la existencia de conciencia. Por otra parte, la experiencia consciente se perdería cuando los centros del córtex parietal sufren algún tipo de trastorno o lesión, unos centros que estarían asociados a la atención. Esto vendría a indicarnos que la conciencia requiere de la atención.

Prinz (2005) insiste en la necesidad de aceptar la presencia de conciencia a partir de los sistemas perceptivos de nivel intermedio sólo cuando la actividad de dichos sistemas es modulada por la atención. Cuando la atención se focaliza sobre algún tipo concreto de dato sensitivo, entonces es cuando la percepción se vuelve consciente.

“La conciencia surge cuando las representaciones perceptivas de nivel intermedio se vuelven disponibles a la memoria a partir de la atención” (Prinz 2005: 386).

Vista de este modo, la atención sería un proceso de selección que se aplica a la información enviada a los sistemas de procesamiento. Así, la conciencia aparecerá cuando las representaciones perceptivas de nivel intermedio se logran procesar a partir de la acción de la atención perceptiva.²¹ Esta teoría sería funcionalista en tanto que la conciencia puede ser

²¹ Una conclusión similar puede extraerse de algunos intentos actuales de aportar una renovada visión neurocientífica de la mente. Decimos ‘renovada’ porque, contra estudios clásicos como los de Crick, persiguen construir una verdadera ciencia del cerebro que sea capaz de aportar respuestas desde un punto de vista biológico y cognitivo a las incógnitas predominantes respecto del estudio de la mente, en general, y de la conciencia, en particular. Frente a la teoría celular clásicamente defendida, estos nuevos estudiosos pretenden establecer una ciencia basada en el fenómeno de la potenciación de amplio término, un proceso neurobiológico que resaltaría el importante papel que la memoria y la atención tienen a la hora de explicar los fenómenos conscientes. Cf. Stoljar and Gold 1998. Esta visión se complementaría con una contundente crítica a la estrategia contraria, a la posición que defendería los conceptos fenomenológicos como elementos de una estructura mental particular intrínsecamente accesible. Cf. Stoljar 2005.

entendida en términos de procesamiento de información computacional. Pero también puede ser entendida como representacional en tanto la conciencia es interpretada como la propiedad de ciertas representaciones mentales (como resultado del procesamiento de la información recibida) de aparecer como disponibles a los sujetos de un cierto modo, en el sentido de ofrecer una explicación causal a las respuestas sensitivas de los organismos.

Desde este punto de vista, la conciencia, para Prinz, no será más que el conjunto de operaciones realizadas por sistemas concretos compuestos por mecanismos neurofuncionales que se encuentran en los organismos que permiten, a partir de la aparición de ciertas representaciones a un nivel intermedio, explicar causalmente las respuestas sensitivas que a un nivel psicológico personal los individuos concretos poseen.

Dentro del representacionalismo intencional fuerte, para Byrne, la experiencia fenoménica es lo mismo que poseer determinados estados con contenidos representacionales. Visto así, dos sujetos tendrán la misma experiencia si tienen el mismo contenido representacional y diferirán en la posesión de experiencias si éstas tienen contenidos representacionales diferentes. Porque, en última instancia, si a nosotros nos parece que el mundo es como se nos presenta, debemos suponer que dicho parecer debe ser representacional. Pensemos en un caso concreto. Imaginemos que estoy viendo un tomate maduro concreto en el estante de una verdulería. Llego a su altura y me quedo fijamente mirándolo. Ladeo un poco la cabeza y lo vuelvo a mirar. Tengo dos experiencias diferentes que tienen un carácter fenoménico distinto, aunque sean respecto del mismo tomate. Si soy un sujeto perceptivo normal, entonces deberé darme cuenta de que lo que estoy teniendo son dos experiencias diferentes con un carácter fenoménico distinto, dado que, numéricamente hablando, son dos experiencias diferentes. Pero dichas experiencias pueden, sin embargo, compartir el mismo contenido. Las dos son experiencias respecto del mismo tomate maduro. Las dos se me pueden presentar como, digamos, experiencias de rojo. Pero las dos deben considerarse diferentes porque difieren en su carácter intencional, porque ambas experiencias representan de modo diferente la misma experiencia de rojo, aunque dicha representación tenga el mismo contenido. Visto de este modo, Byrne estaría defendiendo una teoría intencional fuerte que sostendría que las experiencias no pueden diferir en su carácter fenoménico si no admitimos que también difieren en su contenido representacional, aunque el contenido de dicha representación se refiera al mismo objeto externo (Byrne 2001).²²

2.2.2. Tipos de representacionalismo

Como ya hemos indicado, para una concepción representacional de la conciencia, los *qualia* (el carácter cualitativo de una experiencia consciente) son las propiedades de los objetos representadas en la experiencia y mostrados al sujeto de un modo determinado. Las propiedades

²² Una idea similar puede extraerse del análisis realizado en Schier 2007, cuando nos dice que adoptar un esquema molecular de las representaciones de la experiencia de color permite una mayor sofisticación a la hora de explicar la percepción del color en tanto que, por un lado, permite salvar la inexistencia de propiedades objetivas que una experiencia de color pudiera representar y, por otro lado, permitiría una aproximación más plausible a la dilucidación de los datos psicofísicos de la experiencia de color.

relevantes de las experiencias, por tanto, serán intencionales en tanto que se presentarán de modo directo a la experiencia. Muchos representacionistas consideran que las representaciones (al menos las del color y otras propiedades sensibles consideradas secundarias) son no-conceptuales en algún sentido. Visto así, la mera representación de rojez de un objeto no será suficiente para mostrar el aspecto fenoménico del objeto rojo. La representación deberá ser específicamente una representación visual, producida por un sistema visual normal (o por algún tipo de mecanismo que puede ser, al menos, funcionalmente indistinguible de éste) para poder ser entendida como una experiencia de un cierto tipo. Desde este punto de vista, Dretske (1995) considera que la representación visual misma es sensible y sistémica (relativa al mismo sistema). Sin embargo, para Tye (1995), además de requerir que la experiencia visual sea no-conceptual, es necesario que ésta esté preparada (es decir, permita al sujeto obtener a partir de ella las actitudes proposicionales adecuadas o las respuestas conductuales correctas). Para Lycan (1996), en cambio, aquello que determinaría el modo de una experiencia concreta sería su papel funcional.

Aunque en la anterior sección hemos aportado una primera aproximación a las teorías representacionales más relevantes (priorizando el tipo concreto de representación que defienden), podemos aventurar aquí una clasificación de las teorías representacionales englobándolas en cuatro grupos mayoritarios de diferencias: según su grado de representacionalidad, según persigan la reducción (o no) de los *qualia*, según la caracterización del contenido representacional (como hemos visto) y según el alcance del mismo.

2.2.2.1. Representacionalismo puro, fuerte y débil

Para el representacionalismo puro, el aspecto cualitativo de una experiencia es en sí mismo una representación. Esta representación presenta una cierta clase especial de contenido en tanto que mostraría, tan sólo, representaciones particulares de objetos para los sistemas perceptivos de un sujeto (Thau 2002).

El representacionalismo fuerte, por el contrario, defiende que una representación de un cierto tipo de carácter cualitativo está especificada en términos materiales (Dretske y Tye) o en términos funcionales (Lycan y Prinz) sin necesidad de recurrir, como le sucede al representacionalismo débil, a la existencia de nuevas entidades metafísicamente relevantes. Visto de este modo, entonces, el representacionalismo es la teoría que sostiene que la individuación de un *qualia* requiere o bien de criterios funcionales para ser diferenciado de otro tipo de estados representacionales o bien de criterios físicos para diferenciar experiencias cualitativamente diferentes pero con el mismo contenido intencional.

El representacionalismo débil, por su parte, insiste en que los estados cualitativos tienen contenido representacional, pero en tanto que dicho contenido, en sí mismo, puede considerarse como características internas y subjetivas de los mismos sujetos que los poseen, en tanto entidades metafísicamente nuevas (Block y Chalmers).

2.2.2.2. Modos de caracterizar el contenido de las experiencias

A la hora de explicar la naturaleza metafísica del contenido representacional de los estados fenoménicos podemos diferenciar, en este sentido, entre aquellos que consideran que las experiencias representan propiedades que los objetos físicos realmente no poseen, creando una entidad nueva similar a los datos de los sentidos (Levine), aquellos que defienden que las experiencias representan propiedades disposicionales de los objetos para causar la percepción de determinadas cualidades (Shoemaker) y aquellos que consideran que la experiencia capta directamente las propiedades físicas de los objetos externos (Dretske).

2.2.2.3. Representacionalismo amplio y restringido

Cuando hablamos de actitudes proposicionales (creencias, deseos, etc.), en general, de aquellos estados mentales con un contenido proposicional, es habitual diferenciar entre teorías que defienden que su contenido representacional es amplio y teorías que sostienen que el contenido representacional es restringido. En el primer caso, el contenido es considerado amplio en tanto que no superviene sobre los estados no-relacionales de los sujetos que las poseen porque dicho contenido es externo y está determinado por los objetos en tanto dependientes de los entornos en los que se encuentran (por lo que dos individuos molécula por molécula idénticos, físicamente indistinguibles, podrían tener diferentes contenidos). En el segundo, el contenido es considerado restringido en tanto, consideran, es dependiente de aquello que intrínsecamente podemos encontrar en el poseedor de dichos estados (por lo que dos individuos idénticos físicamente no pueden tener contenidos diferentes).

Del mismo modo, cuando se habla de estados mentales con contenido fenoménico, existen también estas dos mismas posiciones. Por un lado encontramos a aquellos que defienden que el contenido representacional de los estados cualitativos es amplio y perfectamente determinable por las condiciones externas del sujeto que lo posee (Dretske y Lycan, por ejemplo, son seguidores de esta teoría, pero también Prinz, Tye o Byrne). Por otro lado, en cambio, existen defensores del punto de vista restringido, que consideran posible que el contenido representacional de los estados fenoménicos dependa por completo de las características internas del propio sujeto (por ejemplo, Chalmers y Jackson, pero también Block, Levine, Shoemaker, Kriegel o, más recientemente, Farkas 2008).

2.2.2.4. Representacionalismo reductivo y no-reductivo

En las posiciones representacionalistas, también podemos encontrar discrepancias relevantes a la hora de caracterizar el contenido como poseyendo determinadas cualidades físicas. En este sentido, los defensores del reduccionismo, consideran que poseer un determinado estado con contenido representacional es lo mismo que estar en un determinado estado cerebral con ciertas características físicas (Tye), causales (Dretske) o funcionales (Lycan). Pero también encontramos a los defensores del punto de vista contrario, a aquellos que consideran que el contenido representacional de una experiencia tiene un especial modo de presentación que no puede ser reducido a criterios físicos o funcionales (Searle).

2.2.3. Cognición, representación y *qualia*

Es habitual en la literatura reciente sobre conciencia encontrar autores que emplean indistintamente el término *qualia* para referirse al carácter subjetivo de los estados mentales con contenido fenoménico como para referirse, también, a su carácter cualitativo. Como hemos mostrado al comienzo de este capítulo, parece ser un error demasiado común y fatal. Aunque sean dos aspectos del mismo estado, la subjetividad debería diferenciarse del carácter cualitativo, en tanto que resaltarían propiedades diferentes. El primero, indica el modo particular en que la experiencia se presenta a un sujeto concreto. El segundo, indica el contenido cualitativo representado por una determinada experiencia.

2.2.3.1. El carácter cualitativo (representacional) de la experiencia

Deberíamos diferenciar, como hemos dicho, entre el carácter subjetivo de una experiencia (el modo en que un sujeto experimenta una experiencia particular) y el carácter cualitativo de una experiencia (sería una propiedad del *quale* particular, pero no el *quale* mismo) por dos razones fundamentales.

Por un lado, el *quale* concreto de una experiencia podría ser perfectamente explicable en un lenguaje público, mientras que el peculiar carácter cualitativo que una determinada experiencia tiene para un sujeto concreto parece inefable.

Por otro lado, un *quale* concreto podría ocurrir sin que, necesariamente, ningún sujeto fuera consciente de él (pensemos tan sólo en enfermedades o en lesiones cerebrales que provocan casos de visión ciega). Por el contrario, el carácter cualitativo de una determinada experiencia requiere, por definición, de un sujeto que lo experimente.

2.2.3.2. Diferentes usos del término *qualia*

Imagina por un momento tu experiencia visual cuando observas un tomate maduro. Hay algo que subjetivamente se te presenta. La experiencia de ver un tomate maduro es muy diferente de la experiencia de ver, por ejemplo, un plátano. No sólo tienen una forma, un tamaño, un color, etc., diferentes. Sino que también tienen una subjetividad diferente. Ambas experiencias presentan de manera diferente dos objetos distintos. Pero también dos experiencias diferentes del mismo objeto lo pueden presentar de manera distinta, porque son dos experiencias cuantitativamente diferentes que, pensamos, también son cualitativamente diferentes, aunque ambas experiencias puedan representar el mismo objeto. Esta diferencia es una diferencia en cuanto al carácter fenoménico de la experiencia consciente. Por el carácter fenoménico de una experiencia se entiende normalmente el modo en que subjetivamente hablando se nos presentan las cosas. Esto es, el cómo es para mí tener determinadas experiencias (*qualia*). Este sería un modo amplio de considerar el término *qualia*, como aquello que cualitativa y subjetivamente permite representar a los objetos externos de experiencia para un sujeto concreto.

Pero existen otros usos más restringidos del término *qualia*. Piensa en el cuadro de un dálmata (ejemplo extraído de Tye 2005). El cuadro representa un dálmata. Pero aunque nuestra experiencia consciente se centre tan sólo en el dálmata representado, nosotros también somos

conscientes de otras características concretas del dalmata (su tamaño, su forma, su color, su estar moteado, etc.), esto es, nos damos cuenta de todas aquellas características intrínsecas que hacen de aquello representado su ser un dalmata. Estas características no tienen por qué ser representacionales. Entonces, un uso más restringido de *qualia* ha sido también empleado por aquellos que (como los defensores de los datos de los sentidos) consideran que se puede referir al conjunto de características intrínsecas no-representacionales de los objetos que los hacen ser lo que son y que los diferenciarían de los demás objetos del mundo.

En este último sentido, muchos filósofos consideran posible que aquello que podemos referir con el término *qualia* son propiedades intrínsecas conscientemente accesibles de los objetos del mundo, características que no serían físicamente reducibles aunque se considere posible que supervengan a estados físicos. Por ejemplo, en las experiencias visuales es frecuente encontrar autores que afirmen la existencia de un rango visual de la experiencia concreta, esto es, donde la experiencia visual esté determinada por una serie concreta de propiedades específicas de lo visual. Dichas propiedades estarían disponibles de un modo introspectivo, podrían variar aunque no existiera un cambio sustancial en el contenido representacional de las experiencias y serían la contraparte subjetiva visual de las propiedades físicas visuales de los objetos representados en la propia experiencia (Cf. Peacocke 1983 y Block 1990).

Todavía podríamos diferenciar otro uso más restringido del término *qualia*. Aunque los autores que lo emplean persiguen una eliminación del término, consideran que *qualia* refiere a las propiedades intrínsecas a toda experiencia, unas propiedades inefables para el sujeto perceptor, que se constituyen como lo dado en la percepción (Dennett 1988a y 1991) pero que, a pesar de ello, se constituirían como una ilusión al depender por completo de una serie de estados físicos que permitirían su implementación. La dificultad con esta definición sería la posible acusación de epifenomenalismo que ha recibido (véase el apartado 5.2.3. para una ampliación de esta discusión).

De todos modos, no parece ser éste el caso. Todos sabemos a partir de nuestra propia experiencia de la existencia de diferentes tipos de *qualia*. Podemos asumir, sin más, que todos ellos se deben a la existencia de estados físicos que permiten implementar los diferentes tipos de vivencia que disponemos de nuestra experiencia. Esto es, podemos suponer que una explicación naturalista amplia es correcta. Si esto es así, debemos también suponer que la filosofía y la ciencia son capaces de individualizar los diferentes tipos de *qualia* en tanto que emplean los métodos científicos adecuados. Si la ciencia y la filosofía individualizan metodológicamente los diferentes tipos de *qualia*, deberán hacerlo según sus diferentes poderes causales, por lo que podríamos (además) directamente eliminar la acusación de epifenomenalismo que muchas aproximaciones naturalistas han recibido (Megill 2007).

2.2.3.3. ¿Qué estados mentales poseen carácter fenoménico?

Pensemos por un momento qué estados mentales podrían ser considerados como teniendo un carácter fenoménico. Por supuesto, deberíamos incluir todos aquellos estados que se refieran a experiencias perceptivas de todo tipo (las cualidades llamadas habitualmente

secundarias), como ver, oír, oler, etc. Pero también deberíamos incluir sensaciones corporales (como el sentir dolor, picor, escozor, sentir frío o calor, hambre o sed, tener un orgasmo). También sensaciones o reacciones emotivas, como sentir celos, amor, envidia... Por último, debemos incluir también los estados de ánimo, como sentirse aburrido, abatido o estar deprimido. Un buen compendio puede encontrarse en Haugeland 1985: 230-235.

¿Podríamos ampliar esta lista? G. Strawson (1994) argumenta que algunas experiencias proposicionales, como las experiencias de comprender una emisión lingüística o la experiencia de pensar en algo en concreto o recordar algo, también deberían ser consideradas como experiencias con carácter fenoménico. Según su propio punto de vista, las experiencias de esta clase concreta no serían reducibles a experiencias sensitivas o asociadas a imágenes. Su posición parece ser la defensa de estas experiencias de pensamiento como experiencias diferentes por mérito propio. En este sentido dice:

“Toda modalidad sensitiva es una modalidad experiencial, y las experiencias de pensamiento (en las que la comprensión está involucrada de algún modo) son una modalidad experiencial que es independiente de otros tipos de modalidades experienciales” (Strawson 1994: 196).

Visto así, entonces, determinados pensamientos también deberían ser considerados como *qualia*. Una posición semejante defienden Horgan and Tienson 2002.

Pero esta última afirmación es un tanto discutible. Pensemos tan sólo en el aspecto fenoménico de la comprensión derivada de las imágenes lingüísticas, un aspecto que tiene una estructura fonológica y sintáctica relativa a la lengua nativa del sujeto perceptor. Dichas imágenes a menudo son detalladas o modificadas por la entonación y otros elementos externos. Por lo tanto, la posibilidad implícita de Strawson, la idea de un carácter intrínseco, no parecería darse en este tipo de caso. Probemos un ejemplo claro. Parece que Strawson reivindica la posibilidad de que, en casos como la experiencia de un recuerdo, una voz interna nos esté diciendo algo concreto. Nosotros deberíamos oír una voz interna, de algún modo. Y deberíamos sentir algo concreto. Pero esto no sería nada intrínseco al pensamiento concreto del recordar. Sería más bien algo asociado al acto proposicional concreto que supone recordar. En este sentido, no podríamos tratar este tipo de experiencias como teniendo un carácter fenoménico, dado que realmente tendrían un contenido marcadamente proposicional. Deberían ser tratadas como actitudes proposicionales. (Para una diferenciación más exhaustiva entre experiencias fenoménicas y actitudes proposicionales véase el apartado 1.1.).

2.2.4. El representacionalismo como escuela dominante

Como venimos defendiendo desde el principio, en los últimos años el representacionalismo ha ido cada vez más ganando fuerza como la explicación predominante a la hora de dar cuenta de los fenómenos conscientes. Pero, por supuesto, que se haya convertido en la principal escuela de interpretación no quiere decir, primero, que todo defensor del representacionalismo acabe por defender las mismas conclusiones y, segundo, que no existan voces críticas discrepantes con esta visión dominante.

2.2.4.1. Argumentos a favor

Un argumento que pretende defender la posibilidad de que el contenido de toda experiencia con carácter fenoménico es representacional proviene de aquellos que consideran que sólo una aproximación materialista al mismo es capaz de explicar adecuadamente las experiencias conscientes. Si la experiencia sensitiva se nos presentara a partir de elementos de mediación, como por ejemplo los datos de los sentidos, es difícil ver en qué sentido puede defenderse que la percepción pueda estar justificada a partir de la existencia de elementos externos al sujeto perceptor, perdiendo así la existencia de un mundo entorno. Por supuesto, el materialismo también podría defender la existencia de alternativas no-representacionales, como la defensa de una teoría estricta de la identidad de tipos. Y también podría defender alternativas adverbiales, al estilo de Chilsholm o Sellars, en tanto que defenderían una tesis semántica respecto de la forma lógica de las emisiones sensitivas tal que afirmarían la existencia de sensaciones concretas que podrían ser caracterizadas como sensaciones de, por ejemplo, rojo (para el caso de alguien que estuviera viendo un tomate maduro, por ejemplo). Otra posibilidad se asimila al eliminativismo, en tanto que defendería que no existe realmente aquello que podemos identificar como *qualia* (Dennett y Rey son defensores de esta posición).

Un argumento diferente a favor del representacionalismo se desprende de la diferenciación entre aquellas experiencias sensitivas que podríamos denominar verídicas de aquellas que podríamos denominar falsas. Un representacionista defenderá que las condiciones que hacen que una experiencia sea verídica (esto es, que realmente un sujeto perciba como, por ejemplo, de rojo la experiencia de ver un tomate maduro), agotan tanto el contenido representacional como el contenido cualitativo de la misma experiencia. Es decir, una representación de una experiencia será verídica sí y sólo sí en el mundo se dan las condiciones adecuadas para decir de dicha experiencia que es verídica. En caso contrario, se tendrá que decir que la experiencia es falsa porque representa hechos que no se dan en el mundo.

Otro argumento a favor es el argumento de la transparencia (Cf. Harman 1990. Lo hemos desarrollado en la sección 1.4.9.). Decimos que, de manera habitual, podemos atribuir las propiedades que percibimos como pertenecientes a los objetos percibidos en tanto que, parece ser, nuestra experiencia representa los objetos externos como teniendo esas características determinadas. Esta tesis podría ampliarse, también, a las experiencias corporales (Tye y Crane) y a las experiencias emocionales (Tye y Prinz).

Podríamos adoptar también otro argumento a favor del representacionalismo. Podríamos sostener que el mundo parece ser tal y como se presenta a un sujeto. Si el mundo me parece de un determinado modo, no tengo por qué suponer que el mundo no sea de dicho modo (Byrne 2001: 207). Visto así, deberíamos considerar posible que el mundo sea como realmente se nos aparece, por lo que el mundo sería tal y como nos lo representamos (realismo).

2.2.4.2. Argumentos en contra

En la mayoría de casos, mantener una teoría representacional del carácter fenoménico de las experiencias conscientes, como hemos visto, supone tener que admitir también un

realismo respecto de las cualidades secundarias, como el color. Los términos de color, vistos de este modo, se emplean para expresar las propiedades objetivas y públicamente observables que poseen los objetos físicos. Puede explicarse la posesión de dichas características objetivas a partir de términos representacionales que muestren el color o bien a través de términos disposicionales que presenten la disposición de los objetos a despertar determinadas experiencias en los sujetos perceptores. El problema principal viene cuando advertimos la posibilidad de que, al menos en principio, las experiencias de color puedan ser explicadas sin necesidad de requerir una apelación a este tipo de características objetivas. Es el caso de G. Strawson, que considera (como hemos mostrado en la anterior sección) que el color dependería de la visión intrínseca que tuviera cada sujeto en particular de los objetos, por lo que no podría reducirse a meras propiedades físicas del objeto, sino que debería presuponerse como una imagen interna al propio sujeto, implicando un irrealismo respecto del color.

Lo mismo podría aplicarse a casos en los que existen fallos o errores en la percepción. Pensemos en casos como en una alucinación o en una ilusión óptica o acústica, como por ejemplo el caso de un palo que aparece como torcido a partir de la refracción del agua. Si tuviéramos que hacer caso de aquello que vemos, nos dicen los defensores del internismo *à la* Strawson, es decir, si consideráramos cierta la idea de que la experiencia visual se refiere realmente a las propiedades físicas que muestra el objeto físico, entonces deberíamos dar por supuesto que el palo que estamos observando está torcido, cuando realmente no lo está.

Otro argumento que podría afirmarse contra la tesis de la representacionalidad del carácter fenoménico es que la representación puede suceder, también, de manera inconsciente. Esto no supondría un mayor inconveniente de no ser, como defienden los internistas, que toda experiencia debe ser, en principio, consciente o, al menos, capaz de conciencia. Si admitiéramos, como hace el representacionalismo puro, la existencia de estados representacionales inconscientes, entonces dejaríamos una puerta abierta a una falta de intencionalidad consciente en nuestra vida mental. (Los representacionalistas acabarán por afirmar que dichos estados representacionales inconscientes pueden entenderse como estados neurofisiológicos, como hará Searle, o como estados con contenido no-conceptual, como hará Tye, como veremos en los capítulos 3 y 4 respectivamente).

A pesar de las resistencias de las posiciones representacionalistas, las voces críticas han conseguido reivindicar ciertas anomalías en los ejemplos indicados para defender la posibilidad de que el carácter fenoménico de las experiencias conscientes sea representacional. A menudo se considera que el carácter cualitativo de la experiencia consciente es una propiedad representacional que agota la intencionalidad presente en la experiencia. Pero, frente a esta posibilidad, encontramos a todos aquellos que se oponen a esta tesis fuerte en tanto que, consideran, existen estados mentales con carácter fenoménico que no tienen carácter intencional, estados mentales como las sensaciones corporales o los estados de ánimo. En este sentido, Rey (1998: 441) dice:

“Muchos de nosotros nos damos cuenta de que la depresión, la ansiedad, el placer o el orgasmo parecen ser estados referidos a sí mismos y no propiedades que estarían presentes en los objetos”.

Si aceptáramos la teoría representacional, dice, estaríamos confundiendo el carácter cualitativo con las propiedades intencionales de las experiencias conscientes. Serían dos tipos diferentes de *qualia*, dos tipos no asimilables ni identificables.

Los críticos de las teorías representacionales insisten también en que *qualia* diferentes pueden tener el mismo contenido intencional. En este sentido, Peacocke (1983: 12) aporta un ejemplo claro. Imagina, nos dice, que tu experiencia representa dos árboles a distancias diferentes de donde te encuentras, pero como teniendo las mismas características físicas. En este sentido, existe una diferencia cualitativa, porque son dos experiencias numéricamente diferentes cada una con su propio aspecto subjetivo, pero representan ambas el mismo contenido. (Algunos ejemplos similares los aportan trastornos patológicos como la sinestesia o la visión borrosa).

Existen otros muchos contraejemplos para aceptar limpiamente una teoría representacional de la experiencia con carácter fenoménico. En definitiva, todas ellas están vinculadas de alguna u otra forma a la posibilidad de la existencia de discrepancias en la sensación, en el contenido o en el modo particular en que se presentan a un sujeto que, de modos diferentes, pueden tenerse a la hora de suponer cómo representa el mundo una experiencia concreta. A partir de aquí, dejaremos los análisis en manos del recorrido concreto de las tesis de tres famosos representacionistas contemporáneos que vienen a ser representantes de posiciones extremas.

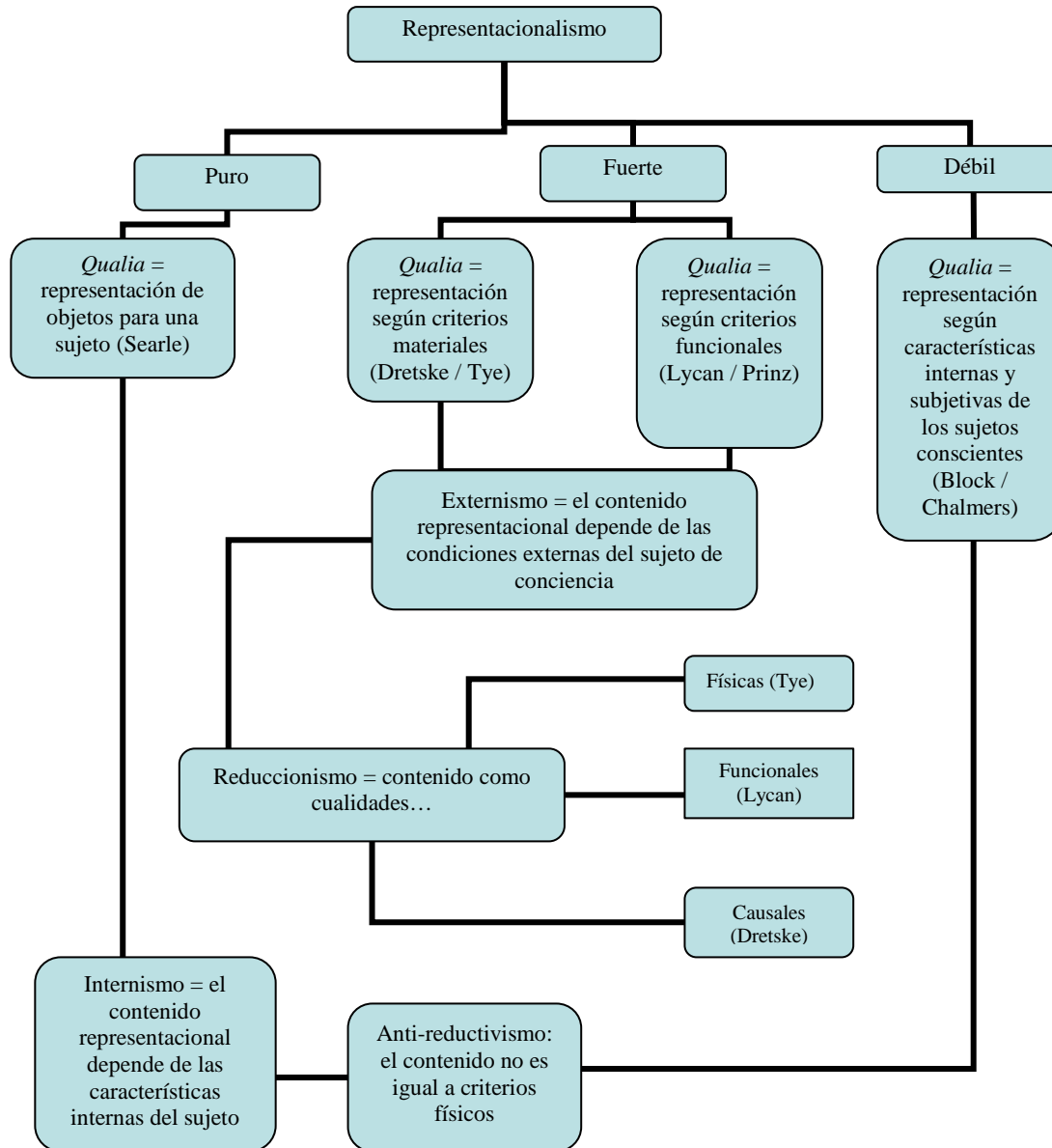
2.3. Recapitulación

La pretensión que animaba este capítulo era la idea general de conciencia. Hemos visto cómo en la literatura científica sobre nuestro tema, suele diferenciarse entre dos sentidos o nociones de conciencia bien diferenciados. Por una parte, suele atenderse a la existencia de una conciencia fenomenológica, aquel tipo de estados mediante los cuales nos damos cuenta de un modo subjetivo de nuestras propias vivencias (el carácter marcadamente cualitativo de las mismas) y que es indistinguible de las experiencias mismas. Por la otra, es habitual también considerar la existencia de una conciencia representacional o por acceso, encargada de aproximarnos a los contenidos informacionales de nuestros estados mentales. A diferencia de la primera noción (entendida de manera causal), esta segunda noción debe entenderse en un sentido funcional capaz de ser empleado en el razonamiento y el control de nuestras acciones. La mayor dificultad que hemos identificado a la hora de entender esta dicotomía era, como hemos advertido, el razonamiento generalizado de la existencia por separado de ambas nociones. Es decir, la mayoría de las aproximaciones actuales sobre la conciencia tan sólo tienen en cuenta una de ambas nociones (o ambas nociones, pero como dos entidades separadas e inconexas) como cierta, por lo que es habitual encontrar errores de fondo en sus conclusiones. La solución más plausible, como hemos postulado, sería confeccionar una definición que respetara las características de ambos tipos de conciencia pero que, sin embargo, no las concibiera como separadas, sino como las dos caras de una misma moneda.

Una vez establecido este esquema general de trabajo, nuestra intención fue mostrar en qué sentido entendían las diferentes escuelas la noción de conciencia de un modo particularizado. En un primer momento, identificamos una separación entre aquellas teorías que afrontaban el estudio de la conciencia fenoménica como una cuestión puramente representacional y la diferenciamos de aquellas aproximaciones que la concebían como un sistema causal cognitivamente accesible sin necesidad de apelar a representaciones. La motivación principal era mostrar las diferencias cruciales entre los intentos representacionales y no-representacionales de descripción y explicación de la conciencia para, en un segundo momento, poder atender críticamente a las conclusiones que sobre nuestra noción más relevancia han alcanzado en los últimos años y, así, poder aprovechar los elementos que, a nuestro juicio, son acertados y poder desechar los errores que ambas concepciones cometen.

Dedicamos una parte importante del capítulo a presentar también por separado las diferencias entre distintos enfoques representacionales. Nuestro interés en dicho momento fue mostrar la existencia de diferentes tipos de representación. En este sentido, dependiendo de la noción de representación que cada aproximación estaba teniendo en cuenta, pudimos comprobar que existía una noción diferente de conciencia que se estaba tomando en consideración. Pudimos, entonces, diferenciar distintos tipos de *qualia* (el carácter cualitativo de una experiencia consciente) según se entendiera en un sentido fuerte o débil el contenido representacional de los mismos, lo que nos permitió también aportar una serie de críticas que permitieron, por un lado, establecer una serie de dificultades en las aproximaciones representacionales de la conciencia pero también, por otro, encontrar una serie de elementos que deberíamos tener en cuenta a la hora de proponer nuestra propia explicación de la conciencia.

Acabamos este capítulo resumiendo en un cuadro explicativo los diferentes tipos de representacionalismo que en la actualidad podemos encontrar en la literatura científica sobre la conciencia.



CAPÍTULO 3. EL DUALISMO REPRESENTACIONAL DE SEARLE

*La naturaleza también me enseña mediante esas sensaciones de dolor, hambre, sed, etc. que no solo estoy hospedado en mi cuerpo como un piloto en un navío, sino que estoy estrechamente unido a él y, por así decir, tan entrelazado con él que me parece que formo con él un todo. Porque si no fuera así, cuando mi cuerpo está herido, yo, que no soy más que un ente pensante, no sentiría dolor, pues sólo percibiría esta herida mediante el entendimiento, de igual modo que el marinero percibe por la vista cuándo algo se ha dañado en su navío.
(Descartes, *Meditaciones metafísicas*)*

*Un suceso mental paradigmático es el juicio perceptual de un ser humano, en el sentido de que hay algo cuasiesférico, cuya superficie es roja, de un diámetro aproximado de un palmo, a una distancia de aproximadamente un palmo ante sus ojos. Este juicio tiene el contenido representacional que le acabamos de atribuir; el juicio sería incorrecto (falso) si no hubiese ante él algo con las características indicadas. Además, siendo un juicio perceptual, posee indudablemente un cierto carácter fenoménico: el juicio de haber sido producido en presencia de ciertas sensaciones visuales, y hay algo específico como lo que es tener sensaciones visuales tales.
(M. García-Carpintero, “Las razones para el dualismo”)*

III. EL DUALISMO REPRESENTACIONAL DE SEARLE

Tras los intentos conductistas y cognitivistas de explicación de lo mental a partir del análisis de la conducta de los individuos y de las conexiones funcionales que permiten la actividad mental, a partir de los años setenta resurge el estudio de la mente por parte de la filosofía. A pesar del importante papel que parece jugar en nuestras vidas, en dicho desarrollo ha existido un claro abandono de la noción de conciencia por no parecer posible integrar dicha noción dentro del marco objetivo de la ciencia natural.

Sin duda es John R. Searle uno de los filósofos que con mayor fuerza ha retomado la lucha y la defensa del papel central de la conciencia en la vida mental de los sujetos, llegando a afirmar el carácter central de la conciencia en una adecuada explicación de la mente dentro de una teoría naturalizada. En este sentido, sostendrá Searle, no existe mente sin conciencia. Pero, a pesar de ello, Searle no reconoce la co-extensividad entre la conciencia y lo mental. También postulará la posibilidad de la existencia de estados mentales inconscientes cuya realidad ontológica consiste, por completo, en la existencia de fenómenos puramente neurofisiológicos. Aunque a la hora de dar cuenta de ellos, al menos a un nivel personal, sea necesario recurrir nuevamente a la noción de conciencia.

Precisamente con la intención de mostrar el modo en que Searle revitaliza el estudio de la conciencia contra el error presente en la terminología mentalista heredada del cartesianismo a la hora de estudiar la mente, proponemos un pequeño recorrido crítico por las tesis más relevantes en defensa de un naturalismo de la mente, en general, y de la conciencia, en particular, que en las últimas décadas ha sostenido nuestro autor. La pretensión general de este capítulo es, como ya indicamos en el prólogo, primero, mostrar los términos en los que Searle concibe su estudio de la conciencia y de la mente desde un naturalismo biológico que pretende dejar de lado las dificultades que tanto los materialistas como los dualistas encuentran a la hora de definir la conciencia por estar equivocados, de base, en su propio estudio, presentando las críticas de Searle a diferentes intentos previos de explicación de la conciencia. Posteriormente, pasamos a analizar pormenorizadamente los argumentos concretos que Searle presenta para defender su propia teoría. En un primer momento, intentaremos rendir cuentas de los supuestos, heredados de sus primeros trabajos clásicos en filosofía del lenguaje, que subyacen al estudio de la mente y de la conciencia: las nociones de redes conceptuales y trasfondo tomarán el papel protagonista. En un segundo momento, presentaremos el programa del naturalismo biológico respecto de la mente y la conciencia como un intento, frustrado, de solución a la incógnita de la vida mental de los sujetos, en tanto que no será capaz de superar las dificultades presentes en su propia doctrina naturalista por acabar defendiendo un dualismo representacional de propiedades que anula toda posible reconciliación entre la conciencia (subjetiva) y la ciencia (objetiva) a la hora de aportar respuestas satisfactorias al clásico problema mente-cuerpo.

Por supuesto, todo esto se llevará a cabo con la intención, como ya hemos hecho patente, de presentar un análisis crítico general de las teorías representacionales de la conciencia que nos permita, por un lado, indicar aquellas dificultades que desde ciertos puntos de vista

debe afrontar toda teoría que pretenda ofrecer una respuesta a la presencia de estados mentales con carácter fenoménico y, por otro, intentar presentar una solución de corte fisicalista a dichas incógnitas.

3.1. El misterio de la conciencia

Desde el punto de vista de Searle, la conciencia (y la mente, al igual que la intencionalidad) pueden (y deben) ser tratadas desde una perspectiva biológica (aunque, como vamos a mostrar posteriormente, Searle considera que actualmente no disponemos de un desarrollo científico adecuado, un desarrollo que ha sido obstaculizado durante décadas por una serie de errores categoriales derivados de los análisis cartesianos). La cuestión principal, aquí, es ver en qué sentido podemos naturalizar la mente. La conciencia, nos dice Searle, no es más que un fenómeno biológico producido a partir de la ontología y la organización estructural y sistemática del sistema nervioso central de los seres vivos. La conciencia es el producto de microprocesos de nivel inferior, sub-doxásticos, que tienen lugar en el cerebro y que se reflejan en los niveles superiores, personales, en el nivel de las actitudes proposicionales. Así, cuando un estado de conciencia se da, existe algo intrínsecamente definitorio en él que permite identificarlo como tal; un ‘algo’ tan sólo cognoscible a un nivel introspectivo, por lo que requiere de la presencia de un sujeto capaz de sentirlo, se hace imprescindible la existencia de un individuo capaz de tener dicha vivencia, y en ese sentido es definible tan sólo como subjetivo.

Vista de este modo, la conciencia aparece como un producto emergente causado por el cerebro. El cerebro, entonces, segregaría conciencia del mismo modo que el páncreas segrega la bilis. Esto es, la conciencia es una propiedad explicable de modo causal por el comportamiento de los elementos constituyentes del cerebro, pero no es una propiedad de ninguno de ellos en particular, sino que tan sólo es definible a partir de las propiedades de dichos elementos en conjunto. A este respecto escribe,

“La conciencia está causada por procesos neuronales de bajo nivel en el cerebro, y es ella misma un rasgo del cerebro... Podemos concebirla como una ‘propiedad emergente’ del cerebro. Es una propiedad que se puede explicar causalmente por la conducta de los elementos del sistema; pero no es una propiedad de ninguno de los elementos individuales, y no se puede explicar simplemente como un agregado de las propiedades de dichos elementos” (Searle 1997: 30).

Vista de este modo, la única manera plausible de abordar el estudio de la conciencia (el único modo *serio* de enfrentarse a ella) sería a través de la investigación científicamente apropiada y dirigida a dilucidar la ontología física, química y biológica de la estructura cerebral que permite la emergencia de la conciencia y de las redes neuronales que componen el cerebro capaz de segregarla.

El problema principal, como nos dirá Searle, es que todavía nuestra ciencia no está en disposición de aportar dicha descripción. Todavía está muy lejano el día en que nuestra ciencia sea capaz de explicar razonablemente dicha estructura neuro-funcional que causa nuestros procesos psíquicos y los *qualia* que los acompañan. En pocas palabras, qué sea la conciencia

todavía es un *misterio*, dado que el estudio científico capaz de llevar a cabo su descripción todavía no es capaz de ofrecer (y no se sabe si algún día será capaz, siquiera, de ello) una explicación del conjunto biológico (y de las estructuras y redes neuronales que lo componen) que es el cerebro y que causa la conciencia.

Para Searle, entonces, vista la imposibilidad de una descripción física de dicho conjunto, no tenemos más remedio que afrontar la conciencia desde el sentido común. La conciencia, así, es todo el conjunto de elementos fenoménicamente relevantes que componen nuestra vida mental y que todos nosotros podemos advertir de una manera subjetiva. Aunque es biológicamente dependiente, la imposibilidad de afrontar su estudio científicamente invalida una posible descripción física, por lo que el modo adecuado de explicarla es apelando a nuestras propias experiencias fenoménicas, subjetivamente dependientes, a nuestro particular, personal y socialmente dependiente modo de sentir nuestras propias vivencias. Así, para Searle, hablar de conciencia

“alude a aquellos estados del sentir y del advertir que típicamente dan comienzo cuando despertamos de un sueño sin sueños y continúan hasta que nos dormimos de nuevo, o caemos en un estado comatoso, o nos morimos, o de un modo u otro quedamos inconscientes” (Searle 1997: 19).

Pero, para Searle, la imposibilidad de la ciencia actual de aportar una explicación física de la ontología del aparato cerebral causante de la conciencia no es el único escollo a salvar. Existe toda una serie de problemas añadidos y tradicionalmente heredados que todo intento de explicación de la conciencia debe salvar. Un problema que se arrastra desde antiguo es el dualismo ontológico cartesiano. Como hemos visto en el capítulo 1, para esta tradición deberíamos diferenciar entre una base física mecánicamente dependiente y un elemento trascendente inmaterial que gobernaría dicha base material. Aunque parezca sorprendente, Searle identifica todavía en la ciencia moderna científicos que consiguen a la vez describir el funcionamiento de dicho sustrato físico mientras defienden la posibilidad de la existencia de un elemento no-físico que ‘insufla’ nuestro cuerpo (como será, lo veremos enseguida, el caso de Edelman o Crick). La única manera real de poder enfrentarnos al estudio de la conciencia, nos dirá Searle, es considerarla como emergente e intentar descifrar los procesos cerebrales encargados de causar la conciencia.

El segundo obstáculo para un desarrollo científico de la conciencia lo podemos encontrar en aquellos que sostienen que hablar de naturalismo biológico (el modo en que Searle denomina a su propia posición) no es más que defender una forma renovada de dualismo de propiedades. Pero, nos dice Searle, hablar de emergencia no supone admitir la aparición de ninguna realidad nueva, sino que

“dicho estado no es una entidad separada de mi cerebro; al contrario, no es sino un rasgo de mi cerebro en el momento presente” (Searle 1997: 21).

Searle identifica un tercer problema en el estudio actual de la conciencia: la explicación de los *qualia*. La conciencia no es más, nos dice Searle, que el conjunto de estados cualitativos

que se nos presenta en los *qualia*. Es decir, la conciencia no es más que un conjunto de vivencias subjetivamente dependiente. El problema aparece, entonces, cuando tenemos que explicar científicamente cómo el complejo sistema biológico que compone el cerebro es capaz de producir dicha experiencia subjetiva. Esto es, el problema aparece cuando surge la necesidad de explicar cómo podemos describir objetivamente la realidad física, química y biológicamente constituida en el sistema nervioso central que permite a los individuos la rica variedad de experiencias fenoménicas que son capaces de sentir de manera subjetiva.

El cuarto problema que Searle considera crucial para poder llegar a eliminar el malentendimiento actual de la ciencia de la conciencia es mitigar la tendencia de las investigaciones recientes de explicar la vida psicológica de los individuos a partir de la metáfora computacional de la mente. Es decir, a partir de la comparación del cerebro y la mente con el *hardware* y el *software*, respectivamente, de un ordenador (al modo en que los actuales estudios en Inteligencia Artificial lo realizan). Searle niega la posibilidad de que la mente sea un programa implementado en la base física del cerebro a partir de su famoso argumento de la habitación china (véase la sección 1.4.5.2.). Desde este punto de vista, los ordenadores son sistemas manipuladores de símbolos que no son capaces de entenderlos. Es decir, un ordenador tan sólo gestiona la combinación necesaria de símbolos pertinente según la entrada requerida a partir de una serie de reglas externas (los programas) que tan sólo tienen un carácter sintáctico, pero carecen por completo de la capacidad semántica necesaria para poder llegar a comprender dicha manipulación, una capacidad que sí poseen los verdaderos seres conscientes. Por tanto, concluirá Searle, la mera manipulación de signos no es justificación suficiente para afirmar la existencia de conciencia. Se vuelve necesaria también la apelación a un componente semántico del que no dispone un ordenador pero que la masa biológica que conforma el cerebro (al menos el nuestro, el humano) sí posee de manera intrínseca (Searle 1980a; Searle 1985: 33-48; Searle 1997: 26-27).

Vamos a detenemos en el análisis que Searle realiza de algunos de los intentos más relevantes de explicación (científica) de la conciencia y que identifica como equivocados por presuponer, al menos, alguno de los cuatro errores anteriormente mencionados. Ello nos permitirá avanzar en nuestra exposición de la propia teoría de Searle y nos dará los materiales necesarios para poder realizar una crítica de sus propios propósitos con la intención final de presentar una crítica general a toda teoría de la conciencia basada en supuestos representacionales.

3.1.1. Francis Crick y la búsqueda científica del alma

Como ya hemos comentado brevemente, Searle considera cierta la hipótesis de partida de Crick. Para ambos, la conciencia está causada de algún modo por las redes neuronales y los procesos de bajo nivel que provocan la aparición de estados mentales. Esto es, la conciencia está causada por la ontología profunda de la estructura física del cerebro humano. Con Crick, Searle también considera que dichas redes neuronales de nivel inferior producen mucho más que mero comportamiento. Producen los *qualia*, las vivencias conscientes que los sujetos humanos

sentimos ante la presencia de determinadas experiencias con carácter fenoménico. Esto vendría a confirmar, entre otras muchas cosas, nos dirá Searle, que el cerebro es mucho más que una simple computadora. Porque, a diferencia de éstas (que tan sólo sirven para realizar una manipulación computacional de símbolos), los cerebros, además, producen vivencias conscientes. Por tanto, concluirán, las redes neuronales producen la vida psíquica de los individuos (Searle 1997: 37; Crick 1989: 3).

Pero Searle, aunque acepta la base ontológica que subyace a la propuesta de Crick, critica varios errores existentes en la presentación de su teoría (tal vez, nos comenta, debidos a su falta de preparación y precisión filosófica).

El principal problema de la teoría de Crick es que parece malinterpretar la naturaleza de los *qualia*. En la posición de Crick siempre aparece una clara referencia a las vivencias fenoménicas de los demás, a sus experiencias sensoriales a partir de los diferentes grados de intensidad de disparo que se dan en las redes neuronales. Pero nunca se refiere al *quale* (en singular), a la vivencia propia. Pero todavía más problemática, si cabe, en la argumentación de Crick es su insistencia constante en presentar un cierto reduccionismo eliminador de las propiedades fenoménicas y de la conciencia. La dificultad principal, aquí, es que dicha explicación reduccionista nunca tendrá éxito. No sólo porque la presentación que aporta su libro es la de una explicación causal de la conciencia, por lo que parece que la teoría de Crick, aunque cierta en su argumento central, peca de algún tipo de inconsistencia a la hora de identificar elementos fenoménicos con elementos físicos (Searle 1997: 38). Sino, todavía más importante, sobre todo porque la naturaleza de las vivencias fenoménicas es irreductible. En este sentido, cuando Crick afirma que “las sensaciones complejas son propiedades emergentes que surgen en el cerebro a partir de las interacciones de sus muchas partes” (Crick 1989: 11), Searle está convencido de que está profundamente equivocado al confundir la naturaleza de los elementos fenoménicos de la mente porque, aunque los disparos eléctricos que se producen en las redes neuronales causan las sensaciones, éstos no son lo mismo que las sensaciones (Searle 1997: 39).

Si lo que Crick pretende hacer, como Searle sostiene, es mostrar que la conciencia es una propiedad emergente del cerebro (esto es, aportar una explicación causal de la aparición de la conciencia), entonces anda mal encaminado apostando por argumentos reduccionistas de la conciencia.

Un importante problema que también se desprende de la argumentación de Crick es su inconsistencia a la hora de mostrar la estructura lógica de dichas redes neuronales que permitirían la aparición de la conciencia. Aquí es cuando Crick hace referencia al denominado ‘problema de la ligadura’ (es decir, al problema relativo a la unidad de la experiencia consciente a pesar de que la información que la compone está dispersa por toda la corteza cerebral y se localiza en numerosas y muy diferentes partes del cerebro. En este sentido es en el que Crick (también en los trabajos conjuntos con Koch) desarrolla la hipótesis de la existencia de una gama de disparos neuronales específica que se produce a 40 hercios que es la responsable de relacionar la actividad de todas las neuronas implicadas en la confección de la imagen consciente. Pero, como nos dirá Searle, que la unidad de la imagen consciente se produzca por

una serie de disparos neuronales aporta una explicación causal de por qué los seres conscientes podemos tener imágenes conscientes, pero no indica nada respecto de la naturaleza propia de dichas imágenes. Los intentos de Crick por ofrecer una explicación biológica distan mucho de aportar una explicación de la ontología profunda de las sensaciones. En este sentido, Crick estará aportando una explicación causal que permitirá definir en términos de emergencia la aparición de la conciencia, pero no aporta ningún argumento que permita arrojar luz sobre la propia naturaleza de las sensaciones. A este respecto comenta

“Aun si las especulaciones de Crick resultaran un cien por cien acertadas, todavía necesitaríamos conocer los mecanismos mediante los que los correlatos neuronales causan sensaciones conscientes y estamos lejos de poder ofrecer una conjetura de ese tipo” (Searle 1997: 42).

3.1.2. Gerald Edelman y la topobiología del cerebro

El enfoque cognitivo de la conciencia aportado por Edelman pretende explicar la naturaleza de la mente, también, a partir de la presencia causal de redes neuronales. Aunque su teoría es mucho más compleja y completa que la de Crick, como nos dirá Searle, Edelman tampoco conseguirá aportar una explicación satisfactoria al problema crucial que éste contempla en el estudio de la conciencia: qué tipo de ontología profunda sub-neuronal permitiría explicarnos que las complejas redes neuronales presentes en nuestro cerebro produzcan realmente el mundo de los *qualia* y todos los complejos matices y fenómenos que pueblan la vida mental humana (Searle 1997: 55).

En la concepción de Edelman es crucial la estructura funcional particular que, considera, tiene el cerebro humano en tanto que base causante de nuestras actividades psíquicas. Dentro de esta estructura, es conveniente resaltar tres nociones fundamentales. La primera es el mapa neuronal. Es un conjunto de neuronas que se activan al unísono en tanto que dependientes de cierto tipo de receptores sensitivos situados en la piel y en la retina. En este sentido, al activarse los receptores sensitivos pertinentes relacionados con un mismo grupo neuronal, permiten accionar diversos mapas neuronales que, al implementarse en conjunto, permiten causar la experiencia psicológica de la imagen generada, esto es, del *quale* concreto (Edelman 1987: 47).

La segunda noción es la de selección neuronal de grupo. La noción de selección se refiere a la progresiva especialización de los diferentes grupos neuronales del cerebro según la progresiva selección de activación de distintos grupos neuronales de acuerdo a la actividad cognitiva pertinente. En este sentido, las neuronas seleccionadas formarán un grupo que se activará bajo un patrón de acción determinado por la característica actividad cognitiva concreta que en cada caso se requiera activar, permitiendo un mayor nivel de especialización neuronal y una mayor flexibilidad mediante la generación de engramas (o mapeados) estándar que ya contendrán codificada la información necesaria y requerida para la activación de aquellos grupos neuronales capaces de causar sensaciones determinadas (Edelman 1987: 56-61).

La tercera noción, el reingreso, se refiere a la capacidad de los diferentes mapeados neuronales de intercambiarse múltiples señales en diversas direcciones (lo que indica su

interconexión) que permitirían un trabajo en equipo entre varios grupos neuronales diferentes a la hora de organizar la generación de una determinada sensación (Edelman 1987: 64).

La teoría que ofrece Edelman, apoyada en estas tres nociones básicas, es una propuesta especulativa del modo en que la mente humana ha ido evolucionando según el planteamiento selectivo darwiniano. Así, Edelman ofrece una reconstrucción que, basada en la capacidad de la memoria, muestra cómo la evolución del cerebro humano fue capaz de pasar de la construcción y activación de los sentidos por motivos puramente de beneficio adaptativo y de supervivencia a la formación de categorizaciones cada vez más complejas. Lo que Edelman pretende demostrar, en última instancia, es cómo el cerebro humano, pasando de una conciencia primaria a una conciencia de nivel superior, evoluciona mediante la configuración y remodelado progresivo de su cartografía global (formada por diferentes, y cada vez más complejas, rutas de mapeado).

El problema ahora, como nos dice Searle, es el mismo que encontrábamos en la propuesta de Crick. Si lo que encontramos es una serie de estructuras y funciones capaces de causar estados cualitativos, ¿cómo pasamos de tener unos a tener los otros? En este sentido, Searle sentencia,

“El problema de explicar los estados cualitativos internos del sentir y del advertir llamados *qualia* no es un aspecto del problema de la conciencia que pudiéramos dejar de lado; es el problema de la conciencia, y *qualia* no es sino un malhadado nombre para la conciencia de todos los estados conscientes” (Searle 1997: 53-54).

Vemos, entonces, que Edelman, como anteriormente Crick, al menos para Searle, no consigue dar cuenta de la ontología profunda de la mente al no aportar una explicación satisfactoria de la naturaleza del carácter fenoménico de las experiencias conscientes.

3.1.3. Roger Penrose: microtúbulos y citoesqueletos

Para Searle, el enfoque de Penrose es radicalmente diferente de los enfoques biologicistas de explicación de la conciencia presentados hasta el momento. Especialmente por dos motivos. El primero, porque Penrose considera que la mente humana realiza ciertas funciones que no son susceptibles de simulación computacional (la mente tiene funciones no-computacionales) y, el segundo, porque Penrose considera que la ontología física y biológica de las neuronas que posibilitan dichas funciones no-computacionales no está tan sólo presente a un nivel funcional en las neuronas y en las redes que conforman (individual o grupalmente consideradas), sino, además, en un nivel neuronal mucho más profundo donde intervienen eventos cuánticos que impedirían la determinación de su computación a nivel macroscópico.

Tanto Searle como Penrose, entonces, estarían de acuerdo en que la conciencia puede explicarse científicamente a partir de los elementos biológicos y físicos que la causan. Ambos niegan también la Inteligencia Artificial fuerte. Pero Searle disiente de Penrose cuando éste niega la posibilidad de la Inteligencia Artificial débil. A partir del Teorema de Gödel, basado en el problema de la imposibilidad de detención de un bucle en la generación de premisas a la hora

de demostrar una verdad matemática,¹ Penrose no tiene más remedio que negar la posibilidad de, tan siquiera, simular la mente humana. Porque la simple simulación se encontraría con serias dificultades sintácticas a la hora de tener que rendir cuentas con las actividades no-computacionales anteriormente identificadas que, en el caso de los humanos, serían una mera cuestión de intuición. En pocas palabras, lo que para un programa de ordenador sería una imposibilidad (demostrar la verdad de ciertos teoremas, por la recursividad de la generación de premisas que necesariamente deben producirse para afirmar la verdad de la proposición anterior), para un ser humano es una mera trivialidad (bastaría con dar un vistazo a la proposición para que el sujeto pudiera entenderla).

Pero, como dirá Searle, es siempre posible pensar en ejemplos que salven esta dificultad. Pensemos en un programa capaz de simular la evolución de la meteorología de una determinada zona geográfica. A nadie se le ocurriría pensar, por supuesto, nos dice Searle, que el ordenador sea el espacio físico donde ocurren los diferentes cambios atmosféricos. Éstos son una simple simulación y, claro está, es una simulación imprecisa basada en los datos que se le han introducido. Pero, si nos damos cuenta, en casos como éste, a los humanos nos ocurre lo mismo. Nuestras previsiones del tiempo son siempre aproximadas y pocas veces acertadas. En este sentido, parece posible concluir que Penrose se equivoca al pretender afirmar la presencia de actividades mentales que no pueden ser simuladas de modo computacional por un ordenador, aunque ellas mismas, como productos de la simulación, no sean reales.

De todos modos, el problema principal de Searle con la teoría de Penrose es que ésta, nos dice, oculta la verdadera problemática de la conciencia bajo problemas todavía más complejos, como son las bases cuánticas de la física o la indeterminación mecánica. Lejos de aclarar el enigma de la conciencia, para Searle, Penrose estaría aumentándolo agregándole misterios adicionales (Searle 1997: 87).

3.1.4. Daniel Dennett y la negación de la conciencia

Cuando analiza la teoría de la conciencia de Dennett, Searle parte de la afirmación de que éste defiende una interpretación fuerte de la mente dentro del marco de la Inteligencia Artificial fuerte. En este sentido, considera que para Dennett la mente humana es idéntica a un ordenador. Si esto es así, se pregunta Dennett, siempre según Searle, ¿qué pinta, entonces, la conciencia? Siguiendo una interpretación radical, lo que Dennett estaría proponiendo es que la conciencia no existe. En una interpretación más suave, Dennett estaría diciendo que la conciencia es algo real, pero que no cumple ningún papel causal en el comportamiento de los

¹ Esta caracterización del Teorema de Gödel, por supuesto, no es la clásica. En primer lugar, existen varios teoremas de Gödel relacionados con esta imposibilidad. El más habitual de ellos parte de la constitución de un lenguaje de primer orden semánticamente cerrado que permite concluir que nunca podremos demostrar en dicho lenguaje que una teoría es a la vez completa y consistente. Otro de los teoremas de Gödel emplea el lenguaje de primer orden creado para la formalización de la aritmética elemental (según la versión de Russell) para concluir que, formulada de tal modo, en la aritmética siempre habría proposiciones (intuitivamente) verdaderas que serían indemostrables desde dentro del mismo sistema. Penrose parece mezclar todas estas interpretaciones y teoremas con otros resultados obtenidos en computación y funciones recursivas utilizando sus resultados de modo partidista para intentar demostrar sus propios resultados y conclusiones. Para una presentación general de los teoremas de incompletud de Gödel, véase Nagel and Newman (2007).

seres conscientes. Afirmaría que la única función de la conciencia sería la de ser el sistema de control del resto de funciones cognitivas y biológicas propias del cerebro y del sistema nervioso central. (Desde mi punto de vista, ninguna de las dos interpretaciones es cierta. Véase capítulo 5).

Tal y como la interpreta Searle, la teoría de Dennett estaría pretendiendo negar la existencia de estados mentales internos, tales como las vivencias, que intuitivamente pueblan el mundo mental de los humanos. En su pretensión de cientificidad, Dennett estaría presentando una teoría tan reductiva que, por no ser congruente con los hechos intuitivos, acaba por negarlos. Así, para Searle, la posición de Dennett es reduccionista (cerca a la teoría conductista) porque explica el comportamiento humano como puramente dependiente de un mero nivel computacional debido a la actividad neuronal y nerviosa del cerebro, dejando reducida la conciencia a mero epifenómeno.

3.1.5. David Chalmers y la mente consciente

La concepción de Chalmers, para Searle, también cabría dentro de la interpretación de la Inteligencia Artificial fuerte. Searle insiste en que Chalmers ofrece un concepto de conciencia concebido como un sistema de información que podría hacernos pensar que un termostato, un circuito de calefacción o un coche son sistemas conscientes. (Chalmers, recordémoslo, es el introductor del concepto de conciencia extendida).² Pero para llegar a esta conclusión, nos dice Searle, tan absurda, Chalmers realiza un enrevesado recorrido por la Inteligencia Artificial, el dualismo de propiedades, el funcionalismo e, incluso, la mecánica cuántica que tan sólo muestra la tan a menudo innecesaria argumentación filosófica en cuestiones de filosofía de la mente (Searle 1997: 149-150). No es de extrañar, entonces, que Searle concluya diciendo que

“algunos libros son importantes no porque resuelvan algún problema o porque lo planteen de un modo que permita vislumbrar un solución, sino porque son sintomáticos de la confusión de una época” (Searle 1997: 149).

3.1.6. Conclusión: cómo convertir un misterio en un problema (para todos)

Para finalizar este apartado, intentaré sintetizar brevemente la posición inicial de Searle, una posición que nos permitirá ir abriendo camino a la presentación de discusiones posteriores y que nos llevará a presentar nuestro análisis crítico, como prometimos al principio, de las propias tesis defendidas por este autor.

² La tesis de la conciencia extendida se debe al clásico trabajo Chalmers and Clark 1998. En este texto, sus autores defiende el conocido como Principio de Paridad: si al enfrentarnos a cierta tarea, una parte del mundo funciona como un proceso que si estuviese en la cabeza no dudaríamos en aceptarlo como una parte del proceso cognitivo, entonces esa parte del mundo es (en ese momento) parte del proceso cognitivo mismo. Esta tesis no tendría sentido si no consideraran posible que todo sistema puede realizar procesos de descarga computacional de datos y tareas en el medio (esto es, que la mente humana es un sistema computacionalmente abierto, que se completa con la colaboración del medio, en el que descarga tantos datos y procesos computacionales como puede) y es capaz de realizar la internalización de procesos de descarga (esto es, que algunos de estos procesos de descarga se internalicen y conviertan en módulos interactivos con las redes neuronales, cuando nuestra mente responde a ellos de manera automática, sin mediar procesos de inferencia o reconocimiento de fiabilidad). Para una ampliación de esta noción, véase Broncano (2006).

Como deja constancia en su polémica contra otros autores, para Searle es un mero hecho empírico que la conciencia es una cuestión de experiencia subjetiva y personal. Pero dicha noción también se encuentra mediada por la acción social, en tanto que advertimos en nosotros igual que en los demás la presencia de determinadas vivencias fenoménicas involucradas en el carácter cualitativo de ciertas sensaciones. Por tanto, si éstos son hechos empíricos será posible que puedan ser explicados por las ciencias. Pero ello no quiere decir que todo sea reducible a mera explicación científica. Aunque es cierto que la mente tiene profundos niveles funcionales anclados en su distribución estructural biológica, nos dice Searle, la conciencia y la actividad consciente aparece a un nivel subjetivo, por lo que debe representar un modo ontológico de ser que torne inviable su explicación a un nivel nomológico propio de las ciencias empíricas. En este sentido, aunque la neurología y la neurociencia permiten establecer ciertas explicaciones generales de cómo las redes neuronales causan la actividad psicológica humana mediante la construcción de profundas redes lógicas de activación y desactivación neuronal, ello no impide la presencia de un nivel todavía más profundo (y, por tanto, inaprensible) que esconda la verdadera naturaleza de los fenómenos conscientes (aunque, como veremos, que no impida la presencia de dicho nivel no demuestra necesariamente su existencia; es más, nos posicionamos claramente en contra de dicha existencia. Véase capítulo 6). A la hora de descubrir esta naturaleza oculta, aunque no sea viable tener presente la Inteligencia Artificial fuerte por aceptar una epistemología científicamente inviable, la Inteligencia Artificial débil puede sernos de gran ayuda, dado que permite la simulación de ciertas actividades cognitivas humanas que, a un nivel muy básico, muestran los funcionamientos internos de interacción entre biología y sujeto.

En su intento por dilucidar el presunto misterio de la conciencia, Searle ha perseguido, por un lado, mostrar que existe una irreductiblemente profunda noción de conciencia que no puede dejar de contemplarse de modo subjetivo, pero que, a la vez, por otro lado, es dependiente de una intuitivamente fuerte concepción científica del mundo (Searle 1992: 85). En este sentido, parece que Searle está intentando ser conciliador al proponer una aproximación compatibilista (Sabatés 1999: 140) en su explicación de la conciencia que tenga en cuenta tanto la teoría atómica de la materia como la teoría evolutiva predominante en biología (Searle 1992: 86).

Pero, como hemos pretendido mostrar, la imagen de la actual filosofía de la mente que Searle tiene en cuenta es extremadamente parcial. Considera que tan sólo habría dos grandes orientaciones que se disputan todo el poder explicativo y definitorio en filosofía de la mente, en general, y de la conciencia, en particular. Por un lado estarían aquellos que, como Dennett o Chalmers, explicarían la mente como simples programas de ordenador y, por otro lado, encontraríamos a todos aquellos que, como Crick, Edelman, Penrose o el propio Searle, consideran que la mente puede ser explicada por las redes neuronales que componen el cerebro. No voy ahora a comentar nada respecto a la sesgada presentación que Searle realiza de las teorías de la conciencia mencionadas. Dicha opinión irá aflorando conforme vayamos ahondando en la teoría de Searle. Lo que sí quiero avanzar aquí es la tremenda confusión que la concepción de Searle presenta cuando analiza las teorías de Dennett y Chalmers. Por un lado, si

consideramos atentamente ambas posiciones, las dos están estrechamente basadas también en los resultados empíricos aportados por la neurobiología actual. Además, por otro lado, intentar unir en un solo grupo teorías tan dispares como la de Dennett (que por supuesto no es reduccionista, aunque sí fisicalista representacional, como pretendo demostrar en el capítulo 5) y la de Chalmers (claro defensor del dualismo de propiedades) tan sólo demuestra, por un lado, estar defendiendo una noción de conciencia totalmente diferente de la presentada por los autores mencionados (muy probablemente todos ellos defiendan conceptos de conciencia diferentes) y, por otro lado, no estar entendiendo en qué términos está llevándose a cabo la actual investigación respecto de la conciencia.

Veamos en qué términos se desarrolla ahora la teoría de la mente de Searle, prestando especial atención a la teoría de la conciencia de nuestro autor, en tanto que, como hemos dicho, una revisión crítica de sus posiciones permitirá obtener una serie de objeciones para rechazar su teoría.

3.2. El naturalismo biológico de Searle

A la hora de afrontar el estudio de la mente, Searle considera como punto de partida la idea de que tanto la intencionalidad como la conciencia, los dos rasgos característicos de lo mental, son elementos biológicos perfectamente explicables desde el marco objetivo de las ciencias físicas. En este sentido, Searle nos presenta un estudio pormenorizado de las condiciones físicas que permiten a la intencionalidad ser afrontada como un rasgo físico de la mente humana y a la conciencia ser analizada en los términos de los estados neurofisiológicos que la causan.

Con la noción de intencionalidad, Searle se refiere a la característica de que muchos de nuestros estados mentales, aunque no todos, consisten en estar dirigidos a (o ser respecto de) objetos y estados del mundo diferentes de ellos mismos. En este sentido es donde se demuestra el carácter representacional de la mente en Searle. Si un estado mental es respecto de algo que se encuentra fuera de él mismo, ese mismo estado mental tiene un contenido que representa en el mundo interno del sujeto al objeto externo al cual se dirige. Así, se plantea un problema crucial en el estudio de la intencionalidad. Después de todo, como nos dice Searle,

“¿cómo pueden ser acerca de algo procesos en mi cerebro que consisten finalmente en ‘átomos en el vacío’? ¿Cómo pueden átomos en el vacío representar algo? Uno se inclina a decir: las cosas y los procesos en el mundo simplemente son... ¿Cómo puede el acerca de algo ser un rasgo intrínseco del mundo?” (Searle 1980a: 340).

El modo en el que se ha intentado solucionar este problema, nos dice Searle, ha sido mediante una serie de intentos de naturalización del contenido intencional. Es decir, se ha intentado explicar la intencionalidad en términos de procesos físicos. Por ejemplo, el funcionalismo, al pensar que el significado no se encuentra en nuestras cabezas (Putnam 1975), consideró plausible la necesidad de estipular todo un conjunto de relaciones reales causales que

vincularán dicho significado (o los estados mentales asociados) con los objetos del mundo. En este sentido,

“si el significado de oraciones como ‘El agua es húmeda’ no puede explicarse en términos de lo que sucede dentro de las cabezas de los hablantes, entonces la creencia de que el agua es húmeda tampoco puede ser un asunto de lo que sucede en sus cabezas” (Searle 1992: 63).

Pero, para Searle, estos intentos de naturalizar el contenido no permiten una explicación satisfactoria de lo que sea el contenido intencional. Por ejemplo, pensemos en el caso de la representación errónea (o problema de la disyunción. Cf. Fodor 1987). ¿Cómo puede alguien equivocarse al representar un objeto externo si depende de tal objeto dicha representación? El problema es que, si algo es una representación de otra cosa, entonces debe mantener relaciones con esa cosa que dificultan la posibilidad de que la representación sea errónea. Es más, todo intento similar de explicación de lo mental fracasará porque se empeña en dejar de lado la intencionalidad. El problema reside en que

“no es posible reducir los contenidos intencionales a algo más porque, si fuera posible, serían algo más, y no son algo más” (Searle 1992: 65).

Para Searle, el modo de aclarar el problema de la intencionalidad, como veremos posteriormente, es describir detalladamente cómo los fenómenos intencionales son causados por procesos biológicos y cómo se realizan en sistemas biológicos, como ocurre, por ejemplo, con la sed:

“Hasta donde sabemos, al menos cierto tipo de sed es causado en el hipotálamo por secuencias de disparos de neuronas. Estos disparos son a su vez causados por la acción de la hormona peptídica angiotensina II en el hipotálamo, y ésta, a su vez, es sintetizada por la renina, que es secretada por los riñones. La sed, al menos la de este tipo, es causada por una serie de eventos en el sistema nervioso central, principalmente en el hipotálamo, y se realiza en el hipotálamo” (Searle 1981: 726).

Al intentar aclarar la noción de conciencia, Searle se refiere a que algunos de nuestros estados mentales son conscientes. Esto que parece circular, es para Searle un hecho evidente porque es

“el hecho central de la existencia específicamente humana, puesto que sin ella todos los demás aspectos específicamente humanos de nuestra existencia serían imposibles” (Searle 1985: 20).

Searle no parece capaz de proporcionar una definición concreta respecto de qué sería la conciencia más allá de mostrar toda una serie de ejemplos:

“Cuando me despierto, después de dormir sin haber soñado, paso a estar consciente, un estado que continúa tanto tiempo como estoy despierto. Cuando voy a dormir, o me ponen bajo una anestesia general, o muero, mis estados conscientes cesan. Si sueño mientras duermo adquiero de nuevo conciencia, aunque las formas de conciencia en el sueño son, en general, de un nivel de intensidad mucho menor que la conciencia ordinaria mientras estamos despiertos. La conciencia puede variar en grado, incluso durante las horas de vigilia, como por ejemplo cuando pasamos de estar completamente despiertos y atentos a estar dormidos y

relajados o simplemente cansados y poco atentos... La conciencia es como un mecanismo de encendido y apagado: un sistema es consciente o no lo es” (Searle 1992: 95).

En este sentido, Searle parece no querer ir más allá de la afirmación, de nuevo circular, que insiste en que un estado mental es consciente cuando sucede en un organismo consciente. Así,

“con conciencia simplemente quiero referirme a aquellos estados subjetivos de sensibilidad o conocimiento que comienzan cuando uno se despierta por la mañana y continúa durante todo el tiempo de vigilia hasta que uno vuelve a dormir sin sueños o entra en un estado de coma o muere o, simplemente, como se suele decir, está inconsciente” (Searle 1990: 635).

En los apartados siguientes pretendemos mostrar los intentos de argumentación que Searle presenta para explicar estas dos nociones desde un punto de vista naturalizado y las posibles críticas que podrían hacersele.

3.2.1. La intencionalidad de la mente

En una serie sucesiva de artículos,³ John R. Searle (un filósofo pionero en el desarrollo de la teoría de los actos de habla, originalmente formulada por el filósofo oxoniense John L. Austin a mediados de los años 50) anticipaba a principios de los años 80 del siglo XX lo que sería su posición respecto de la mente. Producto final de dicha conversión a filósofo de la mente, su libro *Intentionality* (Searle 1983), marcaba un antes y un después en el estudio respecto de la mente. El título nos muestra claramente el cambio de tema. Mediante el estudio de la noción de intencionalidad, Searle pretendía analizar la relación existente entre el lenguaje y la mente, dado que ambos (de un modo u otro) parecen mostrar intencionalidad.

Pero el objetivo de nuestro autor iba, en realidad, mucho más lejos. En última instancia, la pretensión de Searle es ofrecer un estudio general sobre la naturaleza de la mente, sobre su ontología profunda. Partiendo de la base del error que existe en las teorías computacionales de la mente, Searle pretende plantear un análisis de la mente que permita dilucidar, a un nivel profundo y esencial, qué quiere decir que la mente humana es intencional y en qué sentido podemos decir que la mente humana se dirige hacia el mundo.

3.2.1.1. Diferencias entre la intencionalidad de la mente y del lenguaje

Searle comienza insistiendo en la naturaleza intencional de la mente.

“La intencionalidad es aquella propiedad de todo estado y evento mental según la cual le permite dirigirse hacia objetos, estados y hechos del mundo” (Searle 1983: 1).

Esto es, como Searle insiste, es esencial a todo estado mental el ser intencional. En este sentido, es propio de las creencias el creer *algo*; es propio de los deseos desear *algo*; es propio de las esperanzas esperar *algo*.

³ Véase especialmente Searle 1979a y 1979b, 1980a y 1980b, 1981 y 1982.

Pero, como también advierte Searle, aunque todo estado mental es característicamente acerca de algo, muchas veces podemos encontrarnos en estados mentales concretos que, a pesar de ser similares a los anteriores, parecen carecer de esa esencia intencional. Pensemos tan sólo, nos comenta Searle, en aquel individuo que se encuentra en un estado de ansiedad, o en aquel que se encuentra en un estado de miedo. A diferencia de las creencias y los deseos, de los cuales podemos identificar el objeto o hecho al cual se dirigen, en casos como en los de la ansiedad o el miedo, demasiadas veces somos incapaces de determinar aquel objeto concreto (o aquella situación del mundo) hacia la cual se dirigen, y que son causantes de la entrada de dicho individuo en ese cierto estado mental (Searle 1983: 2). Eso sí, aunque existan estados mentales que no son intencionales, aquellos que posean un objeto lo harán de manera interna y esencial (Searle 1983: 5).

Cuando Searle nos habla en estos términos considera que está mostrándonos las características lógicas de la intencionalidad. A este respecto nos dice que:

1. Los estados intencionales tienen un modo psicológico (intencional o representacional), una manera subjetiva de darse como un contenido.

2. El contenido de un estado mental (al igual que el contenido proposicional de un acto de habla) no debe confundirse con el objeto del que es pensamiento, porque el contenido es parte del estado mental.⁴

3. Los estados intencionales (como los actos de habla) tienen una dirección de ajuste que depende de su modo psicológico.⁵

4. Los estados mentales (como los actos de habla) poseen unas condiciones de satisfacción (de éxito).⁶

Los estados intencionales, para Searle, por tanto, son aquellos estados que poseen un contenido intencional presentado de una determinada manera psicológicamente situada, que determinará su dirección de ajuste con el mundo y sus condiciones de satisfacción.

Como hemos comentado ya, para Searle, los estados mentales intencionales son similares a los actos de habla. Ambos poseen intencionalidad, tienen un contenido y un modo psicológico. Incluso ambos tienen una dirección de ajuste con el mundo y unas determinadas

⁴ En este sentido, el contenido (tanto en los actos de habla como en los estados mentales) es intencional, en tanto que se dirige a los objetos físicos pero no puede identificarse con ellos. El mayor inconveniente con dicho contenido es poder explicar en qué sentido podemos decir que las condiciones semánticas que marcan dicho contenido pueden ocurrir sin necesidad de tener que afirmar que ocurren o se encuentran en el interior de la cabeza de los sujetos conscientes. Es decir, en qué sentido podemos naturalizar dicho contenido sin apelar a una dualidad de entidades. Desde este punto de vista, la opción de Searle pasa por intentar describir con el mayor detalle posible cómo esta clase de fenómenos semánticos (ya sean los actos de habla o las experiencias conscientes) son causados por procesos biológicos situados en sistemas biológicos como nosotros.

⁵ La dirección de ajuste hace referencia a la necesidad de que el contenido representacional presente en el sujeto perceptor represente precisamente algo del mundo, esto es, que se ajuste a él. En este sentido, dependiendo del tipo de estados al que nos estemos refiriendo, la dirección de ajuste puede ser de la mente al mundo (cuando, por ejemplo, es la mente la que se adecua al mundo, como en el caso de las creencias o las percepciones visuales), del mundo a la mente (como en el caso de los deseos o las esperanzas, donde se espera que el mundo se adecue a la mente) o puede ser neutro (como en el caso de los estados de ánimo, donde no existe una relación evidente entre mundo y mente).

⁶ Las condiciones de satisfacción son los requisitos necesarios para que podamos decir que se ha dado con éxito un determinado acto de habla o una determinada experiencia. En este caso, se evaluará como feliz o infeliz si se adecua o no un determinado acto de habla o una determinada experiencia a lo que realmente sucede en el mundo.

condiciones de satisfacción. Pero existe una diferencia crucial entre ellos que permitiría diferenciar la intencional del lenguaje de la intencionalidad de la mente. Mientras que la intencionalidad de la mente se da de un modo natural, intrínseco, la intencionalidad del lenguaje es derivada. Esto es, mientras que la intencionalidad de los estados mentales es prioritaria respecto de otras formas de intencionalidad, la representacionalidad presente en el lenguaje es impuesta a los objetos sintácticos que la componen (Searle 1983: vii-viii). En este sentido, para Searle, no puede ser el lenguaje el que imponga la intencionalidad a la mente, sino que es la mente la que permite que el lenguaje sea intencional. Esto es, el significado de las palabras de los lenguajes naturales es algo que la mente transfiere a los elementos sintácticos que componen los diferentes lenguajes. Pero los estados mentales, en sí mismos, tampoco tienen significado. Los estados mentales, simplemente, son esencialmente intencionales (Searle 1983: 27). Incluso los niños y ciertos animales son capaces de tener cierto tipo de estados intencionales (Searle 1991a).

Al analizar la lógica intencional de los estados mentales, Searle está anticipando la ontología profunda de dichos estados. Desde su peculiar punto de vista, entonces, dado que existen sistemas objetivos como los cerebros (humanos) que son capaces de tener estados mentales subjetivos, entonces, es un hecho físico la existencia de sistemas que tienen características mentales. En este sentido, la correcta solución al problema mente-cuerpo no sería negar la existencia de la realidad de los fenómenos mentales, sino apreciar de manera adecuada su naturaleza biológica. Esto es, deberíamos considerar la intencionalidad (y también la conciencia) como rasgos constituyentes de la biología humana como la digestión o la circulación sanguínea (Searle 1983: viii-ix; Searle 1991b).

3.2.1.2. Acción, intencionalidad y percepción

Dentro de la imagen tradicional de la mente, considera Searle, los estados mentales como las percepciones parecen ser más simples que aquellos estados mentales que involucran algún tipo de contenido proposicional. A diferencia de los estados mentales, se insinúa, como las creencias o los deseos, cuyo contenido presupone una mediación proposicional, estados mentales como las percepciones deben ser más simples porque no requieren de dicha mediación. Se dan de un modo directo. En este sentido, el contenido de una experiencia visual no estaría mediado por ningún contenido representacional, por lo que su contenido sería de modo directo el objeto mismo percibido.

El contenido de una experiencia visual, por ejemplo, a diferencia del contenido de una actitud proposicional como, por ejemplo, una creencia, nunca puede ser equivalente a una proposición. Una experiencia visual siempre es tan sólo la experiencia de un objeto, siempre es aquello que permite dilucidar que esto o aquello es el caso (es decir, es el modo verificador de las condiciones finales de satisfacción de la realidad. Cf. Searle 1983: 40). Visto de este modo, según el esquema de Searle, las creencias y las percepciones comparten la misma dirección de ajuste con el mundo. Para ser correctas, ambas deben dirigirse al mundo tal y como es en realidad. Es decir, tienen una dirección de ajuste de los estados al mundo (como se ha

comentado anteriormente que ocurre en determinados actos de habla). Pero, a diferencia de las primeras, las segundas se dirigen de manera directa a él, mientras que éstas están mediadas por un contenido proposicional representacional.

Pero Searle también insiste en que aunque coinciden en tener la misma dirección de ajuste, las creencias y las percepciones difieren en un importantísimo aspecto. Las percepciones requieren siempre que aquello que se percibe sea realmente percibido. Es decir, a diferencia de las creencias, que no son sucesos del mundo, que no necesariamente necesitan dirigirse a sucesos que ocurran en el mundo para que un sujeto pueda creer en ellos, las percepciones requieren necesariamente de la existencia real de aquello que se percibe para poder ser verdaderas.

En este sentido, parece que Searle está dispuesto a aceptar una estrecha conexión causal entre el proceso de percibir y aquello que es percibido. Así, Searle considera como parte esencial de las condiciones de satisfacción de las percepciones que el acto de percibir sea realmente causado por el hecho de la existencia del objeto percibido. Si alguna de estas condiciones no se diera (esto es, si por ejemplo no existiera el objeto percibido o no se activara la percepción ante un objeto), entonces se truncarían las condiciones de satisfacción de la percepción concreta y entonces no habría percepción. O mejor, haría que ésta fracasara. El objeto percibido, por tanto, concluye Searle, debe ser la causa de la percepción. Si el objeto causa la percepción, entonces el contenido de nuestras percepciones dependerá de nuestra experiencia perceptiva. Porque, según insiste Searle, nosotros somos siempre conscientes de nuestras percepciones en tanto que referidas a representaciones internas.⁷

Así, para Searle, toda criatura que puede percibir tiene experiencias perceptivas conscientes que tienen un contenido intencional. Dicho contenido incluye ya las condiciones de satisfacción de la experiencia perceptiva concreta. Dichas condiciones nos dicen qué debe suceder para que una determinada experiencia cuente como una percepción. Una de estas condiciones, si no la más importante, nos dice que el objeto intencional existe y que éste causa la experiencia perceptiva. El sujeto, entonces, percibe el objeto, pero no la experiencia perceptiva misma. Con esta explicación, Searle pretende esquivar las acusaciones que podrían hacerse a una teoría de la percepción vinculada con *sense-data* (véase Colomina 2007 para una revisión de dichos problemas). En este sentido, Searle confiesa estar presentando una propuesta de la percepción cercana al realismo ingenuo en donde las cosas son del modo en

⁷ Una posición crítica podemos encontrarla, por ejemplo, en Armstrong (1991). Allí Armstrong nos dice que lo que percibimos es lo que permite diferenciar unas experiencias de otras, no la misma experiencia perceptiva, por lo que en el caso de diferenciar creencias y percepciones las diferencias deberán ser todavía más acentuadas porque involucran otra clase de percepciones más elaboradas; las actitudes proposicionales tienen por contenido objetos internos (entendidos como representaciones), por lo que no pueden ser entendidas del mismo modo que las percepciones, que se dirigen siempre a objetos situados en el exterior del sujeto. En este sentido nos dice: “no creo que haya ninguna propiedad fenoménica vinculada con el percibir o cualquier otro estado mental de esta clase... La única propiedad involucrada en la percepción es aquella involucrada en el contenido intencional de la percepción. Estas cualidades siempre son cualidades de objetos externos” (Armstrong 1991: 155). La defensa de Searle pasa por considerar errónea la propuesta de Armstrong por ser excesivamente reduccionista en su posición (Searle 1991c). Esta sería la posición de Searle ante el problema, planteado en el apartado 1.4.8., de la transparencia.

como nosotros las percibimos, ayudados para ello del contexto de percepción (como veremos en la siguiente sección).

Si Searle ha presentado la percepción como una de las formas biológicamente primitivas de la intencionalidad, la acción será la otra. En este sentido, dirá Searle, las acciones y las intenciones estarán íntimamente conectadas.

La acción también puede ser analizada, como hemos hecho con la percepción, mediante conceptos familiares a la teoría de los actos de habla. En una acción, podemos perfectamente identificar su dirección de ajuste. Para que exista una acción, ésta debe darse como resultado de la presencia de una intención previa. Cuando acontece una acción, nosotros tenemos una experiencia de la acción. Esta experiencia no es necesaria, dado que es posible actuar intencionalmente sin ser conscientes de ello (del mismo modo que podemos percibir sin ser conscientes de la experiencia misma de percibir), pero no podemos tener una acción sin tener una razón para actuar⁸ (del mismo modo que no podemos tener una percepción si no existe un objeto *interno* de percepción que motive la percepción).

Además, afirma Searle, las acciones y las percepciones tienen auto-referencialidad. Esto es, pueden ser identificadas por sus propias causas. Lo que percibimos cuenta como una percepción cuando el contenido intencional de la percepción nos dice que ésta está causada por las condiciones externas específicas que percibimos. Del mismo modo, una cierta acción contará como tal cuando el contenido intencional de la acción nos diga que la acción está causada por la intención misma. Esta peculiar característica de auto-referencialidad de las percepciones y las acciones las convierten, según Searle, en las formas primitivas de intencionalidad porque, nos dice,

“involucran a los organismos en una relación causal directa con el entorno del cual depende su propia supervivencia” (Searle 1983: 105).

Pero, ¿en qué sentido podemos decir que existe esta relación causal entre las mentes y los cuerpos de los sujetos?

No es un hecho baladí que al final del párrafo anterior remarcásemos la insistencia de Searle en considerar necesaria la presencia de un objeto interno en la mente para poder hablar de una percepción. En este sentido, el mayor inconveniente que parece estar mostrando la teoría de la percepción de Searle es considerar que la introspección puede actuar del mismo modo. Pero, además, también parece existir una marcada contradicción en el propio análisis que Searle realiza de la percepción. Veamos.

La teoría de la percepción de Searle (1983: 40-50) se caracteriza por tres elementos básicos. En primer lugar, la percepción incluiría experiencias perceptivas, esto es, hay estados mentales conscientes que tienen propiedades fenoménicas. En segundo lugar, las experiencias fenoménicas a veces parecen tener contenido intencional, pero no siempre. En tercer lugar, el contenido representacional de las experiencias perceptivas es (esencialmente) un reflexivo del

⁸ No debemos olvidar que Searle (2000) lleva el significativo título de *Razones para actuar*.

ejemplar, esto es, el estado de cosas representado es representado como causante de dicha experiencia perceptiva particular, en concreto. Todo esto implicaría, entonces, la tesis del realismo directo de Searle (una clara aceptación de la tesis de la transparencia, como indicamos anteriormente), la idea de que toda percepción nos indica de modo directo el estado de cosas existente en el mundo.

Pensemos un poco en los argumentos de Searle. Asumir el carácter reflexivo del ejemplar de la percepción supone aceptar también que las propiedades generales de la experiencia perceptiva estén involucradas en la determinación de sus propios contenidos intencionales, lo que nos abocaría a la dificultad de tener que rendir cuentas del modo en que se relaciona el contenido intencional específico de la percepción concreta con las propiedades específicas que permitirían determinarlo. Searle no es capaz de aportar una explicación satisfactoria. Simplemente muestra que eso sucede así, como un hecho intuitivamente cierto. Pero, al menos en este caso, las cosas no parecen funcionar. Porque dado que la teoría de Searle es de carácter descriptivo, fregeano, relacionada con elementos de primera persona, entonces ello implicaría que el sujeto de un estado perceptivo debería tener alguna forma de acceso cognitivo a las experiencias particulares y a su relación causal con los objetos percibidos. De nuevo, Searle no es capaz de aportar una explicación razonable a este hecho, más allá de volver a mostrar el carácter plausible de su posición. Pero no basta este tipo de solución. Necesitamos, además, alguna argumentación (que Searle no aporta nunca) que nos demuestre en qué sentido es viable su explicación.

Como hemos avanzado anteriormente, tal vez el elemento más preocupante de la posición de Searle es una afirmación marcadamente internista que le permite aplicar su teoría perceptiva, también, a la introspección de los propios estados mentales. Si su visión realista directa de la percepción le permite afirmar que nosotros vemos objetos materiales en tanto que disponemos de vivencias perceptivas de nuestras propias experiencias y somos capaces de aprehender la relación causal en las cuales están involucradas, entonces, tenemos un serio problema al pretender que, también de los objetos internos de nuestra mente, de aquellos elementos neurofisiológicos presentes en nuestro cerebro, tengamos un claro acceso consciente a partir de las propias vivencias que permiten mostrar su relación con la experiencia. El problema, de nuevo, aquí, vuelve a ser que Searle no aporta argumentos más allá de la mera afirmación de su plausibilidad. Ante todas estas incógnitas abiertas, Searle nos debe todavía una respuesta razonable.

3.2.1.3. Redes conceptuales y trasfondo en los estados mentales

Como hemos comentado al comienzo de este apartado, la filosofía de la mente de Searle no puede entenderse sin la noción de Trasfondo. Las afirmaciones del propio Searle otorgan un papel crucial a dicho concepto a la hora de explicar la relación entre los elementos de lo mental. Para Searle, el Trasfondo es aquello que permite la interpretación lingüística, es aquello que permite estructurar nuestra conciencia y nuestra experiencia y es aquello que nos predispone a realizar determinados tipos de conducta (Searle 1995: 132-137). En pocas palabras, el

Trasfondo “permite que toda representación tenga lugar” (Searle 1983: 143). Dedicamos esta sección a analizar la noción searleana de Trasfondo.

Del mismo modo que los actos de habla no se dan de un modo aislado, tampoco los estados mentales tienen una naturaleza atómica. Por ese mismo motivo, insiste Searle, no podemos expresar ningún tipo de deseo o creencia sin tener una red conceptual que permita confrontar la naturaleza interna de los estados mentales con el contexto externo hacia el cual se dirigen. La red conceptual que configura el entorno, por tanto, estará compuesta de un número indefinido de estados intencionales que permitirán individuar cada estado mental que un sujeto sea capaz de tener y le permitirá poder dotarlo de sentido. Sólo cuando estos estados intencionales acontecen podemos determinar las condiciones de satisfacción de los estados mentales.

Pero además de las redes conceptuales, en el contexto también existe aquello que Searle denomina el Trasfondo. En Searle (1979b), nuestro autor presenta un ejemplo (anteriormente empleado por Austin) en el cual la oración ‘hay un gato sobre la alfombra, pero yo no lo creo’ puede carecer de significado en tanto que, dice, la expresión ‘hay un gato sobre la alfombra’ es una oración emitida libre de todo contexto (Searle 1979b: 120-121). Pero cuando creemos que el gato está sobre la alfombra, nos dice Searle, asumimos que no tenemos por qué encontrarnos en el lugar en el cual sería fácil comprobar si realmente existe o no un gato posado sobre una alfombra, pero siempre presuponemos que (dado un contexto determinado de uso de la oración) el gato estará sobre la alfombra. Este tipo de emisiones dependientes de contexto conforman un conjunto de emisiones que, por definición, poseen intencionalidad (Searle 1979b: 127). Pues bien, nos dice Searle, este conjunto de emisiones con carga intencional es el Trasfondo sobre el cual se proyecta nuestro lenguaje y toda nuestra vida mental.

Pero entonces surge un pequeño problema con el trasfondo. Como hemos comentado, el trasfondo parece ser pre-intencional, como algo que está ya presente en todo contexto de emisión, por lo que parece difícil indicar el modo en el cual podemos decir que todos nosotros entramos en contacto con él. Searle apuesta claramente por una explicación de este contacto en términos de *know-how*. Esto es, para Searle, nosotros podemos comportarnos (y hablar) como si las mesas fueran sólidas porque creemos que son sólidas (porque sabemos qué es una mesa), pero la solidez de una mesa debe ser contemplada como influyendo de modo directo en nuestro comportamiento cotidiano con las mesas. (¡Imaginemos tan sólo por un momento cómo sería mi mundo si creyera que las mesas no son sólidas!). En este sentido, nosotros tenemos conocimiento de cómo son las mesas, no necesitamos una creencia adicional de que las mesas son sólidas. Así, para Searle,

“el Trasfondo es un conjunto de capacidades mentales no-representacionales que engloban todas las representaciones que tenemos... Debemos saber cómo son las cosas y debemos saber cómo comportarnos ante ellas, pero la clase de saber-cómo involucrada aquí no es, en estos casos, una forma de saber-qué” (Searle 1983: 143).

Por tanto, la influencia del Trasfondo es pre-intencional en tanto que nos indica que nuestro saber-cómo depende de un modo primario de nuestras reacciones habituales ante las diferentes sensaciones y experiencias externas, un saber que es previo a toda creencia y a todo elemento consciente pero que, aún así, sigue siendo normativo. (Una noción similar podemos encontrarla en las reflexiones wittgenstenianas presentes en *Sobre la certeza*).

Hablamos del Trasfondo como contexto en un sentido metafísico. Es algo que está con nosotros en todos los aspectos de nuestra vida cotidiana. Pero además, nos dice Searle, el trasfondo permite que nosotros, seres conscientes, dispongamos de toda una serie adicional de habilidades y capacidades cognitivas que van más allá de aquellas primarias. Dado que es, en último término, el conjunto de todas nuestras reacciones naturales respecto del mundo lo que permite conformar y ampliar dicho trasfondo. En este sentido, aunque suene paradójico, Searle también puede decir que el trasfondo (bueno, los distintos tipos diferentes de trasfondo) está en nuestra cabeza (Searle 1991d: 291). Y lejos de llevarnos a un solipsismo, nos advierte Searle, esta concepción del trasfondo nos permite compartir el trasfondo concreto que cada uno tiene en su propio cerebro a través de las redes conceptuales tejidas y puestas en común para permitir la comunicación y la acción social conjunta (Searle 1995).

Vemos que la noción de Trasfondo enunciada inicialmente por Searle adquiere un papel crucial en su teoría de la mente porque permite mantener, por un lado, sin necesidad de postular la mente como un almacén donde se amontonan todas las creencias subyacentes, una noción de mente como intencional a partir de elementos no-representacionales capaces de aflorar en la conciencia y permitir, por otro lado, la construcción de significaciones representacionales dotadas ellas mismas de intencionalidad intrínseca. Pero esta noción inicial de Trasfondo necesariamente será modificada por la introducción en Searle (1992) del Principio de Conexión (Hermoso y Chacón 2000).

El Principio de Conexión⁹ realiza una presuposición ontológica crucial en la filosofía de la mente de Searle, dado que asegura que

“la adscripción a un sistema de un fenómeno intencional inconsciente implica que dicho fenómeno es, en principio, accesible a la conciencia” (Searle 1990b: 586).

Una vez introducido el Principio de Conexión, el problema de la diferenciación entre los elementos intencionales del Trasfondo se convierte en un verdadero obstáculo insalvable dentro de la formulación original de Trasfondo. Por tanto, si Searle quiere seguir manteniéndolo no tiene más remedio que modificarlo para dar cabida a esta noción de inconsciente. En este sentido, como nos dice Searle, si lo que sucede en el cerebro sólo puede ser o bien conciencia o bien neurofisiología, entonces, al haber negado la posibilidad de la existencia de elementos inconscientes, la red conceptual que engloba holísticamente toda nuestra vida mental no puede más que ser un conglomerado de elementos neurofisiológicos. Según lo expuesto hasta el

⁹ Analizamos en detalle el Principio de Conexión en la sección 3.2.2.3.

momento, de acuerdo con Searle, toda esta estructura neuronal que configuraría nuestro cerebro según la red conceptual, debe formar parte del Trasfondo, en tanto que

“la ontología ocurren de aquellas partes de la Red que son inconscientes es la de una capacidad neurofisiológica, pero el Trasfondo consta enteramente de tales capacidades” (Searle 1992: 194).

En este sentido, el Trasfondo es modificado para asumir, de una forma mucho más débil, una diferenciación sutil entre toda aquella serie de elementos que manifiestamente afloran a un nivel consciente y aquellos elementos neurofisiológicos que, de una manera latente, permanecen ocultos a la conciencia a la espera de que en algún momento consigan elevarse hasta el nivel consciente.

El problema principal con el que aquí se encuentra ahora Searle es con la necesidad de rendir cuentas de la noción de irreductibilidad. Recordemos que para Searle no sólo la mente y la conciencia son irreductibles a meros elementos neurofisiológicos, sino que también es irreductible la noción de intencionalidad. En este sentido, comenta Searle,

“ninguna descripción de hechos objetivos, fisiológicos, de tercera persona, transmitiría el carácter subjetivo, de primera persona, del dolor simplemente porque los rasgos de primera persona son diferentes de los rasgos de tercera persona” (Searle 1992: 127).

Y de la intencionalidad, nos dice Searle, que

“un rasgo aspectual no puede ser exhaustiva o completamente caracterizado solamente en términos de predicados de tercera persona, conductuales o, ni siquiera, neurofisiológicos. Ninguno de ellos es suficiente para proporcionar una caracterización exhaustiva del rasgo aspectual” (Searle 1990b: 587).

Visto desde este punto de vista, Searle parece querer mantener, al mismo tiempo, una noción de inconsciente mucho más débil que la sostenida por la visión ingenua y equivocada de las ciencias cognitivas, pero al mismo tiempo pretende querer mantener la separación conceptual existente entre la perspectiva subjetiva del conocimiento de primera persona y la perspectiva objetiva del conocimiento en tercera persona. Esto es, quiere mantener una separación epistémica sin que ello tenga claras repercusiones en la ontología profunda de los hechos mentales y neurofisiológicos. Pero, como pretendemos mostrar en secciones posteriores, esto es imposible. Porque la irreductibilidad de la conciencia requeriría, también, el reconocimiento de una ontología de lo mental capaz de acomodar también la irreductibilidad de la intencionalidad. Es decir, que la explicación que desde aspectos físicos podamos realizar de la existencia de la mente pueda incorporar los elementos irreductibles que Searle presupone en la intencionalidad. Y esto es algo que Searle parece incapaz de cumplir, por ser una contradicción en los términos.

3.2.1.4. Conclusión: implicaciones del trasfondo y las redes conceptuales en la vida mental de los individuos conscientes

Una primera cuestión que llama la atención de la teoría del trasfondo de Searle es que el punto de vista clásico de cómo las causas y los efectos funcionan en el mundo presupone un principio general de regularidad. Esto es, tenemos la imagen común y habitual de que las cosas tienden a seguir comportándose del mismo modo a como habitualmente se comportan. Por ejemplo, si dejamos caer un lápiz al suelo, éste caerá al suelo impulsado por la fuerza de la gravedad (que sepamos aquí que es a partir de la fuerza de la gravedad como podemos explicar la caída del lápiz al suelo no es importante ahora) y siempre que dejemos caer un lápiz (o cualquier otro objeto) al suelo esperamos que llegue hasta el suelo. (¡Imaginemos por un solo instante cuál sería nuestra sorpresa si, en lugar de estrellarse contra el suelo, el lápiz que dejamos caer permaneciera inmóvil delante de nosotros a la misma altura desde la cual lo habíamos soltado!). Pero, como dice Searle, si investigamos las relaciones causales entre estos dos fenómenos (entre la causa y el efecto), entonces observamos que no es necesario tener ningún tipo de creencia acerca de la regularidad de los hechos del mundo. Es decir, por supuesto todos sabemos que si en un caso como el anterior dejo caer un lápiz al suelo, éste llegará al suelo. Pero no necesito tener la creencia de que eso será así. Entonces, nuestra presuposición de que las cosas del mundo tienen una cierta regularidad causal no requiere de creencias, sino que forma parte de nuestro trasfondo (Searle 1983: 133).

Searle también afirma sin tapujos que su posición es realista, como hemos indicado anteriormente. En este sentido, cuando afirma que el trasfondo engloba todos los hechos intencionales que nosotros mostramos cuando interactuamos con las cosas del mundo, lo que Searle pretende afirmar es que las cosas en el mundo realmente suceden así (Searle 1983: 159). El problema principal es que Searle no es capaz de probar adecuadamente su posición. Porque si nuestras creencias en que las cosas son del modo en que creemos que son pero que estas creencias no residen sobre argumentos que permitan justificar nuestras posiciones y nuestras respuestas, entonces parece difícil poder aportar ninguna prueba razonable para justificar que el mundo realmente sea como se dice que es.

Además, que exista un trasfondo tal que nos diga en todo momento qué tipo concreto de creencia podemos tener o no, incorpora un elevado grado de normatividad. Es decir, el trasfondo necesariamente nos inclinará a tener unas creencias en lugar de otras porque, en situaciones normales, los individuos conscientes pretendemos guiarnos lo mejor posible para alcanzar aquellos objetivos que consideramos prioritarios. Por lo que si el trasfondo nos indica aquello que es pertinente o relevante (aunque no consiga justificarlo), es sumamente importante a la hora de rendir cuentas de la posesión de nuestras creencias. Así, si por ejemplo deseo beber una cerveza y creo que en la nevera hay alguna (y el trasfondo se encarga de asegurar que el suelo desde el sofá donde estoy sentado es sólido, que puedo doblar mis rodillas sin romperlas, que cuando tire de la agarradera del frigorífico ésta no cederá y podré abrirla, que las botellas de cristal no se funden, que la cerveza contenida en la botella tiene un estado líquido, que los vasos no suelen tener agujeros... y un larguísimo etcétera de presuposiciones similares), entonces me

levantaré del sofá e iré a coger una cerveza de la nevera porque he sido capaz de representarme la situación guiado por mis deseos y mis creencias (tanto las conscientes como las inconscientes debidas al Trasfondo). Y este tipo de decisiones o actividades, por supuesto, tendrán lugar a partir de una gran serie de conexiones neuronales establecidas en nuestro cerebro y no por la interiorización de una gran cantidad de memorizaciones previas (Searle 1983: 150). Las cosas en este tipo de casos, y en muchos otros (pensemos tan sólo en el modo en que muchos de nosotros somos capaces de conducir o realizar ciertas actividades cotidianas) parecen funcionar de un modo (casi) automático.

3.2.2. La conciencia como propiedad emergente del cerebro

Como hemos mostrado en el anterior apartado, Searle considera que la intencionalidad que podemos encontrar en el lenguaje es derivada de la intencionalidad que está presente en la mente. En este sentido, los humanos imponemos la intencionalidad que naturalmente encontramos en nuestra mente sobre el lenguaje que hablamos. Así, vemos que el desarrollo de la noción de intencionalidad (y de las nociones afines de trasfondo y red conceptual) permite completar la imagen de Searle sobre el lenguaje, pero también la imagen general que Searle tiene de la mente, como veremos a continuación.

Tal y como Searle la concibe, la mente tiene dos características principales. En primer lugar, es sumamente importante el papel de la conciencia a la hora de estudiar la mente. Los fenómenos mentales conscientes, como el dolor y otros sentimientos asociados a nuestras capacidades cognitivas, deben ser tenidos en cuenta a la hora de estudiar científicamente la mente (Searle 1989: 193-194). Por otro lado, en segundo lugar, no podemos obviar que los estados mentales tienen una irreducible ontología subjetiva (Searle 1992: 19). Es un hecho objetivo que los seres conscientes tenemos experiencias subjetivas, y que dichas experiencias tienen un carácter en primera persona que permiten que tan sólo el individuo concreto que está teniendo dichas experiencias pueda conocerlas de modo directo (punto de vista epistémico).

Contra todos aquellos, como los conductistas lógicos, que consideran que la mente es tan sólo una manera de definir cierto tipo de acceso, que es una mera cuestión de análisis lógico, que las oraciones acerca de la mente pueden ser traducidas a oraciones que involucren tan sólo comportamientos; frente a aquellos que, como los eliminacionistas, reivindican que los estados mentales son idénticos a estados cerebrales, ignorando la realidad de lo mental; contra aquellos que, como los funcionalistas, creen poder traducir las oraciones que contienen términos mentales a meras categorías funcionales; y frente a aquellos que, como los seguidores de la Inteligencia Artificial fuerte, consideran que la mente funciona como un *software* implementado en un ordenador; Searle reivindica la realidad de lo mental, y el papel central de la conciencia, como marcadamente irreductible a un lenguaje fisicalista.

3.2.2.1. El compromiso aristotélico

Lo que parece estar ofreciendo Searle en la presentación de su posición como un naturalismo biológico es una reformulación de la clásica postura aristotélica sobre el problema

mente-cuerpo (Code 1991). Aunque, por supuesto, no fuera formulado de esta forma, para Aristóteles el ser humano se componía tanto de forma como de materia, pero ambos tenían una importante carga causal el uno sobre el otro. En pocas palabras, para Aristóteles, lo psicológico formaba parte de lo físico. Como hemos visto en su polémica contra los teóricos de la Inteligencia Artificial fuerte, frente a éstos que defenderían una primacía de la forma frente a la materia, Searle se inclinaría hacia una supremacía de la materia frente a la forma. En este sentido, para Searle, el problema radicaría en que el computacionalismo estaría reivindicando que los estados mentales pueden ser englobados bajo un nivel meramente formal demasiado elevado como para ser fundamental. La conciencia existiría en un nivel funcional mucho más bajo, a un nivel ontológico primario, que permite que los seres dotados con cerebro humano sean capaces de mostrar una intencionalidad y una conciencia. En pocas palabras, que lo psicológico forma parte de lo físico, pero de algún modo es diferente e irreducible a lo segundo

Pero, por supuesto, también difieren en muchos aspectos. En las siguientes secciones intentaré exponer lo más sucintamente posible los elementos más relevantes de la teoría de la mente de Searle y los principales puntos en los que se apoya su análisis de la conciencia con el fin de mostrar las posibles incongruencias que se presentan dentro de su misma teoría y que, en última instancia, nos obligarán a abandonarla por incoherente (Corcovan 2001: 310, 321), como ya hemos indicado.

3.2.2.2. Las características estructurales de la conciencia

El énfasis de Searle se centra, como estamos viendo, en la idea de que tanto la conciencia como la intencionalidad pueden ser mostradas como fenómenos biológicos, que son meras propiedades físicas del cerebro (Searle 1992: 13). En este sentido, las propiedades mentales no necesitan ser reducidas a propiedades físicas. Son ya, en sí mismas, un tipo de propiedades físicas. En lugar de pensar, por tanto, lo mental como no-físico, Searle estaría sosteniendo que lo mental es un subtipo de lo físico (Searle 1992: 14).

Para Searle, la conciencia puede presentarse de múltiples modos, pero todos ellos comparten una docena de rasgos estructurales (Searle 1992: 136-150; textos 1 y 3 de Searle 2002; Searle 2004: 93-101).

Dentro del sistema ideado por Searle, como comentábamos, existe toda una serie finita de modalidades que se caracterizan principalmente por su ser cualitativas. Es decir, las modalidades conscientes tienen por rasgo principal presentar de un modo particular característico aquello que es su contenido. Esto, entonces, nos indicará ya las dos primeras características estructurales de la conciencia. Por un lado, todo estado consciente se caracteriza por ser una vivencia que presenta cierto tipo peculiar de cualidad al sujeto (Searle 1992: 137-138; Searle 2004: 93-94). Por otro lado, dicha vivencia sólo puede ser presentada al individuo perceptor de un modo subjetivo, es decir, se le presentará sólo a él de un modo concreto (Searle 1992: 140-141; Searle 2004: 94-95). Además, dicha vivencia tendrá por contenido un determinado objeto externo. Es decir, la mayoría de las veces (pero no todas, pensemos tan sólo en el hecho del estado de ansiedad) nuestro estado mental se dirige directamente sobre un

determinado objeto exterior al sujeto (Searle 1992: 139-140; Searle 2004: 96-97). Y dicha vivencia relativa al objeto exterior se presentará de modo unificado. Esto es, nos permite representarnos el objeto externo como un todo sin necesidad de ser particularmente conscientes de todas sus partes (a menos que prestemos atención especial sobre cada una de ellas), respecto de su tamaño, forma, color, etc., a través del tiempo (Searle 1992: 139; Searle 2004: 95-96).

La posibilidad de que nosotros contemplemos los objetos externos, nos dice Searle, es sobre todo debido a que percibimos los objetos como figuras situadas sobre un trasfondo, lo que demuestra la estructura gestáltica (figurativa) de la conciencia (Searle 1992: 141-142; Searle 2004: 100). En este sentido, nuestras percepciones siempre están estructuradas de tal modo que permiten al sujeto atribuirles una forma familiar, lo que indica no sólo la intencionalidad de la conciencia, sino también su carácter estructural (Searle 1992: 142-146). Al estar situadas frente a un trasfondo, las representaciones también cumplen una condición de desbordamiento. Esto es, parece ser el caso que el contenido inmediato de una determinada experiencia fenoménica tiende a rebosar, es decir, a solaparse con otros pensamientos y con otras experiencias (Searle 1992: 146). En este sentido, parece que nuestras experiencias fenoménicas no tendrían sentido sin presuponer su carácter holístico (como muestra el hecho de que Searle mismo apele al ejemplo de Quine de la indeterminación de la traducción para defender esta idea. Cf. Searle 1992: 171).¹⁰

Dentro del campo de la conciencia, debemos diferenciar entre aquellos objetos que son el centro de nuestra atención y aquellos objetos que pertenecen a la periferia de nuestra atención. La posibilidad de realizar esta diferenciación también es un rasgo característico de lo consciente (Searle 1992: 146-148; Searle 2004: 98). Ello parece indicarnos que todo estado de conciencia puede ser localizado dentro de unos determinados límites, no sólo de atención, sino también geográficos y temporales, en tanto que le permite al sujeto saber dónde y qué experiencia concreta está experimentando cada vez (Searle 1992: 148; Searle 2004: 98-99).

Anteriormente hemos mencionado que existen determinadas experiencias que, a pesar de no ser intencionales, sí son conscientes, en el sentido en que interfieren en todo nuestro aparato mental de algún modo determinado al darnos cuenta de su presencia. En este sentido, existen determinados estados de ánimo (como la ansiedad o el decaimiento) que interfieren en nuestra vida consciente porque somos capaces de vernos en dichas situaciones (Searle 1992: 149-150; Searle 2004: 97-98).

A menudo, los estados mentales conscientes tienen una dimensión evaluativa, que se mueve dentro del ámbito aristotélico del placer y del displacer. En este sentido, la experiencia consciente puede ser evaluada según la cantidad de placer o displacer que proporcione (Searle 1992: 98; Searle 2004: 150).

Pero, además, existen estados que influyen de manera activa en la conciencia mientras que existen, también, estados mentales que influyen de modo pasivo en la conciencia (Searle 2004: 99). Así, nos dirá Searle,

¹⁰ La apelación a este argumento quineano también servirá a Searle como contra-argumento a una posible crítica que puede desligarse de un argumento de Block contra su Principio de Conexión. Véase sección 3.2.2.3.

“sólo un ser que puede tener estados intencionales conscientes puede tener estados intencionales, y todo estado intencional inconsciente es, al menos, potencialmente consciente” (Searle 1992: 141).

En este sentido, nos dice Searle, podemos ver una necesaria conexión entre la conciencia y el inconsciente.¹¹ Lo primero sobre lo que debemos llamar la atención, entonces, en el análisis de los rasgos característicos de la conciencia de Searle, es que él siempre habla de una serie de conexiones, no de una sola conexión.

Vista desde este punto de vista, la concepción de la conciencia en Searle nos presenta como seres plenamente reflexivos. Nosotros no podemos sólo tener experiencias como experimentadas subjetivamente, sino que siempre deberemos prestar atención a dichas experiencias para poder darnos cuenta de que las tenemos (Searle 1992: 151-152). Pero, según nos dirá Searle mismo, que esto sea de este modo no quiere decir que requiramos siempre de la introspección para ser conscientes. No necesitamos tener ninguna habilidad especial dentro de la mente para poder ser conscientes (Searle 1992: 153). La mente, por sí misma, es ya reflexiva, por lo que de modo automático presta atención a todo aquello de lo que somos conscientes. Pero, podríamos reclamarle a Searle, ¿cómo puede ser que la mente siempre requiera de esa auto-reflexión para poder ser conscientes de nuestras propias experiencias? En otras palabras, ¿cómo se relaciona la conciencia con todas aquellas experiencias que tan sólo son potencialmente conscientes? ¿Cómo se convierte en consciente una experiencia?

3.2.2.3. El principio de conexión y la ciencia del inconsciente

La primera presentación del principio de conexión aparece en Searle (1989). Nuestro autor indica allí que existe una conexión entre la conciencia y la intencionalidad, afirmando que no pueden existir seres que sean intencionales sin que sean conscientes. Existe una cierta conexión necesaria entre los estados mentales conscientes y la intencionalidad:

“Sólo un ser que sea capaz de tener estados intencionales conscientes puede tener estados intencionales después de todo, y todo estado intencional inconsciente es potencialmente consciente” (Searle 1992: 132).

En este sentido, dentro del sistema de Searle, un estado intencional inconsciente sólo puede ser tal si y sólo si está conectado de algún modo a estados intencionales conscientes.

Para Searle, la ciencia cognitiva siempre ha considerado el inconsciente como poco problemático, porque precisamente siempre ha sido objeto de estudio dentro de dicha disciplina (Searle 1992: 198). La importancia del inconsciente se debe, según Searle, a la gran cantidad de datos que parecen afirmar la presencia de elementos inconscientes en nuestra vida mental. En este sentido, Searle partirá de una afirmación un tanto controvertida:

“la noción de estado mental inconsciente implica accesibilidad a la conciencia. No tenemos noción alguna de inconsciente excepto como aquello que es potencialmente consciente” (Searle 1992: 160).

¹¹ A analizar el Principio de Conexión entre el inconsciente y la conciencia dedicaremos la sección 3.2.2.3.

Así, todo elemento inconsciente es exactamente igual a un elemento consciente pero carente de conciencia. (De nuevo, Searle parece ser poco explicativo a la hora de querer definir el inconsciente, puesto que presupone la noción de conciencia sin haberla explicado previamente).

Los dos rasgos característicos, según Searle, de todo estado inconsciente serían, por un lado, que debe ser tratado por completo como un estado intencional que se da en la mente del sujeto y, por otro, que sus condiciones de satisfacción dependen de un determinado rasgo aspectual. Searle apela a un ejemplo simple para ejemplificar qué pretende decir con estas palabras. Pensemos en mi creencia de que la Torre Eiffel está en París. Esta creencia no está siempre incluida a un nivel consciente. Pero cuando no estoy pensando en ella, tampoco quiere decir que deje de tener dicha creencia. La creencia de que la Torre Eiffel está en París permanece en un nivel inconsciente hasta que aflora a la conciencia. En este sentido, la creencia es genuinamente intencional, es decir, se diferencia claramente de cualquier fenómeno que muestra intencionalidad derivada, en tanto que es un elemento mental. Pero, además, dicha creencia tan sólo es verdadera si se tiene en cuenta sus condiciones de satisfacción como representadas bajo determinados aspectos pero no bajo otros. Es decir, sólo se cumplen sus condiciones de verdad bajo una determinada relación aspectual entre la creencia y el hecho concreto del mundo representado (Searle 1992: 163).

Teniendo como base estos dos rasgos de los estados mentales inconscientes, Searle apela a un argumento, con siete pasos básicos, para defender un Principio de Conexión entre la conciencia y lo inconsciente (Searle 1992: 164-169).

1. Hay una distinción entre la intencionalidad intrínseca (propia de la naturaleza de los estados mentales) y la intencionalidad derivada o como-si (propia de la naturaleza del lenguaje), como dejó clara la distinción presentada en la sección 3.2.1.2. Sólo la intencionalidad intrínseca puede ser considerada como genuina intencionalidad por ser ontológicamente propia de los estados mentales.

2. Los estados mentales inconscientes son intrínsecos porque tienen rasgos aspectuales. Es decir, sus condiciones de satisfacción dependen de que los aspectos relevantes representados en el estado se cumplan.

3. Todo estado intencional intrínseco tiene un rasgo aspectual.

4. Este rasgo aspectual no puede caracterizarse en términos físicos porque presupone siempre una perspectiva subjetiva, dada en primera persona, irreducible a la perspectiva científica en tercera persona.

5. De todos modos, a pesar de lo dicho en el punto anterior, todo estado intencional inconsciente, en tanto inconsciente, consiste por completo en una serie de fenómenos neurofisiológicos.

6. En este sentido, todo estado intencional inconsciente es la noción de un estado que es posible como experiencia consciente futura.

7. La ontología del inconsciente, entonces, consiste en toda una serie de rasgos objetivos del cerebro (explicables desde sus elementos neurofisiológicos) que son, sin embargo, capaces de causar pensamientos subjetivos conscientes ontológicamente irreducibles.

Si el Principio de Conexión de Searle es correcto, por un lado, se muestra la gran importancia de la distinción existente entre intencionalidad intrínseca e intencionalidad derivada. Muchos estados mentales exhiben intencionalidad intrínseca, es su naturaleza. Por otra parte, los estados que exhiben una intencionalidad derivada, se muestran como si fueran intencionales, pero realmente no lo son. Por tanto, el estatuto de ambos estados será siempre diferente. Una intencionalidad derivada puede ser exhibida por casi cualquier cosa, pero una intencionalidad intrínseca sólo puede ser exhibida por estados mentales intencionales. En este sentido, la intencionalidad intrínseca sólo puede ser aplicada a la mente humana. Una característica importante de la intencionalidad intrínseca es que puede ser compartida por varios aspectos de la mente.

En este sentido, la intencionalidad intrínseca también puede ser aplicada a los pensamientos inconscientes en tanto que potencialmente conscientes. Nuestros pensamientos inconscientes no son diferentes de aquellos conscientes. Así, Searle afirma que

“la característica aspectual de la mente no puede ser exclusivamente caracterizada en términos de tercera persona o a partir de predicados neurológicos. Ninguno de ellos será suficiente para agotar dicha característica subjetiva” (Searle 1989: 199; Searle 1992: 157-158).

Pero precisamente este carácter (supuestamente) paradójico del inconsciente (la supuesta incongruencia entre la dependencia de elementos objetivos neurofisiológicos y el carácter irreducible de la ontología de la conciencia) es el que hace al propio Searle tomar precauciones e intentar avanzar a las posibles críticas. Él mismo avanza dos posibles argumentaciones contrarias.

La primera de ellas, dependiente de argumentos funcionalistas como los de Armstrong, parte de la suposición de la existencia de individuos con una distribución funcional adecuada para la percepción de determinadas experiencias pero que, sin embargo, no son capaces de llevarlas a un nivel personal consciente. Los casos más claros son los casos de visión ciega, explicados por Weiskrantz (1986), tal y como mostramos en el apartado 1.4.12. En estos casos, nos dice Searle, el sujeto es capaz de tener la experiencia de nivel personal aunque falle la estructura neurofisiológica porque, y es lo importante, que se pueda producir la experiencia consciente indica que existe a un nivel inconsciente una redefinición del fenómeno mental concreto que permite la aparición de la experiencia a un nivel consciente en tanto que es capaz de reconocer el rasgo aspectual adecuado (Searle 1992: 170). Los casos de visión ciega, como mucho, permiten establecer la existencia de un agujero en la experiencia consciente en tanto que tiene un déficit patológico a nivel neurofisiológico. En estos casos, el contenido intencional es del tipo consciente porque puede implementarse a nivel personal, pero directamente se da a un nivel de percepción inconsciente, aunque existiendo un vacío a nivel consciente (Searle 1995b: 223).

Analicemos detenidamente estos casos. En los casos de visión ciega, como expusimos en la introducción, lo que parece darse son casos en los que el comportamiento intencional de los pacientes de dicha enfermedad está marcadamente influido por la detección de estímulos para los cuales dicho sujeto no está cognitivamente capacitado por tener dañada la zona del sistema nervioso o del cerebro encargada de dichas percepciones. En este sentido, la conducta de los pacientes con visión ciega parece estar causada por un estado intencional cuyo contenido es inaccesible a su conciencia, a pesar de que dichos estados son causalmente eficaces a la hora de producir las acciones intencionales del sujeto-paciente. Si lo que Searle está intentando decir es que los contenidos de los estados perceptivos del paciente son potencialmente conscientes del mismo modo que lo serían los contenidos de los estados perceptivos de un sujeto perceptor normal, entonces, al faltar el elemento neurológico pertinente que permite identificar a un nivel inconsciente la presencia de una determinada percepción, el Principio de Conexión se vuelve totalmente vacuo al no poder precisar el estado real del paciente que es capaz de detectar a un nivel subliminal inconsciente el estado concreto del mundo hacia el cual se dirige, por lo que faltaría el elemento intencional básico requerido a un nivel inconsciente que permitiría la existencia de un contenido intencional a un nivel consciente. La razón principal por la que esto podría ocurrir, suponemos, es que el comportamiento causado por el estado perceptivo neurológicamente establecido que permite afirmar una intencionalidad en el contenido de la percepción concreta tanto en el paciente como en el sujeto perceptor normal no depende tanto de una relación causal como de una relación semántica entre las propiedades concretas del mundo (en tanto relevantes para el sujeto) y la historia cognitiva propia del sujeto perceptor (Jacob 1995: 198). Pero la misma teoría de Searle parece contradecir esta posibilidad. Pensémoslo tan sólo por un instante. Si lo que se pretende es que Searle identifique como potencialmente conscientes los estados perceptivos del sujeto con visión ciega, entonces, para que esto ocurriera, el paciente primero debería curarse, en tanto que la potencialidad tan sólo tiene sentido en tanto que es capaz de llegar al acto de conciencia.

En este sentido, otra importante crítica que puede hacerse al planteamiento inicial de Searle, consistiría en la posibilidad de la existencia de duplicados físicos sin experiencias conscientes (Cf. Block 1978). Nos dice Searle, la posibilidad de la existencia de individuos carentes por completo de elementos intencionales (los duplicados físicos o *zombies* que introdujimos en el apartado 1.4.5.) no es un problema porque su existencia no es posible. Una prueba de ello nos la proporciona, de modo indirecto, el ejemplo de Quine de la indeterminación de la traducción (Quine 1960: cap. 2). Según este argumento, no existen hechos objetivos que especifiquen los rasgos aspectuales de una experiencia determinada, por lo que no es posible la existencia de un sujeto perceptor (pensemos en un *zombie*) que cuando se comporta como-si fuese consciente no fuera realmente consciente, dado que toda intencionalidad mental (por ser intrínseca) requiere de la presencia de elementos aspectuales en el sentido anteriormente expuesto. Si esto es así, entonces, los estados intencionales (los nuestros y los de los *zombies*) tendrían también un cierto contenido intencional determinado por dichos rasgos aspectuales especificados por la vida subjetiva de cada individuo y claramente especificados a un nivel

nerológico, negando así la posibilidad de la existencia de *zombies* y estableciendo la necesaria conexión entre el nivel neuronal y el nivel consciente.

Pero, podríamos decir, este argumento de Searle parece contradictorio. (Gulick 1995: 204:206). Si, como Searle pretende, debemos presuponer la existencia de una conexión necesaria entre la conciencia y la intencionalidad contra todas aquellas corrientes filosóficas que han pretendido una clara separación entre ellas, no parece plausible defender la posibilidad, al menos si se quiere evitar caer en una inconsistencia lógica, de la existencia de estados mentales inconscientes que, al mismo tiempo, son intencionales. Porque la existencia de estados neurofisiológicos intencionales sólo es posible si separamos la conciencia de la intencionalidad, en tanto que la existencia de estados neurológicos pertenecientes a la estructura biológica de los sujetos implica afirmar la imposibilidad de que estos sean mentales porque son claramente contrarios a su especificación en primera persona, por carecer de cualquier posibilidad de ser conocidos mediante la mera introspección (por ejemplo, González-Castán 1999: 87-88. Véase también González-Castán 1992). Pero conocemos de la existencia de estados intencionales inconscientes que no siempre llegan a establecerse a un nivel consciente. Si esto es posible, por un lado, de manera independiente a que exista una necesidad intrínseca entre la conciencia y su contenido, se muestra la existencia de un obstáculo claro para defender el Principio de Conexión de Searle al existir la posibilidad de una desvinculación entre elementos intencionales y elementos de conciencia. Por otro lado, parece plausible la idea (al menos en su posibilidad lógica) de la existencia de individuos funcionalmente idénticos a seres conscientes que, sin embargo, carezcan de toda experiencia fenoménica en tanto serían capaces de estar dotados de estados inconscientes que, a pesar de que nunca llegarán a establecerse ni a aflorar a un nivel personal, estarían dotados de intencionalidad. Y aunque se insista, como hace Searle, en que no se requiere que los estados lleguen a actualizarse como conscientes, sino que tan sólo nos basta con la posibilidad de que puedan llegar a serlo, el punto crucial aquí, como nos dicen Fodor and Lepore (1994), es llegar a poder establecer si se puede o no llegar a estipular de manera cierta si dichos estados inconscientes son capaces de llegar a aflorar a un nivel personal. Visto de este modo, la posibilidad apuntada por Searle de que los estados inconscientes puedan llegar a aflorar a la conciencia tan sólo sería relevante desde un punto de vista modal. En este sentido, y a pesar de las defensas aportadas por la argumentación de Searle (1994), donde simplemente vuelve a repetir la necesidad de dicho Principio de Conexión apelando a la vinculación entre conciencia y contenido, no parece posible que este tipo de recurso naturalizador permita, por un lado, aportar una respuesta clara a la aparente vinculación entre conciencia e intencionalidad y, todavía más importante a la hora de querer defender una solución materialista a la problemática de la conciencia, no parece una respuesta razonable al problema de los duplicados físicos. Por eso mismo, consideramos necesario afrontar ambos problemas desde una perspectiva diferente (véase capítulo 6).

En este sentido, existen una gran cantidad de evidencias empíricas que permitirían afirmar la necesaria existencia de un vínculo reduccionista entre los rasgos físicos de un estado cerebral y los aspectos subjetivos de los estados mentales (Burton 1995: 172-173). Pensemos

tan sólo un momento en los experimentos llevados a cabo por determinados neurobiólogos (tal como indica Paul Churchland 1989) que afirman un nivel de solapamiento entre los estados de nivel primario y los estados de nivel superior. Teniendo en cuenta dichas evidencias, entonces parece necesario apelar más a una reducción de un nivel sobre el otro que a una conexión entre dos tipos diferentes de estado.

3.2.2.4. El papel causal de la mente: ¿el dualismo de propiedades de Searle?

En los últimos años, como hemos mostrado en las páginas anteriores, Searle ha intentado afrontar el viejo problema mente-cuerpo desde una perspectiva diferente a las posiciones materialistas y dualistas clásicas: su denominado naturalismo biológico. Según nuestro autor, esta corriente no es heredera ni del materialismo ni del dualismo, especialmente, porque ambas posiciones todavía estarían involucradas en el manejo de toda una terminología heredada de la filosofía cartesiana que ha oscurecido desde antiguo los intentos de solución del mencionado problema. En este sentido, Searle considera que el problema mente-cuerpo tiene una fácil solución, una solución que ha sido “accesible a cualquier persona con conocimiento desde que comenzó el trabajo serio sobre el cerebro hace aproximadamente un siglo” (Searle 1992: 1), si no hubiera sido por la oscuridad de la terminología heredada. Esta solución pasaría por reconocer que los estados mentales conscientes son el producto de ciertos procesos neurológicos que acontecen en el cerebro humano (y de algunos animales superiores) (Searle 1992: 1). Teniendo en cuenta esta explicación, el problema mente-cuerpo (al menos, el que ha sido formulado en términos clásicos) podría disolverse si somos capaces de dar cuenta de las relaciones causales que ocurren a nivel neurofisiológico que permiten la existencia de fenómenos mentales conscientes.

Los fenómenos mentales, para Searle, por tanto, ocurren a un nivel macro-físico como consecuencia de los fenómenos biológicos y neurológicos que ocurren en el cerebro a un nivel micro-físico. En este sentido, como ocurre con las experiencias conscientes, la conciencia de tener determinadas sensaciones representa en el macro-nivel de las interacciones personales del propio sujeto aquello que acontece a un nivel micro-físico en su cerebro (Searle 1983: 265-271). Las operaciones realizadas en dicho nivel pueden ser explicadas a partir de las relaciones existentes entre las diferentes partes del cerebro. Así, según Searle,

“la conciencia es una característica biológica de los humanos y de ciertos animales con cerebro. Es causada por procesos neurobiológicos y, en este sentido, es una más de la características biológicas de la naturaleza, como la fotosíntesis, la digestión o la mitosis” (Searle 1992: 90).

La única diferencia crucial entre dichos procesos biológicos y el proceso biológico de la conciencia es que, a diferencia de los primeros, la conciencia es ontológicamente irreducible por ser subjetiva.

La imagen que Searle nos presenta de la conciencia, entonces, es una imagen de causa y efecto provocada por los poderes causales de nuestro sistema nervioso central evolucionado y por el desarrollo de los procesos cerebrales internos, capaces de crear seres conscientes como

nosotros. La conciencia (y su subjetividad), por tanto, es una característica emergente de todo aquello que ocurre en nuestros cerebros a un nivel micro-biológico (Searle 1992: 14-15).

Así, dado que la conciencia se entiende como producida por elementos residentes en un nivel micro (representable tan sólo desde una perspectiva en tercera persona), podemos entender los diferentes estados mentales conscientes en primera persona como representados a un nivel superior (macro) por las funciones naturales de nuestro cerebro.

Frente a aquellos que, bien desde posiciones eclécticas como McGinn (1991) o materialistas escépticos como Nagel (1974) consideran que Searle defiende también un dualismo de propiedades que acaba por desconectar lo físico de lo mental, con la fatal consecuencia de postular dos sustancias excluyentes, Searle reivindica que el mundo físico es capaz de manifestarse de modo subjetivo. Para él es un mero hecho natural que cuando el sistema nervioso central de ciertos organismos complejos se desarrolla y evoluciona de la manera adecuada se consiga hacer emerger a un nivel personal los diferentes estados subjetivos de conciencia que los seres conscientes somos capaces de sentir (Searle 1992: 14). Por tanto, concluye Searle, la conciencia es un fenómeno natural.

Aunque Searle incide hasta la saciedad en este punto, consideramos demasiado excluyente sostener la existencia separada de propiedades físicas y mentales que, aunque se implementen en niveles diferentes, puedan a la vez estar conectadas. En primer lugar, es implausible suponer la existencia de dos niveles totalmente separados que, a pesar de ello, están interconectados a un nivel causal. Todavía faltaría por demostrar los mecanismos que permitirían dicha conexión. (Dichos mecanismos se analizan más tarde). Pero, todavía más preocupante, si cabe, es la contradicción en los términos que Searle pretende defender. ¿Cómo podemos sostener que lo físico, definible como aquello objetivo abordable desde la metodología científica que emplea la perspectiva de la tercera persona, puede ser manifestado tan sólo de modo subjetivo, sólo accesible a ello desde la primera persona?

Searle acaba apelando a la existencia en la mente de determinados poderes causales. Los mecanismos biológicos de los niveles inferiores que se dan en nuestro sistema nervioso tienen un claro efecto causal en los niveles superiores. Así, de este modo, determinados cambios físicos pueden generar determinadas creencias, del mismo modo que la detección de determinados cambios físicos por parte de nuestro sistema nervioso a través de los numerosos sistemas cognitivos (por ejemplo, el choque de mi espinilla con la mesita de café o la bajada brusca de temperatura en una habitación) permiten la aparición de determinadas sensaciones (como mi dolor o mi frío) (Searle 1992: 112). Pero también puede ocurrir a la inversa. Mi creencia en que en la nevera hay cerveza y mi deseo de beberme una, harán que vaya a la nevera a coger una cerveza para así poder saciar mi sed. Pero mi sentir frío o mi sentir dolor, probablemente también hará que abra un cajón para coger un jersey que ponerme o que pegue un grito (y refunfuñe entre dientes las virtudes de aquel a quién se le ocurrió construir la mesita a esa altura).

En este sentido, Searle insiste en negar que su posición sea un dualismo de propiedades. Los defensores de dicho dualismo, considera, creen que la mente es una sustancia que opera

dentro de su propia esfera de acción, que no es capaz de producir cambios. Pero como hemos visto, la mente parece también tener un papel causal a la hora de determinar la acción de los sujetos conscientes. Por tanto, dice Searle, su posición no puede asemejarse a un dualismo de propiedades.

Yo mismo, hablando personalmente, he sido acusado de mantener alguna loca doctrina de 'dualismo de propiedades' y 'acceso privilegiado', o de creer en la 'introspección' o en el 'neovitalismo' o incluso en el 'misticismo', aun cuando jamás he apoyado, implícita o explícitamente, ninguno de esos puntos de vista" (Searle 1992: 27-28).

En una fervorosa defensa de su posición, Searle (2002b) ofrece sus propias razones contra el dualismo. Unas razones que, por supuesto, considera que son suficientes para marcar las distancias entre los dualistas y su propia posición. En este sentido, Searle considera la tesis principal del dualismo (la posible existencia de dos sustancias separadas que delimitan, de modo separado, el ámbito de lo físico, por un lado, y el ámbito de lo mental, por otro) como inaceptable (Searle 2002b: 57), porque considera que, en primer lugar, el dualismo de propiedades no consigue comprender cómo el cerebro causa la conciencia y, en segundo lugar, el dualismo de propiedades todavía acepta el obsoleto vocabulario mentalista heredado de la tradición cartesiana (Searle 2002b: 57-58). En este sentido, y como Searle lo entiende, podemos negar el dualismo de propiedades a partir de una serie de críticas dirigidas contra sus argumentaciones básicas.

En primer lugar nos dice Searle, el dualismo acepta la existencia de dos excluyentes entidades metafísicas como constituyentes de la realidad. Por un lado encontramos los fenómenos físicos (objetivos) y, por otro, tenemos los fenómenos mentales (subjetivos). Pero, como dice Searle, esta suposición es simplemente absurda. Nosotros vivimos tan sólo en un mundo, aunque dicho mundo puede tener diferentes tipos de rasgos característicos. En este sentido, todo fenómeno en el mundo depende causalmente de elementos físicos. Así, la conciencia es un fenómeno biológico en tanto que causado por procesos puramente biológicos, físicos. En última instancia, concluye Searle, sólo existe una sustancia: la física. Aunque dentro de dicha entidad existen diferentes modos de expresión, aunque siempre causalmente dependientes de lo físico (Searle 2002b: 59-60).

En segundo término, el dualismo presupone que los estados mentales no son reducibles a estados neurobiológicos, que son algo diferente de éstos. En este sentido, postula la irreductibilidad de lo mental a lo físico. Searle, frente a esta propuesta, es poco crítico. Acepta sin reparos la afirmación dualista de la irreductibilidad de la conciencia, en tanto que presupone una característica distinta de lo físico. Pero, advierte, que no sea físicamente reducible no quiere decir que no sea causalmente dependiente de lo físico. Los estados mentales no son reducibles a términos físicos tal y como pretenden otros proponentes (como los teóricos de la identidad y los eliminacionistas), pero sí que es perfectamente explicable a partir de procesos biológicos que tan sólo pueden darse a un nivel físico (Searle 2002b: 60).

Siguiendo con su propia defensa, en tercer lugar, Searle critica la asunción dualista de la conciencia (y lo mental) como una característica adicional que va más allá de lo meramente físico, como trascendente al mundo biológico. En este sentido, defiende Searle, el dualismo se equivoca de pleno. Porque los estados de la conciencia y los estados cerebrales son las dos caras de una misma moneda. Son la expresión desde diferentes perspectivas (una objetiva y la otra subjetiva) de los mismos procesos neuro-biológicos que tienen lugar en el sistema nervioso central (Searle 2002b: 61).

Para finalizar con su crítica, en un último momento, Searle acusa al dualismo de propiedades de mantener todavía el anticuado vocabulario mentalista cartesiano. En este sentido, nos dice Searle, lo ideal es abandonar dicho vocabulario, porque la misma apelación a la posibilidad de un problema mente-cuerpo, formulado en los términos clásicos, ya nos involucra en la asunción de dicho vocabulario contaminado de terminología dualista. Así, debemos evitar apelar a dicho problema (porque realmente ni tan siquiera existe). Lo ideal es buscar la disolución de dicho pseudo-problema porque, por un lado, promueve la reducción física del mundo acabando por adoptar posiciones epifenomenalistas respecto de la conciencia (si se asume una solución materialista del mismo) y, por otro lado, si nos quedamos vinculados a un dualismo, acaba por defender una explicación física de la realidad incompatible con las actuales conclusiones de la ciencia (un ejemplo claro lo encontramos en Searle 2002b: 61-62).

“Pero si la conciencia no tiene poderes causales más allá de su base neurobiológica, entonces, ¿no estaríamos implicando el epifenomenalismo? No. Compáralo con este ejemplo: la solidez de los pistones del motor de un vehículo no tienen poderes causales adicionales a los que presenta su base molecular, pero ello no significa que dicha solidez sea un epifenómeno (trata tan sólo de fabricar un pistón con agua o con mantequilla). La cuestión subyacente es la siguiente: ¿por qué alguien podría suponer que la reductibilidad causal implica epifenomenalismo si el mundo real es plenamente eficaz causalmente hablando, esto es, sus propiedades de nivel superior son causadas por completo por los fenómenos realizados a un nivel inferior, microfísico? La respuesta aquí sería la siguiente. Porque podría pensar que la conciencia es algo distinto, algo mayor o menor, que su base neuronal. El dualismo tradicional de propiedades piensa que el cerebro nos eleva hasta la conciencia y nos ofrece una imagen de la conciencia generada por el cerebro como la ofrecida por una caldera que genera vapor”.

Para Searle, considerar que la conciencia (o cualquier estado mental) es ontológicamente subjetiva no es una razón suficiente para atribuirle una posición dualista. En este sentido, Searle cree que la subjetividad ontológica presente en la conciencia es inevitable, por lo que la reducción causal de lo mental es suficiente para resolver el problema mente-cuerpo, siendo innecesaria una reducción ontológica, tal y como pretenderían los materialistas (a los que Searle acusará de, históricamente, haber olvidado y negado la importancia del fenómeno de la conciencia) (Searle 1992: 100). Pero parece claro también, dado que no podemos diferenciar entre aquello a lo que accede al ser consciente y el acto mismo del que se es consciente, que el estado mental consciente presupone (y es lo mismo) que el acto mismo de acceso consciente. En este sentido, entonces, parece también claro que dicho acto tan sólo es accesible para mí mismo, por lo que no sería necesario postular ningún principio de causalidad que relacionara estados mentales con neurológicos. Principalmente por una dificultad añadida: la existencia de una tesis causal que relacione ambos tipos de estados presupone tener que

mostrar una serie de datos empíricos que Searle no es capaz de presentar (aunque sea cierto que en neurología son importantísimos los informes y reportes subjetivos proporcionados en primera persona por los propios pacientes), por lo que la tesis (empírica) de la causalidad entre lo mental y lo físico parece devenir gratuita (Olafson 1994: 258; Burton 1995: 164; Pineda 1999: 160).

Llegados a este punto, y para ser justos con nuestro autor, debemos decir que, aunque Searle esté apostando por una solución reduccionista, ello no quiere decir que esté contradiciéndose. Es necesario tener presente que existen diferentes tipos de reducción. Searle (1992: 112ss.) analiza hasta cinco tipos distintos (reducción ontológica, reducción ontológica de propiedades, reducción teórica, reducción lógica y reducción causal), siendo la causal la noción más débil de reducción. Según este concepto, las propiedades causales del objeto reducido son perfectamente explicables en términos de las propiedades causales del objeto reductor. Según Searle, por tanto, al pretender reducir causalmente lo mental a lo físico, lo que está persiguiendo es establecer una reducción de los poderes causales de lo mental como perfectamente explicables en términos de los poderes causales de lo físico sin necesidad de tener que reducir la existencia de uno a la existencia del otro, dado que la ontología de lo físico sería objetiva mientras que la ontología de lo mental es subjetiva. Aunque Searle no lo indique, y por mucho que lo niegue, ésta sigue siendo una posición marcadamente dualista (Pérez Chico 1999: 128-129).

Precisamente por ello, a pesar de los esfuerzos de Searle de desvincularse del dualismo de propiedades, consideramos que sus argumentaciones no son del todo convincentes al mantener, todavía, cierta resonancia mentalista en sus premisas. Si atendemos a la suposición de la existencia de poderes causales en lo mental, podemos observar cómo la posición de Searle sigue manteniendo, al menos en espíritu, la posibilidad de la existencia separada de la mente y el cuerpo. Veamos. Searle parte de la suposición de que nada que no sea un cerebro (humano, en principio, aunque deja abierta la posibilidad de que otros tipos de criaturas con cerebro tengan una noción mucho más débil de conciencia) puede tener una influencia causal sobre la conducta y sobre otros estados mentales. En este sentido, sólo un cerebro es capaz de causar una mente (y su conciencia) por la actividad de determinadas estructuras y procesos neurobiológicos que operan en él. Pero, consideramos, esta posibilidad no indica para nada la imposibilidad de que la mente (y la conciencia) pueda ser causada de otro modo (Warfield 1999: 31). Pensemos por un momento en un termostato. Es capaz de representar (empleando muy ampliamente la noción de representación, aquí) la temperatura que hace en una determinada habitación. En este sentido, podríamos decir que el termostato 'sabe' qué temperatura hace en la habitación. En un sentido laxo, el termostato también puede tener un cierto nivel superior (caracterizado con una noción epistemológicamente débil de 'saber') causado a partir de una serie de procesos físico-químicos que se producen a un nivel primario (la combinación de calor y presión que hace que la columna de mercurio varíe en altura, según el caso). En este caso, podemos concluir, que los poderes causales del cerebro permitan inferir que son los causantes de conciencia en los cerebros humanos no nos dice nada acerca de la posibilidad de que existan otros medios y otros mecanismos causales que permitan la aparición de conciencia.

Pero todavía podríamos ir más allá. Searle acaba también apelando a la posibilidad de que tan sólo el cerebro humano tenga los poderes causales mínimos necesariamente requeridos para alcanzar el umbral de conciencia que permitiría afirmar la existencia de procesos mentales. Pero en este caso, podríamos replicar, tampoco parecen ser así las cosas. Pensemos tan sólo en la posibilidad de que existan seres extraterrestres capaces de estados conscientes (Lewis 1978). Para Searle, si existieran este tipo de seres, nunca podrían llegar a ser capaces de conciencia del mismo modo en que nosotros somos conscientes si no tuvieran una estructura biológica similar a la nuestra (Searle 1984: 41). Pero puede darse el caso, creemos, que existan seres extraterrestres que tengan conciencia y que, sin embargo, sean estructuralmente hablando muy diferentes de los seres humanos. En este sentido, podemos concluir, Searle tampoco es capaz de argumentar suficientemente su tesis de la necesidad causal del cerebro humano para el surgimiento de la conciencia.

Mucho más preocupante si cabe es la concepción emergentista que Searle pretende asociar a la posibilidad causal de los estados cerebrales. Podemos definir claramente una noción de emergencia que necesariamente implique una reducción causal. En este sentido, tan sólo tendríamos que pensar en términos naturales como la temperatura o como el agua. Que la temperatura sea temperatura es causalmente reducible a la presencia de movimiento molecular, y que exista agua es causalmente reducible a la composición química de la sustancia que se encuentra en el vaso enfrente de nosotros. En un caso como el sostenido por Searle, entonces, si consideramos que la conciencia emerge de los procesos cerebrales, debe considerarse como causalmente reducible, por lo que sería perfectamente explicable desde el punto de vista de la reducción de los estados conscientes a los estados biológicos que provocan la aparición de la conciencia y no, como pretende Searle, explicable desde la presencia de una noción autónoma de conciencia totalmente desvinculada de los procesos biológicos que la causan. Si consideramos que la conciencia es emergente de los procesos biológicos que acontecen en el cerebro, entonces, se afirma su identidad causal con procesos neurológicos y no como irreductiblemente subjetiva (Sabatés 1999: 150). Pretender lo contrario supondría una contradicción en los términos.

Pero Searle rechazará todo intento de materialismo reductivo porque no le parece plausible la idea de que toda propiedad mental sea una propiedad física y también bloqueará todo intento funcionalista de explicación de lo mental por considerar errónea la identificación de toda entidad mental con propiedades funcionales múltiplemente realizables por propiedades físicas. El motivo principal de este rechazo es que toda identificación similar entre lo mental y lo físico debe ser considerada como una verdad *a posteriori*. Que una verdad sea *a posteriori* implicaría que sólo pueda explicarse a partir de conceptos diferentes a aquellos que habitualmente maneja el sujeto de conciencia, porque ha sido identificada a partir de propiedades distintas a aquellas que florecen a un nivel personal. Pero si, con Searle, presuponemos que las propiedades mentales son subjetivas (algo normal si consideramos tan sólo el nivel personal en el que se mueve habitualmente nuestro darnos cuenta del mundo), ello nos aboca irremediabilmente al dualismo (por pretender la existencia de dos tipos diferentes de

propiedades), en tanto que la necesidad de tener que explicar las propiedades iniciales a partir de propiedades diferentes (físicas, en este caso) nos obliga a tener que compaginar la explicación de la aparición de la conciencia entre los niveles causales mental y físico. Pero, advierte Searle, si excluimos esa posibilidad, entonces la propiedad característica de lo mental (la subjetividad) queda suspendida y se desvanece toda posible solución del problema mente-cuerpo (Searle 1992: 36-37), porque quedaría por afrontar el aspecto cualitativamente dependiente del sujeto que toda experiencia (al menos en principio) parece tener.

Uno de los motivos principales de Searle para negar la posibilidad materialista de solución, la identificación entre lo mental y lo físico, del problema mente-cuerpo, precisamente es la existencia del hiato explicativo. Esto es, parece que la imposibilidad de explicar la característica subjetiva a partir de términos objetivos hace implausible la tesis de la reducción ontológica. Pero tampoco parece plausible, como pretende Searle, mantener una posición que respete la dualidad ontológica a partir de una relación causal. Por dos motivos principales. En primer lugar, porque parece plausible suponer un criterio epistémico que nos indique que si disponemos de una explicación completa de un fenómeno que no apela para nada a otro agente causal, entonces no requerimos de nada más para que dicho fenómeno acontezca (Jackson and Pettit 1990: 198). Esto es, que si es posible explicar lo mental sin necesidad de apelar a lo físico, entonces parece innecesario postular la relación causal al alto precio de tener que mantener el dualismo de propiedades. Pero, en segundo lugar, todavía más preocupante, porque mantener una teoría como el naturalismo biológico que presupone la posibilidad de la existencia de un hiato explicativo nos conduciría de lleno a mantener una metafísica dualista como la que Searle pretende eliminar (Palmer 1998: 166; Pineda 1999: 163).

Pero, yendo todavía más lejos, aunque diéramos por buena la propuesta searleana de la existencia de una explicación causal de los estados mentales, no debemos olvidar que una relación causal establece una relación diacrónica, es decir, necesariamente existen dos elementos (la causa y el efecto) que están temporalmente distantes (por ejemplo, Kim 1984). Si esto es cierto, entonces, al presuponer que existen relaciones causales entre los procesos neurológicos del cerebro y los procesos mentales conscientes, Searle recae nuevamente en el dualismo al pretender que ambos se den de modo separado aunque se den al mismo tiempo (Pérez Chico 1999: 129-130). Aunque Searle insista en que la relación causal existente entre lo físico y lo mental es sincrónica, donde las entidades relacionadas tan sólo son descripciones distintas de un mismo fenómeno (Searle 1992: 115), que defienda también una diferenciación ontológica entre lo físico y lo mental hace sospechar que nuevamente su posición está extremadamente cercana al dualismo al presuponer una relación interna de interdependencia entre dos tipos epistemológicamente diferentes de aproximaciones al mismo fenómeno (Honderich 1995).

En este sentido, parece que todavía Searle nos debe una explicación de cómo debemos entender esta relación causal existente entre los procesos biológicos y los procesos mentales conscientes en tanto que no parece suficiente, por un lado, rechazar aquello que tradicionalmente se ha tenido como causalidad por ser erróneo y mucho menos vale tan sólo, en

segundo lugar, establecer una declaración inicial de intenciones de abandonar la terminología mentalista heredada. Sobre todo cuando, como hemos visto, parece que dicha intención no se mantiene. Lo que parecería más acertado para disolver el problema mente-cuerpo, y en esto coincido con Pérez Chico (1999: 130), es incluir las discusiones sobre el tema dentro del ámbito de lo epistemológico, dejando resueltos los diversos aspectos ontológicos que la intuición sobre la subjetividad de lo mental nos plantea a partir de una noción material de relación constitutiva entre lo físico y lo mental lo suficientemente fuerte como para vincular la aparición de lo subjetivo a las explicaciones de los procesos físicos que los provocan. Así, parece mucho más razonable suponer que cuando hablamos del cerebro y de la mente nos referimos a la misma entidad ontológica aunque desde niveles diferentes, por lo que no tiene sentido referirse a ellos en términos de emergencia. Pero que tampoco es posible, en sentido estricto, establecer equivalencias ontológicas concretas (Freeman and Skarda 1991), porque las explicaciones causales de lo mental, al menos en el nivel personal en el que se mueven, son extremadamente metafóricas y simples (Pérez Chico 1999: 134-135). Por tanto, la solución a dicha equivalencia pertenece al mundo científico, mientras que en la vida cotidiana las cosas seguirán siendo, como siempre, una cuestión de cotidianidad.

A la vista de las numerosas contradicciones internas implícitas en la posición y defensa de su teoría de la conciencia (y de la mente), podríamos denominar la posición defendida por Searle como un *dualismo cerebral* (Olafson 1994: 255), como un dualismo *biológico* de propiedades (Corcoran 2001: 314) o como un pluralismo de propiedades (Guerrero 2001c: 313), por pretender interconectar mediante mecanismos biológicos el cuerpo y la mente que, según su propio análisis, serían ontológicamente irreconciliables, acabando por proponer un materialismo anómalo que desemboca en un dualismo conceptual (Burton 1995: 166), siendo éste un precio demasiado elevado a pagar para una solución al problema de la conciencia.

El error principal de Searle estribaría, parece ser, en intentar amoldar su concepción de la conciencia a un concepto de realidad demasiado ambiguo. Por un lado, Searle pretende que dicho concepto de realidad sea el que habitualmente tenemos por válido, el sentido cotidiano de realidad. Por otro lado, Searle propone un concepto diferente de realidad, un concepto en el cual lo mental forma parte de la misma, algo diferente del concepto habitual cotidiano. En este sentido, esta ambivalencia entre las dos nociones de realidad va a permitir a Searle sostener la irreductibilidad de lo mental a lo físico en tanto que lo mental tendría ya cabida dentro de la realidad a partir de su concepción de lo real, afirmando que lo mental sería una de sus diferentes propiedades. Sin embargo, *pace* Searle, no se acaba de comprender cómo dos formas diferentes de realidad son capaces de encuadrarse dentro de la misma categoría ontológica. En este sentido, el error de Searle es de base, al pretender aportar una noción de realidad lejana de la habitual.

3.2.2.5. Conclusión: el precio a pagar por el dualismo

Cuenta la anécdota que un buen día Lord Bertrand Russell comentó a un allegado que el sentido común lo inventó John Locke, y que desde entonces era una cosa que sólo han tenido los

ingleses. Haciéndose eco de estas palabras, John R. Searle se ha pretendido constituir como adalid del sentido común en cuestiones de filosofía de la mente, en general, y de la conciencia, en particular (aunque él sea originario de Denver y enseñe en la Universidad de Berkeley, California). Tal vez por ese motivo considera suficiente repetir tantas veces como sea posible que lo único que pretende es reforzar con argumentaciones la imagen general que todos tenemos de la conciencia; una imagen tan generalizada que a menudo resultará en argumentos obvios. Precisamente por ello, los problemas principales de la teorización de Searle son, por un lado, dar por supuesta una imagen manifiesta de la mente y la conciencia en tanto que sostenida por la mayoría de los mortales y, por otro lado, repetir hasta la saciedad, demasiado a menudo sin mayor argumentación que la simple apelación a dicho sentido común, toda una serie de coletillas y afirmaciones de carácter general sobre la mente que tienen una difícil demostración empírica.

Entonces, parece ser que el principal obstáculo para la aceptación de la teoría de Searle se encuentra en su propia teorización. Mejor dicho, en la falta patente de teorización. Pero, por supuesto, no será éste el único de sus inconvenientes.

Pensemos tan sólo por un momento de nuevo en el Principio de Conexión. Podríamos decir que los estados mentales inconscientes están conectados de manera indirecta con la conciencia en tanto que son accesibles a ella en un futuro, en tanto que posibles estados de conciencia. En este sentido, un estado inconsciente (en tanto que neurofisiológicamente definido) será mental en la medida en que es potencialmente consciente. Un estado mental es accesible a la conciencia si y sólo si, según la argumentación del propio Searle, tiene la capacidad de causar o producir estados mentales conscientes y, por tanto, generar en el sujeto de las experiencias fenoménicas la perspectiva subjetiva requerida para su vivencia (Searle 1990: 588; Searle 1992: 160). De todos modos, parecería posible identificar ya aquí un primer inconveniente en la posición de Searle, dado que parece posible inferir que no todo estado cerebral que permite hacer conscientes a los estados mentales es él mismo un estado mental (pensemos tan sólo en el proceso de mielización de los axones, que permite hacer aflorar a nivel consciente ciertos estados mentales pero que él mismo no lo es). En este sentido, parece que Searle está empleando dos nociones diferentes de relación causal, una relativa a la generación de estados mentales a partir de otros estados mentales (como las creencias) y otra relativa a la generación de estados mentales a partir de estados cerebrales, esto es, de estados no-mentales. Es decir, ¿en qué sentido podemos decir que mi creencia en que la Torre Eiffel está en París difiere de la mielización axonal de mi cerebro si ambas son capaces de generar el mismo tipo de estados mentales?

Teniendo en cuenta las críticas presentes en Kim (1995), podríamos decir que Searle confunde entre dos tipos diferentes de causalidad. Podemos caracterizar la noción de Searle (recordemos, la generación de estados mentales a partir de la existencia de estados cerebrales potencialmente conscientes) como una noción causal entendida como superveniencia causal (como enfrentada a la superveniencia constitutiva defendida en Kim 1984 y que hemos expuesto ya al final de la sección 3.2.2.3.). Pero precisamente esta noción de superveniencia causal, a

diferencia de la constitutiva, debería ser analizada en términos disposicionales (véase también Garrett 1995: 211). Desde este punto de vista, los estados mentales inconscientes no serían considerados como causas sino como disposiciones. Éstas se actualizarían en diferentes episodios mentales definidos como estados mentales conscientes en tanto que permitirían a su poseedor tener un acceso de primera persona al contenido fenoménico de dichas vivencias conscientes. La cuestión problemática de dicha noción no es la disposicionalidad en sí, que aunque contradice la inicial posición searleana de establecimiento de una relación causal entre conciencia e intencionalidad, todavía tendría un cierto poder normativo. La mayor dificultad con la que se encuentra dicha argumentación sería tener que rendir cuentas del proceso que permitiría dichas actualizaciones constantes a un nivel consciente y la explicación de los mecanismos cerebrales o físicos que permitirían determinar dicha actualización; una explicación que Searle todavía nos debe.

Porque si, con Searle, admitimos que mi creencia no actualizada de que la Torre Eiffel está en París es una cierta disposición a que se convierta en consciente y ésta no es más que un cierto estado neurofisiológico de mi cerebro, entonces deberemos también admitir que mi creencia no actualizada de que el cumpleaños de Edurne es el día 1 de octubre o de que el Parc Nacional de la Font Roja se encuentra en Alcoy también corresponden a sendos estados neurofisiológicos de mi cerebro, y así para todo estado intencional no actualizado, generando una gran cantidad de estados cerebrales que parecería aglutinar todo nuestro cerebro y que parecería desmentir la negación de Searle de que el cerebro (humano) sea un inventario o un almacén (como hemos mostrado anteriormente) en donde se amontonan estados latentes a la espera de su posterior afloración en la conciencia. Desde esta perspectiva, Searle parecería estar adoptando el mismo punto de vista ingenuo que, denuncia, estaría presente en Freud (Searle 1992: 174-178) al pretender insinuar que el inconsciente tiene todo aquello que caracterizaría a la conciencia, pero sin la conciencia. En este sentido parece irrelevante y superfluo diferenciar entre estados mentales conscientes e inconscientes si lo que vamos a pretender defender es que ambos comparten el mismo estatuto intencional, sobre todo después de ver que lo que acaba por defenderse es, simplemente, una noción mucho más débil que la de causalidad. Así, no habría diferencia entre conciencia e inconsciente más allá de la propia característica de poseer conciencia, por lo que la dicotomía parece derrumbarse (Stigol 1999: 107).

Todavía más preocupante, si cabe, sería la posibilidad abierta al final de la sección 3.2.2.3., donde dejábamos entrever la posibilidad de caracterizar la posición de Searle como un dualismo conceptual. Hemos comentado a lo largo de dicha sección que, aunque parece que la posición de Searle respecto a la naturalización de los estados mentales no parecía diferir sobremanera de una posición dualista de propiedades, y que podríamos aceptar, de manera forzada, la posición de nuestro autor como un dualismo conceptual. Pero ello no consideramos que permita una mejor salida a Searle. Veamos.

Searle se empeña en sostener, por un lado, que la conciencia es una propiedad real, pero irreducible, del mundo y, por otro lado, que la conciencia es uno más de los procesos biológicos del mundo (como la mitosis) y, por tanto, reducible a las características físicas del cerebro.

Visto de esta forma, el argumento de Searle es un dualismo conceptual en tanto presupondría que “en un sentido importante lo físico implicaría lo no-mental y lo mental implicaría lo no-físico” (Searle 1992: 26), postulando la existencia separada de lo mental y lo físico. Así, Searle parece defender la posibilidad de que algunas propiedades físicas pueden darse en un nivel descriptivo superior, el nivel de lo mental, mientras que todas las propiedades mentales se pueden explicar desde el nivel primario de descripción de lo físico. En última instancia, lo Searle está defendiendo es, por una parte, que los estados mentales son ontológicamente subjetivos mientras que, al mismo tiempo, por otra parte, son epistemológicamente objetivos.

Pero esto parece entrar en contradicción con las mismas tesis de Searle, en tanto que él mismo afirma que

“el hecho de que una característica sea mental no implica que no sea física, y el hecho de que una característica sea física no implica que no sea mental” (Searle 1992: 15).

En este sentido, parece que existe un modo en el que podemos decir que la conciencia, en definitiva, sí es reducible, en tanto que se insiste en redefinir los estados mentales como procesos neurofisiológicos, y a éstos como causantes de los estados mentales conscientes (Corcovan 2001: 315). Visto así, parece que Searle está a caballo entre la dualidad conceptual y la reducción. Si se mantiene la dualidad, entonces se corre el riesgo de recaer en los errores del dualismo de propiedades, como vimos en la sección 3.2.2.3. Pero si aceptamos dicha reducción, entonces, al menos desde la posición de Searle, todos los poderes causales de los fenómenos reducidos deben poderse explicar en términos de los poderes causales reductores (Palmer 1998: 163-164). Esto es, si aceptamos la reducción, entonces la conciencia sería reducible a procesos neurobiológicos, por lo que los poderes causales de los estados mentales deberían poderse explicar en términos físicos, algo que Searle no consigue hacer y a lo que se muestra reiteradamente reacio. Lo máximo a lo que podríamos apelar para intentar exonerar a Searle de cometer una defensa de un dualismo de propiedades es asumir que defiende una posición misterianista respecto de la conciencia, en tanto que la falta de una explicación adecuada de los procesos físico-biológicos que permitirían una adecuada imagen de la mente depende de un desarrollo científico no disponible en la actualidad y que, de todos modos, una explicación completa de los mecanismos que operan a nivel primario en la conciencia es inaprensible para el entendimiento humano. En pocas palabras, que lo que parece requerir la posición de Searle, más que toda la serie de reiteradas afirmaciones en la eficacia causal de lo físico sobre lo mental, y viceversa, sería la aceptación de un principio de limitación epistémica que afirmara el cierre cognitivo de la explicación de la conciencia (al estilo anticipado por Northoff and Musholt 2006). Pero, por supuesto, esto lo dejaría irremediadamente abocado a una posición defensora del dualismo conceptual. Al menos así, eso sí, parecería posible entender los intentos de argumentación de Searle como un intento real de explicación y no como, lo que realmente parece, un intento ineficaz de retórica interminablemente repetitivo.

La estrategia de Northoff and Musholt, tras demostrar el intento fallido de Searle de encontrar una tercera vía que se sitúe entre el dualismo y el materialismo y que verdaderamente

permita escapar de la terminología tradicional (tan cartesianamente influenciada), pasa por introducir un nuevo concepto: el concepto de 'limitación auto-epistémica'. Esta noción describe nuestra incapacidad como seres conscientes de poder experimentar de modo directo nuestro propio cerebro como accesible desde la perspectiva de primera persona. Por supuesto, Searle rechaza este tipo de inferencia porque considera que, epistémicamente hablando, las propiedades mentales son propiedades físicas, aunque sea imposible reducirlas desde el plano ontológico. Pero si, como hemos considerado, la argumentación de Searle, por un lado, acaba por asimilar el mismo vocabulario cartesiano que pretende desechar y, por otro, no es capaz de ofrecer una alternativa viable al dualismo y al materialismo, entonces parece necesario que realice una reformulación de sus tesis iniciales para llegar a dar correcta cuenta de la problemática de la separación entre cuerpo y mente, en general, y de la conciencia, en particular.

3.3. Recapitulación

Al comienzo de este capítulo dejamos constancia de nuestra intención de argumentar contra la posibilidad, dejada entreabierto por Searle, de la existencia de una separación entre los estados mentales y los estados físicos. Precisamente nuestra conclusión aquí, como hemos mostrado, ha sido rechazar el naturalismo biológico de Searle en base a dos constataciones básicas. Primera, la teoría de Searle acaba por afirmar un dualismo epistemológico que, aunque permite explicar una relación causal entre lo mental y lo físico, permite la existencia de una irreductibilidad de lo mental a lo físico a partir de la constatación de la existencia de dos niveles diferentes de descripción: uno respecto de los estados físicos y otro respecto de los estados mentales. Segunda, la incongruencia entre su posición (contaminada por el vocabulario mentalista cartesiano que tanto empeño pone en eliminar) y la actual visión física del mundo.

Hemos presentado nuestras reservas a la posibilidad de que la conciencia sea un producto emergente del cerebro, porque por un lado no parece que Searle consiga probar suficientemente que el cerebro segregue conciencia del mismo modo que el páncreas segrega la bilis (esto es, en tanto que los argumentos de Searle no consiguen demostrar que la conciencia es un producto biológico más) y por otro lado la noción de emergencia presupone una noción de causalidad extremadamente débil que dificulta su aceptación, en tanto que la acerca peligrosamente más a posiciones disposicionalista que una verdadera noción de realizabilidad causal.

También nos hemos mostrado reacios a aceptar el Principio de Conexión que permite a Searle argumentar a favor de la existencia de estados cerebrales potencialmente conscientes. Si consideráramos factible la eficacia de dicho principio, como hemos negado, deberíamos aceptar también la posibilidad de la existencia, por un lado, de un nivel inconsciente ingenuamente establecido (al estilo de Freud) y, por otro, de la existencia de la conciencia como un almacén donde se acumularían nuestras potenciales 'herramientas' conscientes biológicamente establecidas. Como hemos mostrado, Searle mismo se encarga de criticar la primera posibilidad y sus propias argumentaciones (aunque persiguen el efecto contrario) le hacen recaer en una

dificultad como la mostrada en la segunda posibilidad. Ello nos ha permitido concluir la inconsistencia (por ser contradictorio) del mismo Principio de Conexión.

Visto lo anterior, no nos ha quedado más remedio que rechazar la propuesta de naturalización de Searle porque, en primer lugar, parece lejano de la realidad aceptar actualmente la existencia de un nivel físico que permita explicar la existencia de estados mentales que, de ningún modo, puedan ser asimilables a éste (no afirmamos que tengan que ser reducibles, tan sólo que puedan ser explicados causalmente). Pero como hemos visto, Searle aceptaría un nivel causal demasiado débil como para poder explicar la existencia de lo mental desde presupuestos físicos (por muy realista que pretenda ser en sus propuestas). En segundo lugar, además, la posición de Searle acaba desembocando en un irremediable dualismo de propiedades, en tanto que acepta de manera clara la posibilidad de dos niveles (aunque sean tan sólo epistémicamente accesibles). En este sentido, lo mental sería epistémicamente diferente de lo físico, contraviniendo las mismas pretensiones searleanas de permanecer anclado en el monismo que sirve de base a su naturalismo biológico.

CAPÍTULO 4. MICHAEL TYE Y LA TEORÍA REPRESENTACIONAL FUERTE DE LA CONCIENCIA

No puede existir un sujeto de experiencias sin experiencias.
(Nicholas Humphrey, *Seeing red*)

El representacionalismo dice que, cuando observas detenidamente un cuadro y sus detalles, estás siendo consciente de algo fuera de ti y de sus cualidades, que experimentas como cualidades de partes del cuadro... Eres consciente de ellas de modo subjetivo, esto es, fenoménicamente.
(Michael Tye, *Consciousness and persons*)

IV. MICHAEL TYE Y LA TEORÍA REPRESENTACIONAL FUERTE DE LA CONCIENCIA

El ejemplo de Michael Tye es la prueba de la existencia de una necesaria conexión entre los estudios filosóficos y los científicos. Educado en la tradición británica (es inglés), estudió Física en la Universidad de Oxford, donde comenzó a sentir predilección por los temas relativos a la separación entre la mente y el cuerpo. Una vez finalizado su doctorado en Oxford, con una tesis entre Filosofía y Física, fue contratado por la Temple University de Philadelphia y actualmente ocupa una plaza de profesor en el Departamento de Filosofía de la Universidad de Texas en Austin. Tye es uno de los máximos defensores, en la actualidad, de las posiciones representacionistas fuertes.

En los últimos años, la teoría representacionista de Michael Tye ha ido radicalizándose paulatinamente y convirtiéndose, cada vez más, en una posición intencionalista que defiende la (necesaria) presencia de una representación de los aspectos relevantes de un objeto externo hacia el cual debe dirigirse nuestras experiencias conscientes. Este externismo respecto de lo mental no supone un problema, pero parece encontrarse con ciertas dificultades cuando se aplica a las sensaciones fenoménicas: parece que no toda experiencia puede ser considerada como representacional y, a menudo, se vuelve difícil encontrar el objeto externo hacia el cual se dirige una sensación determinada.

Siguiendo con su teoría PANIC acerca del contenido de las experiencias de conciencia (Tye 1995 y 2000), Tye (2003, 2008 y 2009) concluye que el contenido de las experiencias fenoménicas (en general)

1. tiene estrechas relaciones con el aparato nervioso y neuronal (según muestran los recientes estudios neurológicos),
2. contiene tanto componentes sensitivos como afectivos (en un sentido evaluativo, pero carente de toda carga moral),
3. permite representar aquellas características relevantes de los objetos (o los objetos enteros) involucrados en las sensaciones y en las experiencias fenoménicas, y
4. es no-conceptual.

Pero la característica principal que Tye resalta del contenido de una experiencia fenoménica es que ésta siempre es

(T) Transparente: esto es,

a) toda experiencia fenoménica está presente de modo inmediato en nuestras vivencias [*awareness*] y, además,

b) la vivencia de cómo es para mí algo [*what is it like*] no es una sensación *de re* de las experiencias o de sus cualidades, sino que es una sensación *de dicto* de aquello que experimentamos como teniendo un determinado carácter fenoménico), porque

- no existe un lugar determinado hacia el que se dirija nuestra atención para sentir dicha experiencia (dado que se daría al unísono con todas las demás experiencias conscientes,

incluidas las corporales, las cualidades primarias y las cualidades secundarias, de modo centralizado),

- permite evaluar los objetos del mundo de modo introspectivo y
- no cabría el error a la hora de individualizar e identificar una determinada sensación como una sensación de algo en particular.

Desde nuestro punto de vista, la exigencia de transparencia es demasiado estricta porque, aunque es cierto que a veces posibilita la explicación respecto de los errores de juicio, impide la posibilidad de error a la hora de explicar esta noción de manera empírica en los estados proposicionales y en los fenoménicos y convierte a la vivencia de la experiencia en virtualmente infalible. Pero sabemos que los casos de error son posibles, dado que, a menudo, el sujeto puede desconocer que dos elementos de pensamiento (términos singulares, experiencias conscientes, emociones, sensaciones, etc.) son co-referenciales y, a pesar de ello, puede referirse a ellos. Es decir, a diferencia de Tye (que aceptará un cierto millianismo a la hora de definir el contenido fenoménico, incluido el emocional), consideramos pertinente, al menos, tener en cuenta un cierto descriptivismo de carácter fregeano a la hora de caracterizar e individualizar el contenido de una experiencia consciente. Y todo ello porque, además de parecernos una ilusión (por ser demasiado restrictivo) el requisito de transparencia, que hechos como los mencionados ocurran nos hace dudar de la posibilidad de que todo contenido fenoménico consciente, incluido el emocional, sea no-conceptual. Porque C1. parece necesario que una explicación adecuada de las experiencias fenoménicas debe dar correcta cuenta del modo en que los seres conscientes nos comportamos en determinadas ocasiones, C2. esas reacciones naturales responden a nuestros intereses, y estos preceden a toda representación y C3. una interpretación estrictamente no-conceptual no parece satisfacer estos requisitos (al dejar fuera ciertos elementos conceptuales presentes en nuestra experiencia, siendo necesario apelar de tanto en tanto a ellos, como veremos en el capítulo final), por lo que se hace necesario un estudio de la experiencia fenoménica que respete la intencionalidad presente en la conciencia en un sentido no sólo representacional.

Precisamente con la intención de mostrar esas carencias en la teoría de Tye, pretendemos en lo que sigue hacer un análisis pormenorizado de la concepción representacional fuerte que nuestro autor actualmente defiende, comenzando por presentar una caracterización del modo en que debe ser entendido su representacionalismo, mostrando sus diferencias con otros representacionalismos afines. Seguidamente, una vez considerado el carácter que en la teoría de Tye adquieren las experiencias con contenido consciente, analizaremos su teoría PANIC, una teoría que persigue presentar todo Contenido de la conciencia como Preparado, Abstracto, No-conceptual e Intencional. Por último, nos detendremos en el análisis de los argumentos que Tye defiende para caracterizar el contenido de las experiencias corporales, las emociones y los estados de ánimo como contenidos PANIC, presentando así de manera completa su teoría de la conciencia y sus soluciones a determinados problemas relevantes en el estudio de la conciencia.

4.1. La teoría intencional de Tye

El representacionalismo, como lo hemos definido anteriormente (véase el capítulo 2, especialmente el apartado 2.2.), es aquella teoría respecto de la mente y la conciencia que considera que el contenido de la experiencia fenoménica está determinado por la relación existente entre cierta clase de representaciones y aquello que se experimenta. Para Tye, estas representaciones no se referirán de modo cerrado a los objetos externos, sino que referirán a aquellas características relevantes desde el punto de vista cognitivo del sujeto perceptor. Así, para el representacionalismo de Tye cuando, por ejemplo, contemplamos un cuadro (con todos sus detalles) estamos experimentando fenoménicamente (o tomando conciencia) de un objeto externo y las características cualidades que éste posee, de tal modo que el contenido de dicha experiencia concuerda con la representación de dicho cuadro a partir de la descripción de sus detalles más relevantes para nosotros y no como una representación del objeto en sí. De este modo, dado un cambio en alguna de dichas cualidades, se modificaría por completo nuestra representación del objeto, transformando la sensación inicial en una sensación diferente.

Ésta sería a grandes rasgos una definición de aquello que Tye considera una explicación representacional de la experiencia consciente. Pero existen diversos tipos de representacionalismo dependiendo de la naturaleza y del papel causal que se otorga al contenido fenoménico.

4.1.1. Clases de representacionalismo

Por una parte, podemos encontrar un representacionalismo *fuerte*. Éste considera que el carácter fenoménico es necesariamente idéntico al contenido representacional que encontramos bajo ciertas circunstancias. Por la otra, encontramos un representacionalismo *débil*, que considera que el carácter fenoménico superviene sobre el contenido representacional que encontramos bajo esas mismas condiciones. Ambos, a su vez, pueden ser considerados como *reductivos* (esto es, tanto el representacionalismo fuerte como el débil pueden considerar que el contenido representacional puede ser traducido *salva veritate* a lenguaje físico o funcional y, por tanto, el carácter fenoménico puede ser igualmente reducido a términos físicos (para el primero) o (para el segundo) sobrevenir a elementos físicos) o como *no-reductivos* (tanto para uno como para el otro, el contenido representacional, aunque pueda tener ciertos aspectos físicos o ser causado por elementos físicos, no puede reducirse a ellos y, por tanto, el contenido de las experiencias es irreductiblemente fenoménico en tanto subjetivo).

También podemos encontrar en la literatura autores que defienden un representacionalismo de corte *externista*. Éste considera que, dado que el contenido fenoménico de los estados conscientes de un agente psicológicamente dotado depende de la presencia de determinados elementos externos a dicho sujeto que permiten confeccionar la historia cognitiva de los individuos (por lo que puede darse el caso que dos individuos microfísica y funcionalmente idénticos pueden diferir en su contenido fenoménico dada una misma situación). Contrapuesto al anterior, encontramos el representacionalismo de corte *internista*, que afirma que nunca podrá darse el caso que dos individuos molecular y funcionalmente idénticos puedan

diferir en su contenido fenoménico en un mismo momento porque éste dependerá de la activación de los mismos estados internos.

Para algunos representacionalismos, el contenido de estos estados fenoménicos es *fregeano*. Esto es, el contenido representacional que caracteriza una experiencia fenoménica es individuado (al menos en parte) por los diferentes modos conceptualizables de presentación de dicha sensación. Para otros tipos de representacionalismo, el contenido será *no-fregeano*: es decir, será no-conceptual y se individualizará a partir de entidades objetivas y sus propiedades.

Una vez diferenciados los tipos de representacionalismo intentaremos exponer brevemente cuál sería aquel que más se asimila a la posición que actualmente está defendiendo Tye.

4.1.2. El intencionalismo de Michael Tye

Para Tye, el carácter fenoménico de las experiencias conscientes es idéntico al contenido representacional que un agente posee cuando se encuentra en un determinado estado sensorial y, a su vez, dicho contenido representacional equivale plenamente a la activación de una determinada estructura fisico-funcional del individuo que lo posee. Esto es, según Tye, el carácter fenoménico es funcionalmente dependiente de la distribución filo-geno-típica del sujeto (de su evolución como individuo perteneciente a una especie y como organismo evolucionado sujeto a un determinado desarrollo histórico-cognitivo) de las experiencias conscientes (Tye 1995: cap. 2, especialmente 56-67). (Para ser justos con el reduccionismo de Tye es conveniente advertir aquí que, aunque defiende una identificación de las representaciones que conforman el contenido de las experiencias fenoménicas con aquellos elementos genotípicos de identificación de particulares, Tye no defiende una reducción de los conceptos fenoménicos que tienen por objeto los tipos, por lo que dichas representaciones tendrían un estatuto contingente de verdad (Tye 2003a), claramente diferenciable de algunas aproximaciones, como la de Jackson (2000) o Chalmers (1996), que las consideraría como verdades *a priori*).

Del mismo modo, para Tye, las experiencias fenoménicas tendrían un contenido no-conceptual porque sería el modo cotidiano de referirse a las experiencias fenoménicas aquel en el que atendemos a las cualidades de los objetos que aparecen a nuestra experiencia y no el tratar con descripciones conceptualmente dependientes. Es decir, nosotros habitualmente nos relacionamos con cosas (del mundo) que poseen cualidades y no con los modos en que son descritas, por lo que también es más fácil suponer que nuestras experiencias conscientes se refieren a dichas cualidades y no a las descripciones que de ellas hacemos (Tye 2000: 60-61). (Sería de resaltar que, aunque actualmente Tye es no-conceptualista puro, en Tye (1995) mantuvo una posición híbrida en la cual consideraba la posibilidad de que, en determinadas circunstancias, el contenido representacional respondiera a un modo determinado de presentación de un objeto y no a las características relevantes que éste posee, aceptando cierto contenido fregeano en las experiencias conscientes. Tye abandonó esta posición intermedia persuadido por Peacocke (2001), que defiende como posibilidad que la experiencia fenoménica

represente al unísono tanto propiedades espaciales objetivas como características objetuales subjetivamente dependientes).

Una vez dicho lo anterior, el contenido sólo puede ser individualizado para Tye de un modo externista, ya que sería el contexto particular con el cual tiene contacto el agente consciente el que estructuraría los componentes de los cuales depende funcionalmente el contenido. Sólo así podríamos explicar cómo dos individuos física-funcionalmente idénticos pueden diferir respecto a su contenido fenoménico dada una misma experiencia consciente, posibilitando así una explicación plausible para los problemas del espectro invertido y la variación estándar. Un caso paradigmático lo encontraremos en la percepción del color.¹

Vistas así las cosas, podemos concluir que la concepción sostenida actualmente por Michael Tye es un representacionalismo fuerte, reductivo, no-conceptual y de corte externista. Pero entonces, ¿por qué decir que su teoría es intencionalista? Como hemos dicho anteriormente, según el representacionalismo de Tye, el contenido de las experiencias fenoménicas lo componen aquellas representaciones de los objetos que tienen por contenido las propiedades objetivas relevantes para el sujeto según las circunstancias externas en las que la experiencia está teniendo lugar. Pero que el contenido sea respecto de objetos (o mejor dicho, represente ciertas cualidades específicas de los objetos) del mundo no nos indica nada más que la existencia de un elemento objetivo que caracteriza e individualiza un determinado estado subjetivo: la vivencia característica de una determinada experiencia fenoménica. Pero no encontramos en esta formulación una referencia clara al modo en que el contenido fenoménico tiene que ser intencional y, por eso mismo, se vuelve necesario atender a la caracterización que Tye hace de este tipo de contenido.

4.2. La teoría PANIC y su aplicación sobre el contenido fenoménico

Llegados a este punto, se hace necesario explicar su teoría PANIC. Este término (seña de identidad desde que Tye lo acuñó en su *Ten problems of consciousness*) es la sigla resultante de unificar los puntos más importantes de su posición, que afirma que el contenido (*Content*) fenoménico de las experiencias conscientes tiene una naturaleza intencional (*Intentional*), es no-conceptual (*Nonconceptual*), es abstracto (*Abstract*) y está preparado (*Poised*)² para causar o provocar otros tipos de estado mental con carácter fenoménico y/o proposicional (Tye 2000: 45-142).

¹ La solución que Tye ofrece al problema de la variación estándar y del espectro invertido será analizada en el apartado tercero. En la sección 4.2.2. analizaremos los argumentos con los que Tye caracteriza el contenido de las experiencias de color.

² El término inglés *poised* tiene un sentido de condicionalidad. Visto de este modo, la caracterización que Tye está haciendo del contenido fenoménico indica que toda experiencia debe estar lista, preparada o dispuesta para permitir la aparición de nuevos estados mentales, sean éstos proposicionales (como en el caso de creencias) o fenoménicos (como en el caso de estados de ánimo). En un texto previo (Cf. Colomina 2008b), traduje esta expresión por el término castellano 'equilibrado'. Sin embargo, esta noción impide apreciar en toda su dimensión la intención que subyace a la propuesta de Tye. Precisamente con la intención de salvaguardar el matiz original otorgado por nuestro autor, me adhiero a la propuesta introducida por Martínez 2008 de traducir este término por la palabra castellana 'preparado' (en el sentido de estar listo para permitir la aparición de otros estados).

4.2.1. La teoría PANIC de Tye

Lo primero que podemos decir de esta teoría respecto de la naturaleza del carácter fenoménico de la experiencia consciente es que considera que éste es transparente. Esto es, que aquello que se siente son las características públicamente observables de las cosas y que cualquiera puede sentir, que dichas características pertenecen al objeto y no a nuestras propias vivencias y que (a pesar de ser una cualidad introspectiva) nuestras vivencias del carácter fenoménico no son respecto de nuestra propia experiencia sino respecto del objeto externo en tanto poseedor de unas ciertas cualidades públicamente observables que encajan en (o pueden ser captadas por) nuestros sistemas sensitivos. Aquí es donde se demuestra la intencionalidad del contenido fenoménico, dado que parece ser el caso que, a pesar de que en primera instancia el contenido parece ser introspectivo (esto es, las características propias del contenido fenoménico parecen ser características internas a la propia experiencia), en realidad, el contenido de las experiencias conscientes se relaciona directamente con las características que poseen los objetos de experiencia.

Al afirmar que el contenido fenoménico es no-conceptual (como ya hemos dicho) Tye quiere significar que éste no requiere de ninguna información previa ni de ninguna red de significados que establezcan una relación categórica entre la representación y su experiencia. Y del mismo modo, cuando dice que el contenido es abstracto, Tye quiere significar que no existen objetos reales concretos que formen parte de las representaciones, sino que éstas serían una abstracción a partir de las cualidades relevantes de objetos de experiencia. El requisito de preparado que debe cumplir el contenido fenoménico se refiere a que éste debe ser de tal modo que sea capaz de, independientemente de la experiencia de la que sea contenido, favorecer a que su poseedor encuentre las salidas cognitivas adecuadas que permitan responder a los estímulos experienciales (esto es, debe ser capaz de crear la creencia adecuada, satisfacer adecuadamente el deseo, etc.).

Necesitamos una aclaración. Aunque Tye la denomine intencionalismo, su posición no tiene nada que ver con intenciones. Simplemente quiere decir que la conciencia siempre representa (o puede mal-representar) objetos, incluidos objetos abstractos sin existencia física. En este punto su posición coincide con el análisis de la conciencia perceptiva de Armstrong (1969). Pero su posición no puede asimilarse a la de otros que, como Anscombe (1957) o Crane (1998), como veremos, defienden la intencionalidad como marco de lo mental. En este sentido, y siguiendo a Byrne (2001), podríamos caracterizar la posición de Tye como un representacionalismo intermodal no-restringido (véase sección 2.2.1.4.). Éste difiere del representacionalismo intramodal en que, a diferencia del primero, el segundo no acepta diferencias en la experiencia fenoménica de los sujetos (negando casos como la Variación Estándar, por ejemplo). Pero también diferirá del representacionalismo intermodal restringido, porque éste tampoco aceptará como posibilidad la existencia de representaciones sobre, por ejemplo, las sensaciones corporales.³

³ He presentado un análisis de la solución de Tye al problema de la Variación Estándar en Colomina (2008a) y se amplía en el apartado tercero de este mismo capítulo. Un análisis de la posición de Tye respecto de las sensaciones

4.2.2. El contenido de la conciencia y los problemas respecto de la percepción

Hemos dicho que para Tye el contenido fenoménico de las experiencias conscientes es una representación preparada, abstracta, no-conceptual e intencionalmente dirigida que se refiere a las cualidades más relevantes para el sujeto de los objetos del mundo, pero de tal modo que configura una representación que en sí misma no es un objeto. Podemos llegar a entender que, de este modo, el contenido representacional requiere siempre la presencia de objetos externos como requisito de adquisición de cierto contenido. Pero entonces, ¿cómo podemos explicar la presencia de experiencias fenoménicas, que habitualmente se consideran puramente introspectivas y subjetivas si no parece posible referirlas a un objeto externo que las acompañe y ocasione? Se hace necesario en este momento matizar la teoría de Tye sobre el contenido.

Observemos, primero, qué ocurre cuando atendemos al color (aunque es transmisible también a otras cualidades ‘secundarias’, como el sabor, el olor, etc.) desde la teoría de Tye. La experiencia de color es uno de esos casos-límite que ponen en aprietos la teoría fenoménica anteriormente esbozada porque, como se ha dicho, la cualidad de color parece ser tan sólo subjetiva, dado que tiende a ser definida (al menos parcialmente) desde su aspecto puramente fenoménico, dado que parece plausible que (a pesar de que el color tiene una realidad objetiva clara) ante una misma experiencia de color pueda darse el caso que a dos individuos les parezca que están viendo dos colores diferentes (la psicología popular está repleta de casos similares). ¿Cómo explicar, entonces, este tipo de casos desde una teoría representacional? Desde una teoría clásica, como es la de Locke, que algo aparezca como, por ejemplo, rojo es interpretado como que ese algo tiene una disposición a ser visto como rojo por sujetos perceptores normales en situaciones adecuadas. La conciencia introspectiva es entendida aquí como un ‘sentido interno’. Pero esta interpretación adolece de una dificultad básica: es circular, dado que explicar la cualidad de algo como siendo rojo a partir de una disposición para ser visto como rojo, presupone que debe ser rojo y que los sujetos disponemos de la capacidad necesaria para captarlo. (Una teoría reciente que acaba por asimilar esta dificultad es la de Jackson (1977) y (2001), que vuelve a una interpretación lockeana al considerar que el color es una cualidad intrínseca y relativa al observador que los objetos ‘despiertan’ en los sujetos, pero sería de suponer que el color permanece en el objeto aunque nosotros dejemos de contemplarlos, por lo que parece implausible que la experiencia de color sea observacionalmente dependiente). Pero desde una teoría como la de Tye, el color se presenta como algo objetivo, como una cualidad invariable de los objetos, y el sistema receptor humano tiene por función captar esa cualidad real y objetiva de los objetos, independientemente de que anteriormente se haya experimentado y de si se dispone de un concepto previo sobre el cual encuadrarlo. Por supuesto, que realmente se llegue a captar dependerá de si la experiencia es relevante para los intereses del sujeto. (Éstas serían básicamente las diferencias existentes entre un representacionalismo dependiente como el

corporales se presenta en la sección 4.2.3. de este capítulo, donde se recogen y amplían argumentos presentes en Colomina (2008b).

de Locke y un representacionalismo no-dependiente como el de Tye, tal y como lo expusimos en el capítulo 2).

Podría objetarse que, algunas veces, parecen darse casos en que un mismo objeto adquiere tonalidades diferentes de color según el modo de iluminación o según el cambio de alguna variable similar. Pero esto no sería un problema si aceptamos que ‘el color de un objeto’ es una expresión vaga (aunque ello lo involucrara con una problemática todavía más compleja, como es la falta de luminosidad, como mostramos en el apartado 1.4.9.) y que puede responder a diferentes descripciones según el momento de presentación de la experiencia (adaptándose, así, a las circunstancias [*biting the bullet*]), pero lo que realmente contaría como el color de dicho objeto, para Tye, sería el espectro de onda lumínica que nuestro sistema visual es capaz de captar. Del mismo modo, el sabor de un determinado objeto no dependerá del sujeto perceptor (aunque sí dependerá de la relevancia que tenga para sus intereses cognitivos), sino de las sustancias que estimulen las papilas gustativas adecuadas y el olor de un objeto dependerá, a su vez, de las sustancias que estimulen las glándulas olfativas pertinentes.

Todavía podría objetársele que las cualidades de color, sabor, olor, etc., son subjetivamente dependientes y que pertenecen al punto de vista del sujeto perceptor, que dependen de cómo se le presenten a él los hechos. Por ejemplo, que el color que presenta un tomate sea rojo dependería de que a nosotros se nos presente como rojo. Pero este caso sería perfectamente salvable porque, como sabemos, no funcionan así las cosas. Este ejemplo sería similar a este otro: ‘está mañana no he visto ningún alcoyano’, pero te has cruzado con Eburne (que es de Alcoy), por lo que la emisión se vuelve falsa, independientemente de que no supiéramos que Eburne nació en Alcoy, porque la verdad de este tipo de enunciados no depende de las creencias particulares de los agentes. Del mismo modo, que el color del tomate sea rojo no depende del modo en que subjetivamente se presente al agente perceptor, sino de que realmente el tomate sea rojo.

Hemos dicho que casos como los de la experiencia del color y otras cualidades secundarias son casos-límite porque, aunque es cierto que a menudo tenemos la impresión de que su contenido fenoménico dependa por completo del sujeto perceptor, siempre podemos acudir a algún sistema objetivo de identificación y catalogación de sensaciones. Por ejemplo, en nuestro caso del color, sabemos que nuestro sistema cognitivo visual está capacitado para distinguir cerca de diez millones de colores diferentes y que esto, en última instancia, dependerá del espectro de visión que nuestro ojo es capaz de captar, pero sólo disponemos de un escaso número de nombres (alrededor de unos 4000) para referirnos a todos ellos (como indican los diccionarios específicos del color, como Maerz y Paul 1950). A pesar de ello, podemos llegar a diferenciarlos todos si hacemos uso de máquinas capaces de discriminar los diferentes espectros lumínicos o acudimos a cualquiera de los innumerables catálogos Pantone. (Un caso similar ocurriría con la sensación de temperatura. Puede dar la impresión de que podemos diferir en el contenido fenoménico de nuestras sensaciones térmicas (nuestro tener frío o calor), pero existen aparatos como los termómetros (o máquinas capaces de medir la velocidad del movimiento molecular de los objetos) capaces de indicar objetivamente la temperatura de un objeto). Pero

existen casos, como el de la experiencia de dolor, en los que la referencia a un objeto externo al que referir nuestro contenido fenoménico se hace mucho más complicada y todavía se ponen peor las cosas cuando tenemos que vérnoslas con los modos objetivos que permitirían demostrar la existencia real de dicho contenido fenoménico.⁴

El problema principal, aquí, con la teoría de Tye es confirmar que el contenido representacional de las experiencias fenoménicas sea realmente inteligible sólo desde una perspectiva no-conceptual (no-fregeana o milliana). ¿Por qué? Principalmente porque la tesis que defiende Tye supondría un cierre sensitivo y cognitivo del mismo modo que lo promueven ciertos filósofos funcionalistas (véase sección 2.1.2.2.), dado que lo que viene a insinuar Tye es que lo que realmente cuenta como una experiencia fenoménica es aquello que es capaz de encajar en nuestro sistema sensitivo, esto es, que nuestra subestructura sea capaz de captar una determinada entrada como un cierto estímulo y que éste sea tratado como una sensación determinada capaz de ocasionar una determinada salida conductual adecuada, volviéndose en su contra problemas tan habituales del internismo funcionalista como el problema de la sobredeterminación (que un mismo estímulo experiencial pueda ser explicado por más de una causa. Una posible salida a este problema se presenta en el apartado 6.1.2.) o el dualismo de propiedades (al presuponer la presencia de algo, aunque se postule como físico, en lo que se debe confiar que salga bien a un nivel subdoxástico que determinará qué se puede considerar como sensación experiencial y de qué tipo de experiencia se trata, como si estuviéramos hablando de un homúnculo).

Otro problema recaería, de nuevo, sobre la presunta incoherencia de la posición representacionista. En el caso del representacionismo fuerte de Tye, afirmar que el contenido fenoménico de una experiencia es idéntico a su contenido representacional (y dado que éste ha sido identificado con una proposición respecto de las características relevantes del objeto al cual se dirige dicha experiencia, el tener esa experiencia), supondría afirmar la identidad entre el contenido de una experiencia y una proposición. ¿Cómo debemos entender el hecho de que ‘cómo qué es’ tener una experiencia pueda ser identificado con una proposición si una proposición siempre es un objeto abstracto mientras que el ‘cómo qué es’ tener una experiencia es la cualidad de dicha experiencia (un término simple)? Parece que este intento vuelve un tanto inconsistentes las tesis representacionistas fuertes. Y aunque siempre podríamos entender su posición como afirmando que para que una experiencia tenga cierto carácter fenoménico sólo debe tener cierto contenido intencional, negando así la posibilidad de identificar carácter con contenido, pero permitiendo afirmar que el hecho de que una experiencia tenga un carácter fenoménico es idéntico a que posea un determinado contenido, el problema, sin embargo, volvería a ser el mismo, porque el contenido es nuevamente entendido de manera proposicional, contradiciendo su propia pretensión de entender todo contenido de experiencia como no-conceptual.

Este inconveniente podría salvarse si, muchas veces, interpretáramos ese contenido pre-conceptual que Tye presupone a las experiencias fenoménicas desde una perspectiva personal

⁴ Cf. Colomina 2008b y sección siguiente.

que tenga en cuenta la denominada actitud intencional, que tenga presente la posibilidad de que todo aquello que sentimos puede ser interpretado desde el punto de vista de las respuestas habituales que los seres conscientes tenemos ante ciertos estímulos (como lo interpretamos en el apartado 6.2.2.).

Pero ello no quiere decir que el contenido de toda experiencia deba ser entendido como proposicional. Nada más lejos de la realidad. El contenido no-conceptual existe (es la carga informacional que nuestros sentidos internos y externos reciben cuando son estimulados por ciertas sensaciones externas, es nuestra distribución biológica evolucionada), sólo que en última instancia debemos tener presente que dicho contenido puede ser conceptualizado (esto es, es conceptualizable) porque puede ser ‘ajustado’ de tal modo que cierta información dada puede ser entendida como una experiencia concreta. Y, por supuesto, todo ello porque disponemos de una distribución subpersonal (a nivel estructural) que prioriza nuestra carga simbólica respecto de la funcional, evitando así abrazar estrechamente el representacionalismo de lo mental y optando por una posición impura que permita una ventaja cognitiva a la hora de explicar las experiencias fenoménicas conscientes (como se matizará en el capítulo final).

Como hemos mostrado desde el principio, uno de los principales motivos que el representacionalismo de Tye tiene para considerar que todos los estados conscientes son intencionales es que éstos parecen cumplir el requisito de transparencia. Es decir, que la reflexión introspectiva sobre una experiencia consciente tan sólo muestra los objetos externos hacia los cuales se dirige dicha experiencia y sus cualidades relevantes para el sujeto, pero nunca muestra las cualidades intrínsecas que caracterizan la experiencia misma. Aunque no pretendemos cuestionar la plausibilidad de esta concepción ahora, si parece un tanto controvertida cuando apelamos a casos concretos en los que los objetos a los que se refieren nuestras experiencias no existen.

Un caso paradigmático que encontramos en los estudios sobre la experiencia fenoménica son los casos de visión borrosa (presentados en la sección 1.4.8.). En los casos de visión borrosa, nosotros podemos observar cualquier objeto de nuestro entorno pero éste siempre aparecerá como borroso. Según Tye, nuestra experiencia representará el objeto de nuestra visión y sus cualidades relevantes, pero nunca nos dirá nada respecto de aquellas cualidades internas que permiten dicha representación (que dependerá por completo del modo en que esté distribuido nuestro sistema cognitivo). Pero parece posible afirmar que, en determinados casos, un individuo puede llegar a ser consciente de algo más que del modo en que representamos el mundo o de los objetos que de él nos presenta nuestra propia experiencia. Es decir, en casos de visión borrosa, cuando a alguien se le presenta algo como borroso, no tiene por qué representarse necesariamente el mundo como borroso. La respuesta tal vez sea simplemente que en determinado momento ve las cosas borrosas, pero no tiene por qué representárselas de ese modo, *pace* las objeciones de Tye (2000: cap. 4). Aunque siempre, no debemos olvidarlo, el sujeto vería el mundo en estos casos cómo si fuese realmente borroso. Esto es, que la visión borrosa y otros fenómenos simplemente pueden formar parte del carácter intencional de la misma experiencia perceptiva, ya que este tipo de fenómenos pueden ser

resultado del modo intencional particular de la misma experiencia, puede ser simplemente el modo particular en que un individuo es consciente del mundo, dependiendo de la propia agudeza de su percepción visual (y no sólo de su contenido) la determinación del carácter fenoménico de una experiencia concreta. Simplemente, un individuo tal es capaz de contemplar el mundo de manera diferente a todos aquellos que contemplamos el mundo de manera no-borrosa.

La simple apelación a la transparencia y la introspección parece escasa para explicar la intencionalidad presente en las experiencias. El problema fundamental aquí es que, en última instancia, el que algo sea representado de un cierto modo puede ser un fenómeno familiar para un determinado individuo; que sea éste el modo habitual en que él contempla el mundo. Que una experiencia pueda ser errónea, o falsa, no quiere decir, en este caso, que el individuo esté equivocado. Simplemente, parece ser que es así como funciona la conciencia a la hora de presentar la experiencia fenoménica (Travis 2004). Para saber por qué el carácter de una experiencia está determinado por la naturaleza intencional, requerimos de una explicación de la conexión intrínsecamente establecida entre lo intencional y lo fenoménico, y no sólo de la simple apelación a aquello aparentemente obvio.

Siguiendo las sugerencias de Crane (2007), sería conveniente comenzar analizando la relación existente entre cómo aparece el mundo al sujeto y el modo en cómo una experiencia lo representa. Estas dos nociones parecen tener un mismo centro común. Ambas describen lo que la idea de intencionalidad captura en un nivel abstracto: aquello que es representado en la mente, un objeto, su contenido. Esta perspectiva, la de cómo aparece el mundo, implica siempre la presencia de un sujeto de experiencia y lo que se representa es el contenido de dicha experiencia. Así, podemos entender que el hecho de que las cosas pueden aparecer de una cierta manera contenga cierto tipo de intencionalidad (derivada, si queremos), la intencionalidad presente en la experiencia misma, salvando el carácter fenoménico de la experiencia sin necesidad de apelar a representaciones y permitiendo una explicación de la unidad de nuestra conciencia y nuestra mente bajo la noción genérica de 'sujeto de experiencia' (o 'punto de vista', si preferimos) experiencialmente situado (una noción en la que insistiremos en el capítulo final).

Vistas así las cosas, en lugar de insistir (como hace el representacionalismo de Tye) en que todo objeto o contenido experiencial tiene un único aspecto evaluativo, parece más plausible suponer que dicho contenido puede presentarse de modos muy diferentes, muchos de ellos de modo afectivo más que evaluativo (como explicamos en el capítulo final). Ello explicaría por qué un mismo aspecto afectivo puede englobar distintas experiencias y diferentes estados mentales, permitiendo una generalización del estado concreto en cuestión al generalizar el aspecto afectivo experimentado inicialmente (como veremos en el siguiente apartado).

4.2.3. El contenido de la conciencia en el estudio de las sensaciones corporales, las emociones y los estados de ánimo

Cuando Tye analiza las experiencias de dolor, comienza por analizar el modo en que nuestro cerebro consigue establecer una imagen de nuestro cuerpo porque, según Tye, el objeto del cual extraemos esa cualidad que representamos como dolor es aquella parte de nuestro cuerpo en tanto que representado por el cerebro que, en el momento justo, nos está doliendo. Que nuestro cerebro sea el creador de una imagen de nuestro cuerpo parece ser la mejor explicación del modo en cómo nosotros lo sentimos si tenemos en cuenta los últimos avances en neurobiología y fisiología cerebral, nos dice Tye. Estos estudios suelen investigar casos concretos en los cuales una cierta alteración en la distribución o funcionalidad cerebral (ya sea por un estado alterado químicamente o a partir de una enfermedad fisiológica o una lesión cerebral) muestra cómo la imagen total de nuestro cuerpo se ve alterada o, incluso, se pierde. Casos paradigmáticos los encontramos en los escritos de Oliver Sacks.

A partir de estos estudios, parece ser el caso que nuestro cerebro consigue articular una única sensación global, una imagen unificada de nuestro cuerpo como sentido al unísono a partir del modo en que sentimos sus diferentes partes. Aquí la introspección sirve de poco, porque la única ayuda que puede ofrecernos es indicarnos aquella sensación que estamos teniendo en un único momento determinado en una parte concreta de nuestro cuerpo y lo que aquí estaría defendiendo Tye es la posibilidad de que se dé una imagen total del cuerpo, lo que presupone una imagen representacional del modo en que sentimos nuestro cuerpo y el espacio que ocupa a través del tiempo. Si conseguimos dar correcta cuenta de este hecho, si conseguimos establecer una adecuada experiencia general de nuestro cuerpo, tendremos las puertas abiertas a la explicación de la unidad fenoménica de nuestras sensaciones corporales (un problema crucial, como indicábamos en el apartado 1.4.14., en el estudio de la conciencia. Una solución la ofrecemos en el apartado 6.2.1.4.).

Hemos dicho que, para Tye, nuestra imagen corporal aparece como unificada, que las sensaciones respecto de las partes de nuestro cuerpo aparecen como unificadas pero, ¿en qué consiste esta unidad? ¿Cómo explicar que las diferentes sensaciones que tenemos y que sentimos a lo largo del tiempo y en partes concretas de nuestro cuerpo formen parte, todas, del mismo cuerpo y de la misma experiencia fenoménica?

Volvamos al caso del dolor. Según la psicología popular, parece que la experiencia de dolor es necesaria y esencialmente privada y subjetiva, que nosotros no podemos sentir vuestros dolores y vosotros no podéis sentir los nuestros. Parece, entonces, que la explicación más plausible sería considerar el dolor como una sensación entendida como un objeto mental (incluso podríamos decir que tiene una cierta funcionalidad biológica, a la hora de llamar la atención del sujeto del dolor sobre un determinado problema físico y/o herida). Pero esta interpretación se encuentra con serias dificultades cuando tiene que aplicarse a determinadas partes del cuerpo. Por ejemplo, tiene muchos problemas cuando tiene que localizarse en algún lugar concreto porque, ¿dónde podemos encontrar el dolor? Si, por ejemplo, tenemos un dolor en la espalda, parece plausible considerar que el dolor está en la espalda. Pero si hemos

considerado que el dolor es un objeto mental no podemos a la vez decir que se encuentra en nuestra espalda, porque deberá pertenecer a nuestra mente y no a un objeto material.

El sentido común todavía podría insinuar que, en casos de experiencia fenoménica, como es el caso del dolor, que éste se encuentre ‘en’, por ejemplo, nuestra espalda debe ser entendido en un sentido especial, en un sentido causal. Esto es, que el dolor y la sensación de dolor de nuestra espalda son una y la misma cosa, independientemente de que no podamos localizar espacialmente el dolor en nuestra espalda. Pero esto tampoco parece muy plausible por dos motivos: el primero, porque de ser así, sólo podríamos conocerlo de modo *a posteriori* a partir de la investigación empírica (algo que no sería posible porque tan sólo podríamos realizarla sobre el cuerpo y nunca sobre la mente), y el segundo, porque el sentido de ‘en’ que emplean es un tanto rebuscado, totalmente alejado del uso ordinario de ‘en’.

Pero no le funcionan las cosas mejor a otros representacionistas que consideran plausible la posibilidad de que el dolor se encuentre ‘en’ el cuerpo porque es una representación que se proyecta sobre él (Maund 2006; Aydede 2006) porque esta respuesta implicaría que es la representación misma lo que realmente se está sintiendo y no el dolor mismo que sentimos. Además, esta concepción también implicaría situaciones divertidas (aunque imposibles) como la siguiente: imaginemos que tenemos un dolor en nuestro pulgar (nos hemos golpeado el dedo con un martillo al intentar clavar un clavo) y que, como acto reflejo, introducimos nuestro dedo en nuestra boca con la intención de calmar el dolor que sentimos. Según esta interpretación, el dolor en tanto representación se proyectaría ‘en’ nuestro dedo, pero como nuestro dedo está ‘en’ nuestra boca, donde realmente tendríamos el dolor es ‘en’ nuestra boca y no ‘en’ nuestro dedo (de Jackson 1977 podemos inferir también esta conclusión). Pero además, esta tesis también fallaría al atender a problemas relativos a ciertas sensaciones de dolor que plantean casos como los ‘miembros fantasma’: muchas personas que han perdido alguno de los miembros de su cuerpo, pongamos por caso una pierna, insisten a menudo que sienten dolor (el caso del picor sería similar) en esa pierna que les falta. Si consideramos que la representación se proyecta ‘en’ la parte del cuerpo afectada, entonces algo falla en este argumento al faltar también la parte del cuerpo ‘en’ la que, se supone, estaría localizada la sensación de dolor. Otro problema con el que se encontraría es con la posibilidad de equivocarnos con la atribución del propio dolor: supongamos que tenemos un dolor en nuestro brazo izquierdo; según esta interpretación, el dolor sería una representación que proyectamos sobre nuestro brazo. Pero imaginemos que nuestro dolor en el brazo no se debe a ninguna herida o anomalía en nuestro brazo, sino a un problema derivado de una disfunción cardíaca (como en el caso de los infartos). Entonces, nuestro dolor en el brazo izquierdo no estaría localizado en nuestro brazo, sino que realmente debería localizarse en nuestro corazón. (Que en un momento concreto el sujeto pudiera equivocarse a la hora de atribuir el dolor o a la hora de identificar dónde tiene dolor no es lo problemático. La dificultad reside, precisamente, en que no se puede localizar de manera estricta el dolor, esto es, no existe una espacialidad física concreta del dolor si estamos presuponiendo tan sólo su existencia *en* la mente). A partir de los anteriores obstáculos, esta interpretación debe ser descartada.

A pesar de las dificultades, Tye considera posible afrontar el problema de la atribución de dolor desde una perspectiva representacional, pero en su caso, como en el caso del color y otras cualidades secundarias, la argumentación a seguir no debe ser respecto a los objetos, sino respecto de las descripciones de los mismos como conjunto representado de cualidades, unas descripciones que deben entenderse como no-conceptuales y, por tanto, previas a todo concepto y experiencia y dependiente de nuestro aparato sensorial de captación de sensaciones. En pocas palabras, dependiente de nuestra distribución sub-doxástica. Desde este punto de vista, nuestra experiencia de dolor debe referir directamente a cualidades, pero no a cualidades de nuestra experiencia misma (hemos dicho anteriormente que aquí la introspección no ayuda), sino a cualidades desagradables en nuestro cuerpo que pertenecen a las regiones afectadas en tanto que somos capaces de captarlas a partir de nuestro aparato sensitivo. Por tanto, debemos atender a esas anomalías corporales y a sus cualidades en tanto que experimentadas (cumpliendo así con el requisito de transparencia) que somos capaces de captar a partir de nuestros mecanismos nerviosos y neuronales (como indica la evidencia neurocientífica). Lo que comprobamos con esto es que las experiencias relativas al propio cuerpo tienen dos dimensiones: por una parte, una componente sensitiva (que se refiere al modo en que se siente) y una componente afectiva (relativa al modo en que esa experiencia se evalúa). (Aunque es cierto que existen casos en que puede perderse la componente afectiva, como en casos de disociación reactiva o en casos de cerebro escindido [*split brain*], donde el paciente pierde toda experiencia aversiva respecto del dolor, aunque sea capaz de seguir sintiéndolo, como indicamos en la sección 6.2.1.4.).

Y el contenido de la experiencia representacional del dolor, en sus dos dimensiones, deberá entenderse como no-conceptual porque ésta no responde a experiencias cognitivas, sino que tiene que ver con elementos pre-conceptuales asociados a nuestro aparato perceptor, a nuestra propia estructura sub-personal. Ello permite explicar el comportamiento de ciertos individuos, como por ejemplo el de un niño de dos meses de edad al que se le ha dado a probar chocolate. En este caso, el niño responde de tal forma que su conducta nos dice que quiere más, porque el chocolate sabe bien, lo que permitiría explicar el comportamiento del niño de 'querer más'. La experiencia gustativa del niño representa un cierto sabor (componente sensitivo) y el niño experimenta que ese sabor 'es bueno' (componente afectivo) sin necesidad de una experimentación previa o de la presencia de algún elemento conceptual o cognitivo. Del mismo modo actuaría un niño pequeño que experimenta por primera vez el dolor: ante la experiencia de dolor (elemento sensitivo), el niño tenderá a evitarla por experimentarla como negativa (elemento afectivo). Esto mismo lo que viene a afirmarnos es que el dolor (como otras experiencias fenoménicas) no es tan sólo una cierta clase de experiencia, sino que también es una cierta cualidad en tanto que experiencialmente representada.

Desde este punto de vista, según Tye, podemos salvar los obstáculos que encontraba la anterior teoría. En el caso de explicar la experiencia de dolor relativa a 'miembros fantasma', el sujeto representaría una cierta experiencia como de dolor y ese dolor sería experimentado como malo, porque al no requerir de un objeto al que referirse podría explicarse su presencia independientemente de la existencia real del miembro. Del mismo modo, ante casos como el

dolor que sentiríamos en el brazo izquierdo a partir de un infarto, como tampoco necesitamos localizar el dolor en un lugar determinado, bastaría tan sólo con afirmar que representamos cierta experiencia como de dolor y que ese dolor se experimentaría como malo. La diferencia aquí es que no se requeriría localizar en nuestra imagen corporal el lugar exacto del dolor, sino que bastaría con saber que estas experiencias representan cualidades diferentes que pueden referirse a partes diferentes de nuestro cuerpo. Esto es, según Tye, que las experiencias fenoménicas representan cualidades (en sentido sensorial y afectivo) de modo *abstracto* y que éstas pueden referirse a determinadas partes del cuerpo. Pero las partes del cuerpo a las que correspondan dependerían ya de una investigación empírica (por eso podríamos tildarla de *externista*), porque que el dolor pueda situarse en un brazo o en una pierna dependerá de que realmente el miembro que tengamos dañado sea un brazo o una pierna.

Atendamos ahora a las experiencias emocionales y a los estados anímicos. Desde la teoría de Tye, las emociones y los estados de ánimo se dirigen a cosas y a personas (esto es verdadero no sólo en el caso de las emociones humanas, sino también de aquellas emociones experimentadas por criaturas de otras especies) de dos modos diferentes (Tye 2008: 31): a través de las experiencias perceptivas o a través de los pensamientos que se tienen acerca de esas mismas experiencias. Pero estos pensamientos no deben entenderse como asociados a creencias. Tan sólo deben entenderse como pensamientos cuyos poseedores pueden tener por verdaderos en el momento adecuado. Cuando sentimos una emoción, la experiencia sensitiva puede representar el objeto hacia el cual se dirige de muy variados modos. Estas experiencias emocionales son muy similares a las descritas en el caso del dolor. Para Tye, las experiencias emocionales son experiencias dirigidas hacia objetos que están fuera del cuerpo del sujeto perceptor y que los representan como teniendo tanto características sensitivas como evaluativas y, también, son causadas (o meramente se siguen) de ciertas disfunciones o anomalías corporales. Los objetos externos son inicialmente representados por un estado perceptivo o un pensamiento no-perceptivo. En otras palabras, el estado emocional representa E, a través de sus características evaluativas, como causando (en algunos casos, como acompañado por) R. La vivencia fenoménica de una experiencia emocional será, entonces, fijada por su contenido representacional: las experiencias emocionales que son representacionalmente sentidas deben ser también iguales a las fenoménicamente sentidas.

Pongamos un ejemplo. Hemos estado tan absortos en nuestra lectura de la tarde que no nos hemos dado cuenta de que ha anochecido. Decidimos volver a casa, pero antes tenemos que pasar por el parking a recoger el coche. Nos encontramos en el parking. Está oscuro. La poca iluminación del aparcamiento no es suficiente para alumbrar toda la zona. Además, casualmente la farola que se encuentra justo encima de mi coche está estropeada. Oímos un ruido extraño. Al girarnos, observamos una extraña silueta, amenazante, que se aproxima. Nos sentimos estremecidos por una angustiosa sensación de miedo que nos obliga a acelerar nuestro paso, sacar presurosamente las llaves de nuestro bolsillo y a meternos rápidamente en el coche. Una vez dentro, a salvo, observamos que aquello de lo que habíamos sentido tanto pánico, en realidad, era un compañero de trabajo que también sale tarde del despacho. ¿Qué nos ha

sucedido? Para Tye, nuestra experiencia se dirige hacia el ruido y la silueta. Esta experiencia ha sido representada como peligrosa, lo que ha provocado nuestra emoción de miedo y nuestra reacción (natural) de tender a ponernos a salvo. Esto no quiere decir que lo que se está experimentando es el miedo, porque el miedo no puede tener contenido conceptual, sino que esa experiencia representa algo como peligroso, lo que ocasiona nuestra emoción de miedo, nuestro pensamiento no-conceptual.

Podríamos objetar que es implausible suponer que las cosas pueden ser peligrosas en sí mismas, sin que nadie las conciba como peligrosas, sin una experiencia sensitiva de las cosas y de sus cualidades. Que las cualidades son sentidas primero y que, tan sólo más tarde, son evaluadas. Pero Tye rechaza este punto de vista porque, nos dice, esas primeras identificaciones son algo natural en el hombre, esas facultades cognitivas pertenecen a nuestra propia estructura sub-personal, como muestran los últimos avances neurobiológicos. Por un lado, como los estudios de Damasio nos demuestran, existe una importante conexión entre la amígdala y la experiencia emocional. Y esta conexión tiene un inmediato valor de supervivencia porque el cerebro está predispuesto a detectar peligros. Por el otro, experimentos realizados sobre pacientes con el cerebro escindido demuestran, también, que existe una conexión entre la experiencia emocional y regiones primitivas identificadas en el subcórte del cerebro.

En otras palabras, desde el punto de vista de Tye, el objetivo básico de la experiencia emocional es ofrecernos el justo valor de las cosas y el impacto que tienen sobre nuestro cuerpo. No necesitamos tan sólo sentir las cosas, sino que, además, necesitamos también conocer el impacto que sobre nosotros provocan, lo que pueden o lo que podrían producir en nosotros. Por tanto, la naturaleza ha instaurado en nosotros (y en otras criaturas) detectores capaces de captar el valor de las cosas que permiten conducirnos de manera adecuada hacia nuestra supervivencia. Pero Tye no está sugiriendo que en todo caso en que sentimos una emoción nosotros sabemos de qué emoción se trata. Sino que lo que nos dice es que nosotros (al igual que otras especies) somos capaces de responder a determinados estímulos externos con una determinada emoción, según el caso. (Por supuesto, el error siempre es posible. No hay razones obvias para negar que las experiencias emocionales se dirijan de manera incorrecta a objetos equivocados, como cuando, por ejemplo, mi miedo se dirige equivocadamente a otras causas diferentes de sus causas reales).

Cuando nosotros tenemos determinadas experiencias, nos dice Tye, podemos decir que éstas pueden despertar ciertas emociones. Si sentimos miedo es porque estamos teniendo una experiencia que representa aspectos de nuestro propio estado corporal y, al mismo tiempo, representa un aspecto evaluativo del objeto o persona hacia la que se dirige nuestro miedo, estableciendo de este modo una conexión causal entre ambos. Nuestra experiencia representa, así, un complejo entramado causal de propiedades tanto internas como externas. Y toda experiencia que represente de manera adecuada dicho complejo causal de propiedades será considerada como una experiencia de miedo. Podemos así decir, según Tye, que este complejo causal es experimentalmente representado como una sensación de miedo. Entendidas de este modo, las emociones serán disposiciones o tendencias a sentir experiencias de un cierto tipo

dependiendo de los estímulos externos adecuados. Las experiencias emocionales, como las entiende Tye, son estados disposicionales que pueden llegar a ser racionalmente justificados (Tye 2008: 40-41).

Podemos objetar que no toda experiencia emocional tiene un contenido no-conceptual. A veces, las experiencias emocionales pueden tener un contenido conceptual. De este modo, cuando siento enfado porque el árbitro ha impuesto un castigo inapropiado a mi equipo favorito, existe un contenido conceptual que acompaña a mi estado. Y la justificación de estar en dicho estado puede adquirir muy distintas formas. La justificación, por tanto, en palabras de Tye, consiste en poder dar razones para entender dicho estado. Este tipo de justificaciones requieren de premisas y conclusiones que tan sólo son posibles mediante la apelación a contenidos conceptuales, por lo que parece que Tye vuelve a contradecirse. Porque, para Tye, este tipo de justificaciones están disponibles para cualquier tipo de experiencia emocional. Sobre la base de las experiencias perceptivas y corporales, una experiencia de enfado, entonces, se generará a partir de procesos justificativos de esta clase. Por supuesto, las experiencias emocionales no requieren siempre la presencia de emociones primarias que justifiquen la adquisición de emociones más elaboradas, sino que en el caso de la presencia de emociones secundarias, éstas tendrían un contenido compuesto tanto por elementos conceptuales como no-conceptuales, unidos a partir de un nexo causal.⁵ Pero los componentes conceptuales, que parecen en este caso ser esenciales a toda emoción secundaria, no deben entenderse más que en un sentido evaluativo. Es decir, para Tye, la presencia en una emoción de una componente evaluativa presupondría, ya, la existencia de una componente evaluativa que uniría, de modo causal, el contenido de una experiencia con nuestra propia estructura cognitiva. Este carácter evaluativo de las experiencias emocionales, por tanto, para Tye, sería parte del carácter fenoménico de las emociones y de los estados de ánimo.

Desde el punto de vista de Tye, por tanto, la tesis de la transparencia también aparece como obvia en el caso de las experiencias emocionales, por dos motivos principales. En primer lugar, porque no existiría ningún lugar concreto hacia el cual se dirige nuestra atención cuando tenemos una experiencia emocional. En segundo, porque, para Tye, algunas de las cualidades representadas por las experiencias emocionales son siempre apreciadas de modo apropiado por el sujeto perceptor. Desde este punto de vista, parece claro que la transparencia permitiría diferenciar entre casos de experiencia emocional y casos de experiencia corporal, mostrando así la complejidad existente en los casos de experiencias emocionales (Tye 2008: 47).

Pero en casos de experiencias emocionales también podría plantearse el mismo problema que anteriormente identificamos en la transparencia de las experiencias corporales. Retomemos el ejemplo del parking. El ruido y la silueta provocaron nuestro miedo, pero realmente nuestro miedo era infundado: nuestra representación de los mismos era errónea. Pero eso no quiere decir que no existiera dicha sensación de miedo. De hecho, tuvimos dicha experiencia, por lo que había algo más en la representación del hecho que su mismo contenido.

⁵ Una teoría similar respecto de las emociones, como hemos mostrado en la sección 2.1.2.2., es la defendida en Prinz (2004).

En ese momento determinado, a pesar de que la experiencia no parecía contener nada fuera de lugar, nosotros sentimos miedo porque en ese momento era el modo particular en el que éramos conscientes de la situación, más allá de su contenido real.

Pongamos un ejemplo todavía más claro. En un divertido episodio, nuestro amigo Bart Simpson comienza a tener un miedo visceral a los payasos. ¿Qué provoca ese miedo en Bart? Objetivamente hablando, parece que nada externo, nada situado en los payasos, parece provocar dicho miedo. Entonces, ¿cómo puede ser que Bart represente como peligroso (lo que provocaría el miedo) a un payaso? Parece que la tesis de Tye es insostenible. Podríamos, sin embargo, afirmar que hay algo más allá de la simple representación, como hemos hecho anteriormente. Bart parece que, más allá de toda representación, ve en el momento determinado de observar un payaso un peligro inminente, esto es, dada la experiencia concreta de contemplar un payaso, Bart es consciente de algo más que del modo en que representa el mundo, es consciente de tener un miedo (visceral, si se quiere) hacia los payasos, de encontrarlos peligrosos. Bart parece contemplar el mundo de manera diferente a todos aquellos que creemos que los payasos son inofensivos.

La simple apelación a la transparencia y la introspección, de nuevo, parece escasa para explicar la intencionalidad presente en las experiencias. Parece que, para saber por qué el carácter de una experiencia está determinado por la naturaleza intencional, requerimos de una explicación de la conexión intrínsecamente establecida entre lo intencional y lo fenoménico, y no sólo de la simple apelación a aquello que tan sólo sentimos en un momento determinado. (De todos modos, la apelación a la transparencia no es el único intento representacionista de explicar el carácter fenoménico de las experiencias conscientes. Byrne (2001) intenta una aproximación diferente al problema pero, por motivos que no analizaremos aquí,⁶ sus argumentos fracasan).

En nuestro anterior ejemplo del parking, nuestro miedo no tenía un objeto claro, pero no carecía de contenido. Tal vez fuera causado por el extraño ruido que oímos, o por la extraña forma de la silueta, o por mi estado anímico tendente a la paranoia (seguramente se deba a una mezcla de todos estos, y otros no-identificados, factores). Esa experiencia de miedo hizo que, en su momento, actuara de forma diferente a como habitualmente lo hago (que caminara más rápido, que me metiera rápidamente la mano en el bolsillo en busca de las llaves del coche, que impulsivamente saltara dentro del habitáculo del vehículo... incluso que mi corazón se acelerara y mis glándulas sudoríparas expulsaran una mayor cantidad de fluido que la habitual). Podemos decir que mi mundo se transformó en el mismo preciso instante en que la sensación de miedo apareció. El resto de mis experiencias estaban condicionadas a esa inicial sensación de miedo, transformando mi mundo. Del mismo modo ocurre con ciertos estados de ánimo: la depresión, por ejemplo. La persona deprimida está sumida en tal estado que el resto de sus experiencias se ven modificadas por ese inicial estado de ánimo. Parece como si 'su mundo' fuese diferente de 'nuestro mundo'.

⁶ Esos argumentos han sido mostrados ya en la sección 2.2.1.4.

Wittgenstein nos dijo que “el mundo del hombre feliz es diferente del mundo del hombre infeliz” (*Tractatus* 6.43). Como hemos indicado, parece ser que Wittgenstein estaba en lo cierto. Desde nuestro punto de vista, el contenido de determinadas experiencias no puede ser entendido de otra manera que desde un determinado punto de vista, el nuestro propio. La noción de cómo me aparecen a mí las cosas cobra fuerza si tenemos en cuenta los ejemplos aportados: tanto en el caso de nuestro miedo visceral como en el caso de la persona deprimida, la emoción y el estado de ánimo determinado tiñen (en la literatura al respecto suele hablarse de ‘colorido’) el resto de nuestras experiencias, haciendo que todas ellas sean vistas desde esa misma perspectiva afectiva. Esta conclusión nos ha permitido tomar distancia de la noción de transparencia tan presente en teorías representacionistas como las de Tye.

De todos modos, parece ser que los últimos avances en neurología apuestan por el importante papel de la amígdala en la producción de estados emocionales y anímicos, por lo que los partidarios de posiciones representacionales consideran que una tendencia cada vez mayor hacia el intencionalismo *à la* Tye permitirá una más rápida explicación reductiva del fenómeno de la conciencia. Posiblemente estos hechos sean ciertos. Pero ello no significa que una tendencia reductiva pueda ser más fructífera. Por un lado, aunque el contenido emocional pueda considerarse (en sentido amplio) como una serie de relaciones causales (en términos informacionales o de co-variación, por ejemplo) con el entorno, parece ser que no existe ese mismo consenso respecto de las bases físicas causantes de la conciencia, un debate en el que nada ayuda el intento naturalista de explicación de la intencionalidad. Por el otro lado, aunque la perspectiva fisicalista fuera cierta, esto no implicaría la verdad del reduccionismo respecto de la conciencia, porque todavía quedaría por explicar aquello que hace intencionales cierto tipo de experiencias conscientes. La teoría reduccionista de Tye intenta salir de esta situación analizando las cualidades específicas del contenido intencional de todo estado consciente. El problema principal, como hemos visto, reside en que Tye sustituye paulatinamente el intento de explicación de la intencionalidad por una explicación en términos de contenido representacional, algo que no es deseable después de todo.

4.3. Los obstáculos de una teoría representacional fuerte respecto de la conciencia

Como venimos comentando desde el principio, la teoría de Michael Tye es explícitamente una teoría representacional de la mente. Según este tipo de teorías, todo estado mental es considerado como un estado representacional. Es habitual encontrar, también, como es el caso de Tye, teorías representacionales que consideran que todo estado representacional es un hecho natural.⁷ Como hemos estado defendiendo, la teoría de Tye podría caber perfectamente en esta categoría, dado que si, como defiende, todo estado mental es un estado representacional y todo estado representacional es un estado natural (pensemos tan sólo en la afirmación que sostiene Tye de que el color simplemente es la representación adecuada de las

⁷ Al realizar esta afirmación estoy pensando especialmente en Dretske 1981 y en Millikan 1984, que consideran que todo estado representacional es un hecho respecto de las relaciones informacionalmente causales (el primero) y las relaciones teleológicas (la segunda) existentes entre los estados físicos del cerebro y los estados del mundo.

condiciones de reflectancia lumínica en un sistema visual normal), entonces todo estado mental será, también, un hecho natural.

Una primera objeción (si se quiere, de carácter general) que puede plantearse a toda teoría representacional es que, en última instancia, existen estados mentales (como los estados fenoménicos) que no parecen ser estados representacionales. Las razones esgrimidas para defender esta posición son, principalmente, de dos tipos. Por un lado, encontramos aquellos que consideran que ciertos estados fenoménicos tienen un carácter diferente al del resto de los estados cualitativos (es decir, existen estados fenoménicos diferentes dependiendo de la experiencia consciente concreta experimentada). Por otro lado, encontramos a aquellos que consideran que cierto tipo de estados mentales tienen, además, cierto carácter fenoménico que los hace diferir de los estados mentales con contenido representacional (es decir, existen estados mentales con carácter relacional, como las actitudes proposicionales, que difieren de estados mentales con contenido fenoménico no-relacional).

A pesar de estas críticas generales, los defensores de las teorías representacionales de la mente consideran que los estados mentales con carácter fenoménico son representacionales, después de todo.

4.3.1. Revisando la teoría PANIC

Un caso concreto, como venimos analizando, lo encontramos en la propuesta de Tye. En su posición, el carácter fenoménico de un estado consciente es una clase de estado con contenido representacional PANIC. Dentro de PANIC, como hemos ya indicado, un estado con carácter fenoménico está preparado para provocar las actitudes proposicionales adecuadas, se refiere a objetos abstractos, es intencional y es no-conceptual. Desde este punto de vista, entonces, dos estados mentales tendrán diferente carácter fenoménico si tienen diferente contenido PANIC y sólo si un estado mental tiene un contenido PANIC podrá ser considerado como un estado con carácter fenoménico y la diferencia, precisamente, entre la existencia de estados con carácter fenoménico y otro tipo de estado es que los primeros representarán determinadas propiedades que no representarán los segundos. Decir que todo estado fenoménico es representacional significa afirmar que, al menos en el caso de Tye, todo estado fenoménico es la instanciación por parte de un sujeto de la representación de un objeto particular como teniendo unas determinadas características. Por lo tanto, en una teoría como la de Tye, que todo estado mental sea tenido como representacional quiere decir que la existencia de toda propiedad mental (sea o no fenoménica) en un sujeto consciente afirma la existencia en dicho sujeto de determinadas propiedades de dicho estado mental en virtud de aquello que las mismas representan.

Dirigiéndonos ya concretamente contra la propuesta de Tye, podríamos decir que no parece plausible que siempre estén disponibles dichas representaciones. Esto es, podríamos dudar de que todo estado mental sea representacional. Por un motivo concreto. Si fuera así, como defiende Tye, entonces dicha teoría debería poder aportar un argumento que permitiera diferenciar entre aquellos estados mentales que poseen carácter fenoménico de aquellos estados

mentales que no lo poseen, dado que la diferencia entre ambos tipos debe también poder presentarse de un modo representacional. Esta diferencia debería poder presentarse cómo la diferencia existente entre las propiedades concretas que ambos tipos de estado representan. Pero parece posible presuponer que dichas propiedades podrían representarse de modo similar sin necesidad de suponer la existencia de un carácter fenoménico concreto (Kriegel 2002: 57), posibilitando la existencia de mundos carentes de propiedades fenoménicas (*mundos-zombie*), por lo que la teoría representacional de Tye se enfrenta ya contra una poderosa objeción.⁸

De todos modos, existen otras objeciones concretas dirigidas contra la propuesta de Tye que merecen ser tenidas en cuenta de un modo separado y ser analizadas en detalle.

4.3.1.1. La condición de preparación del contenido fenoménico

¿Qué es para un contenido estar preparado? Según Tye,

“esta condición es esencialmente funcional. La idea es que las experiencias y los sentimientos, en tanto portadores de carácter fenoménico, juegan un cierto papel funcional. Ellas... están predispuestas (y disponibles) a realizar un impacto directo sobre nuestras creencias y/o nuestros deseos” (Tye 2000: 62).

Es decir, para Tye, toda experiencia fenoménica debe ser capaz de modificar o instanciar creencias y deseos. Por tanto, la diferencia fundamental entre estados con carácter fenoménico y otros tipos de estados mentales, precisamente, es que los primeros están preparados para formar creencias y deseos, mientras que los segundos no lo están.

Contra esta condición de preparación podríamos objetar que existe un fundamental error categórico (Kriegel 2002: 60): parece que Tye confunde la posibilidad de que un determinado estado fenoménico nos conduzca a estados proposicionales con la posibilidad de que dicho estado sea representacional. Es decir, que un estado fenoménico cumpla un papel funcional no quiere decir que se le pueda atribuir un contenido representacional, sino que es el vehículo que permitiría conducirnos hasta la posesión de un estado mental con contenido representacional. Y, aquí, hablar de contenido preparado no hace más que empeorar la situación, dado que parece confundir el estado que está preparado para instanciar actitudes proposicionales con el contenido mismo. Desde este punto de vista, parece que no son las experiencias fenoménicas las que tienen un contenido representacional preparado para instanciar creencias y deseos, sino el estado funcional carente de contenido representacional que permitiría llevarnos a la adquisición de estados mentales con verdadero estado representacional, como serían las creencias y los deseos. Porque si una diferencia que podemos encontrar entre un estado fenoménico y un estado no-fenoménico es la posibilidad de que el primero cumpla la condición de preparación y el segundo no la cumpla, entonces, la diferencia entre ambos estados no es una diferencia respecto de sus propiedades representacionales sino una diferencia respecto de sus propiedades funcionales (Kriegel 2002: 61).

⁸ De todos modos, como mostramos en la sección 4.3.5.2., Tye tiene una respuesta plausible al problema que contra el fisicalismo plantea la posibilidad de la existencia de duplicados físicos carentes de fenomenología.

Pero esta diferencia no sólo debe referirse a actitudes proposicionales. Tye considera que existen otros estados diferentes de los estados con contenido fenoménico que difieren con estos en su condición de estar preparados. Pensemos, nos dice Tye, en cierto tipo de estados (llamémoslos estados-V, por referirse a estados visuales) con contenido no-conceptual que no están preparados, por ejemplo, estados visuales que de modo no-conceptual representan cambios en la intensidad lumínica (Tye 2000: 64-68). Por tanto, para Tye, existen estados visuales que tienen un contenido representacional no-conceptual que, sin embargo, no están preparados. Estos estados diferirán de los estados fenoménicos en que los segundos permiten instanciar creencias y deseos mientras que los primeros no. Pero, entonces, aquí tendríamos el mismo problema, dado que la diferencia entre ambos no puede consistir en una diferencia en su contenido representacional, sino que más bien tiene que presuponerse una diferencia en sus propiedades funcionales, con lo que volvemos a encontrar serias objeciones para considerar plausible la teoría representacional de Tye.

Tye pretende defenderse de objeciones de este tipo a partir de la distinción entre dos modos de entender la condición de preparación. Parece posible que cuando un estado mental con un cierto contenido juega un determinado papel funcional, éste puede o bien preparar la instanciación de determinadas creencias y deseos (por tener la propiedad concreta de hacerlo) o bien puede conducir a un estado capaz de instanciar dichas creencias y dichos deseos (por tener la propiedad de conducir a un estado capaz de realizar dicho papel funcional. Cf. Tye 2002: 63-64). Pero, de nuevo, esta solución no parece hacer justicia al representacionalismo, dado que parece trivializarlo de tal modo que lo hace inconsistente (Kriegel 2002: 62), porque parece que en un momento determinado un mismo estado podría tener o no tener contenido representacional según que la conveniencia de la situación lo requiriera.

Pensemos tan sólo en un caso concreto: el problema de la visión ciega (*blindsight*). Para Tye, los estados de un sujeto con visión ciega no son fenoménicamente conscientes porque no poseen la condición de preparación:

“Ellos [los sujetos con visión ciega] no tienen conciencia fenoménica de la región ciega. Aquello de lo que carecen, según la teoría PANIC, es de la presencia del estado representacional no-conceptual apropiadamente preparado... aquel contenido que está preparado para establecer una diferencia directa en la creencia” (Tye 2000: 63).

Parece que para Tye ésta sería una diferencia en el contenido representacional de los estados cualitativos involucrados. Pensemos por un momento en un estado fenoménico plenamente PANIC (por ejemplo, mi estado de ver un tomate rojo) y el mismo estado en un sujeto con visión ciega. Si la única diferencia entre ellos es que el primero está preparado, pero no así el segundo, la diferencia entre ambos es una diferencia en su estado representacional (dado que los dos estados no representarán las mismas características). Pero también existirá una diferencia respecto de sus características funcionales, ya que el primero permitirá que dicho estado se vuelva consciente, pero no lo permitirá el segundo estado. El problema aquí con la interpretación de Tye es que vuelve ambigua la apelación a la condición de preparación como

criterio de diferenciación entre estados fenoménicos y estados no-fenoménicos. Dado que, de nuevo, parece que la condición de preparación de Tye no parece necesaria para marcar las diferencias entre dichos estados, porque la diferencia concreta en su contenido representacional puede llegar a confundirse con una diferencia en sus características funcionales. Pero más preocupante todavía, si cabe, es la dificultad identificada en la teoría de Tye a la hora de tener que diferenciar entre contenidos conscientes y aquellos que no lo son, dado que al no recaer esta posibilidad en la condición de preparación, como hemos visto, entonces el representacionalismo de Tye parece inconsistente al no poder diferenciar entre estados conscientes y estados sin conciencia.

De todos modos, además de la posibilidad de realizar críticas contra la necesidad de la condición de preparación, también se podrían formular críticas contra su suficiencia (Byrne 2002: 11). Volvamos a considerar el caso de la visión ciega. Un sujeto con visión ciega puede ser capaz de evitar obstáculos interpuestos a lo largo de un pasillo a pesar de tener dañada la parte subcortical de su cerebro (pensamos en un sujeto como el descrito en Mestre *et al.* 1992). A pesar de no tener un acceso consciente a la información que recoge su (dañado) campo visual, parece que ciertas partes de su sistema cerebral son capaces de procesar información de tal modo que juegue un papel causal en su experiencia visual, por lo que el sujeto parece disponer de cuasi-experiencias. Dentro del sistema de Tye, podemos suponer también que estas cuasi-experiencias son abstractas y son no-conceptuales. Pero jamás podremos suponer que están preparadas, dado que está ausente toda posibilidad de un acceso consciente a dichas percepciones. Pero, a pesar de ello, esto no le impide al sujeto con visión ciega esquivar adecuadamente los obstáculos en su camino. Por lo tanto, a pesar de que aparentemente no sea posible, podríamos suponer que realmente estas cuasi-experiencias sí están preparadas porque interfieren en la instanciación de determinadas creencias y determinados deseos que permiten al sujeto proseguir su camino a pesar de la existencia de obstáculos físicos. Por supuesto, dicho sujeto no puede expresar verbalmente estas creencias, pero deberemos presuponer que éstas se manifiestan a través de su comportamiento no-verbal, demostrando así que la condición de preparación, por sí sola, tampoco es suficiente para diferenciar entre estados conscientes y estados no-conscientes, porque como estamos diciendo, a pesar de que pueden interferir causalmente en la conducta del sujeto con visión ciega, éste no dispone de un acceso a las cuasi-experiencias a pesar de que están preparadas para instanciar creencias y deseos (Byrne 2002: 12).

4.3.1.2. Objetos abstractos y contenido fenoménico

Tye afirma que el contenido representacional de los estados mentales es abstracto. Con esto, lo que Tye quiere decir es que

“una condición [para la existencia de estados mentales] es que el contenido relevante sea abstracto, esto es, que dicho contenido no se refiera a objetos particulares concretos. Esto es especialmente esencial para los casos de experiencias alucinatorias...” (Tye 2000: 62).

Esto es, para Tye, la experiencia fenoménica tan sólo puede representar objetos abstractos. El problema principal, de nuevo, con esta condición de la teoría representacional de Tye es que la abstracción parece que no puede ser el requisito necesario que permita diferenciar los estados fenoménicos de aquellos otros estados que no son fenoménicos. Una de las razones principales, como hemos visto, de que Tye afirme la necesidad de la abstracción en el contenido fenoménico es poder rendir cuentas de las experiencias alucinatorias (puedo alucinar a la hora de estar viendo un ovni, por ejemplo) y los casos de representaciones erróneas de experiencias fenoménicas (puedo representarme erróneamente el color de una camisa como malva, cuando en realidad es morada, por ejemplo). Pero, como sabemos, las creencias pueden ser igualmente mal-representadas. Por tanto, si el contenido proposicional de una creencia (que es paradigmáticamente abstracto) puede ser también representado erróneamente, entonces la teoría de Tye parece encontrarse con serias dificultades a la hora de diferenciar aquellos estados con contenido fenoménico de aquellos estados que carecen de él (Kriegel 2002: 58).

Del mismo modo, como indicará por ejemplo Peacocke (1992: 77), el contenido fenoménico puede ser perfectamente dependiente de objetos, dado que puede depender de proto-proposiciones (aunque estas proto-proposiciones sí pueden ser abstractas).⁹ Volvemos a encontrarnos con el mismo problema. Si tanto los estados fenoménicos como los no-fenoménicos pueden ser dependientes de objetos, entonces no existe la necesidad de que su contenido sea abstracto. Y lo que es peor, entonces se dificulta sobremanera la posibilidad de poder diferenciar los unos de los otros.

4.3.1.3. Contenido conceptual frente a contenido no-conceptual en Tye

Para Tye, la diferenciación entre estados con carácter fenoménico y estados que no poseen dicho carácter depende de las condiciones de preparación y de no-conceptualidad del contenido de las experiencias fenoménicas. Como él mismo nos dice,

“la teoría PANIC no implica que las creencias tengan un carácter fenoménico. Un contenido será clasificado como fenoménico sólo si es no-conceptual y está preparado. Las creencias no son no-conceptuales y no están adecuadamente preparadas” (Tye 1995: 138).

En este sentido,

“decir que un contenido mental es no-conceptual significa que el sujeto no necesita poseer ningún tipo de concepto que nosotros, como teóricos, requerimos cuando evaluamos las condiciones de corrección de dicho contenido” (Tye 2000: 62).

⁹ Vamos a presuponer que la opción de Peacocke es dependiente de objetos por ser la interpretación actualmente dominante, aunque nos reservamos la posibilidad de atender a otras opciones de interpretación (como la de Davies 1996: 310) que consideran que la apelación a protoproposiciones por parte de Peacocke no incluye ningún tipo de relación con objetos. En este caso, las protoproposiciones propuestas por Peacocke se tratan de un contenido constituido por una captación no-proposicional de relaciones muy básicas, como la simetría en torno a un determinado eje o escenario (la otra noción básica que le permite a un sujeto reconocer los conceptos perceptivos), que guarda cierta semejanza estructural con una proposición, en tanto que se refiere a la predicación de propiedades de objetos existentes en su entorno.

Como hemos comentado anteriormente, la condición de preparación parece que no permite diferenciar entre estos tipos de estados, por lo que, entonces, deberá depender de la condición de no-conceptualidad la posibilidad de distinguirlos adecuadamente. Principalmente porque las actitudes proposicionales tienen un contenido conceptual.

Según Tye, el contenido de las experiencias fenoménicas debe ser no-conceptual porque un sujeto podría experimentar hoy un determinado tono de rojo y no reconocer el mismo tono de rojo si vuelve a experimentarlo unos pocos días después (Tye 2000: 61). Pero, podemos objetar, el hecho de que un sujeto pueda experimentar un determinado tono de rojo sin necesidad de poseer el concepto apropiado donde colocar dicha experiencia no quiere decir que el contenido de los estados fenoménicos implique necesariamente que éste sea no-conceptual. Si realmente el contenido de un estado con carácter fenoménico es necesariamente no-conceptual, entonces, toda vez que el sujeto se encuentre frente a una muestra de ese determinado tono de rojo, el sujeto debe poder experimentarlo, porque Tye insiste que dicha experiencia dependería de factores externos, en este caso concreto la presencia de las características físicas relativas al mismo color. Pero sabemos que existen casos en los que, a pesar de que el sujeto se encuentra frente a una muestra concreta de ese tono de rojo, el sujeto no es capaz de experimentarlo como un rojo con ese tono concreto, sino con otro diferente, por lo que la afirmación de la no-conceptualidad del contenido fenoménico debe ser rechazada.

Una dificultad todavía mayor a la hora de evaluar la pertinencia del requisito de no-conceptualidad de las experiencias fenoménicas es una ambigüedad presente a la hora de presentar y analizar esa característica. Cuando Tye habla de contenido no-conceptual, a menudo esa caracterización se aplica a los estados mismos, pero en ocasiones se aplica directamente al contenido. Por lo visto, parece que Tye presenta dos interpretaciones diferentes a la hora de hablar de no-conceptualidad. A partir de la interpretación aplicada a los estados, las experiencias mismas tienen contenido no-conceptual. Pero a partir de la interpretación aplicada al contenido, es el contenido el que tiene un carácter no-conceptual. Además, de la ambigüedad presente en el análisis de Tye, ambas interpretaciones tienen, por separado, sus propias dificultades, como intentaremos mostrar a continuación.

Habitualmente, la diferenciación entre conceptualidad y no-conceptualidad suele aplicarse a los estados y no a los contenidos. El propio Tye así introduce esta distinción para aplicarla a estados sub-doxásticos:

“Las sensaciones perceptivas son experimentadas en el sistema conceptual sin que ellas mismas formen parte del sistema. Son estados no-conceptuales o sub-doxásticos” (Tye 1995: 104).

Como hemos comentado ya en el capítulo introductorio, la distinción mencionada especifica que un determinado estado mental con un cierto contenido es no-conceptual si y sólo si alguien que se encuentra en dicho estado no posee ninguno de los conceptos que representan el contenido de dicho estado (Crane 1992: 143). Como aplicada al contenido, sin embargo, la distinción afirma que una determinada representación es no-conceptual si y sólo si es el contenido de algún estado no-conceptual. Pero esto no quiere decir que el contenido no-

conceptual sea una clase especial de contenido, sino tan sólo que si una experiencia perceptiva tiene contenido no-conceptual, entonces la representación que es contenido de dicha experiencia puede ser perfectamente familiar y conocida (al igual que lo son los contenidos de las actitudes proposicionales).

Según la interpretación del contenido, el contenido de una representación, entonces, será no-conceptual si y sólo si no es un pensamiento fregeano (esto es, siempre que no sea una proposición cuyo contenido esté compuesto por sentidos diferentes). Desde esta perspectiva, entonces, el contenido de una actitud proposicional será conceptual (es decir, fregeano), mientras que el contenido de una percepción será no-conceptual (Peacocke 1992: 88).

Así las cosas, Tye acepta la posibilidad de que el contenido de las experiencias fenoménicas sea no-conceptual en este sentido. Pero, como estamos diciendo, el argumento oficial de Tye se refiere a la interpretación que aplica la no-conceptualidad a los estados. Veamos.

Dentro del sistema PANIC, toda actitud proposicional es conceptual y todo estado fenoménico es no-conceptual. La explicación de por qué las actitudes proposicionales carecen de carácter fenoménico, entonces, se explica a partir de la carencia que éstas tienen de contenido no-conceptual. Pero aquí encontramos una dificultad. Parece posible que las creencias conscientes también posean un carácter fenoménico (Cf. Block 1995: 230; Chalmers 1996: 9-10; Peacocke 1999: 205-206; Horgan and Tieson 2002). PANIC afirma que las creencias carecen de contenido no-conceptual a partir de su contingencia, pero como no es un hecho que las creencias de manera esencial carezcan de contenido no-conceptual parece que no es capaz de explicar la diferencia entre creencias y experiencias fenoménicas y que no es capaz de explicar por qué las experiencias tienen un carácter fenoménico.

De todos modos, como venimos insistiendo, a menudo Tye emplea una interpretación de no-conceptual aplicable a contenidos. Según esta interpretación, el contenido no-conceptual es un contenido en el que no están implicados conceptos ni descripciones. En este sentido, el contenido de las actitudes proposicionales (pero también el contenido lingüístico) es fregeano, tiene que ver con descripciones y sentidos. La teoría PANIC intenta explicar por qué las creencias carecen de carácter fenoménico, a diferencia de las experiencias, a partir de la apelación a que las experiencias tienen vehículos representacionales que poseen una estructura topográfica (Tye 1995: 121; Tye 2000: 70-74), a diferencia de las creencias, que poseen vehículos representacionales basados en una estructura oracional y proposicional (Tye 1995: 100). Considerándolo de este modo, Tye es fregeano respecto del contenido de las actitudes proposicionales porque son constituyentes de las proposiciones que son contenido de estas actitudes (Cf. Tye 2000: 18), aunque es no-fregeano en cuanto al contenido de las experiencias fenoménicas (Tye 2000: 28). Tye cree, por tanto, que al menos en algunos casos como en el contenido de, por ejemplo, una creencia, podemos encontrar objetos o propiedades que sean partes constituyentes del mismo, aunque éstos se corresponderán con modos de presentación. Respecto de los conceptos fenoménicos, Tye piensa que “refieren directamente... No existe una separación entre el referente y el pensamiento del sujeto” (Tye 2001: 695). Si, con Tye,

aceptamos que aquello que caracterizará como fenoménico el carácter de una experiencia, a diferencia de una actitud proposicional, es la diferencia en el tipo de contenido, entonces, lo que a Tye todavía le falta es aportar una prueba que permita afirmar, primero, la necesidad de que toda experiencia sea no-conceptual y que toda actitud proposicional sea conceptual (esto es, que una creencia, por ejemplo, no pueda ser fenoménica) y, segundo, explicar aquello que permita fijar una tajante separación entre aquellos estados con carácter fenoménico y aquellos que carecen de dicho carácter, una explicación que Tye no consigue aportar en ningún momento.¹⁰

Por tanto, la diferencia entre estados fenoménicos y estados no-fenoménicos no puede residir en su no-conceptualidad. A menudo, estados mentales que tienen carácter fenoménico también tienen contenido no-conceptual. Pero no es la posesión de contenido no-conceptual lo que otorga a un estado su carácter de fenoménico. Podemos concluir, entonces, que también pueden existir estados fenoménicos que posean un contenido conceptual, y viceversa.

4.3.1.4. La necesidad de la intencionalidad del contenido fenoménico

Dentro de la interpretación de Tye de la conciencia fenoménica, el contenido de toda experiencia consciente debe ser intencional, en el sentido de ser intensional.

“Coincido con Chisholm y Jackson aquí en que ‘ x ve F como S ’ ante el apropiado F como un paradigma de estado fenomenológico. Esta localización es intensional por dos motivos. Primero, puede ser verdad que x vea F como S , aun en ausencia de dicho F en x . Segundo, puede ser verdad que x vea F como S sin que x vea a G como S , aunque F y G sean co-extensivos” (Tye 2000: 54-55).

Al caracterizar un contenido como intensional, normalmente quiere decirse que en contenido es proposicional (lingüísticamente establecido) de un modo que ni es posible su generalización existencial ni su sustitución *salva veritate*. Pero tal y como Tye ha caracterizado el contenido fenoménico, éste debe ser no-conceptual, por lo que no puede ser proposicional. Por tanto, ¿en qué sentido puede afirmar Tye que el contenido no-conceptual de las experiencias con carácter fenoménico es intensional? Lo que parece querer proponer Tye es que la intensionalidad de un determinado contenido fenoménico se refiere directamente a sus características lógicas. Es decir, que del hecho de que alguien perciba un determinado tono de color no pueda estipularse la existencia de un determinado tono de color (evitando así la generalización existencial) y que del hecho de que alguien perciba un determinado tono de color no se sigue que lo que está realmente percibiendo es un determinado porcentaje de espectro

¹⁰ Tye (2002b) se limita a apelar a Peacocke (1992) y a Stalnaker (1998a y 1998b) para defender su caracterización de no-conceptualidad del contenido de los estados fenoménicos, pero no aporta ningún argumento sólido ni innovador que permita reforzar su aplicación en la teoría PANIC. Y aunque es cierto que apela al argumento de la riqueza (Evans 1982: 229 y 125 n.9) aplicado a la interpretación aplicada a los estados que encontramos ya en Dretske (1981: cap. 6) y a la reformulación de Peacocke (1992: 67-68; Peacocke 1998) aplicado a la interpretación aplicada al contenido, hay buenas razones (Heck 2000) para considerar insuficiente la apelación a la riqueza del contenido para defender la propuesta de Tye. La mejor estrategia, en este caso, sería (según mi propio punto de vista) intentar mostrar la suficiencia, además de la necesidad, de la riqueza del contenido como requisito para diferenciar entre aquellos estados con carácter fenoménico y no-fenoménico, algo que los argumentos de Tye no son capaces de mostrar.

lumínico reflejado en su retina (evitando así la posibilidad de la sustitución *salva veritate* de propiedades).

Pero de toda esta explicación parece surgir un claro problema. Si el contenido de una experiencia con carácter fenoménico es intensional y también es intensional el contenido de estados mentales como las actitudes proposicionales. ¿Cómo podemos diferenciarlos? Tye no parece poder aportar una clara respuesta, por lo que la condición de intencionalidad se vuelve problemática.

Tye debería, entonces, reformular su teoría acerca del contenido de las experiencias con carácter fenoménico para intentar demostrar que éste cumple las condiciones PANIC que ha establecido para poder defender, así, un representacionalismo intencional tal y como persigue en los últimos años. De lo contrario, las razones aportadas serían suficientes para hacer peligrar su posición.

4.3.2. El análisis de Tye del problema de la Variación Estándar

Como dejamos constancia en el apartado 1.4.7., uno de los problemas más importantes referentes a los fenómenos de conciencia es el denominado problema de la Variación Estándar.¹¹ Éste se refiere a la posibilidad de que una misma sensación o experiencia pueda parecer diferente a dos individuos *a priori* perceptivamente *normales*. Si fuera cierta esta afirmación (es decir, si se demostrara la existencia de diferente contenido de conciencia en los dos individuos dada una misma experiencia), entonces podría llegar a afirmarse la existencia de un aspecto puramente subjetivo en la conciencia: el contenido fenoménico de una experiencia.

Desde nuestro punto de vista, esta afirmación, además de implausible, es falsa. Para intentar demostrarlo, atenderemos al argumento que recientemente aporta Michael Tye acerca del problema de la Variación Estándar, y que él ejemplifica exclusivamente mediante las experiencias de color (el denominado problema de los colores *reales* o *verdaderos*)¹² y a la reacción que ha producido tanto en reduccionistas como en dualistas. A partir de aquí, presentaremos nuestras propias objeciones y reservas a la aceptación de la (intuitivamente hablando) posibilidad de la subjetividad del contenido fenoménico.

4.3.2.1. Presentando el problema: el rompecabezas del “verdadero” azul

Parece ser que es comúnmente aceptado que la gente varía en su modo de tener experiencias según su edad, su sexo e, incluso, su estatus social. Por ejemplo, parece ser que cuando a alguien, digamos Vicente, se le presenta una cierta muestra de color, digamos la muestra de azul 527 de cualquier catálogo, a él le puede parecer como azul *puro* o *verdadero*;

¹¹ Pero, por supuesto, como hemos mostrado en el capítulo introductorio, no es el único. Conjuntamente con el que en esta sección analizaremos, podríamos considerar el problema del espectro invertido, el problema de los *qualia* ausentes o el problema de la luminosidad, entre otros (todos serán revisados a su debido tiempo). Véase los apartados 4.3.3. y 4.3.4.

¹² Para ser justos, debemos decir que Tye también se interesa por otros tipos de representaciones con contenido no-conceptual, la más relevante de las cuales es la experiencia de dolor, como hemos mostrado en la sección 4.2.3., pero también por las emociones y los estados de ánimo (véase la sección 4.2.4.).

mientras que a otra persona, digamos Miguel Ángel, le puede parecer que la misma muestra de azul 527 pertenece a un azul ligeramente teñido de verde. Intuitivamente, suponemos que al menos uno de los dos debe estar equivocado, dado que no podemos esperar que la muestra de azul 527 sea a la vez “verdadero” azul y azul ligeramente verdoso. Pero si tenemos en cuenta, como nos dice Tye apelando a un clásico trabajo (Block 1999), que no existen clases privilegiadas de receptores (Tye 2006d: 173) entonces, ¿cómo podemos afirmar que uno de los sujetos perceptores anteriormente mencionados se equivoca en su experiencia de la muestra de azul 527 mientras que el otro está en lo cierto?

Este problema, que habitualmente es tenido como un problema referente a la objetividad de tipo fisicalista del color o al modo en que nos representamos el carácter fenoménico de la experiencia de color, aparece como más general en el momento en que atendemos a la problemática de adoptar un punto de vista sobre el diferente papel causal que una determinada experiencia de color tiene sobre dos sujetos distintos pero perceptivamente normales.

Podríamos decir, como haría un fisicalista, que tanto Vicente como Miguel Ángel están equivocados, que la muestra de azul 527 realmente es otro tipo de azul del que ellos perciben, dado que lo que realmente tiene color es el objeto exterior y físico, y éste implementaría el mismo estado mental en los dos sujetos, por lo que ambos deberían ver el mismo (tipo de) color. Si no es así, se debe a que ambos están en un error a la hora de percibir el color de la muestra. Pero, como hemos visto, sí que es posible que dos sujetos tengan reacciones diferentes ante una misma experiencia.

También podríamos decir que nada de lo que podemos observar tiene color, como haría un dualista, ya que el color es algo que nosotros proyectamos sobre el mundo, por lo que es posible que ambos estén en lo cierto. Pero si aceptamos que nada en mi cabeza tiene color (Byrne and Tye 2006),¹³ entonces esta posibilidad debe ser deseada.

Entonces, si ambos no tienen porqué estar equivocados ni ambos tienen porqué llevar razón, ¿en qué sentido podemos decir que un sujeto experimenta una sensación de color?

4.3.2.2. Una primera solución: adaptándose a las circunstancias

Según Tye, los sujetos perceptores hemos sido diseñados por la Madre Naturaleza de tal modo que disponemos de un sistema visual compuesto por varios elementos (entre los que resaltan la retina y el nervio óptico) que permite detectar el color, del mismo modo que podemos decir que los termostatos han sido diseñados para detectar la temperatura a partir de la incorporación en su estructura de un sensor de calor (Tye 2006d: 174). Visto de este modo, podríamos decir que aquellas criaturas que (como los humanos) hemos sido dotados con un sistema de identificación de color, estamos mejor dotados que aquellas criaturas que no disponen de dicho sistema. Pero una vez aceptado que todos nosotros (en condiciones normales) disponemos del mismo sistema perceptivo, ¿no sería incongruente afirmar (como hace Tye) que

¹³ Para una crítica desde una posición dualista a este mismo argumento sobre la representación, véase Pautz (forthcoming a).

en el caso de la Variación Estándar uno de los dos perceptores podría estar equivocado mientras que el otro llevaría razón?

Lo sería si Tye fuera estrictamente funcionalista, dado que asimilaría la causalidad de los estados mentales respecto de su contenido experiencial de un modo estrecho. Pero para una posición representacionista como la que defiende nuestro autor, en el caso del color, nuestro sistema cerebral experiencial no fue diseñado por la evolución de modo cerrado, sino que fue creado de tal modo que pudiera adaptarse individualmente según las condiciones externas (lumínicas, principalmente) en las que estuviera inserto y según el modo particular en que este individuo concreto se desarrollara (Tye 2006d: 175). Podríamos decir que ninguno de ellos tendría porqué estar equivocado o llevar razón.

Si esto es así, entonces podemos observar que la solución de Tye pasaría por aceptar la posibilidad de la diferente experimentación por parte de dos sujetos perceptores en condiciones *normales* apelando a la posible diferencia existente en su sistema visual. Esto les provocaría representarse de modo diferente la misma muestra de color según el modo particular en que su sistema perceptivo se hubiera adaptado a las circunstancias externas, algo que les permitiría poder tener experiencias de color diferentes.¹⁴

Pero entonces, podríamos objetar que dados dos individuos con un sistema *normal* de detección del color y en condiciones *normales* deberían experimentar la misma sensación ante la experiencia perceptiva de la muestra de azul 527, dado que lo que estaría aceptando la solución aportada por Tye es que todo color humanamente experimentable de tipo azul 527 sería aquello que un perceptor con un sistema normal de detección del color y en condiciones normales podría percibir como azul 527 (Cohen, Hardin and McLaughlin 2006: 337) y que todo caso en el que un perceptor no experimente la sensación como ‘experiencia perceptiva de azul 527’ será debido a que realmente no dispone de un sistema perceptivo normal, sino que su sistema perceptivo es a-normal, dado que si dispusiera de un sistema normal, en condiciones normales percibiría el color como azul 527 (*Ibid.*, 337), dado que es altamente implausible que puedan existir diferencias en el contenido fenoménico de dos individuos con un sistema normal de detección de color en condiciones normales ante la misma muestra (*Ibid.*, 338).

Tye aceptará esta crítica, dado que cuando él se refiere a la posibilidad de que un individuo represente incorrectamente la sensación ante una muestra de color, explícitamente estará aceptando la posibilidad de que el sistema de detección de color de dicho individuo sea a-normal dadas las condiciones en las que se encuentre. Porque aquello que puede indicar su malfuncionamiento depende, en gran medida, de una gran cantidad posible de pequeños detalles evolutivos que permitirían hacer pasar un sistema a-normal de detección de color por uno normal (Byrne and Hilbert 2007: 89), con lo que se alterarían las condiciones normales bajo las que podría funcionar la solución de Tye. Por tanto, dado que es extremadamente difícil poder

¹⁴ Como podemos observar, la solución aportada aquí por Tye, que en un primer momento parece asimilar la *actitud intencional* de Daniel C. Dennett, en realidad estaría mucho más cercana a posiciones más realistas, como la de Dretske (1995). Por ejemplo, una de las diferencias más importantes con la postura de Dennett es el diferente punto de vista que ambos tienen respecto de la teleología de la evolución, como se verá más adelante. Pero una correcta atención a esta comparación excedería con creces el espacio de esta nota.

concretar qué podemos considerar como un sistema normal de percepción, dado que biológicamente es difícil precisar el tipo de variación que debe presentar un sistema detector del color para ser considerado como a-normal (una crítica en la que insistirá Pautz (2006: 213-228) y en la que no voy a detenerme ahora),¹⁵ sería necesario especificar en qué sentido puede satisfacerse este requisito (Tye 2006d: 176).

4.3.2.3. Una segunda solución: la alternativa de Tye

Como hemos comentado anteriormente, la posición de Tye acepta (a grandes rasgos) un cierto aspecto bio-funcionalista al considerar que la Madre Naturaleza dota a los perceptores de un sistema visual compuesto diseñado a partir de varios elementos y que permite detectar el color. Pero también aceptaba la gran dificultad que suponía poder identificar un sistema detector del color como normal a partir, precisamente, de dicha complejidad composicional y de la posibilidad de una variación biológica en la distribución de los sistemas perceptivos respecto del color.

Por eso mismo considera una modificación en su argumento inicial. Lo que la Madre Naturaleza estaría haciendo, más que diseñar los sistemas detectores de color de cierto modo, sería ofrecer una serie de instrucciones (imaginamos que biológicas y genéticas) que permitan a los sistemas detectores evolucionar de tal modo que puedan actuar *como* diseñados con espectro tosco (*coarse-grained*) para detectar colores, que según las condiciones ambientales (volvemos a insistir, la mayor parte de ellas referidas a las condiciones lumínicas) pueden actuar *como* diseñados con espectro refinado (*fine-grained*) para detectar colores, del mismo modo que un cronómetro será más refinado si fue diseñado para marcar, además de los segundos, las milésimas (Tye 2006d: 176-177), y que evolutivamente hablando es más ventajoso que actúen del primer modo (dado que permiten una menor matización y una respuesta satisfactoria a una mayor cantidad de casos).

Lo que viene a decir esta solución, aplicada a nuestro caso concreto es que dados los anteriores sujetos perceptores con sistemas de detección del color normales, Vicente y Miguel Ángel, situados ambos ante la experiencia de la misma muestra de color azul 527, pueden representarla de modo diferente, porque uno de ellos dispone de un sistema detector de espectro tosco (lo que le haría representarse el color como azul puro) y el otro dispone de un sistema detector de espectro refinado (permitiendo que pueda representar la misma experiencia de color como azul ligeramente verdoso) y, a pesar de que podemos afirmar que no perciben diferentes colores porque ambos son capaces de percibir el mismo azul, sería conveniente decir que uno de ellos se equivoca a la hora de representar la experiencia del color azul 527 al considerar que representa un azul ligeramente verdoso porque dispone de un sistema detector excesivamente refinado, lo que restringiría la cantidad de casos en los que podría ser empleado satisfactoriamente (Tye 2006d: 177). Esta conclusión, además, permitirá a nuestro autor salvar el externismo respecto de la percepción al admitir que es la representación de un objeto externo

¹⁵ Las críticas que resaltan la importancia de la (posible) variación biológica en el sistema de visión de color aportadas por el primitivismo de Pautz pueden también verse en Pautz (forthcoming b).

(coloreado) lo que varía, al implementar la experiencia de modo diferente según su propio sistema detector, y siendo actualmente imposible por el estado actual de nuestra ciencia poder indicar qué diferenciaría un sistema detector refinado de uno tosco (Tye 2006d: 178).

Podríamos objetar que Tye nos debe todavía una diferenciación entre lo que supone como sistemas de detección con espectro refinado y espectro tosco (Cohen, Hardin and McLaughlin 2006: 337-338), incluso más allá de su apelación a la diferencia existente entre sistemas capaces de distinguir entre colores primarios y secundarios (Tye 2006e: 341-342). Podríamos insistir también en que es implausible asumir que una muestra de color sea representada de modo diferente por dos perceptores normales a pesar de la existencia de diferentes sistemas detectores del color dado que el color no pertenece a los objetos sino a los sujetos perceptores¹⁶ (Cohen, Hardin and McLaughlin 2006: 339), por lo que podríamos decir que realmente la presunta historia evolutiva de nuestro sistema detector no supondría un verdadero obstáculo para la representación correcta del color al carecer de papel causal a la hora de representar el color, siendo posible afirmar que ambos sujetos perceptores están en lo cierto cuando se representan la misma muestra de azul 527 como azul puro (por una parte) y como azul ligeramente verdoso (por la otra) (*Ibid.*, 338).

Aunque es cierto que Tye no especificará las diferencias existentes entre los dos tipos de sistemas detectores del color, sí que insistirá en la falsedad de las críticas citadas, dadas las pruebas evolutivas aportadas por la biología que tienden a afirmar que la correcta percepción del color depende del nivel de refinamiento de los sistemas perceptivos y de la cantidad de casos para los que estén diseñados, permitiendo una mejor adaptabilidad al entorno a aquellos sujetos a los que la Madre Naturaleza haya dotado de un sistema detector que funcione en una mayor cantidad relevante de casos (Tye 2006e: 342-344). Incluso podría objetar que Cohen, Hardin y McLaughlin hacen un mal uso de los resultados de ciertas pruebas biológicas aportadas recientemente por Malkoc *et al.* (2005),¹⁷ dado que dichas pruebas, a lo sumo, podrían indicar la extremada dificultad existente a la hora de establecer la estructura de un sistema detector normal, pero no la imposibilidad de la diferente representación de una misma muestra de color por parte de dos sujetos perceptores normales (Byrne and Hilbert 2007: 90).

4.3.2.4. Conclusión: diferencia sin dualidad de propiedades

Si tomamos literalmente nuestra pregunta inicial, a partir de la solución aportada por Tye deberemos concluir que la experiencia de color no puede ser subjetiva porque responde a la percepción de un objeto físico. Por lo que, dados dos sujetos perceptores normales, ambos representarán del mismo modo la misma muestra de color. Lo que demuestra los casos-tipo de

¹⁶ Véase también Cohen (2004), Hardin (2003) y McLaughlin (2003).

¹⁷ El mencionado estudio de Malcock *et al.* (2005) se refiere a un experimento reciente realizado sobre una muestra de usuarios del lenguaje respecto al empleo que realizan de 34 nombres de color situados en 8 categorías distintas. Siete de dichas categorías podían encuadrarse dentro de un nombre genérico, el azul, pero tan sólo se alcanzó un consenso suficiente amplio sobre dos de dichas categorías para ser catalogadas como azul. Los resultados parecen indicar que existen ciertas diferencias en las categorías de color que son capaces de captar sólo aquellas personas con un sistema categoría refinado.

la Variación Estándar es que existe la posibilidad de que determinados individuos dispongan de un sistema detector del color ‘defectuoso’, en el sentido en que será tan extremadamente refinado y preciso que no permitirá representar correctamente la experiencia de color.

Podríamos objetar a Tye que realmente el problema de la Variación Estándar de color sólo se refiere a la representación sensorial y no a la percepción del color, por lo que la apelación a casos en los que la conciencia falla a la hora de relacionarse con los objetos externos no permite inferir que dicha relación esté basada en la existencia de *entradas* físicas, aunque sean representaciones (Pautz forthcoming c), y que realmente la estrategia de Tye supone una reducción del contenido fenoménico de la conciencia a mera representación.¹⁸ Una crítica a la que nos sumamos en su primera parte, dado que en realidad el posicionamiento de Tye es extremadamente cerrado al considerar que sólo existe un posible concepto representacional correcto dada una determinada sensación de color, lo que supone una atención demasiado estrecha a la noción de contenido. Pero no en su segunda parte, ya que aunque parece salvar la múltiple realización mediante su distinción de sistemas detectores de color a partir de su apelación al refinamiento evolutivo (Tye 2007c: 92-93), en realidad simplemente estaría malinterpretando la teoría evolutiva al presuponerle una teleología (Byrne and Hilbert 2007: 90), lo que parecería presuponer un privilegio (¿dualista?) de lo mental respecto de lo físico.

Pero, y como ya hemos defendido en varias ocasiones, a pesar de que somos críticos con la posición intencionalista de Tye por presuponer que las representaciones actúan causalmente de modo estrecho, suscribimos su externismo dado que consideramos que la percepción del color es objetiva porque depende de representaciones *amplias* que incluyen un aspecto directivo, por un lado, mientras que por otro también incluyen un aspecto descriptivo (Millikan 2005: 166ss.), independientemente del objeto físico representado y del contenido mental que los sujetos perceptores pudieran tener. Por lo que además estaríamos en condiciones de aportar una posible solución al problema de la Variación Estándar del color, dado que lo que contaría como una experiencia de color no sería tanto el objeto físico al que se dirigen los estados mentales ni el contenido representacional de los mismos como la función de este tipo de representaciones, que no sería otra que la de situar al sujeto en el contexto en el que podamos decir correctamente que está teniendo una determinada experiencia de color, algo que dependería de las normas que regulen dicho proceso, permitiendo así incluso salvar la posible acusación de considerar la existencia de sujetos perceptores privilegiados (una crítica que, a pesar de sus esfuerzos, sí parece que puede serle aplicada a Tye al considerar la existencia de sistemas detectores toscos y refinados del color).

4.3.3. Por una nueva versión del contenido de la experiencia fenoménica

Según hemos comentado anteriormente, para Tye, el carácter fenoménico de un estado mental determinado superviene en sus propiedades representacionales. A pesar de las críticas

¹⁸ Pautz no ahonda en esta crítica en los textos tenidos en cuenta aquí, por lo que quién interesado en su respuesta al tratamiento que Tye y el intencionalismo realizan del problema mente-cuerpo puede atender a Pautz (2006b) y a la réplica de Lycan (2006).

presentadas, Tye actualmente considera todavía posible que las experiencias fenoménicas sean representacionales si es posible demostrar que el contenido de las mismas es determinado de tal modo que afecte a nuestras actitudes proposicionales de un modo normativo a partir del establecimiento de las condiciones bajo las que los dispositivos representacionales han evolucionado para poder determinarlo. Precisamente, contra todos aquellos que consideran inviable caracterizar como representacional todo estado mental, Tye cree estar capacitado para argumentar a favor de una diferenciación entre aquellos estados con carácter fenoménico y el resto de estados mentales. En este sentido, para Tye, no toda representación da lugar a un estado fenoménicamente consciente. Es frecuente, nos dice encontrar estados con un contenido representacional que no son conscientes.¹⁹ Pero siempre que se cumplan las condiciones PANIC, dichos estados se volverán conscientes. Veamos.

4.3.3.1. Conciencia y atención

Parece que, según hemos analizado en las anteriores secciones, la teoría de Tye se encuentra con serias dificultades a la hora de poder explicar el modo en que un sujeto es capaz de experimentar sensaciones y percepciones y, especialmente, a la hora de establecer las condiciones adecuadas que permitirían distinguir nítidamente entre aquellos estados mentales con carácter fenoménico de todos aquellos estados mentales que carecen del mismo.

Tye (2009) vuelve a insistir en sus afirmaciones. Su intención, ahora, es defender su posición representacionista frente a las críticas que apelan a la dificultad de diferenciar entre estados fenoménicos y otro tipo de estados mentales. La conciencia fenoménica, nos dice, es una característica de ciertos estados mentales, de aquellos estados mentales que son fenoménicamente conscientes.

“La conciencia fenoménica es una propiedad de aquellos estados mentales que son fenoménicamente conscientes” (Tye 2009: 8).

De un modo intuitivo, la conciencia fenoménica, entonces, requerirá de la presencia de criaturas capaces de conciencia (esto es, de criaturas capaces de captar el carácter fenoménico de dichos estados mentales). Una criatura de este tipo no será capaz de experimentar el carácter fenoménico de estos estados mentales si carece de conciencia, dado que (por definición) experimentar una determinada sensación (o experimentar el determinado carácter fenoménico de un cierto estado mental) presupone la capacidad de poder experimentarlo. En este sentido, que un individuo se encuentre en un determinado estado mental con un cierto carácter

¹⁹ Tan sólo debemos pensar en el estado representacional de un termómetro que marca 25 grados de temperatura o en el estado representacional de un velocímetro que marca 90 kilómetros por hora. De todos modos, y contra aquellos que consideran irrelevantes estos ejemplos por estar lejanos de los estados ‘naturales’, pensemos en los estados representacionales que las amebas o los paramecios tienen cuando evitan chocar entre ellos a partir de la representación de estados a partir de sus peculiares sistemas electro-magnéticos o en las abejas y en sus estados representacionales capaces de comunicar la dirección en la que se encuentra la comida (Tye 2000: cap. 8), aunque carezcan de acceso consciente. También podemos encontrar ejemplos en nosotros mismos, cuando hablamos de experiencias no instanciadas (pensemos en el modo en que muchos de nosotros no nos damos cuenta del ruido de nuestro frigorífico cuando entramos en la cocina a por un vaso de agua).

fenoménico presupone que el individuo experimenta el carácter fenoménico de dicho estado. Es decir, que el experimentar el carácter fenoménico de un estado concreto es idéntico a tener (o estar en) dicho estado.

“Por ejemplo, soy consciente del fuerte estruendo producido a mi izquierda, del silbido de la máquina de capuchino situada tras de mí... Mi perro se da cuenta de los sapos dibujados en su cuenco para el agua... y del hueso de su plato. Como he dicho antes, algunas veces soy consciente de mis propios estados fenoménicamente conscientes. Esto es propio de criaturas conscientes” (Tye 2009: 8).

Contra esta afirmación podríamos aportar un contraejemplo (basado en Block 2001). Existen ciertos casos en los que un determinado individuo puede encontrarse en un estado con carácter fenoménico sin que el sujeto sea consciente de que se encuentra en dicho estado. Pensemos en un ejemplo similar al siguiente: desde su juventud, un determinado individuo ha sido torturado por sus padres en el entorno de su habitación, pintada de un macabro color rojo. Desde ese momento, el sujeto ha formalizado una marcada represión respecto a la imagen de dicha habitación, pero también a todo objeto (pintado o naturalmente teñido de) rojo. Por supuesto, estas imágenes tienen un carácter fenoménico y, además, son perfectamente conscientes, pero el sujeto no es consciente en ningún momento de la relación profundamente reprimida con las imágenes subyacentes referidas a la habitación roja.

Pero esto no es especialmente claro para Tye. Si las imágenes de la habitación roja fueran fenoménicamente conscientes, entonces deberían poder ser experimentadas por el sujeto. Si este no es el caso, como parece sugerir el ejemplo (modificado) de Block, entonces estas imágenes no son fenoménicamente conscientes después de todo. Como Tye mismo indica, que estas imágenes no sean fenoménicamente conscientes implicaría tener que modificar la denominación de dichas experiencias. Por supuesto, pueden existir representaciones profundamente reprimidas en los sujetos conscientes e, incluso, pueden tener un papel causal en su conducta (como muestran los estudios psicológicos actuales), pero ello no demuestra que dichas representaciones deban ser fenoménicamente conscientes, que es el único punto que Tye está reivindicando aquí.

Pensemos en un ejemplo más claro. A menudo, cuando entro en la cocina a por un vaso de agua, no soy consciente del sonido insistente que produce mi frigorífico. Sin embargo, si presto un poco de atención o, por casualidad, escucho el sonido, entonces no puedo dejar de oírlo. En este sentido, muchas veces escuchamos sin oír. Pero cuando mi atención se centra en el sonido que produce mi frigorífico, no puedo dejar de experimentarlo. Es decir, en el momento en que focalizo mi atención sobre el sonido del frigorífico, mi experiencia se vuelve consciente, al atender al carácter fenoménico del sonido experimentado.

De un modo similar de tanto en tanto nos ocurren situaciones similares relacionadas con la percepción visual. Pensemos tan sólo en las ocasiones en las que jugamos a determinados juegos de agudeza visual (casos concretos los encontramos en las famosas viñetas de *¿Dónde está Wally?* o en juegos similares en donde tenemos que descubrir objetos determinados contrapuestos sobre un fondo repleto de formas y siluetas sin definir. Un caso similar lo

encontraríamos en el famoso ejemplo de Wittgenstein del pato-conejo). Entonces, podemos preguntarnos, ¿bajo qué condiciones una experiencia nos hace conscientes de un objeto particular?

En casos como los anteriores, parece que podemos mirar el conjunto completo pero no podemos ver el objeto concreto que buscamos. Pensemos en un ejemplo similar. Visito a un amigo que acaba de comprarse un terrario. En el terrario observo algunos insectos curiosos: algunas orugas, un par de escarabajos rinoceronte, un ciempiés... pero no soy capaz de ver nada más. Mi amigo insiste en que observe de nuevo. Todavía hay más insectos. Atentamente, fijo de nuevo mi mirada sobre el terrario. Me ha parecido detectar un ligero movimiento en uno de los pequeños troncos que sirven de camuflaje a los insectos. ¿Cómo? Un insecto con forma de bastoncillo parece moverse encima de él. Acabo de detectar la presencia de un insecto-palo encima del tronco. ¡Cómo he podido pasarlo por alto anteriormente! Antes he mirado también en aquella zona, pero no he logrado verlo. El movimiento del insecto ha hecho que focalizara mi atención sobre él y, ahora, soy capaz de verlo. Su presencia se ha hecho consciente para mí. Ahora soy capaz de experimentar su presencia. ¿Qué ha ocurrido?

Podríamos decir que siempre he visto el insecto, que siempre ha estado disponible, sólo que al estar camuflado en el tronco no he sido capaz de detectarlo. Pero esto sería correr demasiado. Por una parte, no es tan obvio el modo en que podemos especificar la región relevante del campo visual sobre la que estamos focalizando nuestra atención. Por otra parte, la experiencia en cuestión es una experiencia que representa la presencia de un determinado insecto en un determinado lugar, pero este tipo de explicación no sólo elimina al insecto en sí del contenido de la experiencia (ya que lo confunde con el entorno), sino que también, además, introduce en dicho contenido un elemento arbitrario al establecer una superficie inseparable en el campo visual.

Podríamos poner algunos casos que negaran este tipo de explicación de lo que sería ver un objeto. Pensemos en casos en los que estamos observando objetos lejanos, por ejemplo, una estrella. En este tipo de casos, cambios en el color de la estrella no afectan al color experimentado. Los sujetos que observan las estrellas ven las cosas que están ocurriendo (a menudo, la intensidad de la luz de las estrellas varía y les confiere tonalidades diferentes), pero no por ello se modifica el color de las estrellas que están observando. Lo que esto nos viene a sugerir, nos dice Tye, es que cuando alguien observa un objeto lo hace de un modo determinado, ya que cuando alguien tiene una experiencia visual de un objeto determinado (en este caso, una estrella) tiene una experiencia que representa dicho objeto como siendo de un modo determinado (Tye 2009: 12). En este sentido, la experiencia que está teniendo es una experiencia que contiene el objeto como visto de un cierto modo. Pero, intuitivamente hablando, dicho objeto no es el contenido de su experiencia. Porque la experiencia no es acerca del objeto, sino acerca del ver un objeto de un modo determinado.

Si el objeto fuera el contenido de la experiencia, entonces deberíamos ser conscientes en todo momento del modo en que dicho objeto se nos presenta, algo que, como estamos diciendo, no siempre es el caso. En este sentido, parece que el objeto no se diferencia del trasfondo sobre

el que se sitúa de un modo consciente, por lo que tenemos que buscar una explicación diferente, y satisfactoria, que permita rendir cuentas del modo en que el objeto se presenta a la conciencia. La respuesta más plausible, nos dirá Tye, es suponer que la experiencia de los objetos no nos está disponible de un modo directo, sino que necesitamos echar mano de información colateral para poder llegar a conformar una actitud proposicional respecto del objeto en cuestión (Tye 2009: 12). Es decir, no somos capaces de detectar la presencia de determinados objetos (pensemos en el caso del insecto-palo) si sólo atendemos a nuestra capacidad experiencial. Parece que requerimos de (al menos, la posibilidad de) ciertas actitudes proposicionales, como las creencias, que permitan hacer consciente aquello que aparece en nuestra experiencia.

Pensemos en nuestro ejemplo. Mi amigo me preguntaba si no era capaz de ver nada más que los insectos enumerados al principio, si no era capaz de identificar ningún otro insecto. La respuesta obvia que hemos aportado al principio es un rotundo ‘no, no soy capaz de ver ningún otro insecto’. Estoy expresando mi creencia de que no veo ningún otro insecto. Del mismo modo, en el momento en que he sido capaz de identificar un nuevo insecto, el insecto-palo (porque mi atención se ha focalizado sobre el movimiento del insecto y he sido capaz de identificar sus miembros componentes), mi creencia ha sido modificada y ahora soy capaz de decir que sí veo otro insecto. Es más, ya no puedo dejar de ver ese diminuto insecto-palo moviéndose por el terrario. Por supuesto, la creencia sólo se refiere a la presencia de un determinado objeto con una cierta forma y un determinado color, pero la ausencia de dicha creencia parece que cambia toda la situación, por lo que parece plausible suponer que, al menos en cierta medida, es necesaria para explicar el acceso que a un nivel consciente requieren ciertas experiencias. Parece que lo primero que podríamos decir en casos como en el del insecto-palo es que, de antemano, no es que el insecto no ha sido ya visto, sino que no nos hemos dado cuenta de su presencia. Verlo, lo hemos visto. Aunque no lo hemos identificado porque estaba camuflado en el tronco. Pero lo crucial en este ejemplo es mostrar en qué sentido el habernos dado cuenta de su presencia ha cambiado por completo la situación. (En casos como en la percepción del color parece que la situación se repite. Pensemos en el caso del color marrón del insecto-palo. También el peculiar color del insecto estaba escondido para mí hasta que he sido capaz de tener un acceso consciente a él a partir de su identificación, a pesar de que no es necesario que focalicemos nuestra atención de modo concreto sobre el color peculiar del insecto para ser conscientes del mismo).

En este sentido, parece que atender directamente al carácter fenoménico de la experiencia de un determinado objeto permite que seamos capaces de ser conscientes de la situación completa en la cual se encuentra dicho objeto. Pero para ser conscientes del objeto (como poseedor de características relevantes), entonces, necesitamos poder situarlo de un modo concreto dentro de dicha entera situación. Debemos ser capaces de captarlo como siendo de un modo determinado, y para ello necesitamos apelar a la presencia de determinadas actitudes proposicionales y no sólo al carácter fenoménico de la experiencia consciente.

“Si un estado fenoménicamente consciente de algo es tal que necesito preguntarme ‘¿qué es?, y no sólo sobre las bases de su carácter fenoménico, entonces soy consciente de dicha entidad. Pero si un estado

fenoménicamente consciente no está situado de este modo, entonces no soy consciente de dicho estado” (Tye 2009: 14).

En pocas palabras, no somos capaces de ver cosas si no estamos atendiendo pertinentemente a las mismas. Si tan sólo apelamos directamente al carácter fenoménico de su experiencia, entonces tan sólo podemos decir que las cosas son potencialmente dispuestas para ser vistas, a la espera de que atendamos adecuadamente a las mismas (Tye 2009: 168).²⁰

En este sentido, Block (2007a) sugiere que resultados como los obtenidos por los experimentos de Sperling (véase figura 4.1, en la siguiente página) demuestran que la fenomenología tan sólo sobredetermina la accesibilidad cognitiva a los contenidos de las experiencias conscientes. Pero parece que Block se encuentra con dificultades a la hora de demostrar sus afirmaciones basándose en los datos de dichos experimentos porque los comprende sin ningún tipo de restricción. Me explico. Podemos estar de acuerdo con Block en que los sujetos de los experimentos de Sperling ven las doce letras, pero tan sólo son capaces de dar cuenta de tres o cuatro de ellas, según sea el caso en que focalicen la atención en las columnas o en las filas. También podemos estar de acuerdo con él en que los sujetos de dicho experimento son capaces de ver las letras por separado siempre y cuando atiendan correctamente a cada una de ellas por separado. Pero, y aquí disentimos de Block, lo importante aquí es darnos cuenta de que, para cada caso, aunque los sujetos son capaces de ver todas las letras, para que sean conscientes de todas y cada una de ellas no se requiere tan sólo de un acceso fenoménico (como pretende Block), sino que además se requiere de un acceso cognitivo. Es decir, ante la experiencia de cada letra, los sujetos tienen acceso a un acto cognitivo diferente, y cada acto cognitivo diferente tiene un contenido distinto para cada una de las letras. Por tanto, no basta con tan sólo tener un acceso fenoménico al contenido de la experiencia, sino que, además, cada sujeto requiere de un acceso cognitivo que le permita darse cuenta del contenido concreto de cada acto cognitivo para ser capaz de tener un acceso consciente a cada una de las doce letras por separado.

| | | | |
|---|---|---|---|
| G | J | K | B |
| P | R | T | X |
| D | H | N | V |

Figura 4.1. El experimento de Sperling consiste en mostrar este cuadro a una serie de sujetos durante un periodo medio de 50 milisegundos. El experimento pretendía determinar el tiempo medio que requiere un individuo para poder memorizar las 12 letras. Según los resultados, se estima que hay una memoria sensitiva visual que requiere de, al menos, 300 milisegundos para poder hacerse efectiva por completo.

²⁰ Casos paradigmáticos donde se demostraría la tesis de Tye pueden encontrarse en los experimentos sobre el papel de la atención en la memoria sensorial de Sperling (1960), véase figura 4.1. Pensemos también en la luz ‘permanente’ de nuestro frigorífico: no siempre está presente, sólo cuando abrimos nuestro frigorífico, pero no por ello dejamos de saber que siempre que lo abrimos la luz sigue allí (Tye 2009: 168-171).

Una posible crítica a la que queremos atender aquí es la propuesta en Dretske (2007). En este trabajo, Dretske afirma que, en casos como en los denominados cambios ciegos (esto es, casos en los que se presenta una pequeña diferencia imperceptible a simple vista entre dos imágenes), aunque nosotros vemos las cosas como diferentes, realmente no vemos en qué difieren. Es decir, aunque percibimos ambas imágenes como diferentes, no somos conscientes de que son diferentes. Presenta el ejemplo de dos dibujos que representan sendos muros compuestos por bloques. Ambos son idénticos, pero en el segundo muro se le ha añadido un bloque adicional imperceptible a simple vista (Dretske denomina ‘Sam’ a este bloque adicional). Lo que nos viene a decir Dretske es que somos conscientes de un objeto tan sólo

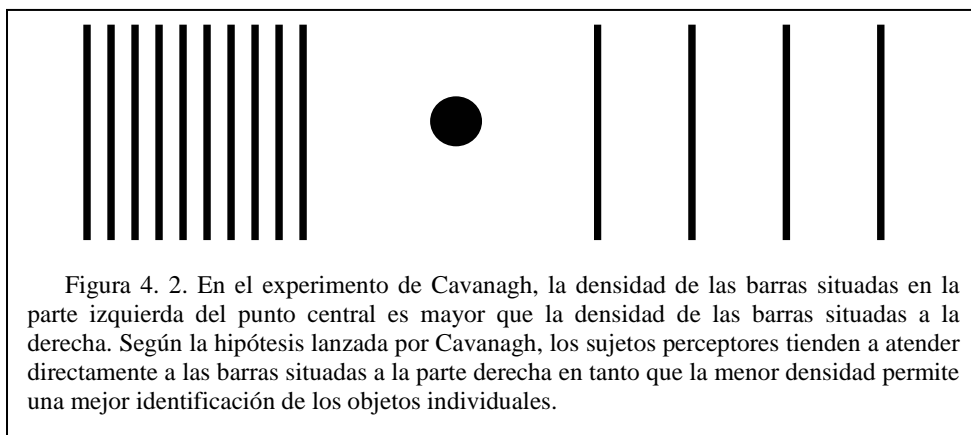
“si nuestra experiencia permite que conozcamos cosas respecto del mismo a partir del modo en que lo observamos” (Dretske 2007: 223).

En este sentido, cuando observamos el segundo muro, no somos capaces de ver tan sólo a Sam, sino que (como en el caso del primer muro) somos conscientes del conjunto compuesto por todos los bloques al unísono. Vemos a Sam junto con los otros bloques conformando un todo, no lo observamos de manera individualizada.

Podemos objetarle a Dretske que, en casos como el anterior, que seamos conscientes o no de Sam depende del modo en que observemos el muro (si lo observamos como un todo colectivo o como un conjunto separado de bloques conformando un muro) y depende del modo en que se nos presente (si tiene la misma forma que los demás bloques o tiene una forma diferente, de diamante por ejemplo), de manera particular, el bloque Sam. Pero, como dice Dretske, podemos saber que los dos muros son diferentes, pero Sam se presenta del mismo modo que los demás bloques (Dretske 2007: 224-225). La dificultad con la respuesta aportada por Dretske parece ser que, de nuevo, parece confundirse la conciencia fenoménica con la conciencia por acceso. Es decir, Dretske puede estar en lo cierto en que podemos conocer ciertas cosas respecto de los objetos cuando las vemos de modo colectivo, pero parece olvidar que al focalizar nuestra atención percibimos algo más que los objetos al unísono. Percibimos el carácter fenoménico que un objeto concreto tiene cuando lo observamos de un cierto modo. Es decir, cuando se nos presenta de modo individual un objeto a cada uno de nosotros, se nos presenta a nuestra conciencia de un modo determinado según nuestra propia historia cognitiva. En el caso del muro que estamos tratando, es cierto que la diferencia es imperceptible a simple vista, pero una vez la hemos localizado, no podemos dejar de percibirla cada vez que observamos el muro. Del mismo modo, cuando descubrimos a Wally en la viñeta, no podemos volver a mirarla sin ser conscientes de la presencia de Wally de una manera inmediata. Ambas percepciones son siempre diferentes ya para nosotros. En ambas, las cosas aparecen de modo consciente porque se nos presentan las cosas de un modo concreto al tener ya la información pertinente: hemos conseguido focalizar nuestra atención, y ello cambia por completo la percepción consciente.

Siguiendo con el argumento, Block (2007b) sugiere la existencia de una evidencia empírica que respaldaría la mayor pertinencia del ver respecto de la atención a la hora de

permitir el acceso consciente de una experiencia, lo que parecería dificultar la idea de que sólo a partir de la atención se puede adquirir conciencia de un objeto a partir de la focalización de la experiencia consciente (como se puede comprobar, la idea es que la noción de objeto sea tenida como una noción demostrativa al estilo defendido por Kaplan 1989). La evidencia empírica a la que Block se refiere son los estudios experimentales llevados a cabo por Cavanagh (1999) en los cuales se demuestra la existencia de limitaciones en la atención (a partir de la saturación existente en la distancia de dos grupos de líneas paralelas) que hacen que tan sólo atendamos al grupo de líneas que parecen actuar como un todo (por ser menor el espacio de separación existente entre ellas. Véase figura 4.2).



Podemos objetar a Block que la supuesta evidencia empírica realmente no demuestra que existan cosas que veamos a las que no somos capaces de atender, mostrando que el ver no es cognitivamente superior al atender. Como Tye nos dice, cierta clase de verbos tienen un carácter colectivo (Tye 2009: 15; Tye forthcoming b). En este sentido, el verbo ‘pensar’ tiene un carácter de colectividad (la prueba está en que puedo pensar en lo buenas que están las palomitas sin pensar en una de ellas en particular o pensar en mis amigos sin tener que pensar en uno de ellos en concreto). La forma verbal ‘ser consciente’ (o ‘darse cuenta’) sería también colectiva en este sentido. Nosotros podemos atender a ambos conjuntos de líneas sin tener que atender a cada una de ellas en particular en el experimento de Cavanagh. El hecho de que el conjunto de líneas de la izquierda esté compuesto de líneas individuales situadas a una distancia menor que las del grupo de la derecha no quiere decir que no pueda ser capaz de ser consciente de cada una de ellas de modo individualizado. Somos conscientes de las líneas. Las vemos. Existe algo hacia lo que mi experiencia se dirige. Mi experiencia es acerca de las líneas. Si permanecemos atentos, podemos distinguirlas una a una. Entonces, las líneas (de modo individualizado) parecen acceder a la conciencia bajo su propio concepto demostrativo. Parece, entonces, plausible pensar que en la parte izquierda existe un todo compuesto por un cierto número de objetos individuales cada uno de los cuales puede hacerse consciente si centramos nuestra atención de modo individualizado (formando demostrativos diferentes). Vemos las líneas como formando un conjunto, pero podemos ser conscientes de la presencia de líneas

diferentes (aunque idénticas) como componiendo dicho conjunto si logramos atender individualizadamente a cada una de ellas.

4.3.3.2. Desplumando la gallina

A partir de las conclusiones extraídas en la anterior sección, Tye intenta aportar una solución plausible a un viejo problema respecto de la percepción (Tye 2009: 16-18; Tye forthcoming a). El problema se refiere a una dificultad que señaló Ryle en el análisis que Ayer (1940) realizó de las experiencias visuales: el problema de la gallina de Guinea (*the speckled hen problem*).

El problema es el siguiente. Imaginemos que ante nosotros tenemos una gallina de Guinea. Vemos claramente una gallina que tiene motas (la principal característica física de la gallina de Guinea), pero no podemos indicar cuál es el número de motas que permiten decir que es una gallina de Guinea. ¿Cuántas motas tiene la gallina? Por supuesto, la cuestión se zanjaría contando una a una las motas de la gallina. Pero lo que aquí nos importa es por qué, de un solo vistazo, no podemos dar cuenta de la cantidad exacta de motas que tiene la gallina. Es decir, de por qué no podemos detenernos en cada una de las diferentes motas sino que nuestra atención capta una gallina moteada, como conjunto.

La respuesta de Ayer (1940) fue indicar una dificultad clara a la hora de determinar la cantidad de motas de la gallina. Para Ayer, por supuesto, existe una visión completa de la gallina, pero no tenemos una imagen definida sobre la cantidad exacta de motas que tiene la gallina. Porque, como bien sabemos, para Ayer, lo que cuenta como una experiencia dependerá de tener un adecuado dato que la gallina presenta a nuestros sentidos. El dato de los sentidos existe, tenemos el dato de una gallina moteada. Pero nuestro *sense-data* no tiene un número definido de motas. Por tanto, no podemos indicar cuántas motas tiene la gallina porque el número de motas que contiene nuestro dato de los sentidos es indeterminado. Pero parece que su conclusión es contradictoria.

Una posible respuesta alternativa es la ofrecida por la teoría adverbial de la percepción propuesta por Chisholm (1942). En este escrito, Chisholm niega la existencia de *sense-data* a partir de la suposición de que el ver incluye ya el sentir un dato de los sentidos, por tanto, es innecesario apelar a ellos para explicar el ver. Si no hay dato de los sentidos, entonces no existiría una indeterminación del objeto (esto es, no existe un número indeterminado de motas sobre la gallina), por lo que el problema de la gallina de Guinea desaparece. Pero las cosas no son tan claras.

Block (2007b) toma este caso como un caso paradigmático de la confrontación entre ver y atender. Para Block, por supuesto, el ver mismo ya contiene en sí todo el poder consciente, por lo que la gallina aparecerá como moteada, aunque no seamos capaces de definir el total de motas que tiene su plumaje. Pero, como insiste Tye, lo que realmente indica el problema de la gallina de Guinea es un error a la hora de comprender la aplicación de los verbos no-distributivos y las propiedades múltiples. Como Tye insiste, lo que permite que la gallina aparezca como moteada es que nuestra experiencia representa a la gallina como moteada. Pero

dicha representación es fenoménicamente indeterminada con respecto al número de motas que el plumaje de la gallina tiene. Porque nuestra experiencia visual no consigue determinar el número exacto de motas que tiene la gallina.

“La razón por la que no podemos enumerar el número de motas de la gallina es que dicha enumeración requeriría que asistiéramos adecuadamente a cada una de las motas por separado... tan sólo vemos las motas. Somos conscientes *de ellas*... aunque podemos formar creencias (y otras actitudes conceptuales) acerca de cada una de ellas por separado, si queremos” (Tye 2009: 17).

Pero la dificultad aquí persiste. Si nosotros vemos por completo la gallina, con buena luz y sin obstáculos visuales (es decir, en condiciones normales), nosotros vemos todas las motas que tiene la gallina. Entonces, ¿cómo puede nuestra experiencia visual estar indeterminada con respecto al número de motas que tiene la gallina sobre su plumaje?

Como Tye insiste, ver algo es ser consciente de estar teniendo una experiencia visual de ello. Por tanto, si somos sujetos capaces de tener actitudes proposicionales, entonces somos conscientes de una determinada cosa cuando podemos identificar sobre las bases de nuestra propia experiencia qué sea esa cosa determinada (Tye 2009: 13-14; Tye forthcoming a). Es decir, ser consciente de algo determinado requiere que la experiencia relevante represente dicha cosa en concreto. El objeto mismo debe ser parte del contenido representacional de dicha experiencia. Por tanto, como indicábamos anteriormente, para Tye, ser consciente de un objeto determinado presupone que seamos capaces de atender a dicho objeto. El ver sólo no basta para ser conscientes de un objeto. Necesitamos, además, prestar atención al objeto concreto para poder ser consciente del mismo.

Volvamos al caso de la gallina de Guinea. Somos conscientes de ella. Existen una serie de motas individuales sobre el plumaje de la gallina que permiten que seamos conscientes de la experiencia de una gallina moteada delante de nosotros. Pero la experiencia visual de las motas de un modo individualizado, como ya ocurría con las líneas del ejemplo de la sección anterior (véase cuadro 4.2), es capaz de formar creencias acerca de ellas mismas de manera individual, siempre y cuando centremos nuestra atención de manera concreta en dichas motas de manera individualizada. Entonces, ¿vemos todas las motas? Depende de lo que entendamos por ‘todas’. No somos conscientes de todas y cada una de las motas cuando somos conscientes de la experiencia de una gallina de Guinea frente a mí. Pero sí vemos todas las motas porque permite representarnos una gallina como moteada. En este sentido, existe un sentido colectivo de ‘todas’ que permite que veamos una gallina de Guinea como moteada. Por tanto, aunque nuestra experiencia visual es indeterminada con respecto al número total de motas que hay sobre el plumaje de la gallina, sería perfectamente aceptable suponer que existe un número concreto de motas que, aunque no podemos ver a simple vista, podemos captar mediante la focalización de nuestra atención de manera individualizada a cada una de ellas, formando así sucesivos conceptos demostrativos con carácter fenoménico.

4.3.3.3. El contenido de la experiencia fenoménica como un demostrativo

Como venimos indicando en este capítulo, para Tye, ver un objeto determinado es percibirlo de modo directo. En este sentido, los objetos percibidos aparecen como componentes de los contenidos de nuestras experiencias perceptivas. Ver un objeto, por tanto, significa que veo el objeto de un determinado modo. Pero si ver un objeto es verlo de un modo determinado, entonces también debemos presuponer que lo experimentamos de un modo determinado. Así, ver un objeto supone tener una experiencia cuyo contenido incluye al objeto percibido como teniendo unas determinadas propiedades aparentes, por lo que la experiencia será verídica si y sólo si el objeto tiene exactamente dichas propiedades. Pero, ¿qué ocurre entonces en casos en los que una experiencia no es verídica; digamos, en casos como las percepciones ilusorias?

Una primera respuesta podría ser la siguiente. En realidad, la situación es la misma, no varía. La experiencia ilusoria también se refiere a la percepción de un determinado objeto, y dicho objeto aparecerá como teniendo unas determinadas características al sujeto perceptor (Travis 2004). Pero, ¿cómo explicar entonces que no se correspondan las propiedades reales del objeto con las propiedades del objeto percibido? Parece más intuitivamente aceptable suponer que, al menos en casos como las ilusiones, la experiencia es inexacta, esto es, que no se corresponde con el objeto real percibido (Tye 2007c: 589; Tye 2009: 77). Pero esto no quiere decir que las experiencias de percepciones verídicas y de experiencias ilusorias no puedan parecer similares. Pensemos tan sólo en la posibilidad de confundir con una fruta real una manzana de plástico que encontramos en un cesto, o en la posibilidad de confundir con un limón real una pastilla de jabón en forma de limón (Austin 1962: 23; Tye 2009: 78). La experiencia de un limón real, por tanto, podría ser indistinguible de la experiencia de una pastilla de jabón en forma de limón. Pero que dos experiencias tengan el mismo carácter fenoménico no quiere decir que tengan el mismo contenido representacional. Dos individuos, uno que experimenta un limón real y otro que experimenta una pastilla de jabón en forma de limón como un limón real, pueden compartir una experiencia con un carácter fenoménico (al menos) similar, pero por supuesto tendrán un contenido representacional diferente, dado que la primera experiencia se refiere a las características de un limón real mientras que la segunda se refiere a las características de una pastilla de jabón en forma de limón. (Recordemos que hemos avanzado una conclusión similar en la solución aportada por Tye al problema de la Variación Estándar).

Es decir, el carácter fenoménico de ambas percepciones puede causar que los sujetos crean que ambas percepciones se refieren a un limón real, pero sólo la primera creencia es verdadera. Por tanto, debe existir algún elemento en el contenido de la experiencia ilusoria que lo haga diferir del contenido de la experiencia verídica, aunque el carácter fenoménico de ambas percepciones sea el mismo. Pero, ¿en qué sentido podemos decir que éste es el caso?

Una propuesta la encontramos en la tesis que afirma que las experiencias no tienen un contenido particular determinado por el objeto experimentado (McGinn 1982: 39). Para McGinn, entonces, los conceptos empleados para caracterizar el contenido de, por ejemplo, las experiencias visuales tan sólo se refieren a conceptos que delimitan sus características físicas tal y como se presentan al observador, sus cualidades secundarias (McGinn 1982: 42). En este

sentido, McGinn supone que las experiencias pueden tener el mismo contenido representacional pero dirigirse a objetos diferentes. Esto es, que objetos diferentes siempre aparecerán como diferentes porque no podemos deducir la identidad de los objetos de experiencia a partir del conocimiento que tenemos de su contenido experiencial, porque sólo conocemos el contenido de una experiencia a través de la introspección, y para ello no es necesario conocer el objeto hacia el cual se dirige nuestra experiencia (McGinn 1982: 38).

También Davies (1992) abraza esta tesis. En su caso, por razones similares a las de McGinn, considera que las experiencias tan sólo tienen un contenido existencial, por lo que existirá un contenido común y compartido tanto por las experiencias verídicas como por las experiencias alucinatorias. Pero, también en este caso, como en el caso de McGinn, parece que si presuponemos la existencia tan sólo de un contenido existencial, puramente fenoménico, en las experiencias, entonces, es posible que tengamos dificultades a la hora de determinar ese contenido en ciertas experiencias verídicas cuando no podamos precisar el objeto al cual se dirigen, o éste sea ambiguo.

Un modo de defensa que los teóricos de dicha tesis sobre el contenido han encontrado para hacer frente a esta objeción es intentar establecer una noción más elaborada de contenido puramente fenoménico. Tanto McGinn como Davies han defendido una teoría en donde las experiencias representarían tan sólo los objetos (o las características) relevantes en cada momento a partir del establecimiento de una relación causal entre la experiencia misma y el contexto de la experiencia, permitiendo así determinar aquellos aspectos relevantes en la experiencia pero sin necesidad de postular ningún rasgo físico del objeto mismo que haga variar el contenido fenoménico de la experiencia, permitiendo tan sólo a la introspección sobre la experiencia misma (y su contenido representacional) establecer las características del objeto percibido. En pocas palabras, lo que pretenden tanto Davies como McGinn es afirmar que las experiencias se refieren a ellas mismas, lo que permitiría establecer de modo concreto cuál sería su contenido.

Pero en este sentido, y como indica Tye, parece posible identificar un error de fondo en esta propuesta. Por un lado, impediría la posibilidad de error en la percepción, porque si consideramos (como hacen Davies y McGinn) que la experiencia se refiere a ella misma, entonces no es posible que representemos de manera errónea ninguna experiencia, dado que la experiencia siempre se referirá a ella misma. Pero sabemos que los errores son posibles. Es siempre posible que al percibir el color de un objeto me parezca un color ‘diferente’ del color que realmente tiene dicho objeto. Pero para que esto ocurra, el objeto mismo debe formar parte (debe ser un componente) del contenido mismo de mi experiencia, algo que no parece posible en la tesis sostenida por ambos autores. En este sentido, la objeción a esta tesis sobre el contenido puramente fenoménico es clara: cuando percibimos un objeto (digamos un tomate maduro), nuestra experiencia visual se dirige directamente sobre el tomate y no sobre nuestra propia experiencia, por lo que no parece adecuado apelar a la experiencia misma para determinar el contenido de dicha experiencia (Tye, 2007c: 593; Tye 2009: 80).

Pero, por otra parte, además, existe un claro inconveniente con la tesis del contenido puramente fenoménico cuando hablamos de casos de percepciones pasadas. Cuando recordamos la percepción de un tomate maduro, parece necesario afirmar una cierta relación causal entre el hecho histórico que hace verdadera la experiencia misma (esto es, el hecho concreto en el pasado en el que observé un tomate maduro) y el estado de memoria que me permite recordar dicha percepción. Si la tesis existencial fuese cierta, entonces la conexión causal entre el hecho y la memoria debería formar parte del contenido de la experiencia, pero éste no es el caso. No necesitamos que aparezca en el contenido de experiencia la explícita relación entre el hecho y la memoria, dado que la memoria misma sería ya, considerada de este modo, por sí misma una clase de conocimiento. No necesitaríamos, entonces, apelar a ninguna relación causal entre ellos para conocer a qué se refiere nuestra experiencia, algo que parece no ser el caso (Tye 2007c: 593; Tye 2009: 80).

En este sentido, parece que la tesis del contenido puramente fenoménico debe ser rechazada. Porque parece el caso que lo que permite que mi experiencia de un tomate maduro sea acerca de un tomate maduro no es cómo se presenta el tomate al observador, sino que exista una determinada relación causal entre la experiencia de un tomate maduro y el hecho contextual de la existencia de un tomate real (aunque dicha relación causal nunca formará parte del contenido de la experiencia misma. Cf. Tye 2007c: 593). En este sentido, Tye parece insinuar que, por un lado, toda experiencia (verídica o alucinatoria) deberá incluir (al menos la referencia a) un objeto y que, por otro lado, el contenido de una experiencia verídica y de una experiencia alucinatoria no puede ser el mismo, en tanto que referirán a contextos totalmente diferentes.

Tye analiza una posible alternativa propuesta por los existencialistas fenomenológicos. Pensemos, nos dice, en dos diferentes experiencias. En la primera de ellas, sentado en el porche de mi casa, leyendo, escucho como en el jardín croar una rana. En el segundo, tal vez bajo la manipulación de un científico malvado que está experimentando con mi cerebro, sentado en mi jardín, leyendo, escucho como croar una rana. En el primer caso, la experiencia es verídica: existe una rana en mi jardín que croa. En el segundo caso, la experiencia es ilusoria: no existe tal rana, aunque yo la oigo croar. Según la alternativa presentada por los existencialistas, en ambos casos el contenido tendría la misma estructura, sólo que en el segundo existiría un vacío allí donde debería representarse un objeto. Es decir, ambos contenidos representan un croar como presentado de un cierto modo al sujeto perceptor (misma estructura), pero en el primero, a diferencia del segundo, existe la representación de un croar real (esto es, la representación tiene una referencia real), mientras que en el segundo, la representación del croar carece de referencia. Tye parece insinuar que, mientras que en la primera percepción existe un objeto concreto que forma parte del contenido de la misma (una rana real croando), en el segundo caso no existe dicho objeto. Parece existir, entonces, un hiato entre el contenido y su objeto. Pero, a pesar de ello, la estructura de ambos contenidos sería la misma. Esta estructura, nos dice Tye, debe entenderse como teniendo un casillero en el cual deberemos introducir el objeto percibido.

En el primer caso, el casillero será ocupado por la rana que croa, mientras que en el segundo caso, el casillero quedará vacío.

En este sentido, si esto es cierto, requerimos de dicho contenido para verificar el valor de verdad de los contenidos de las experiencias. Así, la experiencia primera será verdadera porque existe un objeto en su contenido que tiene las propiedades experimentadas. En el segundo caso, entonces, al no existir un objeto que aparezca en el contenido y que coincida con las propiedades del objeto experimentado, la experiencia debe tenerse por falsa, por ilusoria. Así, en casos como los anteriores donde el demostrativo ‘que’ no se refiere a ningún objeto concreto que tenga las propiedades del objeto percibido, lo que contaría como el contenido de una experiencia sería una forma concreta de estructura en donde aparezca siempre un casillero (por supuesto, Tye está hablando de forma metafórica) que represente a un objeto como percibido, aunque en los casos de experiencias ilusorias o alucinatorias dicho casillero esté vacío.

Con esta solución, lo que está pretendiendo Tye es, por un lado, salvar la posibilidad de que toda experiencia tenga un contenido representacional, aunque sea al precio de tener que afirmar la posibilidad de que, en casos concretos de experiencias falsas, el vacío dejado por el objeto como percibido no pueda ser ocupado por un objeto real. Pero, por otro lado, ello le permitiría afirmar que las experiencias alucinatorias e ilusorias también tienen un contenido, similar en estructura aunque diferente (por no incluir un objeto real) de las experiencias reales, lo que permitirá a Tye salvar su posición de acusaciones existencialistas al negar que el mundo realmente sea tal y como lo percibe una persona con experiencias ilusorias.

Lo que Tye está proponiendo, como nos insiste, es una variante del realismo (Tye 2007c: 589; Tye 2009: 82; Tye forthcoming b). Defiende una clase de disjuntivismo:²¹ Tye sostiene la tesis de que el contenido de la experiencia visual en casos verídicos es diferente (aunque tiene una estructura similar a) del contenido de la experiencia visual en casos de alucinación o ilusión. En este sentido, no parece existir un factor común entre ambos contenidos, el real y el ilusorio. Para cada experiencia, entonces, existe tan sólo un único contenido admisible (aquel que determina el contexto concreto que permite ajustar el objeto percibido a la estructura adecuada, al estilo de los demostrativos, aunque dicho objeto no exista realmente), pero dicho contenido será diferente en la experiencia real y en la experiencia ilusoria porque, como nos indica Tye, en el segundo no existirá el objeto real que debería cumplir las condiciones que representa nuestra experiencia. Pero, ¿esto no sería un inconveniente al impedir que una misma experiencia tenga contenidos diferentes? O al contrario, ¿no impediría que un mismo contenido cause experiencias diferentes?

Lo que las dos experiencias tendrían en común es que ambas poseen un contenido de una cierta clase. Pero dicho contenido tendría componentes diferentes. Lo que Tye nos viene a decir es que, para que una experiencia tenga carácter fenoménico, dicha experiencia debe contener en su contenido (representacional) un cierto conjunto de propiedades. Tye (1995 y

²¹ Una perspectiva actualizada sobre los temas tratados dentro de las posiciones disjuntivistas puede encontrarse en Byrne and Logue (2009).

2000), como hemos mostrado en las secciones anteriores, ha insistido en que las propiedades relevantes de dicha experiencia serán representadas de modo no-conceptual, por lo que el contenido de las mismas deberá estar preparado para causar las respuestas cognitivas adecuadas. En este sentido, lo que Tye estaría defendiendo es un intencionalismo que afirmarí­a que, del mismo modo que las propiedades funcionales son múltiplemente realizables, las propiedades fenoménicas también son múltiplemente realizables si se tiene en cuenta las fuertes propiedades intencionales con las que el contenido es idéntico. Esto quiere decir que el carácter fenoménico de las experiencias reside, entonces, en su posesión de un contenido preparado en el que son representadas de manera no-conceptual toda una clase de propiedades concretas que posee un objeto (físico o abstracto) determinado, aunque éste no exista realmente. Esta perspectiva desemboca en un intencionalismo de segundo orden que permitirá aportar soluciones satisfactorias a problemas clásicos dentro del estudio de la filosofía de la conciencia.

4.3.4. Espectro invertido y *qualia* ausentes en Tye

En el apartado 4.3.2., atendimos a la solución que Tye ofrece de un problema relativo al estudio de la conciencia: la Variación Estándar. En este apartado queremos revisar las propuestas aportadas por Tye, teniendo en cuenta su peculiar visión del contenido de las experiencias con contenido fenoménico, a otros problemas referentes al estudio de la conciencia.

4.3.4.1. El análisis de Tye del problema del espectro invertido

El objetivo de esta sección será mostrar cómo, en escenarios como los presentados por los casos de espectro invertido, la defensa de una noción intencional de contenido compartida por diferentes individuos nos permitiría afirmar, a pesar de la existencia de una diferencia a nivel fenoménico, la existencia de un mismo contenido representacional (como fue descrito en la sección 1.4.6.).

Pensemos en una situación como la siguiente. Imaginemos a alguien, llamémosle Borja, que se encuentra en un estado mental tal que, frente a un determinado objeto, digamos una pelota de playa de color rojo, causa en él la percepción de un objeto verde, aunque ello no impide que aplique correctamente el término ‘rojo’ a dicho objeto, aunque funcionalmente hablando el estado que provoca dicha causalidad es idéntico a aquel estado que en el resto de nosotros causaría que percibiéramos un objeto de color rojo y provocaría que aplicáramos el término ‘rojo’ a dicho objeto. ¿Cómo podría ser el caso que un sujeto como Borja fuese funcionalmente idéntico a todos nosotros y, a pesar de ello, ser capaz de percibir fenoménicamente el objeto de manera diferente, aunque sea capaz de aplicar adecuadamente el término de color (en este caso)? Es decir, ¿serían posibles situaciones de inversión de espectro fenoménico como la descrita en el ejemplo?

Todo parece indicar que no existe una aparente contradicción en que situaciones como las descritas anteriormente sean posibles. Esto es, que exista una diferencia a un nivel fenoménico, aunque exista una identidad a un nivel funcional, no necesariamente debe implicar

una diferencia a un nivel personal. La prueba más clara que tenemos para afirmar dicha situación la encontramos en el propio comportamiento de Borja. Frente a objetos que todos nosotros identificaríamos como ‘rojos’, y que Borja percibe como verdes, es capaz de aplicar correctamente el término ‘rojo’ para describir las situaciones (esto es, es socialmente competente en la atribución de color), a pesar de que a un nivel fenoménico difiere de cualquiera de nosotros, aunque sea funcionalmente idéntico.

Hemos visto una posible solución en un caso de una inversión inter-personal de espectro fenoménico pero, ¿podrían funcionar también las cosas en una inversión intra-personal? Muy posiblemente, sí. Nada parece indicar lo contrario. Imaginemos que Borja, de nuevo, está percibiendo un objeto que fenoménicamente se le aparece como verde aunque es capaz de individualarlo como ‘rojo’. Pero, a diferencia de la situación anterior, en este caso, Borja es sometido a una operación neurológica a manos de un malvado científico que intercambia los contenidos fenoménicos que le permiten la percepción fenoménica de los objetos rojos y verdes. A partir de dicho momento, cuando Borja percibe un objeto rojo, aunque es capaz de percibirlo como verde e identificarlo como ‘rojo’, funcionalmente hablando el estado mental en el que se encuentra sigue siendo idéntico a aquel otro estado mental en el que Borja se encontraba cuando percibía objetos verdes. Del mismo modo que antes, a un nivel personal la situación no parece haber cambiado. Borja es capaz de percibir objetos rojos y de identificarlos como ‘rojos’, aunque su nivel fenoménico sea diferente al de aquella situación en la que se encontraba Borja cuando anteriormente percibía objetos verdes. De nuevo, la inversión parece posible, aunque no es problemática.

En este sentido, parece cierta la posibilidad, al menos, de la existencia de situaciones en las cuales se dé una inversión de espectro fenoménico a la hora de percibir determinadas sensaciones. Parece que, independientemente de la identidad funcional estrecha que exista a un nivel primario (sub-doxástico) entre los sujetos perceptores, es posible definir una noción de *qualia* sobre la cual estos casos de inversión son posibles. Esta noción, según la hemos presentado en el análisis de Tye, no depende de diferencias internas en los estados mentales de los sujetos perceptores sino de diferencias relevantes en las causas externas y en la cadena histórico-causal de aquellos sujetos perceptores que, en condiciones óptimas, están sintiendo una misma experiencia con carácter fenoménico.

“Los *qualia* invertidos pueden estar acompañados de una identidad funcional estrecha. Pero... existe una diferencia funcional amplia por el lado de las entradas: en esta criatura [la que sufre la inversión], el estado sensitivo tiene una causa externa diferente de aquella que tienen las otras criaturas, dadas las condiciones óptimas” (Tye 1995: 206).

A diferencia de análisis funcionalistas como los de Block o de análisis representacionales moderados como los de Shoemaker (1975 y 1994), donde se considera la necesidad de una diferencia interna en los sujetos para que pueda ser posible la existencia de casos de espectro invertido (Cf. Tye 2000: 99-103), en el análisis de Tye de este tipo de casos, primero, se tendrá en cuenta el contenido representacional de la experiencia concreta (un

contenido definido en términos de su teoría PANIC), pero que, segundo, depende no sólo de la distribución funcional de los sujetos perceptores sino, sobre todo, de las características externas de los objetos percibidos y de la peculiar historia cognitiva de cada individuo que permiten que la percepción del objeto concreto pueda ser entendida como correcta en condiciones normales (Tye 2000: 104-116).

La misma solución podría extraerse para casos como los de la tierra invertida de Block (1990). Puede ser perfectamente posible, desde la teoría de Tye, que el mismo individuo que es funcionalmente idéntico en la Tierra Gemela del resto de sus habitantes pueda ser fenoménicamente diferente a ellos cuando percibe un objeto rojo y deba identificarlo como 'rojo', aunque él lo perciba como verde (Tye 1995: 206-207), porque no depende tan sólo de su distribución funcional la capacidad representacional del carácter fenoménico de los estados mentales, sino también de elementos externos como su desarrollo cognitivo particular o los objetos físicos mismos percibidos en situaciones normales.

Desde el punto de vista de Tye, entonces, es posible definir una noción de conciencia fenoménica (y de contenido fenoménico) tal que sea suficientemente robusta como para permitir la posibilidad de la existencia posible de casos de espectro invertido y, además, poder aportar una solución externista a los mismos. Como hemos visto, esta solución pasaría por considerar la posible identidad funcional de los sujetos perceptores más allá de la diferencia fenoménica de los sistemas de percepción, apostando por una evolución histórica y causal de los sistemas cognitivos que permitirían una progresiva adaptación a las circunstancias particulares de cada sujeto de percepción a los hechos perceptivos caso por caso. Esto permitiría que, en escenarios de inversión, la noción intencional de contenido defendida por Tye pueda ser compartida por varios individuos (incluso aquellos que aparecen como fenoménicamente invertidos), permitiendo dejar intacta su percepción a un nivel personal.

4.3.4.2. El problema de los *qualia* ausentes en Tye

La hipótesis de los *qualia* ausentes apela a la posibilidad de la existencia de un sistema que (funcionalmente hablando) sea un duplicado exacto molécula por molécula de un sujeto perceptivo normal (lo que se viene conociendo como *zombies*), capaz de tener estados mentales funcionalmente idénticos a aquellos estados mentales que nosotros tendríamos cuando sentimos una experiencia perceptiva de algún tipo pero que, a pesar de ello, dichos estados carecen de carácter fenoménico. Un caso paradigmático de *qualia* ausentes lo encontramos en el ejemplo de Block (1978) de la nación china.²² Imaginemos que se conecta a todos los habitantes de China mediante equipos de radio de tal modo que simulen la organización funcional que presenta un cerebro, asignándole incluso a cada individuo una función concreta dentro del

²² Chalmers (1996: 320-322) alcanza una conclusión similar pero para propósitos diferentes. Chalmers considera imposibles los casos de *qualia* ausentes porque considera que, en contra de lo que había establecido inicialmente Block, aunque la salida responda perfectamente a la entrada, ello no significaría que el conjunto de individuos chinos que conforman el sistema entraron en un determinado estado mental del mismo modo que decimos que el cerebro se encuentra en un determinado estado mental, por lo que sería conveniente concluir que el cerebro proporciona genuinos estados mentales frente a otros sistemas.

sistema representado. Imaginemos que se les proporciona una entrada sensorial a la que responden funcionalmente con una salida a partir de la computación pertinente de los datos manejados. Aunque la salida responda perfectamente a la entrada, ello no significaría que el conjunto de individuos chinos que conforman el sistema entraron en un determinado estado mental del mismo modo que decimos que el cerebro se encuentra en un determinado estado mental, por lo que sería conveniente concluir que mientras que el cerebro proporciona estados mentales con carácter fenoménico, queda patente que otros sistemas pueden simular estados o reproducirlos pero que éstos carecen de carácter fenoménico.

La razón principal por la que Tye infiere la imposibilidad de la existencia de escenarios de *qualia* ausentes como los anteriormente descritos se debe precisamente a la imposibilidad de la existencia de casos donde se produzca un isomorfismo funcional entre dos sistemas perceptores que, a pesar de ello, presenten una asimetría en su fenomenología experiencial (Tye 2006c: 153-154). Es decir, si tenemos la existencia de dos sistemas de percepción que comparten la misma estructura funcional, todo viene a indicarnos que, dada una misma experiencia, ambos tendrían que entrar en un mismo estado al compartir el mismo contenido representacional. La cuestión aquí no reside en la posibilidad de que ambos compartan el mismo contenido, dado que como ya hemos mostrado hasta la saciedad, en la teoría de Tye la determinación del contenido dependerá tanto de la propia evolución particular del sistema perceptor como de las características físicas relevantes del objeto percibido. Pero, aunque no compartan un mismo contenido fenoménico, ambos sistemas deben tener una cierta fenomenología. Esto es, deben ser capaces de tener una cierta experiencia (Tye 2006c:159). Por tanto, no son posibles casos en los cuales exista un sujeto con una cierta distribución funcional (idéntica a aquella que poseen los seres conscientes) que, sin embargo, carezca de fenomenología. Todo sujeto con una capacidad funcional tal, sin ningún tipo de anomalía y en situaciones normales, estará dotado de fenomenología y entrará en un estado experiencial ante determinados estímulos con carácter fenoménico (Tye 2006c: 162), aunque éstos contenido fenoménicos sean diferentes debido a la peculiar historia cognitiva del sujeto.

4.3.5. Mary, *zombies* y el hiato explicativo: la solución de Tye

Como hemos mostrado al comienzo de este trabajo, uno de los mayores problemas con los que se encuentra todo estudio de la conciencia es la intuición habitual de la existencia de un hiato en el mundo entre aquello físico y aquello mental. Parece ser tenido por cierto que en el mundo existen objetos plenamente físicos (como las piedras, las manzanas o los planetas) y fenómenos irreductiblemente no-físicos (como las experiencias y los sentimientos). Frente a esta posibilidad, aparecen varias opciones. Por un lado, encontramos a todos aquellos que consideran que la reconstrucción física del mundo siempre será incompleta porque dejará fuera hechos no-físicos como los anteriormente descritos (Chalmers 1996; Jackson 1982 y 1993). Por otra parte, existen ciertas consideraciones acerca de la posibilidad de que haya una compatibilidad entre lo físico y lo no-físico (esto es, que nada en el hiato explicativo elimine la posibilidad de una explicación física de los fenómenos no-físicos), pero que (en última instancia) existirá siempre

algo de irreductiblemente subjetivo en las cualidades fenoménicas que impedirá su explicación en términos puramente físicos (Searle 1992). También existen, por otro lado, explicaciones físicas que son escépticas respecto de la posibilidad de reducir aquello no-físico a términos puramente físicos. Entre estos, encontramos teorías que confían en la posibilidad de una ciencia física capaz, con el tiempo, de llegar a obtener una explicación puramente physicalista de fenómenos no-físicos, sólo que dicha explicación está todavía excesivamente lejana (Nagel 1974); pero también encontramos teorías que promulgan un cierre cognitivo respecto de lo no-físico, es decir, que postulan la imposibilidad humana de aportar una explicación en términos físicos de aquellos sucesos de carácter no-físico (McGinn 1991).

La idea principal en la que todos ellos coinciden es que una teoría que pretenda ser explicativamente eficaz debe proporcionar una explicación del (presunto) hiato existente entre aquellos hechos físicos y los fenómenos no-físicos que acontecen en el mundo. Como hemos visto, parece que el peso de la prueba recaería del lado de todos aquellos intentos physicalistas de explicación, puesto que la mera postulación de la separación entre elementos físicos y elementos no-físicos ni tan siquiera debe ser demostrada; tan sólo es intuitivamente supuesta a partir de la suposición de que la fenomenología de los hechos no-físicos de la mente recaen directamente en el ámbito de la introspección (inaccesible desde el punto de vista physicalista en tercera persona).²³ Parece que es todo intento reductivo (como lo denominarán los partidarios del hiato explicativo) o, al menos, physicalista quién debe proporcionar una adecuada explicación del hecho intuitivamente cierto de la existencia de una separación entre lo físico y lo no-físico. Y una teoría que no sea capaz de aportar una explicación satisfactoria de dicha separación, simplemente, deberá ser rechazada.

En lo que resta de apartado voy a intentar presentar los argumentos de Tye para negar la posibilidad real de la existencia de dicho hiato entre los hechos físicos y los fenómenos no-físicos de la mente, en tanto que le permitirán concluir que, aunque aparentemente parece existir un hiato entre las experiencias fenoménicas y los hechos cerebrales que los provocan, realmente dicho hiato no es más que una ilusión conceptual (Tye 1995: 178-182; Tye 1999; Tye 2000: cap. 2). Para ello recorreremos los argumentos de Tye para solucionar dos problemas especialmente importante a la hora de tener que rendir cuentas de la posibilidad de reducir todo fenómeno mental a un hecho físico: la posibilidad de la existencia de duplicados físicos que carezcan de un contenido con carácter fenoménico y la posibilidad epistémica de una separación entre conocimiento físico y conocimiento por introspección (el problema de Mary o argumento del conocimiento).

4.3.5.1. “¡Ay Mary, Mary! ¿Realmente aumenta tu conocimiento?”

Pensemos en Mary como descrita en nuestro apartado 1.4.2. A pesar de los largos años de cautiverio en su habitación mono-cromática, Mary ha llegado a convertirse en la mayor experta mundial sobre los aspectos neurofisiológicos involucrados en la visión del color. Nada

²³ Aunque es cierto que autores como Armstrong elaboran teorías centradas en la noción de introspección, aunque desde un espíritu materialista cercano al funcionalismo, desde las cuales hacer desaparecer el hiato explicativo.

hay, desde el punto de vista físico, que Mary no sepa sobre los mecanismos cerebrales y físicos (de reflectancia lumínica, por ejemplo) que permitan la visión del color. Todo ello lo ha llegado a conocer a pesar de que nunca ha tenido una experiencia real de percepción del color. ¿Podría ser el caso, como sugiere Jackson (1982 y 1986), que una vez liberada de su cautiverio cromático Mary pudiera conocer nuevos hechos ante las experiencias reales de color? Debemos recordar en este momento que la pretensión final de este argumento es mostrar la existencia de un elemento irreductible en el hecho fenoménico de experimentar una sensación con carácter fenoménico, lo que permitiría afirmar la falsedad del fisicalismo. Como lo que prioriza el argumento es mostrar la existencia de hechos que no pueden ser explicados desde una aproximación física, entonces, si pudiéramos bloquear la posibilidad de ese tipo de hechos (esto es, si pudiera demostrarse que los hechos fenoménicos también pueden ser explicados desde un punto de vista físico), podría negarse la posibilidad anticipada por Jackson y demostrar, al menos, que este tipo de argumento contra el fisicalismo carece de peso y, aunque no permita afirmar la verdad del mismo, sí podría al menos remover un pesado obstáculo hacia la creación de una explicación plenamente física de la conciencia.

Desde el punto de vista de Jackson, lo que Mary no conocería es cómo es experimentar el rojo, esto es, lo que no conoce es el contenido fenoménico (tal y como lo hemos definido en las secciones anteriores) del estado de percibir el color rojo. La razón de por qué Mary no conoce este contenido es, por un lado, que ella carece del concepto fenoménico de rojo (después de todo, nunca ha tenido una experiencia de rojo y no sabe cómo es sentir una experiencia de rojo) y, por otro lado, ella no puede aplicar el concepto demostrativo al color representado en la experiencia de rojo (no es capaz de identificar ostensivamente este contenido porque nunca antes ha experimentado el color rojo).²⁴ Parece que, desde un punto de vista conceptual, Mary no conoce lo que es experimentar el color rojo. En este sentido, Tye dirá que, al abandonar su habitación, Mary adquiere un nuevo tipo de conocimiento; un conocimiento mucho más refinado referido al hecho “fenoménico conceptual” de ver un determinado color y al hecho “fenoménico conceptual” relativo a ser capaz de identificar ostensivamente objetos con dicho color (Tye 2009: 129).

Pero debemos recordar que la posición de Tye puede ser cualquier cosa menos conceptual. Para Tye, recordemos, el contenido de una experiencia será siempre no-conceptual. Es decir, nunca referirá a los conceptos establecidos dentro de una comunidad lingüística, por lo que nunca aceptará la verdad de la solución tal y como la presenta Jackson (y que hemos reconstruido en las últimas líneas). Desde la posición de Tye, el estado de experimentar el color rojo puede ser naturalizado como una experiencia PANIC, por lo que Mary conocerá de manera inmediata la esencia PANIC de la experiencia de rojo que está teniendo por primera vez porque ella conoce todos los elementos que, desde el punto de vista físico, son relevantes para la posesión de dicha experiencia. Desde esta perspectiva, entonces, Mary no conocerá nada nuevo.

²⁴ Un concepto fenoménico es aquel que tiene como parte de su contenido el particular modo en que las características determinadas de lo que es sentir una cierta sensación para un sujeto. Por su parte, un concepto demostrativo es aquel que tiene como parte de su contenido el objeto concreto hacia el cual se refiere, en un sentido ostensivo, sea ese objeto el que sea.

Para Mary, en el conocimiento de los elementos físicos que permiten explicar la experiencia de rojo ya estaba contenida la esencia PANIC de la experiencia de rojo, por lo que en el momento en que Mary tiene por primera vez una experiencia de rojo tan sólo tendría que pararse a observar dicha experiencia y compararla con todo aquello que ya conoce para darse cuenta de que lo que está experimentando es una experiencia de rojo, algo que ya conocería de antemano. La reacción más coherente, desde el punto de vista de Mary, a la hora de sentir por primera vez una experiencia de color sería una como la siguiente. “Mira, esto debe ser lo que es tener una experiencia de color rojo: porque actualmente estoy sintiendo el contenido PANIC x que determina que lo que realmente tengo es una percepción de rojo, tal y como me han demostrado con los años mis libros de textos y tal y como han predicho los múltiples experimentos neurológicos que durante años he llevado a cabo en mi habitación monocromática”. En este sentido, parece que Mary no conocerá nada nuevo desde un punto de vista no-conceptual (Tye 1995: 174). Si esto es cierto, entonces, parece que el argumento del conocimiento subyacente a la propuesta de Mary no pone en serias dificultades las explicaciones fisicalistas de Tye porque, aunque es cierto (tal y como Tye admite) que Mary amplía su conocimiento desde un punto de vista conceptual (algo totalmente irrelevante cuando de lo que se trata es de establecer el carácter fenoménico de las experiencias), Mary ya conoce todo aquello que debe conocer para establecer la presencia de un determinado contenido fenoménico presente en su sistema físico-nervioso, por lo que no ampliaría su conocimiento fenomenológico.

El problema con esta defensa de Tye, reside precisamente en un malentendido subyacente a su teoría. Podemos admitir que Mary no conocerá nada nuevo desde un punto de vista conceptual. Pero desde un punto de vista no-conceptual sí que deberá conocer algo nuevo: precisamente cómo es fenoménicamente tener una experiencia de color. Ella nunca antes había experimentado una experiencia de color con su peculiar contenido PANIC, aunque lo conociera de manera teórica, de modo conceptual. Aquí tendríamos dos posibles respuestas. O bien consideramos que no existe una diferencia entre conocer de manera descriptiva (conceptual) un mismo contenido PANIC y ejemplificar (conocer por familiaridad) dicho contenido o bien debemos admitir que sí existe una diferencia. Tye parece decantarse por la segunda solución. Pero el problema se acrecienta a la hora de tener que explicar cómo dicho contenido puede ser conocido como un hecho físico, en tanto que también parece existir una diferencia relevante entre describir dicho hecho y ejemplificarlo.

4.3.5.2. El problema de los *zombies* en Tye

Como vimos en la sección 1.4.5., existe otro argumento propuesto por los partidarios de las aproximaciones internistas y dualistas de la conciencia contra el fisicalismo. Este argumento defiende la posibilidad conceptual de duplicados físicos molécula por molécula idénticos a seres físicos conscientes que, sin embargo, a pesar de ser idénticos desde un punto de vista funcional y físico, tienen la peculiar característica de carecer de cualquier tipo de vida fenoménica. Esto es, cuando uno de estos seres duplicados tiene una experiencia de dolor, se comporta del mismo modo cómo lo haríamos cualquiera de nosotros ante dicha situación e incluso tendría activas las

mismas zonas del sistema nervioso y cerebral que permitirían indicar que la experiencia que tiene es de dolor pero, a diferencia de nosotros, dicho ser duplicado no sentiría nada en absoluto. Carecería del carácter fenoménico de sus experiencias. Como lo diría McGinn, cabe la posibilidad de que existan seres que “carezcan de fenomenología en technicolor” (McGinn 1991: 37). Si se demostrara esta posibilidad, aunque fuera de manera conceptual, entonces esto parecería indicar que los aspectos físicos implicados en la fenomenología consciente no son relevantes a la hora de tener experiencias con contenido fenoménico, por lo que podríamos afirmar la falsedad del fisicalismo.

La estrategia de Tye para defenderse de la acusación de falsedad lanzada por la bomba de intuición que supone la ‘corazonada *zombie*’ (tal y como la bautizó Dennett 2001) es extremadamente sencilla. Tye asume que si se logra afirmar la verdad *metafísica* de la existencia de *zombies*, entonces (pero tan sólo entonces), el fisicalismo correría el peligro de ser falso. Pero la mera posibilidad conceptual de su existencia no implica por sí sola su necesidad metafísica, por lo que deberá descartarse la posibilidad de que el fisicalismo sea falso, aunque sea posible la existencia de duplicados físicos sin fenomenología. Veamos su argumentación.

Según la teoría de la referencia directa (propuesta por Putnam 1975 y Kripke 1980 y 1981), es imposible suponer, aunque tan sólo sea conceptualmente, la posibilidad de la existencia de un mundo en donde exista el agua y ésta no sea H₂O. Porque según la teoría de la referencia directa, los designadores (esto es, los nombres que refieren a los objetos del mundo real, como lo son los nombres para las clases naturales) refieren a sus objetos de manera rígida (es decir, de una manera necesaria, dirigida directamente a su esencia, a aquello que los hace ser lo son y no otra cosa, como lo es su composición química, su origen o su genética) a través de todos los mundos posibles. Y el conjunto de todos los mundos posibles también incluye aquellos mundos conceptualmente imaginables. Por lo tanto, no podríamos concebir un mundo donde el agua dejara de ser H₂O.

Pues bien, nos dice Tye, con los conceptos fenoménicos ocurre lo mismo. Parece ser que aquello que caracterizaría a los conceptos fenoménicos es, precisamente, su ser fenoménicos. Esto es, su carácter introspectivo y en perspectiva, pero cumpliendo las condiciones PANIC anteriormente presentadas. Entonces, nos dice Tye, parece que no sería posible concebir, como tampoco lo es en el caso del agua, un mundo que sea posible, aunque sea de manera conceptual, en donde existan conceptos fenoménicos y éstos no sean introspectivos y en perspectiva, o carezcan de alguna de sus características PANIC, del mismo modo que el agua no puede dejar de ser H₂O. Así, vemos que la posibilidad conceptual de la existencia de *zombies* es ilusoria, porque no existe ni tan siquiera la posibilidad conceptual de la existencia de duplicados que aunque parezcan tener conceptos fenoménicos realmente carezcan de fenomenología. (Recordemos que una conclusión similar es extraída por Tye del análisis que realiza del problema de los *qualia* ausentes. Véase sección 4.3.4.2.). Una de dos, o los duplicados físicos que parecen tener fenomenología realmente la tienen (porque si actúan como teniendo experiencias fenoménicas sería porque realmente las tendrían) o debemos negar la posibilidad de la existencia de *zombies* (dado que es necesario que todo aquel ser que parezca

tener fenomenología, realmente tenga acceso introspectivo y en perspectiva a ella, impidiendo su carencia). Entonces, concluye Tye, la existencia de *zombies* es metafísicamente imposible (Tye 1995: 194-200; Tye 2009: 145-150).

4.3.5.3. ¿Existe eso que podríamos llamar hiato explicativo?

Según hemos estado viendo en las últimas secciones, Tye está firmemente convencido de la verdad del fisicalismo. Su teoría representacionista fuerte es un ejemplo del modo en que, a pesar de defender la existencia de elementos fenoménicos, éstos pueden ser explicados enteramente desde los elementos físicos capaces de causarlos. Pero, entonces, parecería poder replicar los defensores de teorías anti-fisicalistas, ¿cómo podríamos explicar el supuesto privilegio que tenemos los seres conscientes respecto de nuestros propios estados internos? ¿Cómo dar cuenta de esa presunta subjetividad que está presente en nuestras experiencias conscientes?

Tye tiene una respuesta sencilla. Los estados conscientes con carácter fenoménico están presentes en el sujeto en perspectiva. Es decir, para que un estado fenoménico pueda ser comprendido de modo completo es necesario que el sujeto se encuentre situado con respecto a él de un cierto modo, el sujeto debe adoptar un particular punto de vista capaz de permitirle la comprensión total de dicho estado fenoménico

“Un modo de comprender el hiato explicativo entre las experiencias fenoménicas y los estados cerebrales es en términos de producción de estados subjetivos en perspectiva” (Tye 1995: 178).

Entonces, en Tye, la presunta subjetividad presente en los estados fenoménicos no es más que la adopción de una determinada perspectiva por parte del sujeto sobre el contenido de dichas experiencias, una perspectiva que se objetiviza en la adquisición de conceptos fenoménicos en tanto que empleados por la capacidad introspectiva del sujeto respecto de sus propios estados fenoménicos y que configura una determinada concepción de lo que sería para él sentir esa determinada experiencia en un momento determinado

“La subjetividad en perspectiva de los estados fenoménicos depende del carácter en perspectiva de los conceptos fenoménicos, donde los conceptos fenoménicos son aquellos conceptos empleados cuando una persona introspecciona en sus estados fenoménicos y conforma una concepción de lo que sería para ella sentir dichos estados en un momento determinado... los conceptos fenoménicos son conceptos indécicos aplicados a partir de la introspección a los estados fenoménicos” (Tye 2000: 25).

Vistos desde este punto de vista, entonces, los conceptos fenoménicos son conceptualmente irreductibles. Es decir, ningún análisis *a priori* puede dar cuenta de ellos en términos conceptuales (a diferencia de posiciones representacionales como las de Jackson o dualismos como los de Chalmers). En este sentido, comenta Tye, es como debemos entender que los conceptos fenoménicos están situados en perspectiva (Tye 2000: 26).²⁵

²⁵ En este sentido, pueden verse las similitudes de Tye con la posición sostenida en Lewis (1990), que considera que la cuestión de que un sujeto sea capaz de saber cómo es tener una determinada experiencia se zanjaría mediante la

Pero lo que también viene a decirnos Tye es que, dado que la misma posesión de conceptos fenoménicos requiere saber cómo es tener una experiencia, entonces parecería necesario postular la existencia de una conexión directa entre la posesión de determinados conceptos fenoménicos con un determinado papel funcional y los mecanismos cognitivos capaces de causar dichos conceptos fenoménicos. En este sentido, aunque los conceptos fenoménicos siguen siendo considerados como conceptualmente irreductibles, el papel funcional que el concepto fenoménico juega a la hora de determinar el conocimiento del sujeto de una determinada experiencia es, precisamente, situarlo en la perspectiva adecuada que permitiría reconocer una determinada experiencia como la experiencia fenoménica concreta de sentir algo (Tye 1995: 180; Tye 2000: 27).²⁶ Vemos aquí una posible incongruencia en la posición de Tye. Si consideramos, con Tye, que los conceptos fenoménicos son irreductibles, pero que, a su vez, pueden ser explicados desde los mecanismos funcionales capaces de causarlos, ¿no estamos contradiciendo nuestra propuesta? Aún es más. Si consideramos, como hace Tye, que el papel del concepto fenoménico es situar al sujeto en la perspectiva adecuada para que éste consiga entenderlo de modo correcto, ¿no estamos cometiendo una petición de principio al presuponer la presencia de aquello que queremos demostrar?

Aún así, Tye considera que el hiato que parece existir entre los estados físicos del sistema cerebral y los fenómenos mentales es ilusorio. Es decir, para Tye, el hiato explicativo sólo existirá si hay algo inexplicable que necesita explicación. Pero algo necesitará explicación tan sólo si puede ser explicado. Por tanto, dicho hiato no puede ser real, porque como parece cree haber mostrado Tye, los hechos físicos conocidos que describen las ciencias del cerebro actualmente son suficientemente variados y ricos como para poder justificar la existencia de elementos físicos causantes de fenómenos mentales, por lo que dicho hiato se muestra como un cierto estado cognitivo de nivel superior que hace que el sujeto perciba en su propia perspectiva interna la ilusión cognitiva de la existencia de un hiato intransitable entre el mundo introspectivo y el carácter fenoménico de su experiencia.

A esta afirmación podría objetársele, todavía, que parece existir un hiato explicativo entre ciertas propiedades físicas y determinadas cualidades fenoménicas conscientes, como por ejemplo los colores (Shoemaker 1997: 229). En su propia defensa, Tye insinúa, como hemos visto, que el carácter fenoménico en perspectiva (aquello que permite causar la ilusión de subjetividad introspectiva) de una determinada experiencia de color es fijado por el contenido mismo de dicha experiencia de color. Esto es, es fijado por sus condiciones PANIC. Y, como hemos mostrado, dichas condiciones PANIC apuntan a la presencia de elementos no-conceptuales que permiten representar (de modo abstracto) las características físicas relevantes de los objetos externos de tal modo que sean capaces de influir en la creación de actitudes proposicionales adecuadas para la dirección de la vida consciente del sujeto. En este sentido, si

explicación de la adquisición de determinadas habilidades cognitivas por parte del sujeto, una explicación que perfectamente podría ser aportada en términos físicos.

²⁶ Harman (1990) infiere una conclusión similar cuando considera que una persona que es ciega desde su nacimiento (es decir, que siempre ha estado condenada a vivir en un mundo acromático) no puede poseer nunca un concepto de color con el papel funcional requerido para poder decir que el sujeto sabe qué es sentir una percepción de color.

existiera algún tipo de hiato, éste no podría darse nunca de modo interno entre las experiencias y los acontecimientos neuronales del cerebro, sino (a lo sumo) entre los colores mismos y las propiedades físicas que los causan.

“Desde mi punto de vista, las cualidades de las que directamente nos damos cuenta al introspeccionar en nuestras experiencias no son cualidades *de* las experiencias mismas, sino cualidades de las experiencias representadas” (Tye 2000: 40 n.13).

El hiato, si es que existe, deberá ser externo y no interno, por lo que será perfectamente explicable a partir de elementos físicos.

“El hiato es irreal, es una ilusión cognitiva... No existen dos clases de fenómenos naturales: aquel irreductiblemente subjetivo y aquel objetivo. El llamado hiato explicativo deriva de un error clásico al reconocer las características especiales de los conceptos fenoménicos. Estos conceptos tienen un carácter que no sólo explica porqué tenemos nuestras intuiciones, algo importante pero que puede ser explicado en términos físicos (y/o funcionales), sino también explica porqué no podemos confiar siempre en ellas” (Tye 2000: 23).

Lo que tan sólo debemos hacer es encontrar las características (físicas) de la representación visual de los colores que permiten que su naturaleza se exponga de modo completo en la experiencia visual para poder, así, eliminar el presunto hiato existente entre lo físico y lo fenoménico (Tye 2000: 37-39; Byrne 2002b: 19-21).

A esta explicación se podría objetar, todavía, que existe un problema a la hora de definir la naturaleza representacional de los fenómenos conscientes en términos físicos. Porque, si no existe tal hiato entre lo fenoménico y lo físico, porque lo físico causa lo fenoménico, ¿cómo, entonces, causa lo físico lo fenoménico? ¿Lo hace de un modo restringido o lo hace de un modo amplio? Si lo hace de un modo restringido, entonces, las propiedades físicas deben ser identificadas con las propiedades fenoménicas. Esto es, dado un determinado dolor debemos suponer la existencia de una excitación en las fibras C del cerebro, por ejemplo. Esto quiere decir que siempre que exista ese determinado dolor, siempre deberá existir dicha excitación en las células C del cerebro, desmintiendo así la tesis de la múltiple realizabilidad que Tye parecía defender (Noordhof 2003: 542-543). Pero tampoco funcionan mejor las cosas si consideramos que lo físico tiene una causalidad amplia sobre lo fenoménico, tal y como parece defenderlo Tye. Porque aunque es cierto que Tye considera que las propiedades fenoménicas supervienen a las físicas, lo que esto mostraría es la necesidad de la naturaleza representacional de los fenómenos conscientes. Dicha necesidad aparecería, entonces, como una verdad *a priori*. Pero si, como Tye ha insistido, sólo podremos instaurar esta superveniencia a partir de una investigación empírica, la necesidad *a priori* de dicha causalidad equipara el representacionalismo de Tye a otros afines como el de Jackson o a otras posiciones como la de Chalmers extremadamente problemáticas a la hora de dar cuenta de dicha relación (véase apartado 4.2.2.).

No nos engañemos. Estoy firmemente convencido de que el fisicalismo es verdadero y que el hiato explicativo que muchos defienden entre las propiedades físicas y las propiedades

mentales es tan sólo una ilusión. Pero, por los argumentos que hemos recogido en las últimas páginas, considero que el esfuerzo del representacionalismo fuerte de Michael Tye no alcanza a ser fructífero y se queda corto a la hora de mostrar una alternativa real a los planteamientos previamente tenidos en cuenta por aquellos que defienden una visión anti-reductiva de la conciencia. Sería, por tanto, necesario construir argumentaciones reforzadas que, dejando de lado los problemas del representacionalismo, permitieran afrontar los problemas de la conciencia fenoménica desde un punto de vista físico.

4.4. Recapitulación

A lo largo de este capítulo, hemos intentado aproximarnos a la posición que en la actualidad defiende Michael Tye: una intencionalismo representacional que, como nuestro autor nos indica, considera que el contenido de los estados fenoménicos es abstracto, intencional, no-conceptual y está preparado para implementar nuevos estados mentales. Como hemos resaltado, su posición representación presenta muchas ventajas respecto de otras teorías afines, en tanto que respeta un cierto externismo a la hora de considerar el contenido de todo estado mental (el contenido de un estado viene determinado por condiciones externas al sujeto) y permite conservar cierto intencionalismo de lo mental. Pero, como también hemos indicado, el problema principal que presenta su posición es, por un lado, mantener todavía un análisis del contenido fenoménico como representacional (en tanto que siempre dependería de los sistemas perceptivos y cognitivos del sujeto la forja de toda impresión y sensación) y, por otro, postular la necesidad de la transparencia de todo contenido de conciencia (en tanto que dificultaría sobremanera la posibilidad de que un sujeto se representara de manera equivocada un determinado contenido).

Capítulo aparte requeriría la excesiva atención que proporciona a la característica de no-conceptualidad que postula en todo estado mental con carácter fenoménico. Como hemos insinuado desde el comienzo, la exigencia de transparencia es demasiado estricta porque, aunque es cierto que a veces posibilita la explicación respecto de los errores de juicio, tiene problemas a la hora de explicar la noción empírica de error en los estados proposicionales y en los fenoménicos y convierte a la vivencia de la experiencia en virtualmente infalible. Pero sabemos que los casos de error son posibles, dado que, a menudo, el sujeto puede desconocer que dos elementos de pensamiento (términos singulares, experiencias conscientes, emociones, sensaciones, etc.) son co-referenciales y, a pesar de ello, puede referirse a ellos. Es decir, a diferencia de Tye (que aceptará un cierto millianismo a la hora de definir el contenido fenoménico, incluido el emocional), consideramos pertinente, al menos, tener en cuenta un cierto descriptivismo de carácter fregeano a la hora de caracterizar e individualizar el contenido de una experiencia consciente. Y todo ello porque, además de parecernos una ilusión (por ser demasiado restrictivo) el requisito de transparencia, que hechos como los mencionados ocurran nos hace dudar de la posibilidad de que todo contenido fenoménico consciente, incluido el emocional, sea no-conceptual. Porque parece necesario que una explicación adecuada de las experiencias fenoménicas debe dar correcta cuenta del modo en que los seres conscientes nos comportamos en determinadas ocasiones, que esas reacciones naturales responden a nuestros

intereses, y estos preceden a toda representación y que una interpretación estrictamente no-conceptual no parece satisfacer estos requisitos (al dejar fuera ciertos elementos conceptuales presentes en nuestra experiencia, siendo necesario apelar de tanto en tanto a ellos, por lo que se hace necesario un estudio de la experiencia fenoménica que respete la intencionalidad presente en la conciencia, algo que (como hemos intentado demostrar) parece que Tye no consigue.

También hemos atendido a su concepción demostrativa de la percepción, en especial la percepción visual. A partir de los ejemplos que el mismo Tye analiza, hemos concluido con él que es necesario apelar a los mecanismos fisiológicos presentes en el cerebro para poder rendir cuentas de la percepción, donde cobra un papel importantísimo la atención conciente (como demuestra el ejemplo de la gallina de Guinea). También hemos repasado problemas básicos dentro de la filosofía de la conciencia (elencados en el capítulo introductorio) y que resaltan, precisamente, la exigencia de la atención a la hora de aprehender el contenido fenoménico de manera consciente. Pero, a diferencia de Tye, al que matizamos precisamente por ello, no aceptamos la posibilidad de que todo contenido fenoménico sea no-conceptual, en tanto que existen casos (como hemos intentado demostrar) en los que es necesario apelar, al menos, al contexto discurso-semántico en el cual se inserta la experiencia concreta que el sujeto está sintiendo en un momento determinado. Precisamente, como hemos concluido, las reacciones que habitualmente los seres semánticos y conscientes que somos realizamos ante determinadas sensaciones nos indica esto mismo.

CAPÍTULO 5. EL REPRESENTACIONALISMO DÉBIL DE DENNETT

Me doy cuenta de algo y busco una razón para ello, lo que originariamente significa que busco una intención y sobre todo alguien que tenga intenciones, un sujeto, un hacedor: para cada acontecimiento, un título de propiedades; antiguamente se veían intenciones en todos los hechos, ese es nuestro hábito más antiguo. ¿También lo poseen los animales?

(Friedrich Nietzsche, *La gaya ciencia*)

Un individuo planea sus acciones si posee una representación de una meta y de su situación presente, y puede generar representaciones de diversas acciones capaces de dirigir hacia esa meta.
(A. Gulz, *The planning of action as a cognitive and biological phenomenon*)

V. EL REPRESENTACIONALISMO DÉBIL DE DENNETT

El filósofo norteamericano Daniel C. Dennett ve la filosofía como una actividad en la cual los pensadores no pueden hacer otra cosa que lanzar sugerencias y compromisos cognitivos que permitan a las teorías filosóficas explicar los hechos físicos que acontecen en el mundo. Pero, como también nos advierte, la argumentación filosófica siempre estará limitada por la preconcepción y los prejuicios de los individuos que se encargan de su desarrollo. Por eso, la primera y principal tarea de la filosofía debería ser ‘desencantar’ y ‘romper los hechizos’ que persiguen y enmascaran la tradición filosófica.

Una de dichas entidades ‘mágicas’ la encontraríamos en la concepción tradicional sobre la mente. La visión de sentido común que intuitivamente a todos nos viene a la cabeza cuando pensamos en el concepto *mente*, engloba una gran cantidad de errores y de malas comprensiones que la filosofía se ha empeñado muchas veces en mantener. Por tanto, es misión crucial en la filosofía de Dennett devolver a estas entidades la realidad que les corresponde, atendiendo a los hechos pertinentes en cada caso.

Con este espíritu reivindicativo de lo real, de lo físico, Dennett inicia sus andanzas filosóficas de la mano de Quine y Ryle. Sus primeros objetivos fueron crear una visión científica en tercera persona que le permitiera construir una estructura física capaz de solucionar el tradicionalmente enrevesado problema de la mente. Dentro de esta tradición, identifica dos necesidades básicas, como ya nos indica el título de su primer trabajo Dennett (1969):

“primero, construir una teoría del contenido o intencionalidad (un fenómeno que es más fundamental que la conciencia) y, después, construir sobre esos cimientos una teoría de la conciencia” (Dennett 1994: 355).

Desde este peculiar inicio, la pretensión de Dennett es revolucionar la concepción clásica de la mente por tener una preconcepción equivocada de la misma.

“Tan sólo unos pocos filósofos consideran que el contenido y la conciencia son sólo dos más de los temas que conciernen a la filosofía de la mente; la gran mayoría siguen a la tradición en el sentido opuesto: la conciencia, creen, es el fenómeno fundamental, del cual la intencionalidad depende. La diferencia de perspectiva es fundamental, ya que infecta las intuiciones desde las que todo teorizar comienza” (Dennett 1994: 356).

Nuestra educación nos predispone a abordar estas intuiciones de modo tradicional:

“Si tu fidelidad se debe a las ciencias físicas y al punto de vista de la tercera persona, esta disposición al tratamiento del contenido y la intencionalidad pueden parecerle intuitivamente aceptables. Si, en cambio, tu punto de partida se sitúa en la tradición filosófica de la mente y la introspección, entonces esta visión te parecerá un ultraje” (Dennett 1994: 357).

De todos modos, como Dennett se empeñará en decir, todas nuestras intuiciones no son más que manifestaciones de modelos de pensamientos, percepciones y juicios (Dennett 1994: 358), manifestaciones que inevitablemente tendrán una base física.

Dennett será reconocido, sobre todo, por ser el iniciador de muchas de dichas intuiciones, por proponer muchos de dichos modelos, que a modo de metáforas, puzzles, enredos, historias y ejemplos enriquecedoramente contruidos (lo que les ha valido el apelativo de *intuition pumps*, Cf. Dorbolo 2006), un modo que ha permitido revolucionar y renovar las tradiciones filosóficas y sus prácticas con el objeto de clarificar la gran cantidad de confusión conceptual presente en la filosofía moderna y contemporánea. Después de todo,

“una *intuition pump* podría ser la herramienta ideal del filósofo, si nos tomamos en serio una de las mejor conocidas ejemplificaciones de la filosofía. Como dijo Wittgenstein ‘la Filosofía es una lucha [continua] contra el embrujo de nuestro entendimiento por medio de nuestro lenguaje’ (*Investigaciones Filosóficas*, #109)”.

La teoría de la conciencia que sostiene Dennett es sumamente interesante. Vincula directamente la aparición del pensamiento y de la conciencia con la posibilidad de que ambos, en tanto que instrumentos ficcionales pero sumamente útiles, sean producto de la evolución progresiva de sistemas de adaptación al entorno que, generación tras generación han ido afianzándose en las especies hasta desembocar en lo que actualmente conocemos como mente (en el caso de los humanos, pero debemos indicar que, para Dennett, la mente es algo heterogéneo y casi cualquier sistema puede llegar a ser considerado, al menos, *consciente*). Es por ello que no podemos dejar de comenzar este capítulo atendiendo a la concepción general evolucionista que Dennett realiza de la aparición del pensamiento, la representación y, por ende, de la conciencia. En tanto que el germen de todo sistema consciente es el ámbito evolutivo que ha permitido su desarrollo hasta que lo que actualmente es, es sumamente importante prestar atención a esta teoría evolucionista antes de abordar la teoría de la conciencia de Dennett propiamente dicha.

A la hora de atender a la conciencia particular de los sujetos, ya en tanto individuos y no tanto como especie, es importante también para Dennett desmitificar la concepción cartesiana de la mente como un teatro, dejando paso a una interpretación puramente externista del contenido mental que, en un sentido fuerte, tiende a reducir cualquier contenido a elementos puramente físicos, asociados a la neurobiología del cerebro o a la distribución funcional de los sistemas que componen el sistema perceptor en cuestión. Ello nos permitirá inmiscuirnos en la concepción que Dennett tiene sobre la representación mental, entendida en un sentido ficcional, pero que será extremadamente útil a la hora de proporcionar a los sujetos perceptores la información necesaria para, primero, conformarse una adecuada imagen del mundo (lo que le permitiría prever acontecimientos) y, segundo, poder interactuar adecuadamente con el mismo a partir de la elección de la más adecuada acción según el momento lo requiera.

Ello nos llevará a precisar el sentido en que Dennett pretende resolver la supuesta paradoja que (como vimos en el prólogo) se da cuando atendemos a la posibilidad de que un sujeto actúe de manera libre en un mundo totalmente determinado por causas físicas: un compatibilismo *sui generis* que, a pesar de considerar que no siempre un agente puede ser el causante último de su acción (porque puede darse el caso, como así ocurre, de que no dependa

de sí mismo la posibilidad de auto-configurarse como sujeto de acción y porque no siempre es posible la existencia de varios modos de acción para poder elegir entre ellos), no por ello el agente deja de ser responsable de aquello que hace.

5.1. El desarrollo filogenético de la conciencia: evolución e intencionalidad en la filosofía naturalista de Dennett

Tal vez la idea más controvertida de la filosofía de Daniel C. Dennett sea ésta: los seres humanos no somos más que mecanismos dotados de intencionalidad derivada configurada por la evolución de las especies. Esto ha sido considerado por algunos como una afrenta porque nos relega al estatuto de simple artefacto: las acciones del ser humano, así, no responden a sus propios intereses sino que se ven programados por “agentes internos especializados para recibir la información disponible en la periferia del cuerpo” (Dennett 1996a: 102) que constituyen. Es decir, el ser humano queda relegado a un simple compuesto de pequeños sistemas *intencionales* organizados según los propósitos de la evolución, propósitos que por otra parte son ciegos. Sus críticos apuestan por la renuncia a este tipo de intencionalidad: un ser humano dispone de intencionalidad real o intrínseca, de lo contrario no pasaría de ser un mero *zombie* sin voluntad ni decisión. Esta crítica apuesta también por denunciar el carácter tan poco intuitivo de dicha tesis: según el sentido común, los humanos nos sabemos como libres y responsables de nuestros actos, realizamos nuestros movimientos según nuestras propias intenciones.

Los estudios de Dennett, a pesar de sus fines puramente filosóficos, pueden encuadrarse dentro del campo de las ciencias cognitivas. Por su naturaleza multiforme, esta disciplina combina metodologías, teorías y conceptos pertenecientes a diversas esferas científicas, unos ámbitos no siempre fácilmente compatibles metodológicamente pero, nunca, excluyentes. En pocas palabras, el estudio de las tesis de Dennett obliga a que nuestra investigación sea interdisciplinar: nos involucra con terminología filosófica y científica y nos obliga a emplear diversas metodologías (biología, neurología, psicología, etc.) para hacer justicia a la profundidad de los pensamientos transmitidos en los escritos de nuestro autor, profundidad que se irá abriendo paso a medida que avancemos en nuestra investigación y vayamos enfrentándonos a los problemas y los conceptos.

5.1.1. El papel de la evolución en la intencionalidad de la mente

Darwin revolucionó al mundo al introducir la teoría de la evolución. Dicha teoría afirma que todos los seres son producto del proceso de selección natural iniciado no por una Mente Universal ni por una Inteligencia Divina, sino por la acumulación aleatoria de modificaciones ambientales guiadas por algoritmos que permitieron la configuración de un cierto tipo de orden capaz de crear, desarrollar y evolucionar cierto tipo de organismos que consiguieron elevarse hasta la posibilidad de controlar su propio desarrollo.

Este proceso de evolución, en principio arbitrario, acabó por conformar unos organismos con una cierta distribución fenotípica (una cierta organización genética y bioquímica) capaz de albergar mentes primitivas (proto-mentes), mecanismos capaces de

hacerse cargo de sus propias modificaciones. Esto permitiría explicar de modo sencillo el paso del diseño natural al (re-)diseño inteligente controlado por el mismo individuo. Pero, ¿cuál es el problema?

La explicación naturalista (científica) del nacimiento de la mente humana desmitifica el mundo humano, porque esta tesis desmiente la idea de la *folk psychology* (psicología popular) que afirma que el ser humano es libre porque es capaz de causar sus propias acciones (como ya indicamos en el prólogo). En realidad el ser humano no es libre de actuar libremente de este modo porque no elige sus propias acciones, sino que tan sólo responde a la configuración que realiza por él su propia historia cognitiva, lo que restringe su ámbito de acción.

El desarrollo se produce de modo ciego pero no de modo desordenado. Los organismos primigenios son modificados aleatoriamente para probar su resistencia al entorno en que habitan: aquellos diseños que funcionen se potenciarán, los que no funcionen tenderán a desaparecer. En palabras de Henry W. Beecher,

“el diseño al por mayor es de mayor grandeza que el diseño al por menor” (citado en Dennett 1995: 100).

Pero, si el supuesto orden no viene establecido por una intencionalidad superior, ¿cómo pudieron evolucionar los organismos mejor diseñados? Esto es, si negamos la existencia de lo que Locke llamó Mente Universal, ¿a quién podemos atribuir el objeto de la evolución?

5.1.1.1. El relojero ciego

El punto fuerte de la teoría evolucionista (teoría que Dennett hará suya) es la afirmación que dice que la evolución actúa (en palabras de Dawkins 1986) como “un relojero ciego”: la evolución posibilita el caldo de cultivo que permitirá la posibilidad de un auto-desarrollo en los individuos, un desarrollo que no viene delimitado por leyes rígidas ni estrictas que marcan el recorrido, sino que viene delimitado por la imposibilidad de cometer incompatibilidades. Es decir, las partículas bio-químicas que componen todo individuo no pueden mezclarse sin más, sino que suponen una cierta normatividad externa indicada por las enzimas que las componen y que permiten, a la vez que restringen, su posibilidad de unificación en partículas más complejas. Dennett emplea una metáfora: es como si cada partícula tuviera una puerta que sólo puede abrir quien posea la llave adecuada. Esta unificación sí que es aleatoria ya que depende del caso concreto de cada partícula el que llegue a unificarse (o no) permitiendo constituir, así, mejores estructuras.

¿Cómo debemos entender este proceso de unificación y desarrollo? Muchos han supuesto que lo que Dennett (y otros evolucionistas como Ruth G. Millikan)¹ defienden cuando hablan de que este tipo de superación biológica es una determinación física. Pero nada más lejos. Dennett apela a la genealogía para explicar la complejidad de dicho fenómeno: del mismo modo que en ingeniería se recurre a la *ingeniería inversa* para analizar los proyectos de la

¹ Véase especialmente Millikan 1984 y 1993 y lo expuesto en la sección 2.2.1.2.

competencia, también los organismos recurren a esta estrategia (es decir, también los organismos biológicos inspeccionan su estructura para analizar el desarrollo que sea óptimo para su supervivencia), porque los organismos son productos *de diseño* creados indirectamente por el entorno en conjunción con su propia distribución interna (producto de su anterior desarrollo), que no responde más que a los intereses de supervivencia que la evolución depositó ciegamente en el inicio. Los organismos se desarrollan según conviene para sobrevivir. Pero esta declaración corre el peligro de ser malinterpretada (en tanto que puede existir la tentación de ser interpretada en un sentido adaptacionista lamarckiano): lo que Dennett pretende decir es que el organismo no se adapta al ambiente, sino que los organismos se modifican a sí mismos (o son modificados) o se auto-diseñan sólo según el modelo que viene codificado en su genética (que debe ser entendido en un sentido amplio y abierto, y no en sentido restringido y cerrado), y si su nueva estructura reacciona bien en el ambiente se verá potenciada por su funcionalidad. Esto es, los organismos parecen comportarse, a menudo, de tal modo que podemos comprender y explicar dicha conducta como *intencional* (Dennett 1987a: 27).

Esto desmiente toda pretensión de mitificación del proceso por el cual apareció la mente humana. El ser humano es uno más de los organismos evolucionados. Como todas las demás criaturas, también el hombre comenzó sus andanzas a partir de la ordenación delimitada por los algoritmos de la evolución, sólo que su evolución le ha permitido llegar hasta un punto en el cual parece tener cierto tipo de intereses. Pero dichos intereses no son suyos propios sino que vienen determinados por toda la estructura fenotípica que dicho individuo posee, estructura que (recordemos) en un principio fue pasiva y moldeada por la acumulación aleatoria de los algoritmos de la evolución. Por lo tanto, concluye Dennett, la intencionalidad *real* o intrínseca que muchos parecen atribuir al ser humano es una ficción y cualquier intento de fundamentarla no es más que un intento de creación de *ganchos celestes* (Dennett 1995: 102), de fundamentar ideas en hechos trascendentales sin ningún tipo de evidencia empírica. Si algo explica la existencia de la intencionalidad esto es el proceso evolutivo entendido como progresiva elevación de grados de refinamiento en los fenotipos estructurales de los organismos. Dichas elevaciones son realizadas mediante *grúas*² concretas (como capacidades añadidas o recursos aleatorios) que permiten una explicación de las nuevas actitudes y conductas (Dennett 1995: 116), aunque más desmitificada y desencantada, de cómo realmente llegamos a poseer una mente capaz de representar objetos. El empleo de estas grúas debe ser entendido en tanto metáfora que define mejoras en la adaptación de los órganos y de sus derivados. Pero como hemos explicado, dichas representaciones no son directamente elegidas por los individuos sino que son producto de la cognición que posibilita el conjunto evolucionado de su sistema, son producto de la evolución. Por lo tanto, el ser humano (aunque parece que actúe siempre según sus propios deseos y creencias) tan sólo puede disponer de una intencionalidad *derivada*.

² La imagen de las grúas empleada por Dennett es sumamente importante. Al igual que en las construcciones civiles, las grúas no pueden montarse con la ayuda de otras grúas más altas ni a partir de 'ganchos celestes', las grúas convencionales se levantan desde dentro. Esto es, la propia grúa va levantando poco a poco sus partes constituyentes hasta llegar al punto más alto que permite su estructura.

5.1.1.2. Organismos biológicos e Inteligencia Artificial

Dennett (1978b) comienza por denunciar los 25 años durante los que el cognitivismo pretendió ser la única respuesta posible al problema de la mente, porque no estaría claro en qué sentido podemos entender la mente en un modo computacional clásico (entendiendo la mente, como veremos posteriormente, en tanto simulación). Este *Computacionalismo de Alta Iglesia* (Dennett 1987b: 60) pretendía responder de modo mecánico a las pretensiones de simulación de capacidades cognitivas humanas, indicando que los estados mentales y la conducta de los sujetos se debe al procesamiento de fórmulas sintácticas capaces de relacionar de modo causal entradas y salidas de datos de modo unidireccional. Pero esta ortodoxia no es capaz de salvar dos lagunas. En primer lugar, no logra salvar el problema que plantea *el* (llamado) *cuello de botella de von Newmann*³ (“el procesamiento de información simbólica se basa en reglas secuenciales que se aplican una cada vez”, Varela 1990: 55). Si los humanos procesaran las *órdenes* mentales de una en una, el procesamiento de información se ralentizaría, algo que desmiente la intuición que indica que un humano procesa la información muy rápidamente. En segundo lugar, el cognitivismo también considera que “el procesamiento simbólico está localizado” (Valera 1990: 55), que cierta información se encuentra en sitios localizables del cerebro, algo falso porque cuando se dañan ciertas partes del cerebro, el procesamiento de información encuentra vías alternativas de comunicación (como, por ejemplo, la potenciación de las reacciones galvánicas de la piel. Cf. Dennett 1992). Estas lagunas pueden solucionarse si en lugar de establecer sistemas lineales de procesamiento de información se opta por un sistema de procesamiento distribuido en paralelo, un sistema que permite el procesamiento de distintos algoritmos al mismo tiempo.

Los sistemas de procesamiento distribuido en paralelo evitan, además, todo tipo de apelación a representaciones simbólicas (entendidas en términos de un lenguaje de procesamiento computacional o, incluso, muy relevante para nuestro caso, un lenguaje del pensamiento interno al sujeto). Visto de este modo, entonces, se entiende por qué Dennett no está dispuesto a aceptar la posibilidad de la existencia de un lenguaje del pensamiento en tanto que no existiría una base computacional clara que permitiera una entidad sintáctica de tal tipo, sino que se inclinaría por aceptar una serie de ilusiones derivadas del procesamiento de información entrante. Una información entrante que podríamos entender como datos sensitivos puros directamente percibidos a través de los sistemas cognitivos básicos de percepción y reelaborada según su relevancia (en este sentido, un algoritmo deberá ser entendido como una determinada manera de procesar información, en tanto que permitiría simplificar el modo de respuesta comportamental del individuo en cuestión).

³ Este problema lo plantea por primera vez John von Newmann con la intención de reformular la llamada máquina de Turing. Una máquina de Turing es una estructura sintáctica conceptualmente simple capaz de realizar funciones de computación mediante la entrada y salida de datos. Pero dicha máquina sólo realiza un movimiento cada vez, por lo que el proceso es lento para cálculos complicados. Cf. Turing (1937) y von Newmann (1958).

La solución de Dennett al problema de la conciencia pasará por considerar un centro de conciencia no localizado ni centralizado, sino distribuido y complementado por toda la red cerebral y neuronal (Cf. Dennett 1991, especialmente el capítulo 9).⁴

Como hemos dicho, Dennett insiste en la afirmación que defiende al humano como un ser intencional (Dennett 1978a y 1987a). Los humanos son sistemas más o menos complejos de células cuyo comportamiento puede ser analizado, comprendido y explicado como respondiendo a cierto tipo de creencias y deseos (y otros estados mentales); esto es, su conducta puede ser explicada apelando a actitudes proposicionales. El problema surge cuando intentamos unir esta tesis de la existencia de sistemas intencionales con la polémica tesis dennettiana que afirma que los humanos son también artefactos, sólo que desarrollados según los algoritmos de la ley de la evolución. Entonces, como cualquier artefacto, el ser humano tan sólo posee una intencionalidad derivada, lo que obliga a que esos deseos, razones, creencias, etc., a las que parece responder el sujeto en realidad sean considerados como extrínsecos al mismo.

5.1.1.3. ¿Por qué plantea un problema la intencionalidad derivada?

Es famosa la polémica que Dennett sostiene con varios filósofos acerca de la intencionalidad. Como él mismo confiesa (Dennett 1987a: 255), en una de las muchas discusiones que llevó a cabo con Fodor acerca del problema de la representación errónea, se le ocurrió simplificar al máximo las situaciones en las que se llega a un error en la representación de un estado (mental), recurriendo a un sencillo mecanismo expendedor capaz de aceptar monedas. Fodor argumentó que el caso era irrelevante porque de lo que se trataba era de ver cómo es posible el error de representación en seres que poseen intencionalidad original y no en meros artefactos dotados con intencionalidad derivada. Dennett preparó un pequeño escrito (revisado y reeditado como el capítulo 8 de Dennett 1987a) en el que defendía, por un lado, que los artefactos poseen una intencionalidad derivada y, por otro, que los seres humanos no somos más que artefactos modelados según las leyes de la evolución.

Toda una serie de filósofos replicó a su artículo afirmando que el hombre posee intencionalidad *original*: esto es, y en palabras de Dennett, considerando que “nosotros somos significadores sin significado” (Dennett 1987a: 255). Las teorías que defienden la existencia de una intencionalidad original afirman que el genuino agente es aquel capaz de significar por sí mismo todo aquello que hace: ningún artefacto podrá jamás lograrlo, ya que sólo será una herramienta con *el propósito de...*, capaz de hacer *como si...*, fabricado para satisfacer intereses y necesidades programadas en un *software* por un diseñador externo, lo que indica que los objetivos que (aparentemente) presente no serán genuinamente suyos, sino que derivarán de aquellos originalmente instalados por el programador.

Dennett comprendió entonces que el problema de la representación errónea sólo se plantea en aquellas teorías que conciben que los humanos somos los poseedores de dicha intencionalidad original. Para ello, Dennett se apoya en una serie de ejemplos.

⁴ Atenderé a la teoría de la conciencia no-centralizada de Dennett en la sección 5.2.

Concibamos una máquina expendedora de refrescos con un dispositivo dotado de dos tipos de estado (sí/no, por eso es llamado *two-bitser*) capaz de discriminar monedas de 25 centavos de dólar de todo aquello que no lo es. Su intencionalidad es derivada porque sus constructores la diseñaron para ese fin. Cada vez que se introducen 25 centavos de dólar, el dispositivo las acepta y entra en el estado Q, pero cabe la posibilidad de que alguna otra vez pueda entrar en el estado Q cuando se le ha introducido otra cosa que no es una moneda de 25 centavos de dólar o también es posible que no entre en el estado Q a pesar de que se ha introducido la moneda de 25 centavos. Así, es posible que la máquina pueda representar erróneamente el estado Q porque puede *saltar* a pesar de no ser una moneda de 25 centavos de dólar lo que se le introduzca (al igual que puede ser que alguna de las veces que sí se introduzca una moneda de 25 centavos de dólar, el dispositivo no entre en el estado Q). Pero esta situación no plantea ningún problema desde el punto de vista de la intencionalidad porque, como hemos dicho, este dispositivo sólo tiene una intencionalidad derivada. Digamos que tan sólo podemos afirmar que las entradas del dispositivo en el estado Q son equívocas o acertadas si partimos del supuesto de que somos nosotros, los constructores del dispositivo y los usuarios, los que interpretamos que lo son; es decir, si partimos del hecho que afirma que dicho dispositivo discrimina realmente monedas de cuarto de dólar porque es nuestra intención que así sea. Hasta este punto, todos los críticos estarían de acuerdo con Dennett, ya que todos consideran que los artefactos disponen de una intencionalidad derivada.

Consideremos ahora que la misma máquina es llevada a Panamá. Su funcionamiento es idéntico: cuando alguien introduce una moneda, el dispositivo debe detectar que es la moneda adecuada y proceder para expender una lata de refresco. Sólo hay una modificación: en Panamá, en lugar de tener que identificar monedas de 25 centavos de dólar, el dispositivo debe detectar monedas de 25 centavos de balboa. Cada vez que alguien introdujese una moneda de 25 centavos de balboa, el dispositivo *two-bitser* debería entrar en un estado físicamente idéntico a Q. Entonces, ¿por qué negar que cuando se le introducen 25 centavos de balboa no entra en el estado Q, sino que entra en el estado, digamos, QB? La respuesta obvia sería decir que porque no fue esa la intención de los constructores de dicho dispositivo. Los programadores del dispositivo *two-bitser* pretendían construir un dispositivo capaz de discriminar monedas de 25 centavos de dólar y no monedas de 25 centavos de balboa. Pero a pesar de que el dispositivo *two-bitser* no fue creado para discriminar monedas de 25 centavos de balboa, también funciona para este propósito. Esto quiere decir que, desde el punto de vista intrínseco, nada en la constitución del dispositivo que permite identificar monedas de 25 centavos de dólar permite diferenciarlo de un dispositivo construido intencionalmente para discriminar monedas de 25 centavos de balboa, ya que ambos responden satisfactoriamente al propósito de identificar el mismo trozo de metal por peso y tamaño. Es decir, responden adecuadamente a la misma función. En realidad, el estado físico en el que entra el dispositivo cada vez que detecta una moneda, carece totalmente de significado. Para Dennett, el que el dispositivo *two-bitser* trasladado a Panamá pueda situarse en el estado Q o en el estado QB dependería del hecho de que la persona encargada del traslado de la máquina expendedora de bebidas que lo contiene lo

hiciera por la capacidad del dispositivo de identificar monedas de 25 centavos de balboa. Si no fuera así, el estado en el que el dispositivo se situará cada vez que se accionara sería el estado Q; pero si fue ésta su intención, entonces el dispositivo cada vez que *salte* se situará en el estado QB. Se mantiene así la hipótesis de la intencionalidad derivada en los artefactos: en el primer caso, el dispositivo entraría erróneamente en el estado Q cada vez que alguien introdujera en la máquina expendedora una moneda de 25 centavos de balboa; pero en el segundo caso, el dispositivo representaría correctamente el estado QB, pero estaría sujeto a posibles errores al representar dicho estado. Sólo podemos decir que dicho artefacto detector responde adecuadamente a su función porque satisface la intención de los constructores, los propietarios y los usuarios de la máquina expendedora en la que está instalado. Esto es, sólo podemos afirmar que un artefacto muestra intencionalidad si responde satisfactoriamente a los propósitos que se sitúan en su contexto. Hasta aquí todos (tanto Dennett como sus críticos) están de acuerdo. Pero los críticos comenzarán a disentir cuando apliquemos la misma estrategia a otro tipo de organismos no-mecánicos.

Dennett ofrece otro ejemplo. Basándose en los experimentos de Lettvin, Maturana, McCulloch y Pitt (1959) sobre el sistema visual de las ranas, afirma que la indeterminación funcional que facilita la adaptación de los estados cognitivos de las ranas permite una explicación de la evolución de la intencionalidad. Las ranas disponen de un sistema visual adaptable que les permite funcionar de modo eficaz a la hora de alimentarse, porque su percepción de algo identificado como comestible permite activar un cierto estado mental (algo así como ‘veo ahí comida ahora’) que permite disparar la base de su lengua con precisión suficiente como para atraparlo. Esto indica que, a pesar de no concretar qué sea aquello que observa (ya que se dispara el mismo estado si lo que se presenta a su campo visual es una bola negra, una mosca o una mancha), el cerebro de la rana responde adecuadamente a los estímulos visuales adecuados permitiendo activar los mecanismos necesarios para su alimentación. Esto, según Dennett, permite desmitificar el esencialismo de la intención. Si lo que significa dicho estado mental es indeterminado y sólo puede concretarse mediante una intervención externa, entonces la intención (aparente) de una rana de disparar la base de su lengua para alimentarse no puede nunca ser intrínseca, ya que responde a estímulos externos por medio de respuestas internas.

Hemos visto cómo tanto los artefactos contruidos por el hombre como los organismos biológicos menos evolucionados pueden considerarse como sistemas con intencionalidad derivada, y sólo se les puede atribuir un error de representación en tanto que se someten a interpretación por parte de aquellos seres capaces de aplicar la estrategia intencional. Pero, ¿qué ocurre con los humanos?

Los críticos de las tesis dennettianas consideran que un humano no sólo es capaz de representar objetos, o de entrar en estados mentales como Q o QB, sino que son capaces de reconocer en dichos estados un significado, porque dichos estados intrínsecamente significan algo profundo (visto desde un punto de vista internista) o bien porque el significado es algo real con existencia independiente de los humanos particulares (desde un punto de vista externista), lo

que les permite concluir que los humanos tienen una intencionalidad original. Para tratar esta cuestión suelen recurrir al experimento putnamiano de la Tierra Gemela.⁵

Pensemos en un individuo al que mientras duerme es trasladado a la Tierra Gemela. Al despertar, mira por la ventana y al ver un caballo se le activa un estado mental, 'caballo'. Lo que ocurre es que en la Tierra Gemela, *caeteris paribus*, los caballos no son realmente caballos sino algo (molecularmente hablando) diferente, algo que denominaremos *percaballos*. Lo que pretenden estos defensores de la intencionalidad original es que el estado mental que tiene nuestro amigo realmente significa 'caballo' y no 'percaballo'.

Pero según Dennett, el significado es indeterminado y no depende de nosotros mismos. Nosotros no somos más que mecanismos similares a los *two-bitser*, sólo que mucho más sofisticados. Si yo en la Tierra al ver un caballo entro en el estado mental C y mientras duermo soy trasladado a la Tierra Gemela y vuelvo a ver un caballo y vuelvo a entrar en un estado físicamente idéntico a C (sólo que aquí lo llamaremos PC), ello no importa. ¿Por qué? Porque el estado mental se activa de todos modos. Y aunque en la segunda situación no es realmente un caballo sino un *percaballo* lo que veo, de todos modos despierta el mismo tipo de estado. Del mismo modo que el dispositivo *two-bitser* entraba en el estado Q cuando se le introducía una moneda de 25 centavos de dólar como cuando la moneda era de 25 centavos de balboa, ¡no hay nada en la configuración intrínseca (y esto es lo importante) de mi estado mental C que se dispara cuando veo un *percaballo* que lo haga diferir del estado mental de mi doble cuando se encuentra en el estado PC por el mismo hecho! Y lo que indica realmente en qué estado me encuentro depende de la carga semántica del lenguaje que se emplee. Este ejemplo, concluye Dennett, afirma que los humanos también son sistemas intencionales derivados cuya intencionalidad es extrínseca. Como es obvio, los críticos no aceptan este argumento porque relega al individuo humano a mero instrumento. En otras palabras, todos los críticos de la teoría dennettiana consideran altamente contraintuitivas sus tesis, desplazándolo demasiado a menudo hacia posiciones irrealistas respecto a la posibilidad de explicar la intencionalidad.

Dennett recurre a otro ejemplo para intentar convencer a sus críticos. Imaginemos que queremos criogenizarnos para despertar en el año 2406, debemos entonces construir una cámara criogénica capaz de resistir infortunios y eventualidades ya que no podemos confiar en algo externo que garantice su correcta conservación. Necesitamos por ello construir un artefacto con cierta autonomía y autocontrol interno, para lo que deberemos dotarla de objetivos generales, pero definidos, y de la suficiente tecnología como para poder improvisar y poder forjar

⁵ Los experimentos mentales de la Tierra Gemela se introducen por primera vez en la literatura filosófica por Putnam 1975 para defender una posición externista respecto del significado lingüístico. Desde entonces, son utilizados también por todos aquellos que pretenden defender posiciones externistas que, de un modo u otro, conciben un aspecto externo en la atribución de significado a elementos mentales. Las críticas desde posiciones internistas a estos argumentos suelen también hacer uso de contraejemplos empleando la Tierra Gemela, por ser elevado el poder explicativo que proporcionan. En este sentido, una Tierra Gemela es un lugar contrafáctico idéntico a nuestra Tierra donde, *caeteris paribus*, suele existir una pequeña diferencia en relación al significado de algo concreto. En el ejemplo original de Putnam, por ejemplo, la diferencia crucial, que permite afirmar la posición externista del significado, es una diferencia en la composición química del agua que, sin embargo, no tiene relevancia a la hora de que los hablantes puedan referirse a ella de modo normal. Es crucial, también, la existencia de un *döppelganger*, esto es, un doble idéntico molécula por molécula a un sujeto original (Óscar Gemelo, en el ejemplo de Putnam).

estrategias que maximicen las posibilidades de alcanzar dichos objetivos, como la capacidad de evaluar beneficios o discriminar las acciones ineficaces. Un artefacto de este tipo dispone de intencionalidad derivada, porque no perseguirá sus propios objetivos sino los de aquellos que lo construyeron y programaron.

Según Dennett, un artefacto como la cápsula criogénica no difiere demasiado de los seres humanos. Como en el caso de la cámara criogénica, Dennett considera que los humanos también somos sistemas programados intencionalmente por nuestros *genes egoístas* (Dawkins 1989). La evolución de organismos cada vez más complejos no es más que el modo más eficaz que han encontrado los genes para su supervivencia y expansión. La preservación de los genes es la razón de ser del sujeto humano, nos dirá Dennett. Lo que ocurre es que en el caso de los seres humanos, dichos sujetos fueron diseñados de modo que llegaron a ser capaces de ignorar el objetivo primordial y llegar a constituir su propio *summum bonum*. Esto nos indica que la intencionalidad humana es ficticia: se la proporcionó la evolución. Que el humano haya evolucionado hasta el organismo con mente compleja que es hoy en día no es más que una casualidad. Los genes serían, entonces, los verdaderos significadores sin significado.

Los genes realmente se comportan como si tuvieran intencionalidad. Toda su intencionalidad es *como si* realmente tuvieran representaciones propias, pero en realidad no poseen más que una intencionalidad ficticia y derivada. Pero si no son los genes los que dotan de intencionalidad, ¿de dónde proviene ésta? De la propia Naturaleza; es ella la que ciegamente produce los diseños desde los que los genes pueden desarrollarse. Es la evolución la que selecciona un diseño, potenciándolo, y todo ello sin necesidad de representaciones conscientes o inconscientes de las razones para dichas elecciones. Las primarias razones son la optimización de la función a cumplir por dicho diseño. Pero de todos modos podemos caer en la tentación de establecer principios de explicación para dichas elecciones, algo totalmente falso porque las elecciones se realizaron a partir de la funcionalidad de sus movimientos.

¿Por qué la idea que afirma al hombre como un simple artefacto evolucionado dotado de intencionalidad derivada causa tantos problemas? Dennett aporta dos razones. Por un lado, si aceptamos la idea que afirma que somos artefactos con intencionalidad derivada, entonces aquello que pensamos no tiene un significado real porque tan sólo cumplimos la función que nos fue encomendada. El significado sería por tanto una ficción que no depende de nosotros: esto significaría negar la tesis del acceso privilegiado de primera persona a los propios estados (algo a lo que Dennett, como buen quineano, estaría dispuesto). Por otro, si aceptamos que somos artefactos, no podemos tener acceso de ningún tipo porque no existe nada profundo que sea el significado de nuestros estados; todo estado u órgano tan sólo puede tener una significación funcional. Pero no surge ningún tipo de problema si respetamos la teoría darwiniana de la evolución: los indicios nos muestran que no somos más que simples criaturas evolucionadas que surgieron de aquel primigenio *ácido universal* constituido de modo arbitrario por la acumulación de adaptaciones progresivas, por lo que la aparición del significado, aunque sea entendido como una ficción, cumple a la perfección un papel instrumental que permite la

anticipación de las conductas futuras, marcando las pautas de su desarrollo y estableciendo las reglas según las cuales debe realizarse.

5.1.1.4. El verdadero problema de la representación errónea

Tal vez la resistencia más perspicaz a las tesis de Dennett es la planteada por Fred Dretske,⁶ basándose en las ideas que plantea en *El conocimiento y el flujo de información*.⁷ Según Dretske, es necesario que encontremos en el dispositivo *two-bitser* incorporado a la máquina expendedora de bebidas un principio interpretativo que (como sucede con los biólogos respecto a la Naturaleza) nos diga qué funciones le son propias. Dado que Dretske considera posible la atribución de estados con significado natural (unos estados con un contenido estrecho, ya que no puede incluir ningún tipo de error en su activación), para que sea posible que la máquina expendedora con el dispositivo *two-bitser* se equivoque a la hora de entrar en un determinado estado, debemos presuponer que el contenido de los estados en los que pueda encontrarse debe ser un contenido con significado funcional, por lo que dicho significado debe ser externamente adquirido dado que dicho dispositivo puede representarse falsamente sus propios estados (esto es, debe considerarse que el dispositivo *two-bitser* sólo tiene *significado_f*, al estilo de Dretske).

Dennett afirma la incorrección de esta crítica. Teniendo en cuenta el ejemplo de Braintenberg del transductor de simetría, afirma que para la evolución, el que una función pueda a menudo (y esto es lo importante) representar falsamente un contenido es un bajo precio a pagar si ello permite una mayor eficacia por parte de un mecanismo que por otro más funcional pero menos versátil. Por ejemplo, un pez que disponga de un transductor de simetría que pueda llegar a fallar en determinados momentos a la hora de representarse un contenido (por ejemplo, el contenido 'hay un predador cerca') cada vez que algo de cierto tamaño se mueva cerca de él a pesar de que realmente esto no sea un depredador sino que confunda a un pez no-peligroso por un peligro potencial, pero que sea lo suficientemente rápido y ligero como para posibilitar una huida que minimice riesgos, es preferible y tendrá mayores posibilidades de sobrevivir que otro pez que disponga de un transductor de simetría mucho más preciso (capaz de representar correctamente sólo aquello que suponga un peligro) pero que procese la información más lentamente y tenga un peso mayor porque permite una huida mucho más torpe y arriesgada. Esto demuestra que hablar de principios naturales en los que se basa la evolución natural no tiene sentido cuando nos damos cuenta de que lo que se potencia es aquello que sirve para cumplir la función básica de permitir la supervivencia del organismo, en un modo general, por lo que la posibilidad de que dicha potenciación funcione de modo restringido debe ser

⁶ Para una revisión mucho más precisa de esta polémica puede consultarse el excelente trabajo de Miguens (2002).

⁷ Cf. Dretske 1981. Aquí, Dretske distingue entre significado natural o *significado_n* (aquel significado que viene dado de modo directo y que permite excluir cualquier posible error en la representación de lo que significa porque depende de la propia configuración interna de los organismos) y significado funcional o *significado_f* (aquel significado que puede ser atribuido equivocadamente a algo porque responde adecuadamente a las funciones que lo provocan pero que realmente no corresponden a su significación natural).

desestimada, porque dificultaría el nivel básico de supervivencia en beneficio de un nivel funcional secundario que permitiría explicar y evitar la representación errónea.

Dretske contraataca aportando argumentos searleanos. Los artefactos son capaces de manipular símbolos internos, pero no pueden comprender lo que significan; es decir, tienen sintaxis pero no semántica (Searle 1985: 33-48). Los estados deben tener un significado único establecido de una vez por todas y que sea accesible al sujeto, un significado estrechamente vinculado de modo funcional que permita asociar una carga informacional a un determinado contenido (Dretske 1986).

Pero, insiste Dennett, no hay tal significado estrechamente vinculado, sino que sólo existe un significado funcional implementado de modo general: el significado lo establece la Naturaleza porque el estado funcional, aunque está ahí para realizar aquello que debe realizar (por ejemplo, el cuclillo que tira los huevos rivales sin motivo aparente más allá de la maximización de sus propios intereses de supervivencia, intereses que el cuclillo tiene por su estructura fenotípica pero no por sí mismo, sino como producto de la evolución natural), es capaz de modificar su función si con ello encontrara una mayor utilidad, con carácter evolutivamente establecido. Así, un órgano que nació para cumplir una función en particular en un momento determinado puede modificar su función: esto indica que el significado no es estrecho, sino que depende de la significación externa que se le pueda dar a una cierta función. El significado, así entendido, no depende de causas sino de un *contexto causal*.

Fodor cometerá los mismos errores que Dretske, según Dennett, al intentar postular un lenguaje del pensamiento que permitiría asociar el concepto con el significado: no podemos decir que un término mental tiene un significado determinado porque ello dependería del contexto de uso y no de nosotros (Véase sección 2.1.3.1.).

Otro filósofo que defiende la tesis de la existencia de una intencionalidad original es Burge. Partiendo de sus críticas al individualismo, Burge (1985) arguye que el significado de los estados internos de los individuos no puede ser de otro modo de como en realidad es, ya que es su historia cognitiva lo que lo determina. La intencionalidad de los agentes sería, entonces, real: las creencias y los deseos (así como los estados con carácter fenoménico) son reales aunque emplean cláusulas *caeteris paribus* ('si la condición *x* hace que un sujeto experimente una sensación de rojo, entonces, siempre que se dé la mencionada condición *x*, *caeteris paribus*, se experimentará la sensación de rojo'). Así que, según Burge, la intencionalidad original también respondería al contexto causal intencional, aunque lo haría causalmente por razones externas. El ejemplo clásico lo encontramos en Burge (1979). Un hombre va al médico aquejado de un dolor en su rodilla. Insiste en que tiene artritis. El médico lo corrige, insistiendo en que lo que él realmente tiene no es algo que pueda llamarse artritis, sino algo denominado t-artritis. En este sentido, lo habitual sería aceptar la corrección del doctor y comenzar a denominar a su dolor t-artritis, aunque sea indistinguible de lo que él considera artritis (sobre todo si entendemos que el médico es una persona capacitada para realizar dicha corrección).

Las tesis de Burge (como las de Dretske y Fodor) no consiguen llegar al evolucionismo porque prefieren ser realistas acerca del contenido. Pero, según Dennett, su realismo acaba por

depender de la existencia de una intencionalidad intrínseca en la Naturaleza (acaban en “fuertes presunciones de optimalidad que dependen de encontrarle sentido a *lo que la Madre Naturaleza pensaba*” (Dennett 1987a: 275), algo que obliga a otorgar un significado primordial y único (natural) a aquello (órganos, estados, etc.) que sólo lo tienen de modo derivado y funcional a través del proceso adaptativo de la evolución.⁸ Pero que Dennett sostenga esta tesis, no quiere decir que no sostenga una tesis realista sobre el contenido (como veremos más abajo).

Pero Dennett se encuentra con algunos problemas al considerar que la Naturaleza dota de significado funcional y derivado a los estados y órganos. Por ejemplo, ¿cómo podemos identificar el contenido de un estado o un órgano? Dennett responde de modo naturalista: según las leyes de la evolución, que son ciegas y que podemos concretar mediante algoritmos. Esto es, el estado u órgano en cuestión no tiene ningún tipo de significado natural, sino que tropieza de modo accidental con una función (que comienza a considerársele como propia) y que permite dotarlo de un significado funcional y mediante su *implementación* es capaz de reproducirlo e imitarlo. El que mediante la evolución cierto tipo de seres consiguieran poder representarse sus propias razones no es más que algo accidental, aunque extremadamente útil. Por ello, Dennett concluye que nuestro significado es externo y derivado: actuaría del mismo modo que una lista de la compra escrita en papel en tanto que tiene el mismo significado que la misma lista aprendida de memoria, aquel que le otorga *el esquema general de los propósitos*, que incluye aquello que importa para una época, una sociedad o un solo individuo.

Esta intencionalidad puede denominarse real, pero del mismo modo que podemos decir que es real la selección natural de la que se deriva. El problema principal de esta teoría derivada de la intencionalidad es establecer el punto en el que es representada, pero podemos solucionarlo diciendo que los propósitos de la Naturaleza son inexpresados (como hace Millikan), que nosotros somos la cadena y no un eslabón más. En este sentido, la posibilidad de considerar una intencionalidad real aunque derivada, diferenciándola de la posibilidad de la existencia de una intencionalidad originaria, sólo es posible si partimos de la idea de que nuestra significación depende por completo de nosotros mismos, de nuestros propios sistemas representacionales. Nuestra intencionalidad tan sólo sería una ilusión (como veremos en la sección siguiente) que nos permitiría dotar de suficiente sentido a nuestras percepciones. Éstas, a partir del re-procesamiento de la información captada por los sentidos, son capaces de llegar a constituir una imagen del entorno (resaltando aquellos elementos realmente relevantes para el sujeto) que, de un modo causal, interfiere en la conducta y el comportamiento de los organismos.

⁸ De modo similar ocurre en la teoría de la visión de Marr (1982). Éste defiende la eficacia causal de primitivos modelos de intensidad de luz insertos en la retina que funcionan según procesos computacionales y que permiten implementar representaciones tridimensionales del mundo externo de modo independiente del contenido representacional del que disponga el sujeto en cuestión. Pero, y como defiende Dennett, cuando atribuimos contenido a una estructura primitiva en el modelo de visión de Marr sólo podemos defender dicha atribución de contenido si apelamos a que esa es la función que la Madre Naturaleza le asignó (Dennett 1987a: p. 275). Es decir, que dicha estructura primitiva tiene un significado (en el sentido de Grice) que se corresponde con la función que realiza y para la cual fue creada.

Pero entonces, ¿qué decir de los objetos que tienen una función determinada? Que ésta es tan sólo ilusoria. Afirmar que algo tiene la función para la cual fue construido (y que no puede cumplir ningún otro propósito) es equivocado, caeríamos en la falacia intencional (aquella empeñada en considerar que todo tiene un propósito o fin último). La realidad es que la significación no es intrínseca ni natural, sino tan sólo funcional. La función de algo viene dada por su propia historia (cognitiva),⁹ ya que puede ser interpretada de varios modos según su contexto causal. Pero, ¿no corremos el peligro de caer en el adaptacionismo lamarckiano al afirmar que algo se adapta a la función según su conveniencia y su utilidad? No, nos responderá Dennett, porque no estamos hablando de adaptación, sino de instrumentalidad. Algo puede no haber sido diseñado para una función concreta, pero si ahora realiza bien una tarea que no le corresponde, entonces puede ser empleado para dicha función (por ejemplo, el dispositivo *two-bitser*, el pulgar del panda¹⁰ o el pulmón). Es decir, que lo que hace la evolución no es tanto dotar a sus productos de una adaptabilidad, sino que más bien actúa como un ‘chapuzas’: lo que la evolución hace es *bricolaje*. La Naturaleza corta y pega aquello que funciona adecuadamente para alguna función concreta y a veces lo recoloca en el lugar que más provecho puede proporcionar al individuo que lo posee. En este sentido, concluye Dennett, no existe una intencionalidad original, por lo que el realismo de los significados sólo es posible si aceptamos la teoría natural de las funciones como verdadera, naturalizando así el significado.

5.1.2. Tipos de mentes: criaturas creadas

Si partimos del supuesto de la incorrección del dualismo cartesiano, como hace Dennett, ¿cómo podemos explicar la diferencia existente entre los seres humanos y el resto de organismos? Para Descartes éste era un problema con una solución más que obvia: mientras los animales son autómatas perfectamente guiados por mecanismos, el hombre posee un alma capaz de dirigir el cuerpo que habita. Pero si negamos este supuesto, ¿cómo explicar la presunta

⁹ En este punto sí existiría una concordancia entre los puntos de vista de los diferentes funcionalistas que venimos analizando. Para Dretske (1995), la determinación de la función de un determinado estado informacional dependería de cómo esa información es manipulada, procesada y gestionada en la historia del sistema al que pertenece dicho estado. Para Millikan (2001), del mismo modo, la información que llega a los diferentes sistemas de procesamiento se acumula a lo largo del tiempo, lo que permitiría la constitución de diferentes representaciones que, modificándose de continuo, permitirían diferentes tipos de acción. Dennett coincidirá con estos planteamientos. Véase el siguiente apartado.

¹⁰ El ejemplo favorito para estos casos es apelar al pulgar del panda, que incluso llega a dar nombre a un volumen de ensayos (Gould 1980). En este caso, se resalta el papel funcional que el mencionado pulgar del panda ha adquirido a lo largo de su historia evolutiva. Técnicamente hablando, como los demás ungulados, el oso panda no dispone de pulgar prensil. Simplemente dispone de un espolón de hueso, una prolongación de uno de los huesos metacarpianos de las muñecas de sus garras delanteras que, sin embargo, mediante adaptación progresiva, ha ido creciendo y que, ahora, le sirve de gran utilidad para asir y desgarrar los duros brotes de bambú de los cuales se compone su dieta. Pero no es el único ejemplo. Podríamos también apelar a la peculiaridad de *Rheobatrachus silos*, un pequeño anfibio de la familia de ranas acuáticas que habita las charcas del sureste de Queensland (Australia), de gestar a su prole en el estómago. La presencia en sus huevos de una sustancia viscosa, compuesta básicamente por la hormona prostaglandina E₂, permite con su ingesta suprimir la función gástrica del estómago de los ejemplares hembras (encargados de la gestación, en este caso) con el fin de que los ácidos gástricos de su estómago no los perjudiquen y, así, poder finalizar el proceso de incubación. Pero ello no debe hacernos suponer que los huevos han desarrollado dicha hormona con el fin de ser incubados en el estómago de sus progenitores. Más bien, en algún momento de su historia evolutiva, un ejemplar se tragó sus huevos y, por la presencia fortuita de dicha hormona, la gestación pudo completarse en el interior de su estómago (Gould 1991: 347ss.).

superioridad humana? Dennett considera que el ser humano es un organismo evolucionado como todos los demás, que ha sido capaz de *auparse* (como una grúa) a un elevado grado de refinamiento fenotípico que le ha permitido ejercitar ciertas capacidades y competencias que le permiten un mejor desenvolvimiento en el entorno que habita.

Dennett (1996a: 101ss.) propone un marco en el que colocar las diversas opciones de diseño: es la llamada *torre de la generación y la prueba*. Como he indicado anteriormente, según Dennett, la evolución organiza un conjunto de células en diversos modelos a partir del carácter más o menos satisfactorio (*satisfacing*) de sus respuestas: una respuesta adecuada (lo que no significa que sea la única, o la mejor, o la más correcta) es aquella que funciona, y una conducta que funciona se verá potenciada y también se verá reforzada la estructura capaz de ocasionarla, beneficiando por ello sus posibilidades de desarrollo, mejorando las posibilidades de que dicha respuesta satisfactoria vuelva a repetirse y sea así (re-)compensada. La evolución tiene por objetivo la potenciación ciega de aquellos organismos capaces de realizar los mejores movimientos, los más eficaces. Es decir, la evolución potenciará las estructuras de aquellos organismos capaces de *producir futuro*, capaces de posibilitar nuevas vías de acción. Por ello, en cada estrato de la torre podemos colocar multitud de clases de organismos (que incluso llegan a solaparse, en el sentido de compartir el 99,9% de su diseño) capaces de encontrar y realizar movimientos tales que les permitan desenvolverse en su entorno, lo que permitirá la expansión de su población (en un crecimiento horizontal); pero sólo aquellos organismos cuyos movimientos sean más eficaces verán potenciada su estructura de tal modo que serán capaces de sobrevivir y evolucionar (en un crecimiento vertical). Pero, ¿cómo podemos distinguir dichos estratos? Atenderemos ahora a la descripción que el propio Dennett realiza.

5.1.2.1. Las criaturas darwinianas

Este primer estrato de la evolución acoge a aquellos organismos que, ciega y aleatoriamente, fueron generados por selección natural mediante la combinación arbitraria de genes. Aquellos organismos que poseían estructuras mejor diseñadas (que poseían fenotipos estructurales más eficaces) pudieron expandirse y sobrevivir. Dichas estructuras vienen cerradas de antemano y sólo reproducen aquella información codificada genéticamente.

5.1.2.2. Las criaturas skinnerianas

En algunos de esos organismos que fueron capaces de sobrevivir gracias al mejor diseño de sus estructuras, surgió una nueva propiedad: la llamada *plasticidad fenotípica*. Esto quiere decir que aquellos organismos poseedores de dicha propiedad no estaban genéticamente cerrados, como las criaturas darwinianas, sino que permanecían abiertos. Sus estructuras fenotípicas no estaban diseñadas por completo: ciertas partes de ellas podían ajustarse según lo requirieran los sucesos externos, permitiendo así una mayor eficacia en sus respuestas.

Cuando ya no actúa la herencia genética es la modificación abierta y aleatoria de las estructuras fenotípicas la que debe responder satisfactoriamente al estímulo externo. Así, se potenciará aquella conducta que mayor positividad externa proporcione. O mejor aún, los

estímulos positivos externos serán los causantes de la potenciación de ciertas conductas y del refuerzo de la estructura que las sustenta. Una vez que alguna de las primigenias conductas lanzadas ciega y arbitrariamente según los dictados de la evolución ha sido ajustada fenotípicamente por el organismo y es reforzada por el entorno mediante la positividad de su respuesta, la conducta que automáticamente realizará el organismo cada vez que sea necesario será aquella que fue en su momento potenciada, porque es la que mejor responde a la situación dada.

5.1.2.3. Las criaturas popperianas

El ensayo-error es eficaz pero muy arriesgado: puede salir bien siempre que nuestros primeros intentos fenotípicos estructurales sean eficaces, algo poco probable teniendo en cuenta que es mucho más factible encontrar un modo de constitución defectuosa que una efectiva, por lo que es mayor la probabilidad de perecer que de sobrevivir. A ojos vista, será mejor un sistema capaz de pre-seleccionar aquellos comportamientos eficaces y separarlos de aquellos ineficaces minimizando así el alto riesgo de perecer y extinguirse que supone el método de ensayo-error.

Pero, ¿por qué denominar popperianas a estas criaturas? Porque, como dijo Popper de las teorías científicas, esta mejora de diseño “permite que nuestras hipótesis mueran en lugar de morir nosotros” (Dennett 1996a: 109). Los organismos popperianos adquieren la capacidad de modificar ellos mismos sus propias estructuras fenotípicas, adquieren la capacidad de *fijación postnatal del diseño*. En condiciones normales, no hay razón para que el organismo no realice dicha acción, dados los suficientes datos. Pero si esto es así, ¿qué hace del humano una criatura tan especial? La cuestión principal está en ver

“cómo un cerebro es capaz de convertirse en una mente” (Dennett 1995: 615).

5.1.2.4. Las criaturas gregorianas

Llegados a un cierto punto de la cadena evolutiva, aparecen cierto tipo de criaturas capaces de fabricar ciertas herramientas conceptuales que les permiten ir más allá de la mera determinación natural, porque se requiere inteligencia para inventar/fabricar, reconocer y mantener algo como instrumento útil, pero además el uso de dicha herramienta otorga la misma inteligencia a aquellos que la emplean sin haberla creado. Cuanta más información requiera el diseño de una herramienta, mayor inteligencia potencial confiere al usuario. Las herramientas más notables son aquellas denominadas mentales: las palabras.

Las criaturas skinnerianas preguntan “¿Qué voy a hacer a continuación?”; las criaturas popperianas, “¿Qué debo pensar a continuación?”. Las criaturas gregorianas¹¹ se *auñan* a la humanidad en tanto que son capaces de cooperar. Los organismos gregorianos son capaces de

¹¹ Dennett denomina gregorianas a estas criaturas en referencia a W. K. Gregory, famoso paleontólogo de principios del siglo XX que, aunque empañó su carrera con un oscuro incidente respecto de la falsa identificación de unos dientes de pécarí con una nueva especie humanoide (el *Hesperopitecus*), fue de los primeros en comprender que una correcta organización social en los homínidos prehistóricos permitió una sustancial mejora en su supervivencia.

beneficiarse de la experiencia encarnada en las herramientas (especialmente las mentales) que la comunidad social en conjunto ha ido elaborando y transmitiendo a lo largo de los años, lo que les permite reflexionar mejor acerca de lo que deben hacer o pensar a continuación. Les permite aprender “a pensar mejor sobre lo que deben pensar a propósito del siguiente movimiento” (Dennett 1995: 624), consiguiendo así crear una serie de reflexiones internas (meta-representaciones). El organismo gregoriano, así, es capaz de modificar su estructura fenotípica de modo que le permita crear, identificar y generalizar relaciones no existentes en el mundo externo, pero que le serán de gran utilidad por medio del simple uso de conceptos.

La diferencia entre las criaturas (gregorianas) humanas y las criaturas no-humanas consiste en el manejo más eficaz que tienen las primeras de pensamientos concretos referentes a objetos del mundo que se constituyen de forma abstracta como representaciones. Pero dichas representaciones no son individuales, sino que vienen posibilitadas por el entorno social: nadie puede representarse nada más allá que aquello que su entorno le dice mediante sus herramientas mentales. Las criaturas skinnerianas y popperianas pueden actuar tan sólo *como si* tuvieran pensamientos concretos, *como si* reconocieran objetos del mundo, *como si* se escondieran agazapadas tras las ramas de un helecho con la intención de sorprender a su presa. Los dos tipos de intención pueden llevar a cometer errores, pero sólo en el primer caso puede llevar a la representación de una proposición falsa creyendo que es verdadera.

En los animales skinnerianos y popperianos podemos adivinar multitud de lagunas en cuanto a representación y generalización porque no tienen pensamiento (si podemos llegar a denominarlo como tal) más allá de su modo de procesamiento central. En cambio, los organismos gregorianos humanos son capaces de pensamiento aún a pesar de perder parte de sus capacidades localizadas porque su estructura cognitiva no está centralizada sino que se compone de transductores periféricos en los que se delega parte de las competencias con el propósito de permitir al agente economizar tiempo de transacción epistémica.

Los organismos gregorianos somos especiales porque somos capaces de construir hipótesis y representaciones generales acerca de la identidad de los objetos e individuos externos; somos capaces de rastrear objetos según su relevancia en el entorno cambiante y de percatarnos de nuestras acciones porque todo ello lo asociamos a nuestro esquema intencional. Por tanto, sólo aquella criatura gregoriana dotada de herramientas mentales que es capaz de representarse los objetos del mundo, es capaz de ser considerada como *persona*. Pero, ¿cómo es posible que simples organismos evolucionados bajo las leyes de la selección natural hayan llegado a poder construir representaciones de objetos del mundo? ¿Cómo y por qué es posible el pensamiento?

5.1.3. La aparición del pensamiento

En “Conditions of personhood” (Dennett 1978: 267-285), Dennett afirma que para ser considerado persona, un organismo debe ser un sistema intencional de primer orden (aquel sistema de deseos y creencias relacionados con los objetos del mundo que no pueden relacionarse con sus propios deseos y creencias) capaz de dar el salto a un sistema intencional

de segundo orden (aquel sistema de deseos y creencias capaz de establecer relaciones, además de con los objetos del mundo, también con sus propios deseos y creencias), o bien dar el salto a un sistema de tercer orden (aquel sistema capaz de querer que otro crea que quiere algo), o bien dar el salto a un sistema intencional de cuarto orden (aquel sistema capaz de querer que otro crea que él quiere que crea que quiere algo), y así, *ad infinitum*. El paso importante es el salto de los organismos simples a los sistemas de segundo orden, porque este segundo nivel permite que los individuos *compartan notas*, se transmitan información (precisamente en esto consistiría, como hemos visto, el ser un organismo gregoriano), al poder manipular su propia significación. La comunicación permite a cada individuo

“ser beneficiario de las labores cognitivas de los otros, de un modo que nos otorga poderes sin precedentes” (Dennett 1995: 629).

El paso al resto de niveles tan sólo depende de la cantidad de información que sea capaz de manejar un individuo, ya que las operaciones pueden volverse casi automáticas. Pero, ¿cuál es el límite entre aquellas criaturas que son sistemas intencionales de segundo orden y aquellas criaturas que tan sólo muestran intencionalidad aparente?

5.1.3.1. Cuando las palabras hacen cosas con nosotros: la aparición de la representación

Sabemos que algunos animales son capaces de realizar acciones con cierta (aparentemente) intencionalidad. Pero, ¿en qué modo podemos decir que el comportamiento que presentan estos animales es el resultado de una estrategia organizada?

Parece que, en principio, las conductas organizadas requieren de un elevado número de estructuras de control interdependientes coexistiendo en el sistema nervioso, pero la verdad es que no se puede precisar. La respuesta requiere de la apelación a una fuerte presión selectiva que *obligue* al organismo a reorganizar sus estructuras fenotípicas de control de modo que pueda formar otra estructura capaz de llevarlo a la expansión y la supervivencia.

Dicha respuesta obliga al individuo en cuestión a representarse aquello que no funciona adecuadamente en un momento anterior con el fin de interpretar las consecuencias y evaluar, así, las posibles alternativas y cambios a los que someterse. La proliferación de representaciones, nos dice Dennett, obliga a una sofisticación paulatina de los significados y a una continua revisión de las conclusiones, por lo que el organismo se ve obligado a matizar ciertos usos y a restringir ciertas representaciones en modelos organizados, los conceptos, que le permitan ahorrar tiempo y maximizar posibilidades de supervivencia por la mayor rapidez de administración y manipulación de la información. Brevemente, para Dennett, el individuo se ve obligado a construir un lenguaje capaz de interpretar aquello que le ocurre (en relación con su entorno) que le permita, por un lado, explicar lo que sucede a su alrededor y, por otro, concebir soluciones a las dificultades encontradas. El organismo se ve obligado a hablar consigo mismo. (Cuando nos referimos, en este punto de la interpretación dennettiana, a que un individuo habla consigo mismo, en cierto sentido nos referimos a que el individuo en cuestión se ve constreñido a forjar un lenguaje del pensamiento *à la* Fodor que le permita coordinar sus propias

evoluciones internas con las modificaciones externas, ¡sólo que aquí nos referimos a un lenguaje *hablado* del pensamiento!).

La adquisición del lenguaje capacita al individuo para poder auto-atribuirse contenidos, permite al organismo primigenio forjar la capacidad de explicarse sus propios movimientos y desarrollar así la capacidad de evaluarlos. La capacidad del lenguaje, afirma Dennett, desdobra al individuo en dos. Por un lado, tenemos al individuo/agente, que hace, que reacciona, que se mueve impulsado por los estímulos exteriores; por otro, tenemos al individuo/espectador, que interpreta la conducta del primero y posibilita una evaluación según criterios de corrección e incorrección de los movimientos según permitan una mayor maximización de las posibilidades de supervivencia. Es decir, el lenguaje permite al individuo interpretar su propio comportamiento en tanto que permite identificar una conducta como errónea y sustituirla por una más adecuada antes de que la primera se produzca, permitiendo así la anticipación a los propios movimientos y la generación de una serie organizada de conductas posibles. Pero no sólo eso, también permite a un individuo poder interpretar las acciones de los demás, dotándolo de la capacidad suficiente para anticiparse a las acciones futuras de los demás y predecir sus comportamientos y conductas.

5.1.3.2. Imitación, cooperación y egoísmo

Una vez el individuo ha logrado reorganizar su propia estructura mediante la construcción de todo un operativo mental de significados y representaciones mentales que le permiten una mayor agilidad en el tratamiento de información y un mayor grado de adaptación al medio, para Dennett, surge la necesidad de tener que compartir dicha información con los demás individuos con el fin de buscar beneficios en cuanto a la posibilidad de la propia subsistencia. Según Dennett, es en dicho momento cuando aparece *la oportunidad de comunicación*. Cuando una especie entra en disposición del lenguaje, necesariamente la conduce a la producción de conductas de cooperación.

Con la comunicación cooperativa surge la necesidad de representar (mentalmente) la conducta porque aparece la obligación de comunicar explícitamente la intención del comportamiento. Solo que ciertas criaturas están dotadas de *confabulación aproximada*, un sustituto que puede pasar por una representación de la intención capaz de configurar una *lista de opciones* que simulen acciones intencionales pero que sólo se activan según el caso aferente-eferente, dando incluso la sensación de que están bajo su control.

Por ejemplo, el primer ñu en oler al león atacante no puede transmitir directamente a los demás dicha información porque carece del modo representacional adecuado para hacerlo, como tampoco puede engañar a los demás miembros de la manada, ni tan siquiera puede hacer uso fraudulento de dicha información: tan sólo puede echar a correr porque su sistema eferente recibe el estímulo aferente 'Hay un león ahí' o, más generalmente, 'Existe un peligro cerca', que despierta la respuesta adecuada, 'Correr', y ése es el único modo en que los demás ñues pueden hacer uso de la información del primero (indirectamente a partir de la conducta de uno de los miembros de la manada, algo que puede dar la falsa impresión de ser un modo primitivo de lo

que Wittgenstein denominaría seguir la regla). Este salir corriendo puede parecer una acción intencional, parece referirse a la intención ‘Correr para evitar ser una presa’, pero en realidad es una acción derivada porque depende de la determinación de su estructura fenotípica, depende de que su sistema aferente-eferente esté distribuido de éste modo y no depende de su propia voluntad, el que los ñues sean capaces de activar la salida conductual adecuada ante un estímulo externo determinado.

La red conceptual que se requiere para que un animal se muestre *como si* tuviera intencionalidad puede ser completamente diseñada por mecanismos darwinianos o skinnerianos, ya que el aprendizaje ABC¹² permite reforzar mediante condicionamientos ciertos comportamientos operantes. Pero tan sólo una criatura capaz de aprender de su experiencia pasada (de su *historia cognitiva*) puede ser considerada como intencional. Una criatura capaz de forjar estrategias de supervivencia que maximicen sus posibilidades de respuestas satisfactorias no es más que un sistema intencional de primer orden. Pero aquellas criaturas capaces de dar el salto a un sistema de segundo orden serán capaces, también, de forjar verdaderas estrategias intencionales: sólo criaturas capaces de manipular representaciones y transmitir las mediante la constitución de herramientas mentales pueden ser consideradas como verdaderamente intencionales. Dichas criaturas disponen del principio *comando en equipo*: cada individuo es capaz de tener tanto conocimiento del proyecto total como para que todos y cada uno de los miembros del equipo puedan improvisar en vistas a la consecución del objetivo en caso de imprevistos. Esta estrategia de orden superior aparece sólo en aquellas criaturas cuya presión evolutiva les obliga a representar proposiciones, lo que les proporciona la posibilidad de forjar herramientas conceptuales con las que intercambiar información y poder cooperar: aparece aquí el *benegoísmo*.¹³ Cuando no queda más remedio que cooperar o perecer, los organismos no tienen otra salida que forjar estrategias conjuntas que les permitan el más alto grado de supervivencia posible (Dennett 2003: 221-231). En pocas palabras, la presión evolutiva, dice Dennett, obliga a aquellos seres capaces de dar el salto a la manipulación de datos en un segundo orden a construir todo un entramado de símbolos y significados que contengan todo aquello esencial para la supervivencia.

Dennett hace suya una famosa teoría.¹⁴ Clark and Karmiloff-Smith (1993) distinguen dos tipos de criaturas: aquellas que poseen un cerebro con el conocimiento *encarnado* y aquellas criaturas que poseen un cerebro capaz de enriquecerse internamente porque son capaces de volverse a representar el conocimiento que anteriormente se habían representado. Este segundo caso se refiere a criaturas con un cerebro capaz de recuperar memorísticamente aquellas

¹² El aprendizaje ABC (*Asociationism, Behaviorism, Conexionism*) es aquel que permite a un organismo establecer conexiones entre redes comportamentales sencillas mediante la “historia de los refuerzos” que han contribuido a la potenciación de sus conductas satisfactorias.

¹³ El término benegoísmo (*benegoism*) es acuñado por Dennett para referirse al comportamiento que los individuos adquieren cuando se encuentran en una situación límite. Cuando no queda más remedio que cooperar o perecer, los organismos no tienen otra salida que forjar estrategias conjuntas que les permitan el más alto grado de supervivencia posible. El término se refiere a unas palabras que (parece ser) pronunció Benjamin Franklin en la reunión previa a la firma de la Declaración de Independencia de los Estados Unidos. Cf. Dennett (2003: 221ss.).

¹⁴ Una primera defensa de esta teoría puede verse en Dennett (1994). Para una versión revisada puede consultarse Dennett (1997).

experiencias que ya habían tenido. Distinguen entre criaturas que tan sólo responden medio-instintivamente a estímulos y aquellas criaturas que son capaces de emplear estratégicamente aquel conocimiento (representado) que ya habían tenido anteriormente para sus propios fines.

Es decir, Clark y Karmiloff-Smith distinguen entre aquellas criaturas que tienen conocimiento *en* el sistema y aquellas criaturas que tienen conocimiento *para* el sistema. Las primeras son criaturas que responden al condicionamiento operante que les ofrece el entorno; las segundas son criaturas capaces de manipular sus propias representaciones, reproducirlas, combinarlas e, incluso, innovarlas; son capaces de verlas como herramientas (mentales) y de emplearlas como tales para su propio beneficio, y el de su especie. Las segundas son criaturas capaces de manipulación simbólica, como lo sería el ser humano: éste es capaz de guiarse por sus estructuras sintácticas al mismo tiempo que es capaz de comprenderlas, ya que son representaciones simbólicas cargadas de significación e información. Pero, ¿cómo es capaz el ser humano de llegar a comprender? ¿Cómo puede una criatura sintáctica dar el salto a la semántica? ¿Cómo puede un artefacto evolucionado pasar de movimientos sin sentido a significar su conducta? Para Dennett, el hombre adquiere la capacidad de comprender y manejar significados porque es capaz de descargar la mayoría de teorías cognitivas en su propio entorno, proyectando hacia el exterior toda aquella información útil y relevante.

5.1.3.3. Fabricando nuevas herramientas: la irrupción de la semántica

Como hemos dicho, Dennett defiende la teoría que afirma que la evolución de los diseños permite a cierto tipo de criaturas privilegiadas proyectar en símbolos externos la mayor parte de la información funcionalmente guiada. Es decir, adoptan la capacidad de construir *dispositivos periféricos* para almacenar, procesar y volver a representarse cada vez que sea necesario aquellos significados considerados relevantes. El ser humano tiene, así, la capacidad de confeccionar símbolos/herramientas conceptuales que transmiten información y que se colocan en el exterior (constituyendo lenguaje) para su posterior uso. Existe la posibilidad de que cierto tipo de criaturas skinnerianas o popperianas puedan transmitir información mediante este tipo de codificación externa (como por ejemplo el rastro de feromonas que las hormigas emplean para señalar su camino o las feromonas que contiene la orina de ciertos animales para la delimitación de su territorio), pero dicha operación se produce de modo operante y condicionado, no mediante reflexión comprensiva.

En pocas palabras, Dennett defiende una especie de nominalismo instrumental. El ser humano, nos dice, es capaz de etiquetar las cosas, pero no de cualquier modo. Dicho etiquetado se refiere a aquello que es significativo, a aquello capaz de ser encuadrado en sistemas externos de almacenamiento de información. No es necesario que sea lenguaje. Ciertas especies se sirven de marcas externas como, por ejemplo, la posición del sol para orientarse y los restos desenterrados de las semillas ya consumidas (como el cascanueces de Clark 1993). La herencia cultural proporciona autopistas culturales que permiten extender nuestra mente al mundo.

Sólo las criaturas humanas, según Dennett, poseen dos dimensiones de entorno: la interna y la externa. El entorno interno sobrevive precisamente porque es regular: aquellas

estructuras fenotípicas funcionales tienden a perpetuarse y expandirse. En cambio, como el entorno externo está en continua variación, el organismo debe rediseñarse constantemente para poder captarlo. Una criatura necesita para sobrevivir la posibilidad de identificar y discriminar aquello que es relevante de aquello que no lo es. Pero ello sólo es posible mediante el etiquetado de aquello que se considera relevante. Los símbolos externos sólo tienen significado porque se refuerzan con recursos internos que permiten su identificación: el significado sólo tiene importancia porque hay alguien capaz de captarlo. De este modo, las representaciones pueden llegar a convertirse en objetos mentales.

Por ejemplo, en una biblioteca se emplea un código que permite seleccionar y clasificar en secciones temáticas cada uno de los ejemplares de los que se compone, pero ¿qué sentido puede tener esto? También podríamos diseñar un artefacto capaz de clasificar los libros de una biblioteca según las palabras de su título, ¿no? No, porque necesitamos comprender los aspectos importantes que permiten una clasificación lo más correcta o refinada posible, y esos símbolos permiten la rapidez y correcto ajuste del procesamiento de información. Es decir, si la conducta de ciertas criaturas de exportar al entorno externo información ha sobrevivido es porque ello es funcional, ya que permite agilizar el reconocimiento de aspectos y mejorar la estructura cognitiva. Todo ello conlleva el refinamiento de cierto tipo de *tecnologías de re-representación*. Aquí Dennett nos está hablando de meta-representaciones. Se potencia el modo más adecuado y eficaz de clasificar, manipular y reflexionar acerca de las representaciones previas. La estrategia más potenciada es aquella que permite construir herramientas externas cargadas de información y significación pero que a la vez requiere mejorar las propias estructuras fenotípicas con el objeto de reconocer y percibir correctamente dichas representaciones, a pesar del constante cambio exterior.

Es decir, siguiendo el argumento dennettiano, si la conducta de ciertas criaturas de exportar al entorno externo información ha sobrevivido es debido al hecho de que ello es útil, ya que ha permitido agilizar el reconocimiento de aspectos y mejorar la estructura y los sistemas cognitivos de los seres conscientes. Todo ello conlleva el refinamiento de cierto tipo de *tecnologías de re-representación*: se potencia el modo más adecuado y eficaz de clasificar, manipular y reflexionar acerca de las representaciones previas. La estrategia más potenciada siempre será aquella que permita construir herramientas externas cargadas de información y significación pero que a la vez requiere mejorar las propias estructuras fenotípicas con el objeto de reconocer y percibir correctamente dichas representaciones, a pesar del constante cambio exterior.

Esto nos viene a decir que lo necesario para la supervivencia de uno de estos organismos no es conocer todas las características del exterior, sino que en el momento adecuado sea capaz de hacer uso de la información exterior porque ha conseguido modificar sus estructuras cognitivas de modo adecuado, ha sido capaz de prever adecuadamente el próximo movimiento. Ello permite que aquellas criaturas capaces de rediseñarse adecuadamente sean capaces de construir redes conceptuales capaces de relacionar su mundo interno con el externo.

Con la aparición del lenguaje, las criaturas gregorianas se dotan de palabras, y las palabras identifican (como etiquetas) aquella información depositada en el entorno externo. El niño aprende expresiones a medio entender. Escucha sonidos, los recuerda y repite, incluso los comenta. La repetición de palabras las llega a hacer familiares y la familiaridad permite identificar el sentido de una etiqueta. Así se establecen asociaciones y se potencian las ya establecidas.

La contemplación de representaciones permite recordar la significación relevante, aquella que realmente importa para nuestros propios intereses, permitiendo que entendamos la información y permitiéndonos constituir conceptos. Las palabras son, entendidas a la luz de la teoría de Dennett, conceptos primigenios, proto-conceptos, vocalizados y encarnados que podemos manipular para convertirlos en objetos mentales sobre los que deliberar. Pero, ¿cómo saber qué concepto emplear para según que situación? El hombre construye *tecnologías*: toda una red de habilidades estructuradas que permite captar y clasificar su entorno según su propia escala de valores, según sea instrumentalmente relevante y válido para sus propios propósitos. Es aquí donde podemos ver el importante papel de la cooperación que, nos dice Dennett, permite el lenguaje: una vez los organismos han llegado al punto de representarse los objetos del mundo, establecer relaciones entre ellos y comunicarlo a los demás, lo que indica que esto es relevante, se procede al establecimiento de hipótesis que permiten prever el futuro, que permiten anticiparnos a nuestras propias conductas y predecir las conductas de los demás; que nos permite ser conscientes del futuro acontecimiento de los hechos físicos antes de que éstos acontezcan. Es decir, se constituye la ciencia:

“Una vez que tenemos lenguaje –una caja repleta de herramientas mentales- podemos utilizar estas herramientas en la estructura que *genera* y que somete a *prueba*, de manera premeditada y con visión de futuro, conocida como *ciencia*” (Dennet 1995: 626).

La ciencia permite forjar leyes mediante la generalización y establecimiento de relaciones causales que (a pesar de ser ficticias) permiten al hombre tener la ilusión de poder predecir acontecimientos, siendo así de gran utilidad. La ciencia no está pensada para minimizar los errores, sino para que en caso de cometerlos, los errores se cometan en público, permitiendo la puesta en común de dificultades y escollos con el fin de que los demás miembros de la comunidad puedan llegar a una mayor precisión y refinamiento de las hipótesis.

5.1.4. Conclusiones: las dificultades de la teoría evolutiva dennettiana

A lo largo de estas líneas, hemos presentado la teoría dennettiana que afirma que todos los organismos son el resultado de la evolución, marcada por una serie de ciegos procesos de acumulación, lo que les permiten llegar a tener una cierta intencionalidad derivada en tanto que su comportamiento puede ser interpretado desde el punto de vista de la interacción con su entorno. También hemos atendido a las respuestas que desde ciertos sistemas filosóficos ha provocado esta afirmación, en tanto que consideran que la intencionalidad propia de los humanos tan sólo puede ser entendida como intrínseca porque apela a un trasfondo profundo.

Aquí hemos defendido que el concepto ‘derivado’ propuesto por Dennett debe ser entendido en un sentido externista (Acero 2001: 49-54). Con ello no queremos decir que la intencionalidad humana venga asignada por algo o alguien, o que dependa del significado de las expresiones lingüísticas o simbólicas. Pero sí que podemos aceptar que nuestros estados mentales representen sus contenidos por sí mismos en tanto en cuanto se interpretan dentro de un contexto socio-cultural. En este sentido, podemos decir que un estado mental está totalmente indeterminado hasta que no interviene un contexto que lo determine. Y en cuanto indeterminado puede dar la impresión, a menudo, de representar erróneamente un contenido, dependiendo del sistema contextual desde el que se interprete. Esto es, aunque tengamos la aparente ilusión de que todo contenido tenga una carga semánticamente relevante, esto no es más que una ficción útil. Que el contenido mental sea aquí conceptual o no-conceptual no es extremadamente relevante, porque lo que realmente importaría sería, por un lado, que la capacidad de predecir conductas y de anticiparnos a los hechos físicos permita precisamente facilitarnos la tarea de vivir en un mundo físicamente determinado y, por el otro lado, que toda esa capacidad predictiva permita que en un momento determinado el contenido mental de los sujetos conscientes individuales pueda ser conceptualizado, permitiendo así la comunicación y la cooperación necesaria para perseguir dicho objetivo de bienestar.

También hemos podido comprobar cómo la teoría de Dennett está fuertemente marcada por la explicación científica y el ámbito de la psicología cognitiva, lo que restringe el ámbito de su filosofar a un materialismo estricto que, aunque no es reductivo, pretende ser conclusivo. Queremos detenernos brevemente en dos supuestos que subyacen al naturalismo dennettiano acerca del pensamiento colectivo y que persiguen, a nuestro entender, una pretensión de cierre cognitivo, lo que se viene denominando *clausura física del mundo*.

En primer lugar, y aunque estamos de acuerdo con Dennett en que la ocasión del pensamiento viene posibilitada por el lenguaje, no está nada claro el modo en que un lenguaje del pensamiento (el *mentales* a modo de gramática mental universal) puede dotar de la objetividad necesaria a un solo individuo natural como para crear un espacio de aplicación de conceptos. Si tenemos en cuenta que el *mentales* es un lenguaje interno y que es necesaria la interacción de, al menos, dos individuos que puedan poner en relación el comportamiento uno del otro con los cambios externos para abrir dicho espacio, consideramos que es una petición de principio suponer que ambos individuos comparten el mismo lenguaje mental, sobre todo después de haber insistido en que no se requiere de la existencia de un lenguaje interno para que los sistemas de procesamiento funcionen (independientemente que su pertenencia a una misma especie y a una misma cultura básica garantice un mínimo de objetividad). Pero aún aceptando la tesis dennettiana del desdoblamiento inicial del mismo individuo, no vemos cómo es posible la escisión de dicho individuo consciente en un agente capaz de actuar según los cambios del entorno externo y en un interprete capaz de valorar y evaluar dichas acciones en relación al medio si damos por supuesto, y Dennett lo hace (como veremos en el siguiente punto), la unidad de la conciencia y el holismo semántico del pensamiento. Esto es, no vemos cómo de una tesis anti-dualista sobre lo mental como es la dennettiana, capaz de afirmar que todo estado mental

puede identificarse con un estado físico, puede seguirse la tesis que afirma que una parte de dichos estados puede permanecer oculta, por mucho que se apele a una noción de yo como centro de gravedad narrativo (como veremos en el siguiente apartado).

En segundo lugar, si afirmamos, como también hace Dennett, la existencia de vínculos causales entre el contenido mental y el mundo externo no basta con apelar a dicha causalidad para determinar el contenido de un pensamiento, dada la indeterminación del mismo. Pero si aceptamos la tesis del desdoblamiento de la conciencia del individuo, entonces desaparece la indeterminación porque el contenido del pensamiento se refiere a la causa total que provoca las respuestas relevantes en situaciones similares, causa que sería compartida por todas las partes. Pero si aceptamos esto, el problema inicial se transforma ahora en una dificultad mucho mayor al considerar la existencia de una causa total que puede explicar las acciones (*todas* las acciones) a partir de un mismo contenido, violando así el principio de realización múltiple. Esto es, si aceptamos esta noción técnica de contenido, el ámbito de acción viene restringido a la causalidad unidireccional que está presente en la causa total obteniendo tan sólo una vía de ejecución para cualquier contexto, negando así la posibilidad de la aparición de nuevas y distintas vías de acción (cerrando las puertas, como veremos en el apartado 5.4., de la posibilidad de la existencia de posibilidades alternativas de acción que permitan la atribución, entre otros, de responsabilidad moral a los sujetos conscientes).

5.2. El desarrollo ontogenético de la conciencia: el modelo de los borradores múltiples

Como indicamos al comienzo de este capítulo, la teoría de la conciencia propuesta por Dennett tiene una base marcadamente biológica y evolutivamente determinada. Precisamente con la intención de mostrar los términos en los que Dennett entiende esta determinación, hemos presentado en el apartado anterior el desarrollo que para Dennett, desde un punto evolutivo, ha tenido que superar todo ser consciente para poder hablar actualmente de su darse cuenta y de su pensamiento. No podemos acabar de entender el intento naturalista de explicación que de la conciencia fenoménica realiza Dennett si previamente no atendemos, como hemos hecho, al desarrollo evolutivo que de la aparición del pensamiento propone la teoría dennettiana.

Como hemos visto, Dennett parte del carácter instrumental de la actitud intencional (de la posibilidad de que la conducta de cualquier organismo se realice como si respondiera a razones para actuar, razones que permiten anticipar y predecir futuros comportamientos y acontecimientos), de los actuales desarrollos en ciencia cognitiva, desde el desarrollo del programa conexionista (optando por una distribución en paralelo de la estructura computacional del cerebro, aunque la conciencia esté restringida a un uso lineal) hasta los últimos avances en neurociencias (como muestran sus estudios sobre lesiones y trastornos cerebrales), y la interpretación evolutiva para establecer una explicación más o menos plausible de la aparición del pensamiento consciente. Como hemos visto anteriormente, dicha aparición estaba directamente relacionada, también, con la capacidad humana de conseguir darse las propias razones para actuar, de hablar consigo mismo, de conseguir interpretar los propios puntos de

vista subjetivos del modo adecuado. Vemos, así, que la estrategia dennettiana consiste en un intento de construcción de una teoría capaz de describir el modo en que nuestra base biológica es capaz de producir nuestra mente y nuestra conciencia; es decir, el propósito de Dennett es poder construir una teoría naturalizada que explique de modo total y descriptivo la manera en que la mente y su elemento distintivo, la conciencia cualitativa, es causada por el cerebro.

Esta estrategia parte del supuesto de que nuestra propia investigación de la conciencia debe ser guiada por dos realidades. La primera, los datos que desde el punto de vista objetivo, en tercera persona, nos proporciona la ciencia y su método de investigación, a partir del análisis y descripción de procesos físicos. La segunda, los datos fenoménicos que cualitativamente todo individuo consciente experimenta desde el punto de vista subjetivo, en primera persona, que es capaz de aportar los matices necesarios a las sensaciones que todos, como individuos cognitivamente normales, tarde o temprano estamos capacitados para experimentar, en situaciones normales.

Es nuestra intención en este apartado ofrecer una perspectiva del desarrollo que Dennett ha llevado a cabo de su propia teoría de la conciencia fenoménica atendiendo a los dos rasgos distintivos de la misma. Por un lado, atenderemos al método interpretativo dennettiano que permite diferenciarla de otras propuestas naturalistas (la heterofenomenología). Este método posibilita que todo individuo pueda constituirse conscientemente como un conjunto de elementos fenoménicos a partir de su introspección, permitiendo así la creación de un centro de gravedad narrativo con una realidad física directa un tanto peculiar (entendida como ilusoria) que permita sostener una gran cantidad de datos cognitivos a modo de borradores cuyo carácter consciente prioritario siempre estará en continua revisión (a modo del sujeto como haz de percepciones que estipuló Hume). Por otro lado, atenderemos también al anti-mentalismo presente en su crítica a la concepción de la conciencia como un teatro cartesiano. Dennett apuesta por una conciencia sin centros ni jefes compuesta, precisamente, por la gran cantidad de reconstrucciones ficticias debidas a la interpretación continua y sucesiva de los datos captados a través de la experiencia y en la que intervienen los sistemas cognitivos de captación sensorial (tanto internos como externos) y la atención, la memoria y otros elementos similares.

5.2.1. La implausibilidad de la existencia de un teatro cartesiano en la conciencia

Como hemos visto, para Dennett, la conciencia tan sólo puede ser entendida desde el punto de vista natural. Todos aquellos fenómenos con carácter cualitativo que pueblan nuestro campo experiencial deben poderse explicar de algún u otro modo desde la descripción facilitada por las ciencias físicas, dado que todos ellos tan sólo pueden ser el producto de nuestra biología, comenzando por el papel causal de nuestro cerebro y de nuestro sistema nervioso central. Desde esta perspectiva, si entendemos la mente y la conciencia como el producto de una serie de elementos físicos, ¿qué estatuto metafísico le corresponde?

Dennett apuesta por una clara respuesta reductiva. Como la mente, en general, y la conciencia, en particular, tan sólo son el producto de nuestro cerebro, éstas pueden ser reducidas a éste. Lo que para muchos ha supuesto una declaración eliminativista de lo mental o una

posición irrealista respecto de la conciencia (Block 1993; Boghossian 1992), en realidad puede entenderse como una perspectiva instrumentalista de la misma. Lejos de presuponer la irrealidad de la existencia de la mente o de la conciencia, Dennett apuesta por una posición dentro del realismo (Dennett 1978, 1991 y, especialmente, 1995b). Los hechos mentales subjetivos, nos cuenta Dennett, y las experiencias fenoménicas con carácter cualitativo tienen una realidad clara dentro del sistema cognitivo que las genera, sólo que dicha realidad no presupone un trasfondo substancial ni una realidad extra-física trascendente, sino que tan sólo suponen una realidad útil auto-generada, una ficción, que permite al sujeto, desde su propio punto de vista, crear una serie de ilusiones necesarias que lo capaciten como sujeto consciente.

Lejos de situarse en una posición misterianista que apuesta por una suspensión epistémica sobre la explicación nomológica de la conciencia, Dennett aprovecha para denunciar la falta de imaginación que los filósofos suelen tener, una falta de imaginación que muy a menudo confunden con la imposibilidad real de conquistar una verdadera explicación de los hechos. Muchos consideran inexplicable el fenómeno de la conciencia porque consideran que es imposible reducir a una simple explicación científica la gran variedad y cantidad de fenómenos que, todos, somos capaces de experimentar. Dicha imposibilidad puede considerarse de distintas maneras: desde la imposibilidad física, a la metafísica y a la conceptual (como hemos visto en el apartado 1.3.). Todas estas posiciones parecen querer considerar la conciencia como un producto 'mágico' (Dennett 2005: cap. 3), dado que consideran que el hecho de querer explicar tan sólo las conexiones funcionales, la maquina detrás del fantasma (apelando a Ryle), tan sólo permite explicar los problemas 'fáciles' de la conciencia de los que nos habla Chalmers, pero deja intacta la pregunta por el problema 'difícil'. El punto importante aquí, sin embargo, sería otro muy diferente: lo relevante no sería si es o no posible explicar desde un punto de vista físico dicha riqueza, sino que lo verdaderamente increíble es que nosotros mismos somos unos organismos biológicos evolutivamente desarrollados que hemos sido dotados de la capacidad de captar y poseer una gran variedad de experiencias fenoménicas. Es decir, lo importante es llegar a averiguar cómo unos simples organismos evolucionados han llegado a conseguir la vida mental consciente que, de hecho, poseen.

El problema con el que se enfrenta Dennett parte del supuesto de la realidad del pensamiento. Por eso podemos decir que su posición es realista. La cuestión no reside en la pregunta sobre la posibilidad o imposibilidad del pensamiento de los organismos evolucionados que somos, o de las máquinas o de los marcianos (Dennett 2005: 39ss.), sino en la existencia del pensamiento como algo real; es un hecho que dichos organismos pueden pensar. La cuestión principal, en la teoría de Dennett, es llegar a descubrir cómo es posible que de la base física del cerebro se consiga obtener dicho pensamiento, cargado de experiencias fenoménicas con carácter cualitativo.

La estrategia explicativa de Dennett, como hemos visto, parte de una propuesta interpretativa que presupone no sólo un esfuerzo imaginativo, sino también el esfuerzo de dar cabida a la ingente cantidad de pruebas y hechos científicos que avalan, desde el punto de vista

neurológico y fisiológico, el papel causal del cerebro y del sistema nervioso central en la producción de la conciencia.

Siguiendo la propuesta iniciada por Ryle, Dennett apuesta por considerar la inoperancia de la interpretación cartesiana de la mente. Desde esta perspectiva cartesiana, podemos entender la mente como un teatro en el cual se (re)presentan los datos mentales. Todos los estados mentales están presentes de algún modo en dicho teatro, interpretando el papel causal sobre la conducta que les corresponde según las órdenes aportadas por 'un jefe de escena', un centro o núcleo consciente, un yo, que a modo de homúnculo dirige todas las operaciones del cuerpo desde la seguridad de su posición trascendental (tan sólo interconectada con lo físico a partir de la glándula pineal). (Un desarrollo actual de este Teatro Cartesiano ha sido desarrollado por Baars (1997), como ya expusimos en el apartado 1.3.).

La apuesta de Dennett por la disolución de un centro donde la conciencia tendrá lugar, donde emergerá, es crucial en su teoría, dado que lo que interesa en la posición dennettiana (como venimos advirtiendo desde el comienzo del capítulo) no es tanto quién realice la acción sino en qué sentido esas acciones tienen por finalidad, por un lado, la mejor adaptabilidad al entorno (permitiendo la supervivencia del organismo) y, por el otro, la predicción adecuada de conductas y hechos futuros. En este sentido, los hechos mentales, especialmente aquellos con un carácter cualitativo, son un modo especial que los individuos conscientes tienen de relacionarse con su entorno y los datos que les aportan les permiten actuar y comportarse de manera adecuada. A diferencia de la conducta física, este carácter fenoménico que poseen determinados estados mentales conscientes, en tanto entendidos como disposiciones al comportamiento (volviendo Dennett a mostrar su herencia ryleana al defender tesis propias del conductismo lógico) debe tener un origen interno a los sujetos, dado que es la introspección (la visión en perspectiva de la primera persona) la que permite su conocimiento, aunque su descripción pueda llegar a hacerse en términos materiales a partir del método objetivo en tercera persona que aportan las ciencias físicas.

También es un pensamiento habitual en la imagen cartesiana de la conciencia el considerar que toda perspectiva consciente debe tener un sujeto establecido a modo de intérprete, que todo punto de vista subjetivo debe estar dotado de un observador. Si consideramos esta posibilidad, también surge la necesidad de tener que aportar una respuesta a la cuestión de su localización. Si consideramos que toda conciencia presupone la presencia de un sujeto, ¿dónde podemos localizarlo? ¿Dónde podemos encontrar la conciencia? La posibilidad de que exista un punto en el cerebro en el cual situar dicho centro consciente es cada vez más inverosímil, como indican los estudios actuales en neurociencia (Damasio 1994 y 1999; Dehaene and Naccache 2001).

Dennett aporta un ejemplo para rechazar la idea de la existencia de un punto concreto en el cerebro en el cual podemos encontrar la conciencia (Dennett 1991: 132ss.; Dennett 1993; Liz 2001a: 88ss.). Imaginemos que con la intención de saber cuándo una determinada información se ha hecho consciente, intentamos señalar las vías neuronales de transmisión de la información de rojo (para los canales de entrada que transportan información todavía no

consciente) y verde (para los canales que ya han convertido en consciente dicha información). Parece que cuando la información transportada por un canal se haga consciente, debe haber un punto concreto de la infraestructura cerebral (y un momento concreto) en el cual dicho canal pase de estar coloreado en rojo a estarlo de verde. Parece intuitivamente cierto pensar que justo en el punto (y el momento) en el cual hemos observado que dicho canal ha cambiado de color es el lugar donde podemos localizar el lugar material en el cual tiene lugar la conciencia, que dicho punto exacto de la estructura cerebral es el lugar donde la información se hace consciente. Pero, nos dice Dennett, ¿qué nos hace pensar eso? Que los canales de transmisión de información sean verdes o rojos no nos garantiza en absoluto que podamos descubrir un lugar en el cual exactamente la información se convierta en consciente. Dicho cambio, podría ser extremadamente difuso, dado que la conversión del color rojo al color verde de los canales de transmisión de información podría ser sumamente imprevisible y variable.

Lo que nos viene a proponer Dennett, en última instancia, es que abandonemos la ilusión de pensar que existe un lugar concreto en el cerebro en el cual se localiza la conciencia, que dejemos de pensar en la posibilidad de la existencia de un centro en el cual deberá situarse a ese observador interno que parece hacer nuestra vida cualitativa privada y subjetiva. Debemos abandonar, por tanto, la idea de que existe una única corriente de conciencia que acaba derivando en un teatro cartesiano (Dennett 1992: 115-125, 267-270).

5.2.2. La heterofenomenología y el modelo de los borradores múltiples

El modelo de los borradores múltiples es la alternativa de Dennett a la propuesta realizada por el Teatro Cartesiano. Siguiendo con su idea de la inexistencia de un núcleo de la conciencia, Dennett afirma que toda nuestra actividad mental se realiza en el cerebro según indica el programa conexionista: esto es, en el cerebro existen una serie de procesos analógicos establecidos en paralelo (siguiendo con la metáfora computacional) en los cuales la información recibida por los sistemas perceptivos del sistema se procesan y reprocesan varias veces de manera continua, siempre sometidos a perpetua revisión. Esta operación cerebral permite la existencia de múltiples borradores, a modo de interpretaciones de dichos datos sensoriales, que los sistemas de procesamiento elaboran para que las respuestas motoras y lingüísticas puedan hacer uso de ellos. Precisamente el modo característico según el cual acabe empleándose dicha información es lo que dará lugar a la aparición de nuestra conciencia.

No existe ningún tipo de jerarquía establecida que privilegie uno de los borradores respecto de los demás. Cualquiera de dichos borradores puede acabar alcanzando el nivel consciente. Que sea uno u otro es algo primeramente aleatorio, y no habrá un lugar concreto donde emergerá y se hará consciente, precisamente por lo variable y borroso que es el mencionado 'umbral de conciencia'. Estos borradores son representaciones simples que, al ser reelaboradas, se van tornando cada vez representaciones más refinadas y amplificadas, siendo procesadas de continuo en los variados y complejos flujos de actividad neuronal que se encuentran repartidos por todo el cerebro hasta que uno de ellos alcanza la categoría de representación consciente (Dennett 1991: 126ss.). Pero no debemos creer que lo que acaba

presentándose a nuestra conciencia es un producto acabado. Todo lo contrario. La representación que acaba por convertirse en consciente se da en un momento en el que es capaz de traspasar los niveles de discriminación consciente. Pero en todo momento esa representación puede dejar paso a cualquier otra cambiando, por tanto, el nivel consciente por completo porque continuamente co-existen en los diferentes estratos cerebrales procesos de revisión (o de re-edición, como indica Dennett) que refinan dichas representaciones simples.

Todas estas representaciones son el conjunto más o menos unificado y revisado de pasadas percepciones. Estas percepciones no son siempre conscientes, pero pueden llegar a adquirir un nivel consciente, siempre y cuando la revisión que los distintos procesamientos cerebrales la hagan tan sumamente interesante que consiga superar nuestro nivel de discriminación. Que lo hagan o no depende de si son capaces de alcanzar el nivel de la conciencia. Dicho nivel no es privilegiado, ni es un observador inobservado, sino que ¡es la conciencia misma!

5.2.2.1. Dos tipos de revisiones de la experiencia: Orwell vs. Stalin

Según Dennett, existen dos modos en los que una determinada experiencia consciente puede haber sido modificada y revisada: estos dos tipos diferentes reciben el nombre de revisiones orwellianas y revisiones estalinianas (Dennett 1991: 129ss.; Dennett and Kinsbourne 1992). Para reconstruir qué sean estos dos tipos de revisiones que Dennett identifica vamos a emplear un ejemplo que él mismo nos describe. Hemos salido a pasear por el parque en una soleada mañana de domingo. Nos cruzamos de pronto con una mujer de pelo corto y gafas cuando nos disponemos a sentarnos para disfrutar tranquilamente de un café en la terraza del bar del parque que, airada, está levantándose de su asiento y se dispone a marcharse. Cuando, unos pocos minutos más tarde nos encontramos con un amigo en ese mismo lugar y nos pregunta si hemos visto a una hermosa mujer con el pelo corto y gafas con la que había quedado en la terraza del bar nosotros, sinceramente, le decimos que acabamos de ver marcharse a una hermosa mujer con gafas, pero que ¡ésta llevaba el pelo largo! ¿Por qué esta confusión? Sinceramente, le hemos respondido a nuestro amigo que no, que a quien hemos visto marcharse era a una mujer totalmente diferente, una mujer con el pelo largo. Pero, sin embargo, sí la hemos visto. La mujer con la que nuestro amigo tenía una cita era la mujer con el pelo corto y con gafas a la que hemos visto abandonar pocos instantes antes la misma terraza en la que ahora nos encontramos tomando nuestro *bel'café*. Hacemos un poco de memoria y recordamos que, poco después de salir de casa, nos hemos cruzado con Cristina, y ella sí tiene el pelo largo. ¿Tal vez hemos confundido a la mujer del pelo con gafas con Cristina? ¿Cómo, si a Cristina la vemos todos los días? ¿Cómo hemos podido mezclar el pelo largo de Cristina con la imagen de la mujer con gafas que estaba sentada en la terraza del bar? Parece que habría dos posibles respuestas.

La primera sería la revisión orwelliana (Dennett toma este nombre porque el tipo de revisión propuesta es similar al tipo de procedimiento de falsificación de la realidad que el Ministro de la Verdad de la novela *1984* de Orwell emplea para interferir en la mente de los

ciudadanos). Este tipo de revisión modificaría la información que ya es consciente en un momento determinado. Se comenzaría dudando de las experiencias que se han tenido hasta tal punto que en un momento determinado dejamos de creer en aquello que recordábamos haber percibido para acabar creyendo algo que en realidad no ha sucedido. Esto es, en nuestra conciencia comenzaría a desaparecer la creencia verdadera de haber visto a una mujer con el pelo corto y gafas sentada en la terraza del bar para acabar falsamente creyendo que a quien hemos visto es a una mujer con el pelo largo, porque nuestra percepción anterior de Cristina (que sí tiene el pelo largo) ha interferido de tal modo en nuestra memoria que acabamos por recordar falsa, aunque sinceramente, que a quien hemos visto es a una mujer con el pelo largo. Lo que nos ha jugado una mala pasada, en esta ocasión, es nuestra memoria, que ha mezclado percepciones diferentes verdaderas en una única percepción falsa.

La segunda clase posible de revisión propuesta por Dennett es la revisión estaliniana (por ser similar al procedimiento empleado durante los juicios políticos llevados a cabo por los tribunales rusos durante el mandato de Josef Stalin). Durante estas revisiones, lo que se modifica es la información percibida antes de que se haya hecho consciente. Nuestra percepción nos hace equivocadamente ver una mujer con el pelo largo y con gafas sentada en la terraza del bar porque nuestra memoria ha hecho que en dicha percepción interfiriera nuestra anterior percepción de una mujer con el pelo largo (nuestra Cristina). Aquí, nuestra memoria no nos ha jugado malas pasadas. Hemos recordado muy bien haber visto una mujer con el pelo largo. Pero ese mismo recuerdo ha hecho que miráramos mal, haciéndonos creer que esa mujer de pelo largo realmente estaba sentada en la terraza del bar cuando, en realidad, no era el caso.

Desde una visión de la conciencia como Teatro Cartesiano, debería ser posible e imprescindible poder discernir nítidamente cuándo ha ocurrido la mezcla de recuerdos y percepciones. Deberíamos poder decidir de qué modo se ha producido la confusión, si hemos mezclado recuerdos o hemos fallado a la hora de percibir. Porque tan sólo así podríamos llegar a dilucidar cuándo ha tenido lugar la experiencia consciente, si antes o después de la interferencia entre recuerdos y percepciones. Pero, como nos dice Dennett (1991: 137-139), nunca podremos llegar a determinar el momento exacto en el que esto ocurre, independientemente de la cantidad de información que lleguemos a manejar respecto del funcionamiento del cerebro de un ser consciente, porque la diferenciación entre los dos tipos de revisiones (aquellas que obedecen a recuerdos y aquellas que lo hacen a percepciones) que dan lugar a las experiencias conscientes no depende del conocimiento de los mecanismos físicos del cerebro; ni siquiera obedece a la posibilidad de poder detectar nítidamente desde el punto de vista introspectivo de la primera persona el momento exacto en el que se produce el error. Porque, precisamente, ¡el error mismo forma parte del modo en que un ser llega a hacerse consciente!

Si todo esto es cierto, como pretende Dennett, ¿cómo es posible que, a pesar de la existencia de dos modos diferentes de revisión de la información perceptiva, podamos llegar a ser conscientes?

5.2.2.2. La heterofenomenología como método para la conciencia

A pesar de que en un momento determinado es posible que en nuestro nivel consciente surja el error perceptivo o se (re)produzca un recuerdo falso, lo habitual es que lleguen al nivel consciente reconstrucciones de percepciones verdaderas que reproduzcan la realidad de los hechos observados. Esto es, lo habitual es que aquellos datos que acaben por hacerse conscientes sean representaciones (débiles) de la carga informacional que los sujetos perceptivos captan.

Como hemos visto anteriormente, el sujeto consciente recibe de los sentidos toda una serie de datos que acaban por transformarse en información que pueda ser procesada. Dicha información es procesada en los sistemas adecuados para rendir cuentas (desde el punto de vista físico y funcional) de esos datos y promover, así, las diferentes conductas. En este sentido, y hablando a un nivel introspectivo, cada individuo procede a interpretar de modo personal los datos que sus propios sentidos le permiten captar.

Esta afirmación parece dejar en el aire la posibilidad de que, en un momento determinado, podamos llegar a tener una ciencia objetiva sobre la conciencia. Es en este preciso momento cuando Dennett (al comienzo de su *Consciousness Explained*) expone el método de la heterofenomenología. La heterofenomenología es, en principio, un método que se propone para estudiar la conciencia. Consiste básicamente en poner sobre la mesa todos aquellos datos disponibles en primera y en tercera persona sobre la mente de los individuos concretos para intentar armar una especie de rompecabezas lo más adecuado y comprensible posible, aunque al final del proceso siempre puede sobrar alguna pieza, capaz de explicar en qué consiste la conciencia. Pero a lo largo del libro, vemos que también se concibe como un instrumento que todo ser es capaz de emplear para sí mismo para llegar a alcanzar un nivel consciente.

En este sentido, una vez el sujeto ha captado los datos que su sistema perceptivo está capacitado para captar, dicha información sufre un proceso de discriminación aplicado por los propios intereses que el sujeto tiene (a un nivel funcional). Es decir, toda esa carga informacional es filtrada a través de un método interpretativo (la heterofenomenología así entendida), una especie de cuenta-historias intra-personal que procede a describirse para sí mismo el modo en que los datos pueden ser entendidos, configurando así un borrador de cómo debe ser entendida una experiencia concreta (desde el punto de vista de la primera persona) y cómo puede ser respondida de manera comportamental (desde el punto de vista de la tercera persona). En este sentido concreto, con la información obtenida de los sentidos, todos los sujetos realizan interpretaciones personales que, en un sentido u otro, cuando lleguen a ser conscientes, adquirirán el estatuto de disposiciones a la acción, unas predisposiciones a comportarse de una determinada manera que tienen un origen interno al propio sujeto perceptor y que acaban por ayudarlo a conformar su nivel de conciencia.

Si la entendemos de este modo, entonces, la intención de Dennett al aportar una propuesta metodológica de tipo heterofenomenológicamente interpretativo parece ser un intento de explicación del carácter de la experiencia fenomenológica de los sistemas perceptivos (a un nivel puramente cognitivo) que tenga en cuenta tanto, por un lado, el nivel puramente físico (un

nivel caracterizable como el *hardware* de un ordenador) que está distribuido en paralelo y tiene la capacidad de procesar adecuadamente la información aportada por los sentidos y, por otro lado, un nivel funcional que aporta una competencia interpretativa a los agentes que les permite de modo más o menos adecuado producir múltiples borradores de lo que serían sus propias percepciones y que, también, posibilita la fijación del contenido experiencial a partir de multitud de micro-juicios que le permiten ir (de un modo disposicional) configurando las posibles respuestas conductuales.

Si, como hemos visto, Dennett está en lo cierto y la experiencia se procesa de este modo, entonces ambos niveles que conforman la interpretación se encuentran a un nivel sub-personal. Si esto es cierto, entonces tanto el procesamiento de información como el proceso interpretativo dependen de un nivel funcional, lo que relega al estatuto de mera ficción a las experiencias cualitativamente clasificables que son reconstruidas a partir de la información aportada por los sentidos. Si las experiencias mentales con carácter cualitativo no tienen una realidad, entonces ¿cómo debemos entenderlas?

5.2.3. Quineando los *qualia*

La opción de quinear los *qualia* (las experiencias con contenido cualitativo o vivencias) se refiere al hecho de ‘negar con firmeza la existencia o la importancia de algo real o significativo’ (Dennett 1988: 226). Siguiendo con su intento de destrucción de la posibilidad, siquiera, de poder entender la mente al modo cartesiano, lo que propone Dennett es eliminar la posibilidad de la conciencia fenoménica a partir de la negación de la existencia de vivencias porque, si éstas existieran, nos dice, deberían ser propiedades mentales (subjetivas e introspectivas) con respecto a cuya representación el sujeto consciente no siempre tendría un acceso epistémico claro, rompiendo la necesidad de la especial autoridad que debería proporcionarle el acceso de la primera persona. (Otro intento de eliminación de las vivencias lo encontramos en Harman 1989). Si, desde el punto de vista dennettiano, debemos abandonar la posibilidad de la existencia de experiencias cualitativamente accesibles, entonces la presencia de experiencias con contenido fenoménico deben tener un papel más normativo que epistémico, al estilo de todas aquellas características de lo mental que gozan de un contenido intencional (como hemos indicado en la sección anterior). Si esto es así, ¿en qué sentido debe entenderse esta normatividad?

Primero atendamos a qué se refiere Dennett cuando habla de la eliminación de las vivencias. Atendamos al siguiente ejemplo (Dennett 1988: 231ss.). Chase y Sanborn son dos empleados de la cadena de calidad de una empresa que tuesta, muele y envasa café. El trabajo de ambos es probar diferentes muestras de café durante todo el día. Como es de esperar, ambos han perdido la sensación de placer que, para muchos de nosotros, supone tomar una buena taza de excelente café, ya que es lo único que hacen a lo largo de todo el día. Un día, durante el desayuno, ambos comentan su situación. Chase cree que la pérdida de la sensación placentera de probar el café se debe a que durante los años su especial capacidad para distinguir diferentes tipos de café se ha refinado tanto que cada vez le cuesta más encontrar un tipo de café que

realmente le satisfaga y que, precisamente por eso, ya no encuentra placer en las catas de café que realiza a lo largo del día. Sanborn, en cambio, considera que no es su especial capacidad de diferenciar cafés lo que ha cambiado, refinándose, sino que es el sabor del café el que debe haber cambiado, un hecho que viene indicado porque el sabor del café ya no le produce la misma sensación placentera que le provocaba cuando comenzó a trabajar para la empresa. Para Dennett, sin embargo, aunque el acceso a la experiencia que ambos tienen permite indicar qué tipo de sensación es la que experimentan, ello no les indica nada respecto a aquello que realmente causa dicha sensación concreta. Esto es, aunque ambos son conscientes de la misma sensación, ninguno de ellos es capaz de determinar la causa por la cual han dejado de experimentarla como placentera. Visto así, aunque no exista una diferencia en la sensación experimentada, nada impide decir que pueda existir una diferencia real en el modo en que dicha sensación es causada. Cuando atendemos a esta posibilidad de tener sensaciones, no existe una separación (lógica) entre la apreciación del sabor del café (en este caso) y la peculiar reacción placentera o no que dicha acción presupone. Lo que indicará, en último término, cual de los dos está realmente en lo cierto a la hora de determinar la causa de su sensación es cual de las dos creencias está justificada realmente por los hechos del mundo, si la de Chase o la de Sanborn. Es decir, lo que indique (como en los casos de Variación Estándar analizados en el capítulo sobre Tye, véase la sección 4.3.2.) cual de ambos está en lo cierto es la existencia de una situación en el mundo (el hecho físico) que permita verificar la creencia de alguno de ellos.¹⁵

Esta situación nos lleva a la tesis de la eliminación de las vivencias. Si en un momento determinado podemos prescindir del acceso a la conciencia desde una perspectiva en primera persona (es decir, si podemos prescindir de la pretensión de infalibilidad, de privacidad y de incorregibilidad de las experiencias cualitativas) y, aún así, no existe ninguna diferencia en cuanto a la sensación que nos aporta la percepción de un determinado sabor de café (por ejemplo), entonces podemos prescindir de la presencia de dicha sensación placentera porque, en última instancia, tendría un papel causalmente irrelevante a la hora de explicar la conducta del sujeto (lo que le ha valido en numerosas ocasiones a Dennett la acusación de defender cierto tipo de epifenomenalismo (Cf. Sanfélix 2003: 326 n. 164).

Pero la propuesta de Dennett debe ser entendida desde otro punto de vista.¹⁶ Lo que Dennett persigue no es la eliminación del aspecto consciente y fenoménico de los *qualia* (esto

¹⁵ Una idea similar es la que, desde una posición externista, emplea la parte final de García-Carpintero 2001 para, por un lado, socavar la pertinencia de una teoría totalmente fundacionalista del conocimiento con un carácter marcadamente internista como es la teoría cartesiana y, por otra parte, salvaguardar las razones que ofrece Loar 1997 para afirmar que las vivencias subjetivas son propiedades de objetos materiales esencialmente reconocibles por sus propiedades observables.

¹⁶ Ante la tesis que sigue, se podría objetar que para que la eliminación dennettiana de las vivencias fuera, siquiera, plausible desde este punto de vista, debería presuponerse que la concepción cartesiana de los *qualia* (defendida por Block, Chalmers o Searle, entre otros) es la caracterización ortodoxa de los mismos. Pero tal cosa, como hemos visto en el capítulo 2, no parece ser el caso. Una tesis similar sostiene García-Carpintero 2001. De todos modos, y a diferencia de él, aunque es cierto que Dennett emplea a menudo la tesis eliminacionista respecto de las vivencias como recurso estilístico para desmitificar las 'entidades' filosóficas tradicionales (Cf. Dennett 1988 o, más recientemente, Dennett 2006a), considero que cuando Dennett apela a la eliminación de los *qualia* lo que persigue es, en última instancia, hacer ver que diferentes posiciones representacionales son insostenibles. Precisamente esta idea

es, de su aparición a un nivel personal, algo que estaría garantizado como ficción), sino que lo que defiende es la eliminación de las vivencias en un sentido epistémico, intelectual.¹⁷ Esto es, en un sentido más o menos laxo, lo que Dennett estaría defendiendo es que, en última instancia, los *qualia* deben ser eliminados (como entidades metafísicamente relevantes) porque el sistema cognitivo de los seres perceptivos no tiene por qué llegar a ser consciente del modo en que las diferentes sensaciones son causadas, sino que tan sólo le vale con llegar a ser consciente de la sensaciones mismas. Que lo único de lo que puede ser consciente es aquello que puede llegar a ocupar una posición dentro del nivel personal del sujeto (de acción). Esto es, si hemos caracterizado a las vivencias como aquellos elementos experienciales que a un nivel físico-funcional, por un lado, y a un nivel interpretativo que depende también de sistemas de procesamiento de información distribuidos en paralelo a lo largo y ancho de todo nuestro cerebro, por otro, podemos concluir que no tiene por qué existir ningún tipo de acceso epistémico a dichos estados sub-doxásticos, por lo que la realidad de dichos estados a un nivel epistémico es irrelevante para explicar la posesión de una determinada sensación y, lo que es más importante (como veremos en el apartado 5.4.), la acción del sujeto. Lo que nos interesa ahora es cómo puede tener ese nivel sub-personal un papel causal en la conducta del sujeto (para evitar así la acusación de epifenomenalismo) y, lo que sería más importante, por qué dicho nivel tiene una normatividad respecto del comportamiento de los sujetos.

5.2.4. El papel normativo de los estados mentales (incluidos los fenoménicos)

Un hecho que parece carecer de cualquier duda es la presencia, al menos entendida en un sentido débil, de cierto tipo de *qualia*; en un sentido en el cual podemos decir que existen ciertos tipos de estados de los cuales somos conscientes por familiaridad (todos los seres conscientes parece que tenemos percepciones y propiocepciones con un determinado contenido fenoménico). Sin embargo, la presencia de estados cerebrales y funcionales no permitiría garantizar la existencia de estados intencionales, por lo que es necesario postular, al menos, la existencia de estados con un contenido representacional determinado que garantice la posibilidad de que los sujetos conscientes seamos capaces de actos que denoten dichas vivencias (Sellars 1963; Shoemaker 1984), en tanto que ¿son los *qualia* los que hacen que merezca la pena vivir! (comentario personal de Sellars a Dennett. Véase Dennett 1991: 395 o Dennett 2006b: 110).

Vistas así las cosas, el acto de eliminación de las vivencias, entendidas como el acceso subjetivo al carácter cualitativo de ciertos estados mentales, parece indicar que éstas son irrelevantes para explicar la aparición de la conciencia. En este sentido, el único modo de entender qué sea la conciencia se vería relegado al mero hecho de que un ser tuviera estados con un cierto contenido representacional cuyo contenido, a su vez (estableciendo, así, una indeseable regresión al infinito) tan sólo concerniera a sus propios estados (cognitivos y

parece salir reforzada de la lectura de aquellos textos que García-Carpintero 2001 emplea para defender lo contrario (Cf. Dennett 1979 y 1991).

¹⁷ Para la distinción entre los diferentes sentidos de *qualia*, véase la sección 2.2.3.2.

funcionales. Cf. Block 1993: 182-183). Es decir, parece que el único sentido de conciencia relevante para Dennett sería una especie de ‘conciencia de’, una conciencia puramente reflexiva, intencional, auto-dirigida tan sólo a los propios estados funcionales, como diferenciada de una conciencia por acceso capaz de poseer un contenido inferencialmente disponible como posibilidad racional de control de la acción (Block 1995). Pero desde el punto de vista de Dennett, ello no tiene un claro inconveniente.

Hemos visto que, en la teoría de Dennett, todo estado mental (especialmente aquellos estados mentales que tienen un contenido con carácter fenoménico, a los que hemos caracterizado como vivencias) debe tener un contenido objetivo, físico, en tanto que dicho estado mental es producido por un determinado estado de cosas (esto es, por ciertos objetos y sus propiedades físicamente observables). En este sentido, dentro de la teoría defendida por Dennett, la práctica de atribuirnos estados mentales a nosotros mismos y a nuestros semejantes (esto es, el acto mismo de la actitud intencional) implica presuponer que el contenido de dichos estados es compartido por todo ser consciente, presuponiendo también que todos aquellos seres a los que se atribuye dicho estado mental comparten el mismo propósito, promoviendo de este modo la acción a seguir. Vistas así las cosas, lo que cabría preguntarse es en qué sentido podemos decir que estaría determinado el contenido objetivo de dichos estados mentales, dado que, si concebimos que dicho contenido es amplio, corremos el riesgo de caer en la indeterminación y en la vaguedad de nuestros conceptos mentales. Pero si concebimos dicho contenido como reducido (*narrow*), como hace Dennett, al considerar que, aunque existe la posibilidad de la múltiple realizabilidad, todo contenido está determinado por un estado cerebro-funcional concreto (aunque ese estado no tenga una localización clara dentro de la arquitectura cerebral), entonces caeríamos de nuevo en la indeterminación del contenido objetivo de los estados mentales porque tendríamos que demostrar en qué sentido dichos estados cerebrales intuitivamente poseen dichos contenidos (Stalnaker 1989).

Tal vez, la teoría dennettiana de la conciencia no deba ser entendida como una teoría eliminacionista, con implicaciones epifenomenalistas, sino que deba ser entendida de un modo en el cual, primero, exista la posibilidad de mantener cierto tipo de realidad respecto de la existencia de vivencias pero, segundo, al mismo tiempo se supedite la posesión de contenido representacional de los estados funcionales y cerebrales a la direccionalidad de los mismos hacia estados superiores.

5.3. La teoría dennettiana como una teoría representacional de Orden Superior

Al haber apostado por la posibilidad de la existencia en un nivel sub-doxástico de dos instancias, aquella que procesa de manera múltiple la información y aquella que permite la captación de las percepciones, la posibilidad de la aparición de la conciencia queda a un nivel diferente dentro de la teoría de Dennett. Parece que se salvaguarda la posibilidad de la existencia de un nivel sub-personal que es base física de la aparición de la conciencia, pero ¿qué ocurre a nivel personal? Desde la posición de Dennett, parece claro que tan sólo podremos decir que un sistema cognitivo como el anteriormente descrito alcanza el nivel de la conciencia

cuando llega a tener estados mentales de nivel superior (*higer-order mental states*), esto es, cuando el sistema sea capaz de tener estados mentales que se refieran (que tengan por contenido) a otros estados mentales. Es decir, para que aparezca el nivel de la conciencia, el sujeto perceptivo debe dejar de tener percepciones (o interpretaciones de las percepciones) y debe comenzar a mostrar creencias respecto de dichas percepciones, debe sustituir el nivel sintáctico del sistema cognitivo por el nivel semántico de las actitudes proposicionales respecto de los estados mentales (como la discriminación, la identificación y la percepción), porque precisamente, para Dennett, sólo los sistemas representacionales (aquellos sistemas que muestran esta clase de intencionalidad) son capaces de alcanzar la conciencia.

La presencia de estos estados de nivel superior (o de segundo orden) posibilita también que la teoría dennettiana sea una teoría disposicional. Como hemos comentado, las creencias son estados de segundo orden que se refieren a estados primarios (como la percepción de algo frente a nosotros). A pesar de que estas creencias no tienen por qué ser explícitas, necesariamente deberán tener un papel explicativo en la conducta del sujeto, dado que marcan las pautas a seguir para satisfacer los estados de orden superior. Así, dada una creencia respecto de una percepción, se establecerá una representación que permitirá la acción del sujeto que posee dicho estado superior.

¿Cómo ocurre esta adquisición de estados de nivel superior? ¿En qué sentido la teoría de Dennett es una teoría de Orden Superior y en qué se diferencia de otras teorías similares? ¿Permite esta normatividad de los estados superiores explicar realmente la aparición de la conciencia?

5.3.1. La teoría descriptiva de Dennett frente a otras teorías de Orden Superior

Tal vez la mejor manera de entender el modo en que Dennett postula su teoría es comparándola con otras teorías que, de uno u otro modo, han pretendido también dejar constancia de la aparición del pensamiento consciente a partir de la atribución a los sujetos de los estados mentales de un contenido intencional que incluyera otros estados conscientes.

Dichas soluciones han sido de lo más variadas, dependiendo del modo en que entendían cada una de ellas el concepto de representación o la noción de intencionalidad. Lo que tal vez sea un rasgo común de todas ellas es un intento sistemático (tal vez también sintomático) de naturalizar en mayor o menor medida dichas nociones, intentando aportar explicaciones físicas a sus propuestas con la finalidad de cerrar problemas respecto de la aparición de pensamiento consciente.

Atenderemos en las próximas sub-secciones a una exposición y comparación de las teorías de Orden Superior (HOT, del inglés *Higer-Order Theory*) atendiendo especialmente al papel que otorgan a las representaciones y a la solución que intentan ofrecer a la problemática relativa a la presencia de determinado contenido fenoménico que puede ser tachado de no-conceptual.

5.3.1.1. Teorías de Primer Orden (FOT) y Teorías de Orden Superior (HOT)

Una primera diferenciación que deberíamos hacer, ampliando la explicación realizada en la sección 2.1.2.4., se establece entre aquellas teorías de Primer Orden (*First-Order Theory*, o FOT) de las teorías de Orden Superior. Una teoría de Primer Orden es aquella que pretende caracterizar todas las propiedades fenoménicas de la experiencia (las vivencias, aquello que es para alguien sentir una determinada experiencia) en términos de contenido representacional de la experiencia. Presentada de este modo, entonces, para una teoría de Primer Orden (o representacional) aquello que diferenciaría una experiencia de, digamos, verde y otra experiencia de, por ejemplo, azul sería la diferencia existente entre las propiedades (físicas) representadas por ambas experiencias (Kirk 1994; Dretske 1995; Tye 1995 y 2000). De modo similar, la diferencia entre un dolor y un picor puede ser explicada, en términos representacionales, en base a las diferentes propiedades físicas representadas (como trastornos o anomalías) situadas y localizadas en regiones concretas del propio cuerpo del sujeto de experiencia (James 1884; James 1890: Vol. II, cap. 25). Así entendida, será una teoría de Primer Orden toda aquella teoría respecto de la conciencia fenoménica que considere que una experiencia con carácter cualitativo esté *preparada (poised)* para generar un impacto tal en las actitudes proposicionales (incluyendo los procesos de razonamiento práctico) del sujeto que sea capaz de modificar su comportamiento, que varíe su nivel personal (Cf. Tye 1995).

Desde una perspectiva FOT, la explicación de la conciencia fenoménica consiste básicamente en mostrar la transición que se da desde organismos con un (más o menos) variado repertorio de comportamientos reflejos (más o menos) ocasionados por las características básicas de su entorno (como, por ejemplo, las amebas) a organismos cuyos actos reflejos innatos son acciones guiadas por informaciones cuasi-perceptivas entrantes, entendidas como estados cognitivos que despiertan salidas comportamentales pre-establecidas (un claro ejemplo lo encontraríamos en las abejas de la miel, como afirma el capítulo final de Tye 2000, o en las orugas, como comenta Reber 1997) y, posteriormente, a organismos que poseen un cierto nivel de aprendizaje de acciones también guiado por informaciones cuasi-perceptivas (como ciertos pájaros que han conseguido aprender a construir sus nidos aprovechando los salientes de las fachadas) que, a través de un proceso evolutivo estándar, han conseguido llegar a convertirse en organismos cuya información perceptiva les ha permitido comenzar a disponer de pensamientos y razonamientos conceptuales básicos (como cierto tipo de primates superiores capaces de emplear herramientas fabricadas por sí mismos, como pequeñas varillas introducidas por los orificios de los hormigueros destinadas a ‘pescar’ termitas). El paso a un nivel representacional humano, visto desde este punto de vista, estaría supeditado a la cantidad de información que nuestro sistema perceptivo estaría capacitado para procesar permitiendo, así, conformar cada vez más complicados razonamientos.

Como hemos comentado anteriormente, una teoría de Orden Superior es aquella que considera que la aparición de la conciencia fenoménica se debe a la posibilidad del sujeto de representar estados mentales cuyo contenido es (la representación de) otros estados mentales. Es decir, explican la aparición de la conciencia a partir de la intención de ciertos estados mentales

de dirigirse a otros estados mentales. Desde este punto de vista, podemos diferenciar distintas perspectivas HOT según el modo en que se entiende esa direccionalidad.

Desde una perspectiva en sentido interno (*inner sense*), una experiencia consciente aparece cuando los estados perceptivos de primer orden (los sistemas perceptivos de nivel sub-personal capacitados para procesar la información del entorno) son *sondeados* por una facultad interna (una facultad de resonancia y ascendencia kantiana, similar a la introspección que permitía la apercepción trascendental) capaz de producir los estados de segundo orden a partir del 'darse cuenta' de la presencia de estados primarios (Armstrong 1968; Lycan 1996).

Desde otro punto de vista de acuerdo con el cual la conciencia fenoménica aparece cuando el estado perceptivo primario es captado por un estado superior de modo adecuado, podemos diferenciar entre aquellas teorías, llamadas actualistas, que defienden la presencia actual de un determinado estado de segundo orden acerca de un estado perceptivo primario que posibilita la aparición de la experiencia consciente (Rosenthal 1986; Gennaro 1996) y aquellas teorías, llamadas disposicionales, que priorizan la disponibilidad de un estado perceptivo determinado para un cierto estado de nivel superior que permitiría la aparición de la experiencia consciente (Carruthers 1996), tal y como vimos en la sección 2.1.2.4.

Por otra parte, también podemos diferenciar aquellas teorías, como la de Dennett, que consideran que, desde una perspectiva descriptivista respecto de los estados de segundo orden, resulta plausible la aparición de las experiencias conscientes a partir de la descripción (lingüísticamente formulada) de los estados mentales de un determinado sujeto y del papel que éstos realizan en su pensamiento y en su conducta.

Las ventajas de las teorías (representacionales) de Orden Superior respecto de las teorías (representacionales) de Primer Orden parecen hacerse patentes cuando atendemos a determinados problemas concretos de la conciencia, como la posibilidad del espectro invertido o de los *qualia* ausentes (expuestos en el apartado 1.4.), porque desde el primer punto de vista, al poder prescindir de una apelación a aquellas propiedades no-representacionales de la experiencia consciente, caracterizadas por su acceso introspectivo, podemos simplificar en recursos cognitivos a la hora de tener que ofrecer una explicación de la conciencia fenoménica y del carácter cualitativo de las experiencias, evitando así cualquier mención de estados o propiedades físicas (con poder representacional) internas a la propia mente.

Pero, ¿realmente son tan deseables las explicaciones ofrecidas por las teorías de Orden Superior?

5.3.1.2. Las dificultades de las Teorías de Orden Superior

La mayor dificultad que encontramos al tratar con las HOT del sentido interno es con la complejidad computacional que requiere la explicación de los monitores internos de detección postulados. Todo estado interno de detección podría ser reducido a una explicación física, pero para ello se requeriría precisar aquellos eventos físicos del cerebro que posibilitan las detecciones y las salidas comportamentales de los diferentes sistemas perceptivos postulados. Pero además, es difícil explicar cómo un detector interno como los mencionados es capaz de

detectar la presencia de una experiencia y ser capaz, a la vez, de entenderla como una experiencia. Parece que estas teorías se verían obligadas a tener que identificar las bases físicas de realización de las experiencias en el cerebro humano y (re)construir, a partir de las bases ofrecidas por la presencia de información de carácter físico, los requisitos de representación de la experiencia que estos estados físicos son capaces de implementar. Esta acción, sin embargo, parecería tornar inevitable que la detección generase experiencias de nivel superior, dado que se requeriría de un nivel más elevado y sofisticado que el puramente ofrecido por el nivel sistémico de detección para poder rendir cuentas de las experiencias mismas.

Vistas así las cosas, parecería que una HOT del sentido interno permitiría refinar las percepciones primarias hasta un grado tal que pudiera integrar todo el contenido perceptivo en una única experiencia, explicando así la unidad fenoménica y del objeto (Lycan 1996). El problema radicaría ahora en demostrar cómo los estados secundarios se integran en el proceso, sobre todo después de ver que, en realidad, lo que postulan las HOT del sentido interno es la presencia de un nivel de simulación mental que haga posible la atribución de estados mentales a los sujetos (incluidas las auto-atribuciones). A pesar de tener grandes similitudes, no deberíamos confundir esta propuesta con la ofrecida por las teorías simulacionistas de la mente, como la de Goldman (1993), que consideran que la perspectiva interna es un producto evolucionado con la finalidad primaria de 'leer la mente' de los demás (y la propia), dado que lo que defienden las HOT del sentido interno es un nivel donde se generan simulaciones que permitan la generación y la anticipación a las conductas mediante la estipulación de predicciones.

Si la idea principal de una teoría del sentido interno, como hemos visto, afirma que nosotros tenemos tan sólo un acceso introspectivo a, al menos, algunos de nuestros estados mentales, ¿cómo se supone que podemos, entonces, llegar a generar simulaciones de la vida mental de los otros seres conscientes tales que nos permitan anticiparnos a sus movimientos mediante la creación de predicciones de su propio comportamiento? Parece más plausible rechazar esta teoría que intentar ofrecer una explicación razonable. Por varios motivos.

En primer lugar, si concebimos (como bien hacen las teorías reductivas) que nuestras habilidades cognitivas son implementadas a partir de una teoría (en un sentido laxo, entendida como una configuración jerárquica de nivel de captación de información) respecto de la estructura y la funcionalidad de nuestra mente (Cf. Fodor 1987), entonces parece que toda teoría debe incluir conceptos que permitan entender dichas habilidades desde un punto de vista semánticamente relevante (Cf. Carruthers 1996b), por lo que una mera simulación tendría una laguna insalvable a la hora de dar cuenta de ese nivel intencional aplicado a los otros porque existiría una separación concreta entre los niveles de unos y otros. Es decir, la simple simulación por parte de un sujeto consciente de las capacidades cognitivas de otro sujeto tan sólo tendría un sentido práctico de inferencia, pudiendo ser totalmente equivocado que realmente el otro sujeto disponga de dicho nivel cognitivo (esto es, entrando en la falacia de suponer que todo sistema pueda ser entendido como intencional, al aplicar indiscriminadamente la actitud intencional). Pero, además, en segundo lugar, aunque consideremos la posibilidad de la inexistencia de dicho nivel semánticamente relevante, una teoría del sentido interno tendría

dificultades a la hora de tener que rendir cuentas de cierto tipo de propuestas evolutivas. Por ejemplo, consideremos un organismo que dispone de un sistema primitivo sensible a la luz (como parece el caso de los primeros estadios evolutivos de los ojos. Cf. Llinás 2003). La explicación de la evolución de dicho proto-sistema hasta alcanzar la sofisticación actual del sistema visual de un ojo parece perfectamente explicable desde un punto de vista puramente mecánico, a partir de las continuas adaptaciones a los cambios de luz, al aprendizaje de movimientos, etc. Desde el punto de vista de las teorías del sentido interno, dicha evolución requeriría una facultad de simulación que, de un modo u otro, necesitaría de un cierto nivel de conceptualización de sus propias entradas y salidas, un nivel conceptual irrelevante para la explicación de dicha evolución.

Vista desde este punto de vista, una teoría del sentido interno parece implausible, porque es algo contingente que nosotros, seres conscientes complejos, podamos haber evolucionado del modo en que lo hemos hecho. Podríamos haber evolucionado de un modo diferente, sin necesidad de haber alcanzado el nivel de poder representarnos nuestros propios estados mentales, por lo que, en definitiva, todavía parece más irrelevante si cabe la posibilidad de la existencia de un nivel de experiencias que conceptualicen otras experiencias. Parece mucho más plausible, y no sólo por una mera cuestión de economía evolutiva, pensar que el hecho de tener una experiencia ya nos acredita para saber que tenemos dicha experiencia, sin necesidad de tener que apelar a un nivel experiencial conceptualmente elevado que permitiera tener dichas experiencias.

Hemos comentado también la existencia de teorías actualistas de nivel superior. Estas teorías apuestan por explicar que un estado no se vuelve consciente hasta que llega a ser plenamente *sentido*, es decir, consideran que para tener una determinada sensación consciente se requiere de un cierto tipo de estado reflexivo superior sobre dicho estado que indique al sujeto que se encuentra teniendo determinada experiencia (Rosenthal 1997). Pero éste parece no ser el caso. Parece demasiado pedir que sobre todo estado con contenido fenoménico deba actuar otro estado intencional que permita entenderlo como perceptivamente relevante y, por tanto, elevarlo hasta el nivel de conciencia. Imaginemos, por ejemplo, que un buen día, estando en un bar tomando una cerveza, un amigo despistado me quema la mano con un cigarrillo. Retiro mi mano ante la sensación de intenso calor que provoca mi sensación de dolor. ¿He sido consciente desde el principio de que he tenido un dolor? Parece que poco importa, desde un punto de vista neutro. Lo importante parece ser que, en un momento determinado, ante la sensación de calor que el cigarrillo ha provocado en mi piel, la reacción ha sido retirar la mano del foco de calor. Desde una perspectiva actualista, la reacción de apartar la mano se explicaría del siguiente modo: (la piel de) mi mano, al entrar en contacto con el calor del cigarrillo, ha permitido que tuviera una determinada experiencia sensitiva, una experiencia sobre la cual ha intervenido una creencia (la creencia de que me estoy quemando) y un deseo (el deseo de evitar todo daño), lo que ha permitido que dicha experiencia adquiriera el estatuto de sensación consciente de ‘dolor’, permitiendo así que pudiera retirar la mano. La respuesta obvia sería la siguiente: ¿no es un tanto rebuscado pretender que toda esa cantidad de información se haya ejecutado de modo tan

rápido? Aunque fuera cierto que nuestro sistema nervioso ejecutara tan rápidamente dicha información, ello presupondría también un enorme gasto evolutivo. La respuesta más plausible sería suponer que, ante la sensación de calor sobre mi piel, mi mano se repliega para evitar el daño mayor. Sólo después de realizada dicha acción, puedo llegar a ser consciente de que tengo dolor porque me he quemado. De otro modo, parecería como si siempre tuviera que ser consciente de la sensación desagradable para poder ponerme a salvo, algo que no concuerda con los principios de la supervivencia. En este sentido, la posibilidad de que siempre debamos tener una experiencia de nivel superior sobre nuestra experiencia primaria para explicar la reacción de la retirada de mi mano parece que caería en saco roto cuando no existe dicho reconocimiento por parte del sujeto. Incluso dejaría sin explicar hechos tan concretos como la reacción de ciertos animales (incluidos los humanos) de huir frente a la mera posibilidad de la existencia de un peligro cercano. Entendido de este modo, parece más plausible decir que, aunque en todo momento cabe la posibilidad de reflexionar sobre la sensación primaria a partir de estados intencionales secundarios, lo realmente importante es la *disponibilidad* de la existencia de un estado fenoménico primario que permita explicar la reacción de retirar mi mano de la fuente de calor, una disponibilidad puesta de manifiesto cuando después nos damos cuenta de que sentimos dolor porque nos hemos quemado.

Desde el punto de vista de las teorías disposicionales de nivel superior, la conciencia fenoménica es un sub-producto (indirectamente creado a partir de la evolución, sin que ésta capacidad fenoménica fuera la función propia por la cual evolucionó, como una *exadaptación*, al modo en que Gould and Lewontin (1979) lo entienden) que emerge por el procesamiento de los varios sistemas sensitivos que generan una representación completa del entorno, un proceso de manipulación de la información que permite hacer disponibles toda una serie de sub-sistemas que generan una gran variedad de conceptos, razonamientos, planificaciones, pensamientos y creencias que permiten la acción de los organismos (esto es, los predispone para actuar según un modo más o menos determinado dependiendo de la entrada sensitiva que se tenga en cada momento) sin necesidad de apelar a un sentido interno que sondee cada una de las experiencias sensitivas de los diferentes sistemas (Carruthers 1996 y 1998). El problema más claro a la hora de enfrentarnos con esta concepción de la posibilidad de la conciencia reside en el hecho de tener que explicar la modularidad previamente concebida. Si entendemos la presencia de módulos, aunque sea desde un punto de vista funcional (como tienden a hacer sus partidarios, apelando a la posibilidad de que el contenido de dichos estados dependa de la capacidad de los sistemas productores de representaciones de *consumirlos*, en el sentido de Millikan 1984), entonces parece como si la presencia de representaciones bajo las cuales los organismos acaban actuando dependiera realmente de la presencia de múltiples homúnculos que deciden por el propio organismo cómo actuar en cada momento según la entrada sensitiva pertinente. Ello supondría que dicha capacidad sistémica indicaría la especialización de los diferentes módulos, impidiendo de este modo la posibilidad de la existencia de varios modos diferentes de acción que pudieran ser realizados a partir de la misma entrada. Aunque nos apoyáramos en los descubrimientos evolutivos recientes respecto de la especialización del cerebro, como por

ejemplo en los últimos experimentos sobre la posibilidad de la existencia de un órgano del lenguaje o los descubrimientos respecto de la especialización de diferentes partes del cerebro,¹⁸ como suelen hacer los partidarios de esta teoría (Cf. Carruthers 2000), precisamente son esas mismas interpretaciones las que nos dan buenas razones para negar la realidad de la explicación disposicional (recuérdese los ejemplos del pulgar del panda o del disruptor de simetría a los que atendimos en la sección 5.1.3.).

5.3.1.3. La Teoría Descripcionista de Orden Superior de Dennett

Podríamos calificar la posición de Dennett como una teoría de orden superior basándonos en los siguientes tres aspectos de su concepción.

En primer lugar, dentro del sistema dennettiano, los homínidos son entendidos como organismos evolucionados hacia la posesión de una amplia variedad de sistemas de procesamiento de la información especializados en competencias más o menos particulares y determinadas organizadas de un modo interno como siguiendo las pautas de las arquitecturas conexionistas en paralelo. Estos sistemas especializados, a diferencia de lo que ocurre en las teorías disposicionalistas, son entendidos como sistemas de intercambio informacional que cooperan entre ellos, permitiendo la aparición de procesadores capaces de crear ‘herramientas’ que permitan organizar de manera externa la información vital respecto del mundo en el cual se mueve el organismo, estableciendo pautas de conducta que permiten su supervivencia y expansión genética. Estos sistemas no operan de modo cerrado, no son considerados módulos, por lo que siempre es posible la redistribución de competencias y la manipulación errónea de información, con la consecuente acción fallida. Estos sistemas de procesamiento operan de modo independiente, pero tienden a establecer una coherencia comportamental a partir de un equilibrio proporcionado a nivel sistémico global por una emergencia consciente de unidad.

En segundo lugar, la posibilidad de que los homínidos hayan conseguido adquirir una herramienta evolucionada como es el lenguaje les permite el intercambio interpersonal de información. Dicho intercambio ha permitido, además, la sofisticación de cierto tipo de sistemas de procesamiento de información que han supuesto una ventaja respecto a las demás especies, consolidando así formas de intercambio y cooperación a partir de la posibilidad de establecer representaciones mentales de las propias experiencias.

Por último, la posibilidad de que los homínidos hayamos comenzado a tener la capacidad de *contarnos historias* a nosotros mismos, de hablar con nosotros mismos, permitió la evolución de un sistema interno de habla (la máquina joyceana)¹⁹ que, a modo de procesador virtual, engloba de manera sistemática y organiza toda la arquitectura cerebral, permitiendo así

¹⁸ Tal vez la propuesta más ambiciosa, desde Chomsky, de identificar la existencia de un órgano del lenguaje es la que encontramos en Lightfoot (1999) y en Anderson and Lightfoot (2000). Sobre otros desarrollos (modulares) de la especialización del cerebro, véase Samuels (2000) y Wynn (2000). Para una crítica de esta concepción modular que intenta identificar un órgano del lenguaje, véase Colomina (2008).

¹⁹ Dennett denomina máquina joyceana a la secuencia de contenidos mentales conscientes que inundan la vida mental de los sujetos, aquello que W. James denominó flujo de conciencia, en honor de las excelentes descripciones que el escritor irlandés James Joyce realiza de ellas en sus novelas. Cf. Dennett 1992: 227.

la continua interconexión y la posibilidad de alcanzar cierto tipo de datos físico-funcionales, a través de una interpretación constante, un nivel personal que lo diferencia de la mera mecanización y lo dota de un nivel semántico relevante para el individuo.

La diferencia principal con otras teorías de orden superior puede encontrarse en un dato importante que en el sistema de Dennett es crucial a la hora de entender la posibilidad de que una experiencia llegue a hacerse consciente. Es totalmente innecesario para que una experiencia llegue a ser consciente postular la presencia de un pensamiento (la existencia de una actitud proposicional) entre dicha experiencia y nuestra disposición a transmitirla de modo lingüístico o a actuar según la misma (Dennett 1991: cap. 10, especialmente 327ss.). Evitando esta clase de intermediarios, se demuestra que las capacidades lingüísticas han tenido una evolución prioritaria y que no es necesario estipular la existencia de un contenido conceptual a un nivel sub-personal permitiendo, así, la presencia de sistemas independientes de manipulación sintáctica de la información. Pero precisamente esta misma afirmación le ha valido la crítica de recaer en el disposicionalismo (Cf. Rosenthal 2000).

Dennett, sin embargo, niega la existencia de pensamientos realmente contruidos de modo independiente al lenguaje, por lo que no es posible la aparición de ninguna clase de pensamiento de orden superior de modo independiente del mismo, lo que lo aleja de posiciones disposicionales (Carruthers 2000: 271-272). Desde este punto de vista, los contenidos de los estados mentales de orden superior tan sólo pueden ser formulados de modo lingüístico (unos contenidos que los individuos conscientes pueden estructurar individualmente como estados causalmente efectivos a través del sistema interpretativo que permite la estrategia intencional postulada en Dennett 1987).

Tal vez, a estas alturas, el problema principal de la teoría de Dennett no sea tanto la imposibilidad de la existencia de estados de nivel superior estructurados e individualizados de manera independiente de la capacidad lingüística de los organismos (al menos, cuando nos referimos a organismos gregorianos que han conseguido alcanzar un nivel de construcción semántica), sino la imposibilidad de mostrar cómo existen capacidades cognitivas de manera independiente de dicha capacidad (como ocurriría en el resto de estratos identificados por Dennett. Cf. Carruthers 2000). Parece ser que apelar a la posibilidad de un nivel interpretacional (el nivel que el mismo Dennett 1991: 132 denomina *operacionalismo* de primera persona), a modo de actitud intencional, sólo es válido cuando los organismos han conseguido elevarse al nivel de intercambio semántico, pero que a un nivel más cognitivamente dependiente de estructuras biológicas sin contenido conceptual, dicha interpretación se hace sumamente difícil, por lo que la presencia de un nivel interpretativo en dichos organismos se aplica tan sólo desde una perspectiva de tercera persona, impidiendo así cualquier posibilidad de niveles superiores de pensamiento en dichos organismos, haciendo tambalearse la teoría dennettiana.

5.3.2. Críticas a la teoría de Orden Superior de Dennett a partir de consideraciones biológicas y neurofisiológicas

Tal vez una de las razones más influyentes por las que la teoría de Dennett ha conseguido alcanzar tanto éxito, es la apelación a las nuevas y más recientes teorías neurocientíficas para explicar, por una parte, la evolución de la posibilidad de pensamiento en los seres humanos y, por otra, la simplicidad con la que su estrategia del modelo de los múltiples borradores permite describir la presencia de un nivel interpretacional de la conciencia individual. Pero, como venimos insistiendo, ¿es esto suficiente para garantizar la realidad de una propuesta filosófica?

Desde la aparición de sus escritos sobre conciencia y evolución (especialmente Dennett 1991 y Dennett 1995), Dennett ha recibido una gran cantidad de críticas por el empleo interesado, precisamente, de los recursos que utiliza para desarrollar su propuesta. En esta sección vamos a revisar algunas críticas contundentes a estos usos que, de ser ciertas, podrían poner en peligro la consistencia interna de la teoría dennettiana.

5.3.2.1. “Pero, la evolución no funciona de ese modo”

Como hemos visto a lo largo del apartado 5.1., para Dennett, del mismo modo que los ordenadores han sido diseñados por sus programadores, nosotros hemos sido diseñados por la selección natural (Dennett 1995: 401-427). Esto es, del mismo modo que la intencionalidad de los ordenadores es extrínseca al ser el producto de la intención de los constructores de dichos mecanismos, la intencionalidad humana también es extrínseca porque hemos sido programados por nuestros genes egoístas (Dennett 1990: 59). Por tanto, Dennett acaba por postular una simetría entre la intencionalidad de las máquinas y la nuestra propia, dado que considera a ambas como derivadas.

Tal vez la crítica más contundente que se le pueda hacer a esta visión dennettiana de la intencionalidad es que, si consideramos la intencionalidad humana como derivada, ésta debe ser explicada de manera completa y necesaria por la acción de la selección natural. Sabemos que no bastaría con afirmar la inexistencia de una intencionalidad propia de la selección natural porque el propio Dennett así lo considera. La selección natural actuaría de modo aleatorio, sin ningún tipo de guía ni ningún tipo de teleología intrínseca que permitiría seguir la posibilidad de un plan. Pero, entonces, ¿cómo explicar la aparición de dicha intencionalidad?

Para Dennett, la existencia de ciertos rasgos que permiten postular una intencionalidad en los organismos se debe a una explicación adaptativa. El producto de dicha adaptación es el proceso que, funcionalmente hablando, mejor ha respondido a las exigencias de adaptación. Pero parece ser que nunca podríamos, como Dennett insiste, postular a la misma selección natural como aquello que permitiría explicar el proceso adaptativo, dado que recaeríamos en una petición de principio (Dretske 1990). Sería necesario, entonces, diferenciar entre la propia selección adaptativa y los objetivos de la misma. En este sentido, sería más plausible suponer que, si un organismo posee una cierta característica, entonces puede ser explicada a partir de la existencia de procesos selectivos que han potenciado la presencia de dicha característica, pero

no, como pretende Dennett, suponer que existen procesos selectivos que permiten la aparición de una determinada característica, permitiendo así la adaptación del organismo al entorno.

Pero aún suponiendo que Dennett tuviera razón y que el proceso adaptativo permitiera explicar de este modo la aparición de ciertas características en los organismos, todavía se podría criticar su posición por presuponer una cierta acumulación adaptativa. Parece ser el caso que, contraviniendo toda posibilidad de que los repetidos ciclos de interacción con el ambiente permitan la aparición de determinadas estructuras que se replicarán en las generaciones sucesivas, en realidad, la selección natural actúa de un modo no-acumulativo (Sober 1995), por lo que la explicación de la aparición de determinadas características en los organismos no se debe a su prolongada exposición de determinadas condiciones ambientales sino al resultado de determinadas mutaciones transmitidas a nivel genético a las generaciones futuras, dejando como irrelevantemente explicativa la apelación a la selección natural, dado que la desaparición de todos aquellos individuos que no poseen la característica relevante del organismo no ofrece ninguna explicación causal de su supervivencia (Cummins 1975: 750-751; Sober 1995: 393). Por tanto, podríamos concluir, la estrategia dennettiana se basa sobre una controvertida interpretación de la selección natural, entendida como adaptacionismo.

Pero todavía podríamos considerar, como hace Dennett, que aunque se pueda asimilar su posición con un adaptacionismo, la explicación ofrecida por la acumulación adaptativa no responde a ningún tipo de funcionalidad intrínseca. Supongamos que pueda ser verdad que la selección natural pueda jugar un papel explicativo en la adquisición de ciertas características por parte de determinados organismos. ¿Podríamos asegurar, entonces, que el argumento dennettiano funciona? ¿Cómo podemos explicar la individuación de dichas características?

Parece razonable admitir que existirían tres tipos diferentes de explicación que permitirían rendir cuentas de dicha individuación:

1. la presencia de criterios funcionales,
2. la presencia de criterios morfológicos, y
3. la presencia de criterios nomológicos.

¿Puede la explicación de Dennett dar cuenta de alguno de dichos criterios?

Analicemos un caso que permita ejemplificar los tres diferentes tipos de criterios de individuación: la característica de ser un corazón. Desde criterios funcionales, podríamos decir que

“ser un corazón puede explicarse como aquella entidad que tiene la función normal evolucionada de bombear sangre en el sistema circulatorio de un cuerpo” (Burge 1989: 312).

Desde la posición de Dennett, no podríamos ofrecer una explicación similar porque, si es la selección natural la que permitiría explicar dichas características funcionales (una explicación extrínseca), entonces, decir que es la capacidad de bombeo de sangre (una característica interna) lo que hace que un corazón sea un corazón presupone que se potencia la capacidad de bombeo de sangre desde la misma selección, lo que deja sin explicar la capacidad misma del bombeo y, lo que sería más importante, la característica de ser un corazón, porque no

existe ningún rasgo extrínseco a la capacidad misma de bombear sangre que permita explicar la evolución por selección natural de lo que sea un corazón.

Desde criterios morfológicos, podemos decir que el que una determinada entidad sea un ejemplar de algo se debe a que tenga el tipo de característica que permita indicar qué sea un ejemplar tal (Neander 1991: 180). Es decir, que un corazón sea un corazón vendría determinado por tener un determinado tamaño, un determinado color, estar compuesto de un determinado tejido, etc. Parece que la teoría del propio Dennett imposibilita este tipo de explicación al determinar que las características relevantes no son de este tipo. Pero se podría insistir, todavía, que la característica funcional también depende de este tipo de criterios morfológicos, indicando que un corazón sería aquella entidad que juega un determinado papel causal (y que toda aquella entidad que no juega dicho papel causal será de otro tipo. Cf. Cummins 2002). Pero, como también el propio Dennett podría decir, este tipo de explicación morfológica basada en propiedades funcionales dejaría sin explicación la posibilidad de la presencia de entidades que realmente podrían catalogarse como tales a pesar de tener una malformación, como por ejemplo la posibilidad de que existan corazones que, en un momento determinado, aunque sigan siendo corazones, tengan algún tipo de disfunción.

Desde criterios nomológicos, podríamos decir que el que dos ejemplares de algo pertenezcan al mismo tipo se debe a la existencia en ambos de la codificación de los mismos genes, en un sentido al que Millikan (1984: 23) ha denominado *establecimiento reproductivo de familia*.²⁰ Esto es, los criterios que permiten establecer la individuación de determinadas características dependen de un modo interno de la presencia de replicadores genéticos capaces de transportar información de generación en generación y no dependen, como parece que Dennett pretende al apelar a la selección natural, de criterios externos.

Si Dennett no es capaz de salvar estas tres explicaciones que permiten rendir cuentas con la explicación de la individuación de características, como parece ser el caso, entonces parece que su teoría se tambalea. Apelar a la noción de selección natural parece que no se sigue de las actuales investigaciones en genética evolutiva, por lo que Dennett no es capaz de cumplir sus propios requisitos. La selección natural parece que es un elemento necesario para poder explicar la aparición de la intencionalidad en la mente humana, pero no parece ser suficiente, porque para ello Dennett debería poder indicar el modo en que aparece en la mente dicha característica (Nanay 2006). Si, como parece, Dennett no pudiera explicar esta aparición de modo extrínseco, entonces deberíamos concluir que, aunque la intencionalidad presente en los ordenadores sí es derivada, la intencionalidad presente en la mente humana debe ser intrínseca y perfectamente explicable desde criterios funcionales, morfológicos y nomológicos de

²⁰ Esta noción general engloba bajo el mismo tipo a ejemplares que han evolucionado o se han desarrollado gracias a la codificación de los mismos rasgos genéticos. Ello no presupone que los replicadores que permiten dicha evolución sean siempre genes, ya que varios factores pueden jugar un papel de replicador dependiendo del tipo de familia al que pertenezca el individuo determinado del que se habla. Esta diferenciación entre genes y otros tipos de replicadores permite a Millikan atender a recientes descubrimientos en la herencia genética, dado que parece el caso que otros mecanismos diferentes de aquellos que permiten la copia de las moléculas del ADN podrían explicar la transmisión de información genética de una generación a otra.

individuación de características, tal y como permiten actualmente los diferentes estudios en biología evolutiva.

5.3.2.2. Algunos datos desde las neurociencias

La idea principal que subyace a la teoría de Dennett es que la explicación de la conciencia puede llegar a obtenerse algún día desde la actividad científica de las ciencias experimentales del cerebro. La teoría dennettiana de los borradores múltiples parte del supuesto de la inexistencia de un módulo principal dedicado únicamente a la generación de lo que conocemos como conciencia (como hemos visto en la crítica que Dennett plantea a la posibilidad de la conciencia como un Teatro Cartesiano) y la inexistencia de criterios anatómicos y/o funcionales que permitan determinar la información que llega a alcanzar el nivel consciente, considerando que es una mera cuestión de azar el que alguna de las discriminaciones que el continuo proceso de reinterpretación de datos cognitivos llevadas a cabo por el cerebro llegue a dominar (momentáneamente) la conciencia.

Ahora bien, si ésta es la forma en que el cerebro produce nuestra conciencia, parece claro que podría existir una serie de dificultades derivadas. En primer lugar, la inexistencia de una jerarquía cerebral que permita generar la conciencia, parece que deja en la más absoluta anarquía la posibilidad de que los procesamientos se produzcan, siendo incapaces por sí solos de establecer un orden concreto de interpretación de datos, dificultando la posibilidad (como pretendería Dennett) de afirmar que la conciencia sería un estado adaptativo del cerebro que actuaría en (el cerebro de) los organismos como un eje de control de la conducta, permitiendo las anticipaciones de acontecimientos. Parece, de todos modos, cierto que existen en un momento concreto una gran cantidad de procesos trabajando en paralelo en el cerebro que emplean el mismo contenido, realizando reinterpretaciones continuamente y añadiendo y modificando los datos que componen dicha información. Pero parece que no sería adaptativamente explicable que cualquiera de ellos pudiera llegar a ser consciente en un momento concreto, hasta el punto de llegar a tener la capacidad de controlar explícitamente el comportamiento del organismo. Sería más plausible suponer que dichos procesos de análisis y reinterpretación se organizan de modo jerárquico, permitiendo así que progresivamente se vayan elaborando las interpretaciones que, respecto de la información cognitivamente adquirida, el organismo posee, permitiendo al organismo alcanzar cada vez un mayor nivel de complejidad hasta llegar a convertirse en consciente. De todos modos, aunque este proceso fuera cierto, desde los datos aportados recientemente por las ciencias del cerebro, parece claro que la mayor parte de las representaciones cerebrales no son capaces de controlar la conducta de los organismos porque no consiguen alcanzar la cantidad de información necesaria (es decir, carecen del contenido requerido) para que el organismo sea capaz de responder de manera más o menos adaptativa al entorno en el que habita.

Parece, por tanto, bastante más plausible suponer que existen múltiples estructuras cerebrales que son capaces de representar la información cognitiva que el organismo posee pero a un nivel más elevado de elaboración (esto es, siendo ya conscientes), aunque sea verdad que

existan zonas cerebrales que procesen ese mismo contenido sin que nunca llegue a hacerse consciente, siendo la cantidad de información que son capaces de representar lo que distinguiría unas de otras. Por lo tanto, parece más plausible concluir que sí existen criterios anatómicos para afirmar la existencia a un nivel consciente de información representacional que permiten, a su vez, a éstas (las representaciones) de un nivel funcional explicar de modo causal el control de la conducta del organismo consciente (Ruz, Tudela y Acero 2002: 92-93). Pero, ¿es esto cierto?

Tradicionalmente, dentro de las neurociencias, se diferencia en el cerebro entre regiones primarias y secundarias de procesamiento de la información, distinguiendo tan sólo la modalidad perceptiva de la misma, siendo el procesamiento de las unidades secundarias bastante más complejo que el proceso llevado a cabo por las primarias. Toda esta información se unifica y se combina en unas zonas terciarias denominadas de asociación, situadas en la corteza cerebral, habitualmente considerada como procesadora multimodal de la información, es decir, combinatoria de elementos informacionales procedentes de modalidades perceptivas diferentes. Dicha conjunción multimodal parece que sería necesaria para el control de la conducta, según las restricciones indicadas por el entorno del organismo, producido por los objetivos (adaptativos) del individuo (Kupfermann 1991). Se suele tener por cierta esta secuencia jerárquica entre unidades primarias, secundarias y terciarias para explicar, por ejemplo, los diferentes estadios de procesamiento de la información a partir del cual el sistema perceptivo de los organismos es capaz de crear preceptos (Cf. Marr 1982), pero también es verdad que actualmente se considera que las interacciones neuronales no son tan simples como se pensaba, dado que (al menos en el caso de la visión) parecen existir múltiples conexiones neuronales de reentrada de la información que reenvían la información ya procesada, o con un nivel de procesamiento superior, a unidades de procesamiento más básicas o, al menos, que dichas unidades superiores de procesamiento consiguen influir en las áreas cerebrales de procesamiento primario (Cf. Edelman 1989 y Churchland, Ramachandran and Sejnowski 1994).

Si fuera cierto que existen este tipo de reentradas a un nivel neuronal, entonces también parece cierto que existen zonas del cerebro que duplican las discriminaciones interpretativas que Dennett defiende, contraviniendo así las leyes del ahorro cognitivo. Suele apelarse, sin embargo, a una redistribución del trabajo que permite, por un lado, negar esa duplicación indeseada (e indeseable) y, por otro, permite explicar cómo existe un proceso de continua re-elaboración de la información cognitiva de modo tal que cada vez se adecue más a las representaciones necesarias y requeridas para la producción de una conducta coherente con las restricciones del entorno. Pero, como Dennett bien afirma, entonces surge el problema de tener que dilucidar cuáles de dichas representaciones son capaces de llegar al nivel consciente y, en la medida en que esto es posible, cuál sería el criterio que permitiría discriminar positivamente unas (las que pueden llegar a ser conscientes) de las otras (las que no llegarán a ser conscientes). La respuesta obvia parece ser apelar al diferente nivel de elaboración del que ambos tipos de representaciones disponen. Tan sólo las representaciones más elaboradas estarían en *disposición* de llegar a alcanzar el nivel consciente y de ser capaces de controlar la conducta del organismo, dado que la información elaborada por dichas representaciones permite conferir un determinado grado de

funcionalidad a la información anatómicamente representada (en tanto que parece ser que ciertas regiones cerebrales tienen restringido tal potencial de representación) capaz de emerger a la conciencia, permitiendo su papel causal sobre el comportamiento del organismo.

Pero es probable que este tipo de explicación no sea válida. Primero, porque no existe un rasgo anatómico que permita conferir mayor nivel de elaboración a ciertas representaciones en detrimento de otras (Dennett y Kinsbourne 1991: 238) y, segundo, porque de lo que aquí se trata no es de determinar el nivel disposicional que permite la acción, sino de establecer el papel causal que cierta información, cuando llega al nivel de la conciencia, es capaz de tener en la conducta de los organismos.

Uno de los principales argumentos que encuentra Dennett para negar la posibilidad de la existencia de zonas cerebrales determinadas capaces de procesar estados representacionales conscientes es la aparición evolutivamente reciente de la conciencia a un nivel filogenético. Sólo si la conciencia fuera evolutivamente lejana, podría darse un nivel adaptativo tal que existieran zonas específicas del cerebro que (a modo de módulos) permitieran dicha representación. Pero dada la reciente aparición de la conciencia, parece más plausible suponer que la conciencia se genera en zonas del cerebro inicialmente desarrolladas en la evolución para otros objetivos y que, a través de sucesivas adaptaciones, han llegado a generar conciencia. El problema principal en esta concepción de la aparición de la conciencia de la teoría de Dennett la encontramos cuando atendemos a estudios recientes sobre fisiología del cerebro, ya que el análisis de la conciencia llevado a cabo por Dennett (especialmente el que desarrolla en Dennett 1991) se refiere explícitamente a estados auto-conscientes de nivel superior, dejando sin explicación aquellos niveles carente de auto-conciencia que se dan a un nivel primario.

Por ejemplo, estudios como los de Zeki (1990 y 1992) parecen sostener la afirmación de que existen, al menos para el nivel de la percepción del color, zonas concretas del cerebro que son capaces de representar a nivel consciente los diferentes colores a partir del fenómeno de la constancia del color, lo que permitiría al organismo dotarse de una imagen estable de su entorno. De manera similar, Churchland (2005) defiende la posibilidad de la existencia de vectores neuronales cuya activación codificaría (permitiendo, así, explicar) las diferentes sensaciones de color. Lo que vendrían a presuponer dichos estudios es que tan sólo existen cierto tipo de células neuronales, localizadas en zonas del cerebro concretas, que permitirían determinadas actividades mentales en estado consciente. Es decir, que el sistema cognitivo del organismo no detecta de modo directo la información externa, sino que el cerebro construye una imagen del entorno a partir de múltiples procesos que permiten al organismo adaptarse a su medio. Las operaciones y procesamientos llevados a cabo por los diferentes grupos neuronales que pueblan las diferentes regiones cerebrales nunca llegan (ni podrían llegar a hacerlo) a emerger a nivel consciente; tan sólo la imagen global y coherente que es representación del entorno (también del interno) es la que llega a alcanzar la conciencia, sirviendo así de base para el control del comportamiento del organismo. La posibilidad de la existencia de una única representación a nivel consciente permitiría, por un lado, explicar la compleja aunque coherente interpretación que de la realidad realizan los organismos y, por otro, reducir al mínimo la

posibilidad de realizar acciones indeseadas, permitiendo además que (unidas a la memoria) dichas conductas puedan ser revisadas, reevaluadas y juzgadas para realizar acciones posteriores (Ramachandran and Hirstein 1997). Además, la presencia de una única representación permitiría la explicación del *espacio de trabajo global* estipulado por el propio Dennett para evitar la posibilidad de errores en la acción, pero no como un espacio virtual. Porque parece existir cierto sustrato físico en la corteza prefrontal dorsolateral y en la corteza cingulada anterior en el que podemos encontrar procesadores modulares específicos que permitirían la activación y desactivación de procesamientos de información concretos capaces de implementar la conducta pertinente en cada caso (Dehane and Naccache 2001), permitiendo incluso la integración informacional necesaria para explicar de manera cognitivamente reductiva fenómenos como el rellenado de información en el punto ciego (Ramachandran and Hirstein 1997).

Visto desde este punto de vista, parece necesaria la presencia de ciertas estructuras cerebrales cuya actividad pueda ser consciente a diferencia de otras áreas del cerebro cuyo procesamiento, aunque necesario también, no llega al nivel de la conciencia. Dichas áreas conscientes permiten representar de manera estable una imagen del entorno del organismo a partir del procesamiento de la gran cantidad desorganizada de información que los sistemas cognitivos captan.

5.3.2.3. Revisando la metáfora dennettiana

De la anterior sección se desprende una gran cantidad de datos que, a nivel neuronal, permiten articular una imagen diferente de la conciencia como funcionalmente dependiente de áreas localizadas y especializadas del cerebro que, a través del procesamiento de la información cognitiva, elevan a un nivel personal una determinada representación del mundo. Pero, ¿invalida este hecho la propuesta dennettiana? Veamos.

Para Dennett, son factores aprendidos de modo cultural (la mayoría de ellos establecidos lingüísticamente de manera directa o indirecta) los que nos permiten controlar nuestro comportamiento en un grado tal que seamos competentes y responsables de nuestras acciones. Tal vez, el error de Dennett sea el haber prestado una indebida atención a las estructuras nerviosas que permiten dicho control de la conducta, en parte debido a la extrema complejidad de su organización (Ruz, Tudela y Acero 2002: 106). Por tanto, parece más plausible concluir que no existe la anarquía cerebral postulada por Dennett. Esto es, por supuesto la máquina joyceana (como la hemos definido en la nota 105) que Dennett afirma como medio de control se instala en nuestros cerebros de manera cultural y lingüística, pero la materialización de dicha estructura se realiza en zonas específicas y especializadas del mismo. Nuestra conducta, por tanto, sería el resultado de dicha organización cerebral, debida tanto a factores biológicos como culturalmente dependientes, permitiendo que las zonas cerebrales sean moldeadas con suma facilidad (en gran parte debido a la suma plasticidad del cerebro en el nacimiento) a partir de la gran cantidad de información recibida de modo cognitivo a través de los diferentes sistemas perceptivos, permitiendo la especialización de diferentes regiones con la

capacidad suficiente y necesaria para generar representaciones de nivel superior con el suficiente grado de abstracción (de ahí que muchos consideren estas representaciones débiles como *abstracta*)²¹ como para que sean capaces de controlar el comportamiento del organismo en vistas a cumplir sus objetivos internamente establecidos.

Desde la propuesta de Dennett, una vez atendidas las críticas, estas expectativas y objetivos que permiten la acción del organismo deben tener un papel funcionalmente relevante en la activación de comportamientos, pero no, al menos, de un modo estrecho. Es decir, en un momento determinado, dada una misma situación debe poder ser el caso que el organismo actúe de modo diferente; que dado el mismo objetivo, el sujeto pueda actuar de manera diferente cada vez. Si, como hemos dicho, consideramos la presencia de objetivos y expectativas como *abstracta* (como elementos categoriales con contenido intensionalmente dependiente), entonces, no nos queda más remedio que devolver la posición de Dennett dentro de las vías del realismo al establecer la existencia de modelos racionales de acción dentro de las codificaciones procesadas en tanto respuestas que tienen capacidad de activación ante representaciones complejas del entorno, permitiendo así a los organismos establecer una serie de capacidades intencionales como objetivos de acción diseñadas por la selección natural, lo que nos permitiría volver a hablar de la presencia de una intencionalidad derivada en los organismos biológicamente evolucionados.

5.4. Addenda: la acción moral como paradigma en Dennett

Sentimos que nuestras acciones están sujetas a nuestra voluntad en la mayoría de las ocasiones, e imaginamos sentir que la voluntad misma no está sujeta a nada; porque cuando al negarla, somos llevados a intentarlo, sentimos que se mueve fácilmente por todas partes, y produce una imagen de sí misma incluso en ese lado donde no se estableció (Hume 1739: 408).

Al comienzo de este trabajo, en el prólogo, realizamos una contextualización histórica de nuestro tema. Intentábamos, sobre todo, situar el tema de la conciencia en un marco de desarrollo histórico que permitiera ver el tránsito de la anterior asimilación de la conciencia con la pre-existencia de un alma que posibilitaba, entre otras cosas, la presencia en los sujetos de una serie de categorizaciones que le ofrecían una imagen del mundo, donde se incluía el mundo interior de la propia mente, hasta el punto de permitirle confeccionar una imagen de su propia personalidad individual. Asociado directamente a esta problemática clásica encontrábamos también el problema relativo a la libertad. Dentro de este tema, la principal preocupación era ver en qué sentido podíamos decir que las acciones particulares de los hombres concretos tienen sentido desde su propio libre albedrío, incluidas dentro de un mundo determinado por una normatividad influida por la Divinidad. A este respecto, son multitud los autores que, de un

²¹ La diferenciación entre objetos abstractos y objetos concretos es antigua, pero la discusión en términos actuales es relativamente reciente. Podríamos considerar que un objeto abstracto (en el sentido en que Dennett lo emplea, derivado del usado ya por Reichenbach en sus escritos sobre Filosofía de la Matemática) es 'un punto matemático... definible en términos de fuerzas físicas y otras propiedades no-materiales' (Dennett 1991b: 96) mientras que un objeto concreto es algo con realidad física. Para una ampliación de estas definiciones, véase Rosen (2001).

modo u otro, han intentado salvar la libertad humana de dicho determinismo a partir de la apelación a la bondad divina. Pero también son muchos los autores que, en contra, postulan la existencia de una determinación en la que no tiene cabida la posibilidad de que un sujeto pueda ser libre, dada la restricción que impone el mundo entorno. Estos argumentos suelen apelar a la pre-determinación del propio sujeto para ser de un modo concreto, un modo que no puede ser modificado de ninguna de las maneras, incluso impidiendo la posibilidad de que las acciones dependan del propio individuo, recayendo la responsabilidad de sus propios actos en la determinación existente. Teniendo esto presente, vemos que el problema de casar la presunta determinación del mundo externo con la ilusión que los sujetos conscientes tenemos de ser libres (al menos, la ilusión de poder decidir por nosotros mismos cómo ser y cómo actuar) es un tema recurrente dentro de la historia del pensamiento.

Uno de los argumentos escépticos más fuertes, por lo que respecta a la cuestión de la atribución de responsabilidad moral, que se ha formulado en el campo de la actual filosofía de la acción, sería el Argumento Básico de Galen Strawson (Cf. Strawson 1986 y 1994). La exigencia strawsoniana de una responsabilidad por ser lo que se es, por los rasgos constitutivos del agente, para poder dar cuenta de la responsabilidad moral por las propias acciones, convierte tal atribución en algo, aparentemente, injustificable. Nuestra intención en este punto será rebatir el argumento strawsoniano, o al menos ponerlo en entredicho y, para ello, daremos primero unos pasos preliminares, que nos permitirán, como conclusión, afirmar la posibilidad de tal responsabilidad.

En un primer momento, pues, mediante una breve ojeada a la historia de la filosofía trataremos de establecer cuál puede ser el origen de la noción de individuo que está en la base del Argumento Básico, y que es la que dota de toda su fuerza a la argumentación escéptica strawsoniana. En segunda instancia, estableceremos ese mismo Argumento, que se pretende válido independientemente de compromisos con distinciones entre determinismo e indeterminismo, para, en tercer lugar, exponer sucintamente los intentos de respuesta que desde posiciones compatibilistas e incompatibilistas se han venido, o podrían venirse, formulando, haciendo especial hincapié en la propuesta de Daniel C. Dennett, que nos permitirá, aun criticando la del propio Dennett, proponer una concepción de sujeto que dé cuenta de todas las limitaciones anteriores y, en tanto que histórica y socialmente configurada, polémica con el esquema de la propuesta strawsoniana.

5.4.1. La noción de individuo en el Argumento Básico de G. Strawson

La imagen que del agente se desprende de la tradición dominante en el campo de la filosofía de la acción actual, sobre el que se vierten las cuestiones relativas a la libertad (entendida como voluntad libre) y a la responsabilidad moral (que tiene a la libertad como su presupuesto), no es la única posible.

La noción de individuo, de su libertad y responsabilidad moral asociadas, que subyace a esta tradición, podría tener sus orígenes más remotos en el pensamiento de un autor escasamente citado y conocido por ella, el pensador medieval Pedro Abelardo.

Éste tendría una tesis fundamental para el temprano *surgimiento del individuo* que caracterizó al siglo XII (una idea de marcada raigambre en el voluntarismo agustiniano): la moralidad de un acto se define sólo por la intención de quien lo realiza, lo cual enlaza bien con la noción teológica cristiana que anima su concepto de voluntariedad.

El pecado como consentimiento interno a la inclinación a realizar una acción mala, implicaría siempre, según se desprende de la doctrina abelardiana, un suceso psicológico específico, una decisión y, en tanto que rechazo consciente de la voluntad divina, sería un acto de plena afirmación de la propia individualidad. Este carácter interno y aislado de la decisión, el *consensus*, implicaría, alejándose de otros modelos como el aristotélico, por un lado, la separación entre el carácter moral o *ethos* y la decisión moral: el *consensus* no es necesitado por ninguna inclinación antecedente. Y, por otro lado, el corte entre la decisión moral y la acción externa. En ese espacio vacío entre el *carácter* y la *acción*, estaría la decisión moral y, por tanto, el individuo.

Fáciles de establecer son los vínculos genealógicos entre la concepción abelardiana del sujeto y el individuo atomizado, reducido a mera estructura decisoria con creencias, deseos y valores (los 'creseos'), desconectados éstos (y vaciados de contenido), en beneficio, en principio, de una apropiada atribución de responsabilidad, de la sociedad y de la historia (entendidos éstos más bien como amenazas condicionantes que como condiciones de posibilidad), en la filosofía de la acción contemporánea.

Así pues, en un mundo en el que, en virtud de los avances científicos, no hay proposición general más ampliamente compartida, con mayor apoyo inductivo y empírico, que la de que todos los sucesos, entendiendo suceso en sentido amplio, tienen explicación, cabe preguntarse por el lugar del agente, instancia de la voluntariedad, de la decisión y de la responsabilidad. Si por explicación se entiende la explicitación de las circunstancias que hicieron necesario el suceso acaecido, y si tales sucesos, en un mundo naturalizado, se insertarían en cadenas o secuencias causales (epistemológicas pero, de modo especial, también ontológicas), donde un estado de cosas causa otro, que a su vez es causado por uno anterior, y así sucesivamente, ¿en qué se diferenciaría el mero suceder, o causalidad transeúnte, del causar la acción del agente, o causalidad inmanente?

Y la respuesta a esta pregunta sería que, precisamente, el agente se diferenciaría el mero suceder en el cumplimiento de las condiciones de la acción y voluntad libres, que darían cuenta de la subsiguiente atribución de responsabilidad, en la estela de la tradición dominante, y que serían esencialmente dos: que esté en poder del agente el realizar o no la acción; que esa 'capacidad' resalte sobre un amplio trasfondo de deseos, de posibilidades no actualizadas, pero actualizables (el Principio de Posibilidades Alternativas) y que el agente constituya el origen de esa acción, deseo, volición, y no sea otra cosa ajena a él la que tenga el 'control' (el Principio de la Originación, de la Auto-configuración o de la Responsabilidad Última).

Sin embargo, más allá del rotulo *concepción heredada del determinismo*, que se desprendería de la caracterización que del mundo naturalizado he hecho, donde las acciones y elecciones humanas se entenderían como efectos de las secuencias causales en que, más

ampliamente, se integran todos los sucesos del mundo (Honderich 2002), el problema tal como lo plantean los sucesores del determinismo fuerte clásico sería el de una cuestión de incompatibilidad generalizada. La libre voluntad implicaría verdadera responsabilidad y tan incompatible sería con el determinismo como con el indeterminismo, siendo por tanto inviable su defensa.

Así pues, según se explicita en uno de los argumentos más relevantes en este sentido, el Argumento Básico de Galen Strawson (1986 y 1994), no seríamos responsables moralmente de nuestras acciones, sea el determinismo verdadero o falso, dado que nadie podría cumplir el requisito de Responsabilidad Última (que postulamos como una de las dos posibles condiciones de la responsabilidad moral). Las acciones libres, que impliquen verdadera responsabilidad lo serían en virtud de una razón del agente que las causa; esto es, en función de cómo es uno mentalmente. Pero no querríamos, ésta es la baza de Strawson, que esas cadenas causales que llevan de nuestros deseos, creencias y valores a nuestras voliciones y acciones, fueran nuestras de modo azaroso, sin bases, por suerte (entendida como *lack of control*, no como *chance*, y del tipo de la que podríamos denominar suerte constitutiva), ni en virtud de algún tipo de control externo al agente (Williams 1993).

Por supuesto que cabe, como arguye Dennett, frente a las exigencias de ultimidad o control absoluto, de auto-determinación como auto-creación de corte teológico, apelar a la 'complejidad' del agente, a su ser sensible a razones, a su capacidad para la revisión y alteración del carácter basándose en la reflexión. Pero esto sería quizá tomarse las cosas con demasiada ligereza o pretender recluir en el campo de las meras abstracciones teóricas lo que constituye un problema metafísico que sí afectaría a nuestras conductas e instituciones y que los avances de la genética, neurología, sociología, etc., plantean como cercano y cotidiano.

Sin ultimidad, por un lado, ni el mecanismo sensible a razones mencionado por Dennett (y argumentado por los semicompatibilistas como Fischer and Ravizza 1998) habría sido adquirido por medios controlables por el agente, ni su efectiva aplicación de tales capacidades en circunstancias específicas se ejercería bajo su control. Y, por otro lado, los intentos de alcanzar la ultimidad por vía indeterminista chocarían con el requisito de la racionalidad, de la elección conforme a rasgos propios.

5. 4. 2. Las respuestas al Argumento Básico de Strawson

En cuanto a las posibles respuestas al Argumento Básico, una tradición filosófica que nos concierne, todas las respuestas (compatibilistas, incompatibilistas libertaristas-pragmátistas y revisionistas) resultan insatisfactorias por lo que respecta a la pregunta por el origen, condición de la atribución de responsabilidad de los agentes individuales, y en gran medida por la razón ya apuntada del tipo de caracterización de éstos, que los deja inermes ante planteamientos como el de Strawson.

Compatibilistas clásicos (como Hume o Ayer), o de nuevo cuño, con su compatibilismo refinado (jerárquicos, como Frankfurt; evaluativos o de fuente múltiple, como Watson; de la razón, como Wolf), se caracterizan por un postulado de libertad implicando verdadera

responsabilidad en un mundo determinista y, excepto quizá Wolf (que frente a los otros acentuará el componente normativo), por la cada vez más afinada búsqueda de un yo real o profundo, origen de las cadenas causales, de las razones, que concluyen en acciones de las que pueden responsabilizarse.

El problema, sin embargo, es que, como se encargaría de mostrar Frankfurt (1982), si bien el compatibilismo clásico acertaría a distinguir, según los compatibilistas, en qué consiste la libertad de acción, descuidaría el libre albedrío, cayendo en complicaciones varias y víctima de diversos contraejemplos claros, como la dificultad para evitar atribuir responsabilidad moral a los animales, que al menos en algunas de sus especies más desarrolladas parecen, en tanto que poseedores de deseos y creencias, capaces de desarrollar acciones libres (al modo compatibilista); o, aún más grave, la incapacidad en que se encontraría para explicar casos de experiencia de falta de libertad en virtud de deseos no compulsivos.

La concepción estructural y jerárquica de Frankfurt parecería sortear esas dificultades, atendiendo a una libertad, con la atribución de responsabilidad subsiguiente, entendida como un querer lo que se quiere hacer. El agente frankfurtiano poseería deseos de hacer cosas, o deseos de primer orden, y los deseos que le movieran efectivamente a la acción constituirían su voluntad. Sin embargo, y aquí estaría el grado de complejidad que lo distinguiría de niños y animales, además poseería un segundo nivel de deseos, de segundo orden, o actitudes reflexivas y evaluativas, cuyo objeto serían los deseos de primer orden, constituyendo *su volición de segundo orden*, deseos de que ciertos deseos nos conduzcan a actuar, una subclase de éstos.

La propuesta frankfurtiana, sin embargo, ha sido objeto de importantes críticas, como la que Watson (1987) realiza, en el seno de su propia tradición. Y es que, entre otras observaciones críticas posibles, se puede (remedando el Argumento Básico) cuestionar que la adición de niveles del mismo tipo de estados mentales (deseos, especialmente) logre otra cosa que desplazar jerárquicamente la *wantonidad*.

En ese sentido resulta interesante la propuesta de Wolf (1990), que caracteriza en estos términos críticos las propuestas compatibilistas observadas, denominándolas punto de vista del *yo real*, y propone la suya propia, asimismo compatibilista, *el punto de vista de la razón*. A las argumentaciones vistas opone, pues, Wolf algunas observaciones en la línea apuntada (y que, asimismo, recoge el propio Watson 1987), como la importancia de la historia y del entorno.

El problema es que la perspectiva wolfiana, pese a plantear la libertad y la responsabilidad como funciones de la capacidad para actuar de acuerdo con la razón, esto es, como poder de reconocer lo Verdadero y lo Bueno, recaería en problemas semejantes a los que pretende criticar y superar. Pues si bien se optaría en su perspectiva por una renuncia a la responsabilidad metafísica de la propia constitución del yo, decantándose por una responsabilidad moral, con la capacidad de auto-revisión y corrección como corolarios, de nuevo el control normativo, de lo correcto, aun cuando no se hubiera reducido a una cuestión de complejidad psicológica, se vería aquejado por el problema del origen.

Aparentemente cerrada la vía compatibilista, la opción incompatibilista, en una de sus más sutiles formulaciones, la de Kane (1996), se presenta como alternativa deseable al intento

de responder al reto del Argumento de Strawson. La perspectiva incompatibilista rechaza la posibilidad de defender la libertad y la responsabilidad al tiempo que se afirma el determinismo, tal como lo hemos expuesto, apostando por la indeterminación y, en muchos casos, por una instancia decisoria, desligada de los estados mentales del agente, amenazados por ese determinismo invasor, en tanto que seres en el mundo natural, y dadas las cadenas causales en que se enlazan tales estados.

Sucintamente expuesta, cabe decir de la posición de Kane que, sorteando algunas de las objeciones clásicas que se le hacen al incompatibilismo y consciente de los límites expuestos respecto de la anterior línea argumentativa, propone contemplar la exigencia de ultimidad en su carácter radical, así como la condición finita humana, planteando la exigencia, si es que ha de existir algún tipo de responsabilidad y libertad valiosa, de que el agente realizase algunas acciones libres indeterminadas, en el sentido de una posibilidad categórica, entre alternativas igualmente asequibles dadas unas mismas circunstancias.

Pero si las razones o motivos nos inclinan a elegir A en lugar de B, siendo A el resultado razonable de la deliberación, parecería que la elección de B, en exactamente las mismas circunstancias, si posible sería irracional. Para evitarlo y defender esta posibilidad, que algunas veces debería tener el agente en su mano, esto es, la de poder ir configurándose mediante ese tipo de decisiones y acciones, postula Kane la racionalidad plural, la posibilidad de que existan razones diversas, por ejemplo morales y no morales, que nos obliguen a escoger, dada la historia psicológica del sujeto, una opción antes que otra.

Teniendo en cuenta esta salvedad, lo que Kane propone es que en determinados momentos el agente elige, en virtud de una intención general, no *a priori*, no determinada por rasgos y circunstancias anteriores (lo que de nuevo recuerda la caracterización abelardiana del *consensus*), de un modo u otro, con libertad, reforzando con ello sus rasgos morales u otros.

Sin embargo, y dejando de lado la compleja discusión relativa a las múltiples críticas que el libertarismo de Kane ha recibido, baste hacer hincapié en las dudas que suscita una posición que, en última instancia se entiende como elección sin bases, y trata así de evitar el regreso al infinito del Argumento Básico, que acepta en su formulación como reto. Cabe preguntarse si resulta plausible una autodeterminación hecha con independencia de las propias creencias y razones, que se reduce a constatar en el momento de la elección, que ésa, cualquiera que fuese la decisión tomada, era la opción preferida, porque, de hecho, tautológicamente fue la preferida y realizada.

Así pues, los ya tradicionales argumentos compatibilistas e incompatibilistas, se encuentran en dificultades al tratar de responder al Argumento Básico. Dificultades compartidas, pese a la diferencia de posiciones en el espectro del campo de la acción, que permiten especular sobre la común condición que les aboca al fracaso ante retos como el que plantea Galen Strawson.

Y es que, ciertamente, mantienen todas estas corrientes supuestos comunes que, ya hemos parcialmente adelantado y denunciado, en tanto que no asumidos como tales serían presupuestos o pre-juicios. La concepción puramente estructural del agente, desentendiéndose de

contenidos, la atención y el intento de dar cuenta de la libertad en términos de procedimientos antes que de actitudes, el formalismo, pues, unido a (en general) una desatención radical a la dimensión histórica y social del agente (en todo equiparable así al pecador abelardiano), caracterizan a estas diversas corrientes.

5.4.3. Dennett sobre libertad y responsabilidad

Llegados a este punto, queremos detenernos en los intentos de solución que el filósofo Daniel C. Dennett ofrece de las dificultades apuntadas en las dos anteriores secciones de este apartado. Como hemos tenido ocasión de mostrar, la respuesta adecuada a los problemas de la Auto-configuración Personal y del Principio de Posibilidades Alternativas, los dos pilares sobre los que (parece) se sustenta la posibilidad de atribución de la responsabilidad (moral) de nuestras acciones, pasaría por el rechazo o, al menos, la reformulación de la anterior concepción del sujeto como un agente-mónada marcado teo-teleológicamente y la formulación de una nueva noción de yo en la que se tenga en cuenta tanto la inclusión del individuo dentro del mundo físico como el respeto de su libre voluntad.

Dentro de tal intento, que podemos denominar compatibilista, podemos encuadrar la teoría dennettiana del sujeto. Dicha tentativa consta de dos momentos: por una parte, una vertiente filogenética en la que Dennett intenta incluir al ser humano dentro de la cadena evolutiva; por otra, una vertiente ontogenética en la que ofrece una nueva visión del sujeto como capaz de auto-atribuirse a sí mismo razones para actuar.

Así pues, desde el primer punto de vista, Dennett propone un marco en el que colocar las diversas opciones de diseño: es la llamada torre de la generación y la prueba (como hemos mostrado en la sección 5.1.2. Cf. Dennett 1996: 101ss). La evolución organiza un conjunto de células en diversos modelos según lo adecuado de sus respuestas: una buena respuesta (lo que no significa que sea la única, o la mejor, o la más correcta) es aquella que funciona, y una conducta que funciona se verá potenciada y también se verá reforzada la estructura capaz de ocasionarla. Por ello, en cada nuevo estrato de la torre podemos colocar multitud de clases de organismos que pueden desenvolverse en su entorno, lo que permitirá la expansión de su población (crecimiento horizontal); pero sólo aquellos organismos cuyos movimientos son más eficaces verán potenciada su estructura de tal modo que podrán sobrevivir y evolucionar (crecimiento vertical).

El organismo humano, así, puede modificar su propia estructura fenotípica de modo que le permita crear, identificar y generalizar relaciones no existentes en el mundo externo. La diferencia entre las criaturas humanas y las criaturas no-humanas consistiría pues en el manejo más eficaz que tienen las primeras de pensamientos concretos referentes a objetos del mundo que se constituyen de forma abstracta como representaciones.

Los organismos humanos, así, son privilegiados porque la evolución les ha dotado del poder de construir hipótesis y representaciones generales acerca de la identidad de los objetos e individuos externos; podemos rastrear objetos según su relevancia en el entorno cambiante y percatarnos de nuestras acciones porque todo ello lo asociamos a nuestro esquema intencional.

Pero, si desde el punto de vista filogenético es la aparición del yo, como soporte de agencialidad marcado por las herramientas cognitivas y lingüísticas que él mismo se ha construido, el que permite al individuo evolucionado dotarse de sus propias razones de acción (lo que le permite *auparse* hasta la libertad), desde el punto de vista ontogenético lo que nos interesa es ver el modo en que el lenguaje permite al yo individual ser considerado como un agente al que se le puede atribuir responsabilidad moral.

Siguiendo la argumentación dennettiana, “un yo es, sobre todo, un *locus* de auto-control” (Dennett 1984: 81). Pero por otra parte, también considera Dennett que la responsabilidad implica una gradación, ya que depende de la capacidad evaluativa del individuo el que nuestras actitudes proposicionales sean revisadas, criticadas y modificadas según el modo en que cada uno quiere ser, una revisión que, obviamente, presupone una atención especial al entorno socio-político en el que está inmiscuido el agente.

Alguien podría criticar la postura dennettiana afirmando que tan sólo es cuestión de suerte el que un individuo esté más capacitado que otros para auto-reflexionar sobre sus propias evaluaciones, por lo que depende de la fortuna de nacer con mayor capacidad de racionalidad el llegar a ser considerado como una agencialidad responsable. Pero Dennett no es partícipe de esta idea, antes bien para este autor todos los seres humanos estarían igualmente capacitados para controlar sus propias acciones, por lo que aunque alguien esté mayormente dotado que otros (en cierto sentido), dado el largo proceso que se requiere para la construcción del propio yo, todos ellos tienen (en principio) las mismas posibilidades de llegar a adquirir la responsabilidad (moral) sobre sus acciones.

Del desarrollo de esta concepción de la acción moral, Dennett extrae su concepto de yo y es que, en un momento determinado, la teoría dennettiana se ve en la necesidad de dotar de cierta forma a toda la enorme cantidad de deliberaciones y decisiones que sobre sus propias acciones realiza un determinado organismo. Lejos de suponer un Teatro Cartesiano en el que se representan dichas deliberaciones y sin admitir un único punto en el que se coordinan dichas meta-reflexiones, Dennett postula la existencia de lo que denomina un ‘centro de gravedad narrativo’ donde se dará cita todo aquello que nos puede ser atribuido en tanto agentes de acción.

Sin embargo, una primera crítica surgiría a raíz de esta idea y es que, si el yo se considera como una creación, como una ficción, parecería adolecer de cierta debilidad metafísica y tendría serios problemas para soportar el peso de la responsabilidad moral.

Pero la respuesta de Dennett es que precisamente dicha debilidad es su virtud, ya que aunque el yo sea una ficción, cualquiera estaría orgulloso de haber sido su creador, ya que éste englobaría todo aquello que se ha hecho, se ha creado, se ha inventado e, incluso, se ha fingido, por lo que en cualquier momento podríamos crear uno nuevo, sustituyendo así al anterior, lo que nos permitiría, así, adecuarnos a las nuevas circunstancias.

Pese a todo, la debilidad ontológica seguiría presente al afirmar el mismo Dennett que, en el proceso de constitución del yo, el sujeto no dispone de ningún tipo de entidad directriz, que no existe un Jefe de la Mente que a modo de capataz organice las diferentes narraciones que

van surgiendo sino que, muy al contrario, se da una explosión (inconsciente) de identidades y personalidades (entendidas como narraciones pasadas) que se van desarrollando, configurando y abandonando conforme una de ellas cobra cierto protagonismo y se va imponiendo sobre las demás en tanto es elegida (o simplemente aparece como más atractiva).

A primera vista, Dennett parece que ha sido capaz de incluir dentro de la cadena determinista un cierto tipo de libertad, porque parece que su noción de yo capacita para, en última instancia, elegir nuestra propia personalidad en el sentido en que parece permitir que seamos nosotros los que nos demos nuestras propias razones para ser como somos, en tanto que son nuestras auto-evaluaciones las que permiten incorporar cierto tipo de meta-reflexiones a nuestra configuración personal. Sin embargo, desde el punto de vista del determinismo radical cabría plantearse una duda y es que, si hemos considerado que el lenguaje es producto de una serie causal determinada por la ley de la evolución, entonces parece que el propio lenguaje (que es el requisito básico para llegar a dotarnos de libertad, según la teoría dennettiana) es un mecanismo automático creado por las leyes físicas (en tanto que no puede ser improvisado por ser el vínculo histórico entre una determinada sociedad y su propio pasado) para construirnos a nosotros y no un instrumento que nosotros construimos para que nos permita auto-configurarnos como agentes responsables.

Una respuesta de tipo dennettiano podría afirmar que el lenguaje, en tanto que tradición recibida y heredada pero sometida a revisión y crítica, posibilita una nueva configuración de conceptos y categorías, por lo que en realidad permite tomar distancias de las originales razones y creencias que por medio de él pudieran transmitirse, admitiendo así una nueva cosmovisión que diera lugar a la nueva generación de narraciones no-determinadas. Pero también existe una crítica más débil, de circularidad, que afirma que si tenemos por cierta la tesis de la constitución narrativa del yo, entonces deberíamos presuponer una cierta entidad previa que narre todo el proceso de construcción y que, por tanto, estaría fuera de nuestro control o bien tendríamos que suponer que el yo emerge de una encarnizada lucha a muerte entre *yoes* que van surgiendo según acontece la historia cognitiva del sujeto, sin necesidad de una entidad previa, pero dejando al lenguaje como un elemento que se auto-gestiona. La solución dennettiana, sin embargo, es clara: lo que ocurre en el momento previo a la configuración del yo es lo segundo, pero el lenguaje no es un elemento auto-gestionado, sino que es un elemento democrático que cualquier narración o sujeto tiene a mano para poder auto-constituirse como tal.

Dada esta explicación, más o menos satisfactoria, del carácter ontológico del yo, todavía queda pendiente la acusación de debilidad. ¿Puede el yo, tal y como lo ha tejido Dennett, soportar el peso de la responsabilidad moral? Si, por una parte, la ficcionalidad del yo parece debilitar la base sobre la que podemos referir la atribución de responsabilidad moral y si, por otra, la misma ficcionalidad permite una mayor y más radical auto-evaluación que un yo sustancial y limitado, entonces parece que la problemática de la ontología del yo y de la posibilidad de su auto-configuración dentro de un mundo determinista se mueve entre el establecimiento de un yo menos ficcionalizado pero capaz de soportar el peso de la

responsabilidad moral o la concesión de un yo ficcionalizado que, a cambio de un mayor auto-control, ceda un determinado grado de responsabilidad.

La posición compatibilista de Dennett es capaz de superar la segunda opción, pero debe enfrentarse con la primera, ya que un centro de gravedad narrativo no garantiza el suficiente nivel de auto-gestión del yo necesario para soportar la carga de la responsabilidad moral. Parece que el complemento que requiere la configuración de un agente capaz de originar sus acciones y decisiones a partir de sus propios deseos y creencias es que dicho principio de autonomía descansa sobre un lecho de posibles alternativas de acción todavía no realizadas pero que podrían serlo en el momento de la acción (o elección) del agente, el denominado Principio de Posibilidades Alternativas (PPA).

Para que una acción pueda ser catalogada como libre y voluntaria el agente, parece, debe ser capaz de poder explicarla a partir de las razones que lo movieron a actuar del modo en que lo hizo, razones que actuarían como guías de la acción humana, incluidas dentro del mundo, con poder causal.

Si atendemos a su desarrollo ontogenético, como aquí hemos hecho, la cuestión principal acerca de la responsabilidad moral de un agente es optimista. Porque a pesar de que los deseos y creencias que nos mueven a actuar, según Dennett, son producto de una acumulación e interiorización de normas y convenciones sociales, esto no nos debe hacer desesperar en nuestro intento de actuar libremente dentro de un mundo concebido como determinista. Pero, ¿cómo es posible que en un mundo dominado por la explicación física de los hechos tenga cabida la acción llevada a cabo desde la vida mental de las máquinas (humanas)? Esto es, si concebimos que todo hecho del mundo tiene una explicación física, ¿cómo podemos atribuir poder causal a lo mental?²² Dennett concibe al ser humano como intencional; esto es, explica las actitudes comportamentales de un agente desde sus propios propósitos y fines. Aunque dicha intencionalidad es concebida como derivada (Dennett 1987): lo que los agentes suponen como sus intereses en realidad son fines debidos a la acumulación aleatoria y arbitraria de modificaciones ambientales indicadas por algoritmos que permitieron la configuración de un cierto tipo de orden, una rutina, capaz de crear, desarrollar y hacer evolucionar cierto tipo de organismos que consiguieron elevarse hasta la posibilidad de controlar su propio desarrollo (vista desde el punto de vista filogenético).

Es decir, todo agente es un artefacto, incluidos los seres humanos. La única diferencia existente entre la inteligencia humana y algún otro tipo de inteligencia (artificial) es que esta última ha sido creada por la primera, mientras que la inteligencia humana es producto de nuestros genes egoístas (como vimos en la sección 5.1.1.2. Cf. Dennett 1995). Pero todos responden al mismo principio: su intencionalidad responde a los intereses de quien los diseñó.

²² Aquí estamos haciendo una explícita referencia al principio de exclusión de Kim. Este principio nos viene a decir que, dado que toda causalidad en el mundo es física, todo aquello que no responda a causas físicas queda fuera de esta causalidad. Del mismo modo, toda causa física debe producir efectos físicos, por lo que de una causa física no puede desligarse un efecto que no sea físico. Por dicho motivo, si entendemos lo mental como algo diferente a lo físico, no puede ser causal ni puede ser efecto de causas físicas.

Dennett afirma que la cultura humana es producto de la evolución memética.²³ En la evolución del ser humano como organismo biológico, apareció en un determinado momento una serie de mecanismos capaces de adaptarse al medio según los elementos externos (culturales) presentes, mecanismos necesarios para la supervivencia de la especie. Dichos mecanismos fueron incorporados a la estructura fenotípica de la especie, como elementos que permiten la acción. Así, lo que en un principio parecían elecciones libres de los individuos, en realidad son respuestas de los mecanismos implementadores de la acción, mecanismos que a pesar de ser flexibles y adaptables a los contextos, tienen una explicación física.

Sin embargo, la dificultad añadida surge cuando se pretende la negación del mencionado PPA. Tendemos a pensar que somos libres y que en un momento dado podríamos haber elegido actuar de modo diferente a como realmente actuamos, por lo que cuando no podemos hacer otra cosa consideramos que se ha dado un cierto tipo de determinismo puntual que impide la posibilidad de actuar de otro modo. Dennett niega el PPA porque, por un lado, no existen evidencias que indiquen que una persona podría, o puede, haber hecho otra cosa que lo que realmente hace y porque, por otra, la responsabilidad del agente respecto de la acción no remite aunque éste no pudiera haber actuado de otro modo. Parece que la diferencia entre poder haber hecho otra cosa y no poder haberla hecho es el determinismo, pero no es así para Dennett, para quien la responsabilidad es exactamente la misma si estamos determinados como si no lo estamos, porque en un momento dado nosotros hicimos lo que hicimos guiados por nuestro diseño.

Actuamos como actuamos porque tenemos los contenidos mentales que tenemos, pero ello no supone una reducción de la responsabilidad de nuestra acción porque responde a nuestra propia historia cognitiva, la acción nos pertenece por estar causada por los contenidos cognitivos del agente. Es una ilusión pensar que si estamos determinados no somos responsables de lo que hacemos porque es nuestra estructura la que causa las acciones que realizamos y nuestra estructura depende de nuestra historia cognitiva y del modo en que consideremos las enseñanzas. Para Dennett, el PPA es una ilusión (memética) porque no podemos nunca hacer otra cosa que la que realmente hicimos, pero no por ello dejamos de ser responsables de lo que hacemos. Parece que la solución para la toma de decisiones es estar diseñados para poder actuar del modo en que debemos hacerlo, como hemos mostrado anteriormente.

Pero aquí Dennett parece suponer la existencia de una causa completa para toda acción. Es decir, supone que dada una determinada acción existe un, y sólo un, modo único factible de explicación. Pero, si todo hecho físico, en este caso una acción humana, puede ser explicado a partir de una causa física, entonces cualquier pretensión de causalidad de lo mental quedaría

²³ La teoría de los memes es una teoría análoga a la de los genes que Dawkins (1989) desarrolla para dar cuenta de cierto tipo de ideas que permanecen a través de generaciones, del mismo modo que los genes permiten la transferencia de ciertas características físicas a unos individuos a sus descendientes. Estos memes son descritos como paquetes discretos de información que se transmiten por herencia (semántica, no ya genética) de generación en generación, a modo de ideas preestablecidas.

anulada.²⁴ Cuando negamos la posibilidad de actuar de otro modo distinto a como en realidad actuamos, damos por supuesto que no tenemos otra salida que realizar la acción que hacemos; esto es, que estamos determinados.

Cuando hablamos de acciones correctas no importa en absoluto que no pudiéramos hacer otra cosa, porque damos por supuesto que lo que cabe hacer en dicha situación es lo que haría cualquiera de nosotros en el mismo caso. Pero no ocurre así con acciones que consideramos incorrectas. Cuando por lo que nos preguntamos es por la censura de un cierto acto sí tiene importancia el poder llegar a determinar que un cierto agente no tenía más opciones que realizar el acto que definitivamente hizo. Porque si llegáramos a la conclusión de que el agente tenía otras posibilidades de acción, entonces cabe la posibilidad de que el agente se equivocara al hacer lo que hizo y, por lo tanto, ser un acto censurable. Si, por el contrario, demostramos que no es posible que el agente actuara de modo diferente, entonces demostramos que el agente estaba obligado a actuar como lo hizo y desaparece (o, por lo menos, disminuye) la responsabilidad de sus actos.

Vemos ahora que lo que realmente le ocurre a Dennett es que, en su intento por desligarse de la antigua noción de yo como agente-mónada, se queda todavía a un paso, ya que aunque es consciente de que no existe una finalidad desde el punto de vista filogenético, se queda corto al no conseguir romper definitivamente con la tradición metafísica. Requerimos de una ruptura todavía más radical que la dennettiana para lograr dar con un concepto humano de responsabilidad moral. Al dar demasiada importancia al marco naturalista y evolutivo y descuidar el marco social y cultural, Dennett no tiene en cuenta el modo en que el sujeto se entrena moralmente para formar parte del mundo al que pertenece; esto es, no tiene en cuenta la capacitación para llegar a ser considerados como seres responsables.

Lo que se trataría es de hacer referencia al marco socio-cultural como un proveedor de mecanismos reflexivos y de mecanismos no-reflexivos (al modo en que lo defienden Fischer and Ravizza 1998)²⁵ en vistas a conseguir lo que la sociedad realmente conlleva: la consecución de la mayor cantidad posible de agentes responsables moralmente. Lo que sería deseable es establecer una concepción del yo que tuviera en cuenta ambos marcos (el naturalista y el socio-cultural) en tanto que permitiría que el agente tuviera responsabilidad moral de sus actos, aunque no fuera *causa sui*, dado que los condicionantes sociales e históricos (los estados mentales que escaparían a su control) serían entendidos no como propiedades de un agente esencialmente determinado, sino como condiciones de configuración del individuo y posibilidades abiertas a la deliberación y actuación de un sujeto socio-histórico (lo que dotaría del suficiente autocontrol al sujeto) que posibilita un mayor o menor grado de desarrollo moral. Es decir, sería conveniente instaurar una nueva imagen del sujeto en la que no aparezca tan

²⁴ Como nos decía el Principio de Exclusión de Kim, expuesto en la nota 108.

²⁵ En este volumen, Fischer y Ravizza pretenden desarrollar una visión unificada respecto de la responsabilidad moral de nuestras acciones, teniendo especialmente en cuenta las omisiones y modificaciones que los casos al estilo de Frankfurt, que acaba por afirmar que lo máximo a lo que puede aspirar un agente para ser moralmente responsable de sus acciones es a demostrarse como verdadera guía de su propia acción, a pesar de nuestra existencia en un mundo determinado, una guía que requiere de la inclusión del agente dentro del flujo de configuración como sujeto que permite una determinada sociedad.

separado de los factores sociales, históricos y políticos que lo capacitan y constituyen como agente moral y que le permiten ser lo que es.

5.5. Recapitulación

Al comienzo de este capítulo advertimos la necesidad de atender, por una parte, a la teoría evolucionista que Dennett sostiene sobre la aparición filogenético de la conciencia y de la mente como un paso previo para atender, por otra parte, al desarrollo ontogenético de la conciencia de los sistemas particulares. Expusimos con el mayor detalle posible la interpretación dennettiana de la famosa idea darwiniana de la evolución, en tanto que presupondría la aparición de la conciencia como un suceso biológico desarrollado a partir de la acumulación aleatoria de ventajas adaptativas que se han ido sucediendo a lo largo de generaciones de sistemas de percepción que han culminado en aquello que en la actualidad podemos denominar mente. También atendimos a los diferentes grados de desarrollo que, dependiendo del tipo de atencionalidad que dichas mentes son capaces de obtener, podemos diferenciar a lo largo de dicha historia evolutiva, desde las criaturas que tan sólo son capaces de emplear sistemas simples de detección (las criaturas darwinianas) hasta alcanzar el grado de las criaturas capaces de manipular símbolos semánticos a partir de sus características semánticas y representacionales (las criaturas gregorianas). Al exponer dicha teoría, nos detuvimos y remarcamos su carácter externista, en tanto que los sistemas evolucionados con mayor capacidad consciente son capaces de llegar a tener contenidos mentales a partir de la asimilación de la información que les llega del exterior a partir de sus sistemas de captación sensitiva, una información que es reconvertida en representaciones que pueden ser empleadas por dichos sistemas para confeccionar las mejores estrategias de acción, reservando dichas representaciones en un nivel de practicidad extrema.

Una vez establecidas las pautas del desarrollo filogenético de la conciencia, repasamos la teoría de la conciencia de Dennett. Cuando mostramos los puntos fuertes de su teoría, vimos cómo Dennett está empeñado en negar la existencia de un nivel puramente mental a partir de la negación del Teatro Cartesiano clásicamente establecido. A diferencia de dicho teatro, para Dennett la conciencia no es un escenario donde aparecen de una vez por todas los contenidos experienciales que el sujeto percibe o experimenta sino que, al contrario, no existe tal escenario. En su lugar, encontraríamos una serie de procesadores de la información cognitiva que, mediante su continua revisión, irían confeccionando una historia cada vez más compleja de las relaciones posibles entre el sujeto y el mundo. Dichas historias o narraciones elaboradas por los sistemas de procesamiento de la información enviada por los canales y sistemas sensorio-perceptivos pugnan entre sí para alcanzar la conciencia. Pero ninguno de ellas es capaz de estipularse como vencedora absoluta, dado que no existe ningún nivel de privilegio en tanto que no existe en el cerebro (sede de los procesadores de información) ninguna jerarquía establecida que decida qué debe ser considerado como prioritario, llegando incluso a poder imponerse una información no del todo cierto (o incluso completamente errónea) que se convierta en la relevante para el sujeto.

Precisamente la parte que más preocupación ha despertado en nuestro análisis crítico es aquella en la que hemos expuesto la posibilidad de poder interpretar la teoría dennettiana como una teoría de Orden Superior, en tanto que la existencia de un primer nivel primario que posibilite, tan sólo mediante la manipulación de representaciones (por muy útiles que sean), la aparición de un segundo nivel de estados (e incluso un nivel concreto de acción), parece querer sostener la idea de la existencia de un nivel puramente disposicional que, por lado (como indican las actuales interpretaciones de la teoría evolutiva), anularían por completo la aparición de pensamiento y estados con contenido fenoménico y, por otro (como indican los últimos avances en neurología), complicarían demasiado la labor de los sistemas perceptivos y de los procesadores de información. Hemos matizado, entonces, la teoría dennettiana para que, aunque es cierto que algunas veces podamos identificar el contenido fenoménico con un cierto nivel representacional, sea coherente con la existencia de un nivel no-representacional de captación y actuación previo en tanto que respondería al modo habitual a los estímulos presentes mediante las conductas que todo ser consciente habitualmente realizaría.

Concluimos dicho capítulo con un análisis de las dificultades que la teoría dennettiana sobre atribución de la autoridad de la acción a los sujetos, en tanto que parece que el nivel de agencialidad que el nivel ficcional de representaciones que sustentan su autoconfiguración como agente es demasiado débil como para soportar el peso metafísico que exigiría el soporte de la acción.

CAPÍTULO 6. HACIA UNA TEORÍA NO SÓLO REPRESENTACIONAL DE LA CONCIENCIA

Al tratar de caracterizar la cognición humana como un caso especial de cognición de los primates, el reto es hallar la pequeña diferencia que constituyó la gran diferencia; un pequeño cambio o conjunto de cambios que transformaron el proceso en aspectos fundamentales. En nuestra opinión, el único candidato posible para dicha adaptación o conjunto de adaptaciones es uno que cambiara la cognición de una especie básicamente individual a otra básicamente social y colectiva. La cognición humana es lo que es porque los seres humanos están adaptados para la vida cultural y todo lo que entraña.
(M. Tomasello y J. Call, *Primate cognition*)

Comprender a otra persona es proyectar sus sentimientos en nosotros mismos, producirlos en nosotros por medio de la imitación... Un sentimiento similar aparece en nosotros a consecuencia de una antigua asociación entre el movimiento y la sensación. Volcamos nuestras capacidades en comprender los sentimientos de los demás en un alto estado de perfección y siempre estamos practicando de modo involuntario dichas capacidades en presencia de los demás.
(Friedrich Nietzsche, *La gaya ciencia*)

VI. HACIA UNA TEORÍA NO SÓLO REPRESENTACIONAL DE LA CONCIENCIA

A lo largo de los anteriores capítulos hemos dejado constancia de varias dificultades que una aproximación representacional de la filosofía de la mente tiene a la hora de afrontar la problemática de la conciencia.

El objetivo principal de este capítulo final es desarrollar la idea de que la conciencia, en última medida, es un mecanismo de anticipación que permite al sujeto prever los acontecimientos futuros mediante la percepción directa de los elementos que conforman el mundo y la realidad que lo circunda, evitando en todo momento los inconvenientes anteriormente remarcados en toda imagen representacional de la mente y permitiendo exponer brevemente en qué consiste esta ventaja evolutiva que, en último término, permite la supervivencia y desarrollo de las capacidades cognitivas del individuo. En pocas palabras, la conciencia debería concebirse como un ingenio tecnológico creado evolutivamente del cual pueden beneficiarse en última instancia todos los seres capaces de darse cuenta de ciertas propiedades de los objetos y del mundo externo y que, en mayor o menor medida, son relevantes para su propia supervivencia. Esta idea partirá del supuesto materialista que afirma la identificación de, al menos, algunos elementos mentales con ciertos elementos neurobiológicos. Por dicho motivo, pensamos que es mucho más adecuado mantener una posición fisicalista respecto de la conciencia (en tanto que pretende defender una cierta visión de cómo es la realidad), porque parece mucho más coherente con la actual imagen que la ciencia tiene del mundo y de los hechos que en él se desarrollan. Nuestra posición también pretende ser realista, al defender (la posibilidad, al menos) que el mundo sea tal y como aparece en un primer momento al sujeto.

6.1. Conciencia e intención: el contenido de los estados fenoménicos

Como ha quedado patente a partir del desarrollo de los anteriores capítulos, es habitual presuponer que los estados mentales poseen dos características distintivas. Por un lado, encontraríamos que los estados mentales se caracterizan por ser representacionales, algunos de los cuales se caracterizarían por poseer intencionalidad. En este sentido, los estados mentales intencionales estarían directamente dirigidos a objetos externos que todo sujeto sería capaz de identificar y que, de una u otra forma, están representados (esto es, son contenido) por los estados mentales pertinentes. Por otro lado, ciertos estados mentales tendrían la propiedad de ser fenoménicos, esto es, de tener un carácter fenomenológico caracterizado por aquello que Nagel (1974) identificaría con algo como lo que sería para un sujeto concreto estar (o poseer) dicho estado (esto es, experimentarlo). La mayoría de teorías representacionales, como hemos mostrado en los anteriores capítulos, acaban por identificar las representaciones con las intenciones. Demasiado a menudo, es un supuesto que podemos encontrar en sus argumentaciones. El problema principal que, hemos admitido, tendrían las teorías representacionales de la conciencia sería, precisamente, presuponer que todo estado mental

paradigmático debe poseer ambas propiedades al unísono. Es decir, la mayor dificultad de una explicación representacional de la conciencia se encontraría en el hecho de tener que describir de qué modo todo estado mental consciente tiene, por un lado, un carácter intencional y, por otro, tiene también un carácter fenoménico. Pero no sólo debemos detenernos ahí. También deberemos poder averiguar si prescindir del carácter fenoménico que muestran determinados estados mentales, como ciertas de las posiciones analizadas hacen, es realmente posible a la hora de poder identificar la conciencia.

En casos concretos de determinados estados mentales, pongamos por caso tener una creencia, vinculados directamente con la posesión de un determinado contenido proposicional, la propiedad de la intencionalidad es clara. Pensemos tan sólo en nuestra creencia en que va a llover. Cuando un individuo concreto se forma dicha creencia, de manera directa se relaciona con (la posibilidad concreta de) que llueva. Visto de este modo, la creencia posee unas condiciones semánticas concretas, una serie de características representacionales específicas que permitirán indicar el valor de verdad concreto de nuestra creencia (su ser verdadera o falsa). Pero la posesión de dicha creencia también está directamente relacionada con la posesión (o aparición consciente) de otras muchas creencias e intenciones. Probablemente, si un individuo cree que llueva (o en la posibilidad de que llueva), también tendrá la intención de permanecer seco, por lo que aflorará a su conciencia la creencia (normalmente situada a un nivel potencial) de que cubierto por un paraguas es menos probable que acabe empapado, que debe coger el abrigo largo (y no el corto), que hoy debería ponerse las botas altas (en lugar de los mocasines), etc. En este sentido, es pertinente considerar que las actitudes proposicionales tendrían una serie peculiar de características.

6.1.1. Intención, representación y el carácter fenoménico

Es comúnmente aceptada la idea de la naturaleza relacional de los estados mentales. Es decir, los estados mentales suelen ser considerados como las relaciones existentes entre agentes psicológicamente dotados y sus contenidos. Esta relación es claramente patente si atendemos a actitudes proposicionales como las creencias (podemos explicar la creencia de David de tener un buen trabajo como la relación existente entre David como sujeto mental consciente y el contenido proposicional ‘tener un buen trabajo’), los deseos (podemos explicar el deseo de Jordi de tomar una cerveza como la relación existente entre Jordi como sujeto mental consciente y el contenido proposicional ‘tomar una cerveza’) o las percepciones (podemos explicar que Manolo perciba que un tomate es rojo como la relación existente entre Manolo como sujeto mental consciente y el contenido proposicional ‘que el tomate es rojo’).

Podemos caracterizar los contenidos de dichos estados mentales como elementos marcadamente particulares porque:

- 1) El hecho de que alguien posea un estado mental característico con un determinado contenido tan sólo demuestra la existencia de ese alguien y no indica nada respecto de la existencia de lo que es expresado por dicho contenido (es decir, un individuo puede tener perfectamente una percepción respecto de algo que es irreal, que es ilusorio).

2) El hecho de que alguien posea un determinado estado mental con un contenido característico no nos dice nada respecto de la verdad o falsedad de dicho contenido (es decir, alguien puede poseer una creencia y demostrarse que es falsa).

3) El hecho de que alguien posea un determinado estado mental con un cierto contenido depende por entero de la perspectiva subjetiva del agente (es decir, alguien puede querer tomar una cerveza y no poder hacerlo porque su nevera está vacía, dejando insatisfecho su deseo).

4) El hecho de que alguien posea un determinado estado mental con un contenido determinado ejerce un poder causal sobre la acción del sujeto (es decir, tener un determinado deseo D incita al sujeto a cubrirlo y, si se tiene la creencia de que haciendo A se puede cumplir D, es muy probable que el sujeto acabe haciendo A).

Pero también es cierto que, además de ocupar un lugar dentro de la explicación de la conducta, los contenidos mentales tienen también un carácter marcadamente normativo. Si los contenidos mentales tienen un papel causal respecto de la acción consciente (o inconsciente) deben poseer también un carácter de explicación nomológica general. Pero si, como hemos dicho, los contenidos mentales poseen unas condiciones de verdad que permiten indicar cuándo se satisface o no se satisface un determinado contenido mental según su ajuste a las circunstancias, entonces nos encontramos con el problema de la imposibilidad de establecer leyes generales de explicación de cómo lo mental tiene una relación causal con lo físico.

Es cuando surge esta necesidad de establecer leyes naturales que expliquen esta relación entre estado mental y contenido cuando encontramos el programa de naturalización del contenido, empeñado en demostrar la existencia de una serie de estados físico-corporales que permitirían determinar el estado mental en el que se encuentra el sujeto en cada momento. Para esta concepción, al menos cuando es entendida de modo internista, la presencia de determinados estados mentales es idéntica a la posesión de determinados estados físicos (cerebrales y funcionales). El contenido de dichos estados físicos permitiría determinar de un modo necesario y suficiente el papel causal de dicho estado mental. Porque dicho programa de naturalización afirma que el significado de dichos estados se estipula a partir de una serie de representaciones de ciertas características físicas estrechas, demostrando así la vinculación natural entre la mente y la intencionalidad. El problema con el naturalismo de lo mental lo encontramos cuando atendemos a la posibilidad de tener que rendir cuentas de un estado mental con un contenido erróneo (como en el caso de la creencia falsa, el deseo insatisfecho o la percepción ilusoria): ¿qué ocurre entonces? ¿Cómo explicar esta posibilidad si vinculamos la posesión de un estado mental a la representación de un determinado estado físico?

Habría dos modos diferentes de afrontar este problema. Por un lado, encontramos a aquellos que consideran que todo estado mental es esencialmente diferente de un estado físico y que el elemento relacional de la mente con su contenido depende de propiedades intrínsecas a la mente que se dan de un modo totalmente independiente de las condiciones externas, aunque estas condiciones podrían causarlo. Por otro lado, encontramos a aquellos que consideran que la introspección no puede dar cuenta de las condiciones de verdad del contenido mental dada la

restrictividad de las propiedades intrínsecas, por lo que optan por explicar dicho contenido como contextualmente dependiente de los hechos externos a la mente.

Pero todavía quedaría un ulterior punto a tratar respecto de la naturaleza del contenido mental. Cuando atendemos a actitudes proposicionales, como mi creencia en que va a llover, o mi deseo de beber una cerveza, parece que el contenido de dichos estados mentales puede ser expresado a través de elementos conceptuales (esto es, entrarían dentro de las categorías culturalmente dependientes que vehiculan nuestros significados y que, a menudo, requieren un conocimiento previo de aquello a lo que se refieren). Pero parece que cuando atendemos a ciertos contenidos mentales fenoménicos (esto es, contenidos ya no tan claramente definidos como conceptualmente dependientes sino vinculados directamente con nuestra peculiar distribución subjetiva, con cómo nos aparecen a nosotros las cosas) como mi sensación de dolor o mi orgasmo (y otros tipos de estados conscientes caracterizados por su vivencia o toma de conciencia [*awareness*] y otros casos-límite, como mi ver el color rojo), el modo en que éstos se relacionan con sus respectivos estados mentales parece cambiar con respecto a las actitudes proposicionales, dado que en las segundas la relación con un elemento objetivo parece estar garantizada (aunque sabemos que, a menudo, dicha relación puede estar equivocada, como hemos comentado) mientras que aquel elemento proposicional con el cual se relacionarían los estados mentales fenoménicos parece ser más difícil de precisar o, incluso, como algunos pensadores actualmente sostienen (llegando a ser considerada la posición dominante, como ya hemos comentado a la hora de analizar las teorías de los partidarios de la existencia del contenido no-conceptual en el apartado 1.2. y la sección 2.1.3.), puede llegar a ser totalmente inexistente. De este modo, ni tan siquiera podríamos estar seguros del todo de que existiera una relación proposicional. ¿Cómo debemos, entonces, tratar esa relación?

6.1.2. Los problemas del fisicalismo causal

Intuitivamente hablando, tenemos la impresión de que aquello que hacemos es fruto por entero de nuestra propia y libre voluntad, que nuestras creencias y deseos tienen relevancia causal y que tenemos posibilidades alternativas a aquello que de hecho acabamos haciendo. Pero, al mismo tiempo, convivimos con una imagen científica del mundo que nos dice todo lo contrario, esto es, que aquello que nosotros realizamos depende no de nuestro libre albedrío, sino de las leyes físicas imperantes y del transcurso anterior de los acontecimientos. Este punto de vista dominante hoy día se ha denominado fisicalismo causal, y se basa en dos supuestos principales: por una parte, la tesis de la clausura causal del mundo físico (que todo efecto físico puede explicarse completamente apelando sólo a términos físicos) y, por otra, la idea de la causa total o de la exclusión explicativa (que para todo acontecimiento existe una, y sólo una, causa física que permite explicarlo).¹

Una aplicación concreta de la tesis de la clausura causal del mundo físico sobre los estados mentales podría adquirir la siguiente formulación:

¹ Esta sería una simplicación del Principio de Exclusión de Kim, que ya comentamos en la nota 108.

(1) Los estados mentales son propiedades funcionales que vienen explicadas por su propio contenido.

Esto es, las propiedades funcionales que permitirían definir los estados mentales, sólo pueden ser explicadas a partir del contenido de tales estados mentales (que debemos entender en términos físicos). Así, tendríamos que para cada contenido, siempre existiría una explicación completa en términos de leyes físicas, por lo que en definitiva podríamos reducir el aspecto mental de nuestros deseos y creencias al aspecto puramente físico. Pero, como podríamos nosotros mismos predecir, una crítica posible a este principio resaltaría que no se detiene en el elemento de causalidad, relevante para la explicación de clausura propuesta. Más bien, el principio tal y como está formulado apela a una clausura explicativa, siendo tan sólo aplicable en términos epistemológicos.

Pero lo que nos interesaría resaltar aquí es que existe una posible aplicación de dicho Principio de clausura, en tanto que aplicado a estados mentales, que es causal. Esto es, en una segunda formulación:

(1*) Todo hecho tiene una causa física, por lo que si se acepta la existencia de estados mentales, éstos deben ser efecto de una causa física.

Como hemos comentado en el apartado 5.1., Daniel C. Dennett afronta esta tesis del siguiente modo. No es verdad que podamos dar una explicación funcional de los estados mentales, por la simple razón de que éstos últimos no existen. Pensar que podemos explicar en términos mentales una acción es ilusorio porque confundimos los mecanismos biológicamente evolucionados de los que disponemos con propiedades funcionales determinadas por su contenido. Los mecanismos biológicos tienen por objetivo la supervivencia, por lo que entender que un determinado contenido restringe la funcionalidad de una propiedad es no atender a la influencia que el contexto aporta a la hora de determinar tal funcionalidad de una propiedad.

De este modo, concluíamos con Dennett, no podríamos afirmar que el contenido esté restringiendo la aplicabilidad de la propiedad porque en realidad lo que nos indica la evolución es que la funcionalidad de una propiedad viene determinada por su aplicabilidad en un contexto o en otro (por ejemplo, la funcionalidad del pulmón vendría propiciada por su capacidad de oxigenación del torrente sanguíneo). Visto de este modo, entonces, el eliminativismo ficcionalista de Dennett estaría acorde con la posición fiscalista que este Principio de Clausura Física promulga (tal y como insistimos en la sección 5.1.4.), en tanto que estaría de acuerdo en que, si algo no se ajusta a sus principios de causalidad físicas (en este caso, la existencia de estados mentales fenoménicos), simplemente deberíamos concluir que no existen.

En cuanto a la aplicación de la segunda tesis fiscalista a los estados mentales podríamos decir que:

(2) La sobredeterminación causal de los eventos físicos ha de evitarse a toda costa (esto es, la existencia conjunta de dos o más explicaciones completas e independientes de un mismo hecho).

Es decir, que dado un determinado suceso, existiría una explicación en términos físicos que diera cuenta de dicho evento y, siempre que se repitiera, la explicación podría recomponerse

en los mismos términos. La presencia de una añadida explicación en términos mentales de ese mismo suceso, esto es, apelando a deseos y creencias u otro tipo de estado, sería sobredeterminación, superflua repetición de una explicación ya completa y además básica, primaria, por lo que esta segunda podría, con facilidad, depotenciarse un tanto en favor de la primera.

Dennett intenta dar cuenta de las dificultades de esta tesis y, para ello, desmitifica la existencia de vínculos causales entre lo mental y lo físico. Sin embargo, su eliminación de lo mental otorga un excesivo protagonismo a la determinación biológico-ambiental y este es el punto en el que no podemos sino distanciarnos de él (como veremos un poco más tarde y ya avanzamos en la sección 5.1.4.), porque la eliminación no siempre es recomendable y, además, es altamente contra-intuitiva cuando apelamos a aquellos elementos mentales con un carácter cualitativo.

Así pues, aunque Dennett afirma que es precisamente esa posibilidad de evolución la que dota de plasticidad a las propiedades (con lo que parece que tenemos una posibilidad de explicar de modo múltiple una misma propiedad, según hayan sido sus vías de evolución), nosotros creemos que su teoría acaba (por ello mismo) en cierto tipo de exclusión causal (en el sentido en que ya indicamos que lo entiende Kim), similar a la que afecta al fisicalismo causal cuando presupone la posibilidad de la múltiple realizabilidad de lo mental sobre lo físico y, en última instancia, en una posición que no deja espacio para la explicación causal de la conducta.

De esta manera, si con Dennett, aceptamos la eliminación de todas las propiedades mentales en favor de propiedades físicas, entonces el contenido de las propiedades funcionales se referiría a una sola causa total (física) que provocaría las respuestas relevantes en situaciones similares. Esto es, si aceptáramos la noción técnica de contenido ofrecida por Dennett, a nuestro juicio, y conocidas las posiciones fisicalistas, lo que resultaría es que el ámbito de la acción quedaría restringido a la causalidad unidireccional que está presente en la causa total, obteniendo tan sólo una vía de ejecución para cualquier contexto, negando así la posibilidad de la aparición de nuevas y distintas vías de acción o de nuevas entidades (como los *qualia*) y, en suma, de algo que plausiblemente con nuestras intuiciones primeras pudiésemos denominar conciencia de aquello que realizamos y sentimos (hasta el punto en que parecería totalmente irrelevante si somos o no conscientes de aquello que estamos sintiendo o haciendo).

Una posible salida a estas dificultades, expuesta por diversos autores en los últimos tiempos, pasaría por considerar que un mismo contenido (mental) puede llevarse a cabo dependiendo de las posibles variaciones del contexto externo, con lo que podríamos eliminar gran parte de las mencionadas restricciones físicas tan problemáticas para la posición de Dennett. Incluso el mismo Dennett intentará realizar una defensa similar de su posición cuando introduce el ejemplo del dispositivo two-bitser de detección de monedas de 25 centavos de dólar instalada en una máquina expendedora de latas de refresco (como analizamos en la sección 5.1.1.4.). Pero, como hemos indicado y como se desprende de lo anterior, los estados mentales no podrían identificarse de modo independiente en tanto propiedades físicas, y de este modo la tesis de la intrinsicidad funcional se tambalearía ya que podríamos diseñar contextos en lo que

una misma implementación causara dos conductas totalmente diferentes (como vimos en la sección 5.1.1.4., cuando la máquina expendedora era transportada a Panamá) sin incurrir tampoco en sobredeterminación, al emborronarse la causa total.

Sin embargo, tampoco esta propuesta esta libre de objeciones, y es que entendemos que es cuestionable que un solo estado mental sea capaz de determinar contextos de realización, ya que, como podemos argüir apelando a las tesis del holismo de lo mental, los estados mentales suelen darse acompañados de ciertos estados que pueden ser caracterizados como intencionales (en tanto que aceptaríamos la tesis de la interconexión entre la intencionalidad y lo fenoménico. Cf. Pautz 2008). Por esta razón, nada parece asegurarnos que dicha disparidad de estados no acabara dando lugar a un contexto en el que, lejos de favorecer la múltiple realizabilidad de las propiedades mentales, se determinara nuevamente una sola vía conductual al determinar demasiado estrictamente las pautas funcionales concretas de los estados involucrados, dando al traste con la postulación de la relevancia causal independiente de lo mental sobre lo físico.

Para poder hacer frente a esta dificultad, creemos que es necesario apelar a una noción de *qualia* en la que no exista una separación tan tajante entre lo mental y lo físico, y en la que se tenga en cuenta el papel de la evolución a la hora de elaborar un trasfondo biológico que permita explicar el modo en que nuestro conocimiento crece a partir de nuestras acciones y reacciones naturales a la presencia de este tipo de estado con carácter cualitativo. Acciones y reacciones éstas que tenemos en tanto que somos los animales que somos, y con las que se relacionan nuestras habilidades y capacidades cognitivas (y nuestros conceptos asociados de modo empírico a ellas), de tal modo que se torna insostenible la dicotomía tradicional entre lo natural y lo cultural, habilitándose, al mismo tiempo, un lugar para la conciencia en un mundo físico aceptable para las tesis naturalistas del mundo moderno.

6.1.3. ¿Tienen eficacia causal los estados mentales con carácter fenoménico?

Como hemos comentando, parece existir una diferencia crucial entre los estados mentales con carácter fenoménico y el resto de estados mentales. Es decir, parece posible insistir en que no está del todo claro en qué sentido podemos afirmar la existencia de un poder causal en los estados mentales fenoménicos. Veamos. Al igual que las actitudes proposicionales, parece acertado presuponer que existe una relación entre los seres que poseen una cierta realidad psicológica con carácter fenoménico y los contenidos de dichos estados. La diferencia estriba en que no parece darse esta relación en un sentido proposicional. Es decir, ante la existencia de una experiencia de dolor, parece que no podemos decir que exista una relación proposicional entre el sujeto que tiene dolor y el contenido 'Tener dolor aquí ahora', o ante el estado de ánimo de abatimiento parece que no podemos decir que exista una relación entre el sujeto consciente y el contenido proposicional 'Estoy abatido ahora', o ante un orgasmo parece que no podemos decir que exista una relación entre el sujeto del placer y el contenido proposicional 'Estoy sumamente excitado aquí y ahora'. Parece que los estados fenoménicos de dolor, abatimiento y de orgasmo pueden darse en un sujeto sin que, necesariamente, exista un contenido proposicional que permita indicar al sujeto aquello que está teniendo, por un lado, y sin que, lo que sería más

importante, éste sea consciente de que se encuentra en dichos estados, por otro. Visto de este modo, los estados mentales con carácter fenoménico tienen una naturaleza diferente de aquellos estados relacionales cuyo contenido puede ser considerado como proposicional.

Pero todavía más problemático parece explicar en qué sentido dichos estados mentales con carácter fenoménico pueden demostrarse como hechos físicos. Si, consideramos, no somos capaces de poder identificar, al menos, la supuesta eficacia causal de los estados mentales con causas físicas (esto es, si no conseguimos identificar las propiedades mentales con propiedades físicas), entonces deberíamos estar dispuestos a perder la supuesta eficacia causal de nuestros estados mentales (Liz 1995). Pero, si tenemos presente casos concretos de estados mentales con (aparentemente) un determinado poder causal, podemos llegar a determinar una posible respuesta (Cf. Liz 2000: 133-163). Veamos.

A menudo, decimos que determinadas acciones han sido causalmente llevadas a cabo a partir de la existencia de determinados estados mentales. Pensemos, por ejemplo, en un sujeto que, tras sentir un fuerte pinchazo en una de sus ruedas, se lleva la mano a la mejilla y emite un grito de sufrimiento. Parece ser que el estado mental que le indica dolor es la causa de la acción del sujeto. Pero más bien, podríamos decir, es la situación física de dolor la que, de un modo *reflejo*, ocasiona la acción, obligando al sujeto a llevarse la mano a la mejilla y a emitir un grito (digamos, ésta es la manera habitual en que todo ser consciente de un estado de dolor actúa *por instinto*). Vista de este modo, la inicial suposición (que la acción está causalmente ocasionada por la presencia de un estado mental) puede ser desechada en tanto ha quedado constancia de la existencia de una serie de efectos físicos que, realmente, dependían de causas físicas (la existencia de dolor físico) y no de estados mentales (de la experiencia concreta de dolor).

Pero en determinados casos, no parece tan fácil determinar qué tipo de propiedades físicas definirían ciertas propiedades mentales. Pensemos por ejemplo en nuestro estado de nerviosismo ante una primera cita con esa persona que tanto nos gusta. ¿A qué puede deberse dicho nerviosismo? ¿Con qué estado físico podemos identificarlo? ¿Somos conscientes de dicho nerviosismo, siquiera? Es más, ¿podríamos asegurar que dicho estado de nerviosismo actúa de manera causal sobre algunas de nuestras acciones y estados? Atendamos un poco a dicha situación. Miguel Ángel es un tanto tímido. Por fin ha conseguido que aquella chica que tanto le gusta acceda a ir con él al cine. Minutos antes de la hora convenida, Miguel Ángel aparece en el lugar determinado. Comienzan a sudarle de manera imperceptible las manos, su pulso se acelera, sus labios comienzan a temblar (y otra serie de cosas que podríamos atribuir a su estado de ansiedad) según se va acercando el momento de la cita. Cuando aparece la persona citada, el estado de nerviosismo se acentúa hasta tal grado que hace que el saludo pertinente sea un tanto brusco (casi no consigue darle un beso en la mejilla porque su nerviosismo le ha impedido predecir el lado hacia el cual debía girar su cabeza, haciendo que por poco sus cabezas no llegaran a chocar), su voz entrecortada emite un susurro casi imperceptible, etc. ¿Cómo explicar que las sucesivas acciones de Miguel Ángel estén tan sumamente condicionadas por la presencia de su estado de nerviosismo y ansiedad? Debemos identificar dicho estado de nerviosismo con un determinado estado físico (o con un conjunto de propiedades físicas) que

permitan su aparición en la vida mental de Miguel Ángel. Lamentablemente, no disponemos actualmente de dicha identificación. Pero debemos ser pacientes y confiar en que en un futuro la investigación empírica pueda contribuir en la clarificación conceptual relativa a nuestros propios estados mentales y sea capaz de proporcionar una explicación determinada a esta misma cuestión (García-Carpintero 1998: 181). Mientras tanto tendremos que conformarnos con presuponer una posible respuesta.

Como en el anterior caso del dolor dejaba constancia, para poder explicar el poder causal de un estado mental debemos presuponer la existencia de determinadas propiedades físicas con poder causal que puedan ser identificadas con los estados mentales capaces de causar determinadas conductas. En el caso de la respuesta refleja al dolor, la apelación a causas físicas era bastante fácil: de manera biológica, nuestro cuerpo está programado para reaccionar prácticamente por su cuenta para procurar nuestra propia supervivencia (piénsese en casos concretos como nuestro salir huyendo ante casos de peligro inmediato o nuestro acto de retirar la mano cuando notamos que nuestro vaso de leche está excesivamente caliente). Aquí tenemos un claro caso de eliminación de lo mental en beneficio de la explicación física. Pero no parece posible poder realizar la misma eliminación en el caso de la existencia del estado de nerviosismo presente en nuestro amigo Miguel Ángel. Pero ello no quiere decir que no podamos identificar dicho estado de nerviosismo con una serie de elementos físicos que sean los causantes de su conducta. Ilustremos este punto, de nuevo, con otro ejemplo (rescatado de la sección 4.2.3.). Hemos estado tan absortos en nuestra lectura de la tarde que no nos hemos dado cuenta de que ha anochecido. Decidimos volver a casa, pero antes tenemos que pasar por el parking a recoger el coche. Nos encontramos en el parking. Está oscuro. La poca iluminación del aparcamiento no es suficiente para alumbrar toda la zona. Además, casualmente, la farola que se encuentra justo encima de mi coche está estropeada. Oímos un ruido extraño. Al girarnos, observamos una extraña silueta, amenazante, que se aproxima. Nos sentimos estremecidos por una angustiada sensación (una mezcla de miedo y ansiedad) que nos obliga a acelerar nuestro paso, sacar presurosamente las llaves de nuestro bolsillo y a meternos rápidamente en el coche. Una vez dentro, a salvo, observamos que aquello de lo que habíamos sentido tanto pánico, en realidad, era un compañero de trabajo que también sale tarde del despacho. El ruido y la silueta provocaron nuestro miedo, aunque nuestro miedo fuese infundado. En ese momento determinado, a pesar de que la experiencia no parecía contener nada fuera de lugar, nosotros sentimos miedo porque en ese momento era el modo particular en el que éramos conscientes de la situación, más allá de su contenido real. Pero, como hemos comentado anteriormente, para que exista un poder causal, el estado mental concreto de miedo debe responder a exigencias materiales concretas (aunque todavía no dispongamos de dicha explicación en términos físicos). La presencia concreta de dicha sensación (en tanto identificable con determinadas propiedades físicas) permite explicar toda la cadena causal de acciones consecutivas que, de una manera deliberada, cumplían la exigencia instintiva de supervivencia (aunque no de modo eliminativo, la presencia del estado de miedo permite explicar de modo reductivo la acción). En otras ocasiones, como hemos comentado en la

sección anterior, la presencia de determinados estados mentales permite una explicación alternativa (a partir de la tesis de la múltiple realizabilidad) de la realidad física, en tanto que proporciona nuevas explicaciones causales que, aunque dependiente de causas físicas, permiten ofrecer una imagen no-jerarquizada del mundo capaz de aportar nuevos hechos físicos, nuevas propiedades físicas, identificables con estados mentales múltiplemente realizables.²

Pero, entonces, podemos decir, ¿qué ocurre con todos aquellos estados mentales que no parecen tener un poder causal? ¿Cómo podemos explicar su presencia en la mente? Como hemos mostrado anteriormente, no todo estado mental parece tener una directa relación con nuestro comportamiento. Pero no por ello debemos decir que pasan desapercibidos para nosotros. Pensemos en la inmensa cantidad de sensaciones que tenemos a lo largo del día. Muchas de ellas se tienen por un simple instante. Otras perduran durante algún tiempo. En todo momento, nos inundan multitud de sensaciones, todas al unísono. Pensemos tan sólo cómo cambia nuestro mundo cuando despertamos. Durante el sueño, nuestra experiencia se basa tan sólo en aquello que soñamos. Pero cuando despertamos, parece que recobramos la conciencia (no quiero decir que cuando estamos dormidos no seamos conscientes; más bien podríamos decir que nuestra conciencia entra en un letargo). Cuando despertamos, sentimos la calidez de nuestras camas, arropados por las sábanas. Notamos también el confort de haber descansado, la comodidad de nuestra almohada, etc. Cuando abrimos los ojos, nos asaltan multitud de nuevas sensaciones. Experimentamos el mundo externo. Observamos los objetos de nuestro dormitorio, repleto de formas, tamaños y colores. Experimentamos el calor y la luminosidad de los rayos del sol entrando por la ventana y golpeando nuestra piel. Escuchamos el ligero trinar de los pájaros. Y un larguísimo etcétera. Todos estos estados mentales con carácter fenoménico no parecen tener una eficacia causal muy elevada, pero son sumamente importantes en nuestras vidas cotidianas principalmente porque, hagamos lo que hagamos, siempre están presentes en nosotros de alguna o de otra manera.

Llegados a este punto nos encontramos con una dificultad. Si la conciencia no parece ser eficazmente causal en todo momento, entonces podríamos decir que la conciencia (al menos en estos casos que hemos estado comentando) podría ser considerada como un epifenómeno. Pero nos equivocariamos. Éste no parecería ser el caso. Es cierto que los elementos de conciencia no siempre están implicados directamente en la creación de una acción. Tenemos muchas sensaciones diferentes que, de ningún modo, causan nuestra conducta o tienen un papel a la hora de explicar la aparición de nuevos estados mentales. Pensemos en el siguiente ejemplo.

² Al aceptar la tesis de la posibilidad de la existencia de nuevas propiedades físicas y la posibilidad de la aparición de cierta novedad física que permita dar cuenta de la realidad, aceptamos la tesis defendida por Liz (2007) sobre la posibilidad de la existencia de un fisicalismo explicativo que no caiga en los problemas de la exclusión causal y de la clausura física. Lo realmente crucial de esta posición es que en el mundo físico continuamente surgen nuevos poderes causales físicos que ninguna teoría física supuestamente completa es capaz de tipificar. Dichos poderes causales son físicos pero tan sólo describibles mediante lenguajes no-fisicalistas (como un lenguaje intencional como el lenguaje ordinario, repleto de referencias cualitativas y/o demostrativas) o lenguajes técnicos impuros como el lenguaje de la ingeniería. De todos modos, no debemos olvidar que aquello que se defina como físico debe responder a los criterios identificados en la nota 119, algo que permitiría identificar todo objeto del mundo como constituido por propiedades (o entidades) físicas primarias y permitiría atribuir su comportamiento a los poderes causales de las propiedades físicas primarias que lo componen (aceptando, así, la noción de superveniencia causal identificada por Kim 1984).

Entramos en una cafetería con la intención de tomar un café. Elegimos de la carta un café keniaata, descrito como intenso, de aroma penetrante y sabor ligeramente afrutado. Cuando nos disponemos a beberlo, podemos sentir cómo el aroma del café invade nuestro sentido del olfato. Es un aroma intenso. Cuando mis papilas gustativas entran en contacto con el café, estalla en mi paladar un concierto indescriptible de sensaciones. El café es delicioso, con un sabor fuerte, con un toque que podríamos describir como cítrico. La débil amargura del café queda disimulada tras la sensación de frescura que le proporciona esta acidez, pero en su justa medida. Raramente podríamos decir que todas estas sensaciones actúan causalmente sobre alguna de mis acciones o estados posteriores, pero nunca podríamos decir que nos ha dejado indiferentes. Por supuesto, el saborear el café seguramente ha modificado mis rasgos faciales. Habrá hecho que entornara los ojos, que esbozara una sonrisa de satisfacción, etc. Pero como hemos dicho, estos hechos deberían considerarse más propiamente como consecuencias directamente relacionadas con nuestra distribución fenotípica, con el modo instintivo y habitual en cómo reaccionamos ante determinados estímulos, por ser tan sólo físicos los mecanismos que los producen. Lo que no podemos negar es que el sabor y el aroma del café estaban presentes en mí. Aparecieron justo en el momento en que sentí que el café invadía mis sentidos. La explicación de por qué puedo sentir el aroma y apreciar el sabor de un café se debe a la presencia de glándulas en mi nariz que permiten el olfato y a papilas gustativas que permiten captar el gusto que, en contacto con el café (que transporta determinadas sustancias químicas), permite despertar dichas sensaciones en el momento en que el cerebro recibe la orden de segregar determinadas sustancias encargadas de informarnos de lo que está ocurriendo. Podríamos decir que éstas han inundado por un momento mi vida consciente. Han dominado mi estado consciente haciendo que, por un simple instante, me olvidara de todo y me centrara en disfrutar del café. Posiblemente las sensaciones que me han proporcionado la degustación del café me ha reconfortado. Incluso es posible que me recordaran algún momento importante de mi vida en la que, de algún modo, estuviera presente un café (a modo de la magdalena de Proust). Visto así, aunque no podamos afirmar la eficacia causal de dichos estímulos, la mera presencia en la mente tampoco permite considerarla como un epifenómeno. Las sensaciones mencionadas han existido, por supuesto. Aparecieron de la nada y, posiblemente, desaparezcan pronto. (Realmente tenemos multitud de sensaciones similares a lo largo del día). No han tenido un papel relevante a la hora de explicar mis conductas posteriores o la aparición de otros estados mentales (incluso, podemos decir, que posiblemente no tienen un papel causal concreto). Pero ello no quiere decir que no tengan ninguno. Estas vivencias permiten que recuerde hechos anteriores, que me forme creencias respecto del mundo o, simplemente, como en el caso del café, *colorear* por un momento toda mi vida mental, permitiendo que (aunque tan sólo fuese por un instante) mi mundo se transformara. En definitiva, la presencia de elementos de conciencia como los mencionados parece tener un pobre papel causal. Pero su presencia es sumamente importante para nosotros. (Pensemos por un momento tan sólo cómo variaría nuestro mundo si dejásemos de sentir y ser conscientes de todas aquellas cosas de las que podemos llegar a serlo). En última instancia, el carácter

cualitativo de estos estados mentales es lo que colorea nuestro mundo y hace que merezca la pena vivir la vida.

6.2. ¿Por qué deberíamos abandonar la teoría de la representación en favor de una teoría intencional de la conciencia?

La idea central de este apartado es ofrecer una imagen más o menos completa de una teoría alternativa de la conciencia a aquellas posiciones que hemos analizado críticamente a lo largo del presente trabajo. Partiremos de la consideración crucial de que tan sólo una perspectiva fisicalista permite dar cuenta satisfactoriamente de una explicación coherente (con nuestro mundo mental) y consistente (con los avances neuro-científicos) de la conciencia. Ello nos permitirá ofrecer una respuesta acertada a los que hemos identificado como problemas fáciles de la conciencia (duplicados físicos, argumento del conocimiento, espectro invertido, etc.) que nos posibilitará las herramientas requeridas para aportar una solución materialista al problema fuerte de la conciencia: a la posibilidad de explicar en términos físicos la conciencia sin necesidad de apelar al temido hiato explicativo derivado del viejo problema mente-cuerpo.

6.2.1. Los problemas fáciles, revisados

Como ya hemos dejado constancia al principio de este trabajo (véase apartado 1.4.5.), lo que parecería encontrarse detrás del argumento respecto de la existencia de los ‘zombies’ es la intuición de la psicología popular que afirma que la conciencia humana es irreducible a lo físico, porque siempre nos faltará algo a la hora de saber qué experiencia consciente está teniendo alguien; esto es, siempre nos faltará el *realmente estar* en ese mismo estado consciente (vertiente ontológica) y ello porque siempre parece que falta algo cuando observamos a alguien que parece estar en un estado consciente, siempre nos faltará el *realmente saber* en qué estado consciente se encuentra (vertiente epistémica). Esto supondría un refinamiento de la antigua teoría cartesiana de la conciencia que afirma que sólo podemos saber qué se siente al encontrarse en un determinado estado si realmente se tiene y que tan sólo podemos tener un acceso a la experiencia consciente desde la autoridad que nos proporciona la primera persona (una teoría que afirmaría la existencia de determinados estados mentales como directamente dependiente de la existencia de determinadas propiedades internas a los sujetos). Esto nos obligaría a diferenciar dos vertientes dentro del internismo dualista neo-cartesiano.

a) Versión metafísica del dualismo: *What is it like to be...* El experimento mental de Nagel (1974) ofrece la explicación más adecuada de aquello que defienden los internistas desde el punto de vista ontológico. Por mucho que nosotros pretendamos saber *qué se sentiría al percibir como lo hace un murciélago*, nunca podremos llegar a conocerlo porque jamás podremos llegar a sentir del mismo modo en el que siente un murciélago sus propios estados conscientes, y no sólo porque carecemos de la distribución filogenética necesaria, sino porque siempre quedaría fuera de nuestro alcance las componentes cualitativas y subjetivas que caracterizarían ese *qué se sentiría* al estar en el estado consciente de un murciélago.

b) Versión epistemológica del dualismo: el argumento del conocimiento. Imaginemos a Mary, una muchacha que ha crecido en una habitación carente de toda impresión cromática excepto el negro, el blanco y el espectro de grises que va de uno al otro (tal y como fue descrita en el apartado 1.4.2). Con los años, Mary ha llegado a saber *todo* aquello que desde el plano físico puede saberse sobre el color sin haber tenido nunca una experiencia cromática más allá del mencionado espectro en blanco y negro. ¿Qué ocurriría si un día se le permite observar directamente una manzana roja? Según Jackson (1986), lo que ocurriría sería que, a pesar de conocer todo aquello que desde el plano físico se requiere sobre la experiencia del color, Mary conocería algo nuevo, aquello que se siente realmente cuando uno tiene una experiencia de color. En pocas palabras, que lo que distinguiría a un estado consciente de otro que no lo es sería que el individuo que lo padece realmente sabe qué se siente al tener dicho estado consciente (Jackson 1982).

Como ha quedado patente, los desvelos dualistas pretenden salvaguardar la conciencia de cualquier interpretación puramente fisicalista, especialmente en su versión funcionalista, porque consideran que, aunque es posible tener un acceso externo (físico-funcional) a los estados conscientes de los individuos, la experiencia de la conciencia fenoménica no puede reducirse tan sólo a un estado funcional, sino que debe existir algo más no-funcional con lo que pueda identificarse. La dificultad, de todos modos, estaría también en la posibilidad, propuesta por muchos fisicalistas, de pretender identificar dichos estados con ciertas propiedades representacionales (como venimos mostrando a lo largo de todo este trabajo). Consideramos posible ofrecer una respuesta adecuada que, a la vez que respeta el realismo sobre los aspectos cualitativos de la experiencia consciente, permite aportar una (posible) solución de corte materialista a estas incógnitas respecto de la conciencia.

6.2.1.1. Ni aún viviendo en un mundo *zombie*

Cuando afrontamos el problema de la conciencia desde una perspectiva materialista, lo primero que nos llama la atención de la corazonada ‘zombie’ es su engañosa simplicidad, en sus dos dimensiones. Desde la vertiente metafísica, nos dice que para poder tener realmente un estado consciente no vale con poder actuar *como si* se tuvieran estados conscientes, sino que debemos tenerlos (como hace, por ejemplo, Searle, como vimos en la sección 3.2.2.3.). Desde la vertiente epistemológica, afirma que para saber en qué consiste estar en un determinado estado consciente, no nos vale con conocer qué debe darse en el plano físico y funcional, sino que también debemos tener o haber tenido dicha experiencia consciente. Pero lo que aquí realmente queda patente es su condición de argumento falaz, dado que el dualista parece emplear dos sentidos diferentes de lo que sea saber qué es estar en un estado mental. El primero de ellos presupondría que para conocer un estado consciente necesitaríamos conocerlo bajo un modo determinado de darse (sería un sentido intensional). El segundo, un sentido extensional, postularía que para conocer qué es un estado consciente, sólo requerimos conocer alguno de los posibles modos en los que puede presentarse, y que no sería muy relevante el modo en que

realmente acabara por darse. La premisa de los argumentos dualistas (esto es, la existencia posible de ser físicamente idénticos a seres conscientes que carecerían de estados fenoménicos) está formulada en el sentido intencional (es decir, para filósofos como Chalmers, que defienden la existencia conceptual de los *zombies*, sólo tendríamos acceso a un estado consciente si conocemos de modo estrecho aquellos mecanismos capaces de implementarlo), mientras que la conclusión (esto es, que dicha existencia de *zombies* permite inferir la existencia de algo intrínseco en dichos estados fenoménicos que los hace inaprensibles por una ciencia física) supondría el paso ilegítimo a su versión extensional (esto es, se contraviene la anterior afirmación al considerar que, en el caso de los *zombies*, podrían encontrarse en un determinado estado mental aunque carecieran del mecanismo necesario para determinar en qué tipo de estado mental se encuentran, contradiciendo su propia noción estrecha de activación). Tal vez esto se deba a que es el único modo que encuentra el dualismo para poder defender su existencia. De todos modos, lo importante para su conclusión sería lograr convencernos de cómo podrían darse ambos sentidos de modo independiente sin presuponerse el uno al otro (Balog 1999 y 2000).³ Pero, pueden replicar los dualistas, como nos muestran los análisis modales sobre contingencia y necesidad planteados por Kripke (véase la sección 4.3.5.2.), parece que este tipo de argumento materialista sería superficial, porque concebiría como contingente algo que es necesario. A saber, tenemos la aparente impresión de que el sentido intencional es algo que puede darse o no, cuando realmente es necesario que se dé, independientemente del mundo concebible en el que suceda, y una prueba de ello sería que siempre es posible imaginar la existencia de algo que es coherente.

La estrategia materialista pasaría por diferenciar el qué del cómo para así poder identificar las experiencias conscientes con estados físico-neurobiológicos, algo que consideraría necesario y no contingente. Pero la diferencia estibaría no en ese sentido, sino en el modo diferente de acceso. Porque el modo en que podemos acceder a los segundos es indirecto, mientras que el acceso a las primeras es siempre directo, por lo que no habría buenas razones para considerar como real la ilusión de la contingencia en el caso de los estados mentales. La razón sería que en el caso de los estados mentales no es posible encontrar un argumento que permita demostrar la existencia de su reducción física, por lo que se concluye que las experiencias (en tanto fenoménicas) no son estados físicos, como demuestra la posibilidad de que pudiera darse un estado mental sin una implementación física o un estado

³ Aunque los argumentos esgrimidos en estos textos por Balog pretenden defender al fisicalismo de los ataques llevados a cabo por Chalmers y Jackson, muy similares por tanto a otras defensas de corte materialista como, por ejemplo, el llevado a cabo por Stoljar (2005), tenemos que realizar una pequeña matización. Balog es una ferviente defensora de la estrategia fenoménica conceptual (Cf. Balog forthcoming). La mencionada estrategia identifica la existencia de determinados conceptos fenoménicos como conceptos particulares de tipo sensitivo que contendrían la caracterización propia de una determinada experiencia perceptiva (es decir, Balog defiende la existencia de contenido conceptual en los estados mentales con carácter fenoménico). Por el contrario, argumentos como los defendidos por Stoljar parten del supuesto contrario. Consideran implausible la existencia de dichos conceptos fenoménicos, en tanto que una explicación puramente funcional, dependiente del contraste con ciertos elementos neuro-biológicos, es suficiente para explicar la existencia de elementos con carácter (presuntamente) fenomenológico. Aquí hemos ofrecido tan sólo una caracterización mínima de esta polémica.

físico sin una experiencia consciente, como demostraría el caso de los ‘zombies’ (Chalmers 1996: 194-198).

Vista la superficial infructuosidad del primer argumento materialista, debemos afrontar el problema desde una perspectiva diferente de la que aporta la impresión de contingencia y que, a la vez, permita una explicación de los estados conscientes en su doble dimensión, en tanto que estados físico-funcionales y en tanto que experiencias conscientes.

Una posible solución, planteada actualmente en relación al argumento del conocimiento, pasaría por considerar que lo que sucedería con Mary es que ésta no posee elementos experienciales que le permitan conocer qué se siente al percibir en color. Todo sería nuevo para ella. Pero eso no sería del todo cierto. Mary ya ha tenido experiencias cromáticas previas (aunque sólo fueran en blanco y negro), por lo que ya ha experimentado qué se siente al tener experiencias cromáticas (aunque sólo sea desde el punto de vista conceptual que le permite conocer el rol funcional de los *qualia* relativos al color) y además conoce todo aquello que desde el punto de vista físico-funcional puede conocerse acerca del color. Por lo tanto, en el momento en que Mary sintiera una sensación similar a las que anteriormente ha tenido cuando experimentaba situaciones cromáticamente relevantes, tan sólo debería acudir a los datos empíricamente extraíbles para saber qué experiencia cromática concreta está teniendo en dicho momento. Es decir, lo que Mary tan sólo debe hacer es aplicar la capacidad de discriminación, reconocimiento e imaginación de sensaciones que sus conocimientos respecto de los aspectos físico-funcionales de los *qualia* de color presentan ante cualquier experiencia cromática real a la que sea expuesta, dado que dispone de la suficiente *familiaridad* (en un sentido russelliano, en tanto que todo conocimiento experiencial consciente presupone un conocimiento básico en términos de relación conceptual lingüísticamente establecida por considerar que siempre hemos estado en contacto con experiencias de este tipo) para ser capaz de representarse adecuadamente toda experiencia consciente de color, aunque también entraría aquí en juego la posibilidad de que se representara erróneamente otra sensación diferente (una posibilidad a-problemática desde este punto de vista, dado que la posibilidad de que Mary también pueda estar equivocada no depende de la representación misma, sino de la capacidad de Mary de realizarla).⁴ Visto desde este punto de vista, no es posible que un ser pueda actuar de tal modo que parezca discriminar,

⁴ Una visión de este argumento es la ofrecida por Lewis (1990). Aunque Lewis aquí tan sólo pretende decir, basándose en ideas de Nemirow (1990), que el análisis reduccionista deja fuera ‘algo más’ que parecería tener la experiencia consciente, ese elemento de presunta subjetividad que reivindican los internistas y los dualistas, pero que dicho elemento no sería más que una habilidad cognitiva de identificación perfectamente explicable desde el punto de vista físico-funcional. Una interpretación diferente es la que ofrece Loar (1997), decantándose por una interpretación internista. Para él, ese elemento representacional dependería de elementos funcionales, pero no podríamos reducirlo a ese aspecto relacional, dado que es esencial a la experiencia consciente el mantener ese acceso desde la primera persona, ese acceso directo, como contrapuesto al indirecto en tercera persona característico de las ciencias. En mediación de estos argumentos, aparece el de García-Carpintero (2000 y 2001), que apuesta por una interpretación relacional de la capacidad discriminativa en términos de historia conceptual que puede entenderse desde un punto de vista funcional, pero cuyo contenido depende enteramente del contexto en el cual venga implicado el sujeto que desarrolle su historia cognitiva. La diferencia crucial con los demás argumentos es que el suyo no acepta esa mediación de familiaridad por presuponer un conocimiento básico ya conceptualizado (porque, aunque mínima, la familiaridad presupone una relación previa con representaciones), sino que García-Carpintero se refiere a elementos pre-lingüísticos, en un sentido funcional, dado que deberíamos atender aquí no a un nivel personal, sino a un nivel sub-personal.

reconocer o imaginar una determinada sensación consciente, porque ello supondría ya la presencia de una representación y, por tanto, la existencia de una experiencia consciente, con contenido intencional.⁵

Un problema que podemos encontrar en esta clase de argumentos es el que pone de manifiesto Block (1995): es demasiado pedir el establecimiento *a priori* de elementos relacionales a nivel sub-personal que permitan la identificación de las experiencias conscientes, dado que primero necesitaríamos tener dichas experiencias conscientes para que, posteriormente, se dieran en un nivel funcional. Pero eso es algo que tan sólo podríamos descubrir *a posteriori*, dado que no es el caso que los *qualia* sean constitutivamente relacionales. Por supuesto, hablar aquí de elementos relacionales presupone identificar las experiencias conscientes con propiedades a las que tendríamos acceso en tercera persona. Pero esto no sería un problema si aceptamos que dicha identificación es posible establecerla tan sólo de modo empírico, ya que perderían su presunto estatuto constitutivamente relacional al tener que ser explicados en el seno de una total teoría física. Así, en el caso de la corazonada ‘zombie’, si los elementos experienciales de nivel personal de la conciencia pueden establecerse empíricamente como idénticos a ciertos estados cerebro-funcionales de nivel sub-personal, entonces parece no haber problema en eliminar (como dice Dennett, se trataría de ‘quinearlos’) los primeros en favor de los segundos (Dennett 1988), por lo que en realidad, todos seríamos ‘zombies’ carentes de verdaderos elementos experienciales dado que éstos no existirían. Sólo existirían los estados físicos capaces de causar los *qualia*, todo sería reducible a ellos (Polger 2000). Los *qualia* particulares que cada cual tendría, aunque sumamente importantes a la hora de ‘colorear’ el completo mundo mental de cada individuo concreto, no tendrían una actividad causal específica. Pero sí tendrían un importante papel a la hora de establecer, como estamos diciendo, una correcta relación entre el sujeto concreto de percepciones y el mundo tal y como le aparece. Con esta afirmación conseguimos dos objetivos básicos. Por un lado, conseguimos dar cuenta de la posibilidad materialista de salvaguardar la existencia relacional de una serie concreta de elementos vivenciales presentes en todo sujeto consciente (elementos que le permiten ver el mundo desde una determinada perspectiva) y, por otro lado, evitamos la acusación de epifenomenalismo al admitir un papel crucial de los *qualia* sobre el ‘coloreado’ mental de los sujetos.

Por muy descorazonadora que parezca esta última respuesta, no debemos ser alarmistas. Argumentos fiscalistas como el que se está pretendiendo defender parecen querer reducir todo a elementos funcionales y físicos desde un punto de vista constitutivo de carácter biológico (en el

⁵ Pero no ocurriría lo mismo con otros problemas relativos al mundo de la percepción, como el espectro invertido o la variación estándar. Por ejemplo, vista la posibilidad de que la capacidad de representación pueda fallar de modo individualizado (en tanto que permite crearse creencias falsas respecto del mundo), es perfectamente posible desde este punto de vista que dos individuos perfectamente idénticos molécula por molécula tengan la misma capacidad representativa pero que, sin embargo, tan sólo uno de ellos tenga una percepción correcta, mientras que el otro ha fallado en su representación (como hemos visto en el apartado 4.3.2.). La solución a problemas como el espectro invertido pasaría no tanto por considerar un error a la hora de representar una determinada experiencia de color sino, como veremos un poco más adelante, en diferenciar los distintos papeles que las experiencias concretas de color de dos individuos (supuestamente) cromáticamente invertidos tienen en sus respectivas producciones de percepciones de color (tal y como indicaría, por ejemplo, Broackes 2007).

sentido evolutivo adecuado), pero tan sólo a un nivel sub-doxástico. Pero esta eliminación de las experiencias conscientes no supondría un problema, al menos en una teoría de este tipo, porque tendríamos de un argumento capaz de salvar la impresión intuitivamente aceptada de que existen los *qualia*: la actitud intencional. Desde este punto de vista, enteramente en un nivel personal, todos actuaríamos *como si* tuviésemos estados conscientes, aunque realmente tan sólo responderíamos a nuestra distribución geno-filo-típica. Incluso más, desde este punto de vista, todo artefacto que parezca actuar de modo intencional, puede ser tratado como consciente porque, en definitiva, ¡todos somos artefactos construidos! La diferencia es que los ordenadores y demás artilugios son construidos por los humanos, mientras que los humanos somos artefactos moldeados por la evolución de nuestra especie.

Al comienzo de esta sección hemos advertido que la corazonada ‘zombie’ era un argumento más lógico que empíricamente posible que pretendía servir a ciertos intentos internistas de explicación de la conciencia para salvaguardar un aspecto presuntamente esencial de la experiencia consciente: el acceso privilegiado en primera persona y la imposibilidad aparente de sentir cualitativamente la misma sensación que los otros cuando están en un determinado estado consciente permitían afirmar un elemento irreductible a aspectos físicos en la experiencia consciente que la caracterizan.

Los argumentos materialistas esbozados parecen negar dicho elemento, por dos motivos principales. En primer lugar, parece que la presunta verdad intuitiva esbozada por Nagel sólo es una ilusión proporcionada por nuestra psicología popular porque, como muestran los estudios de Akins (1993) acerca de la fisiología de los murciélagos, podemos tener un acceso en tercera persona que permita explicar el funcionamiento de dichas experiencias conscientes como estados funcionales y físicos, siempre suponiendo un método descriptivo capaz de dar cuenta de los elementos aparentemente subjetivos de la conciencia.⁶ El problema ahora consistiría en cómo explicar la intuición que todos tenemos de que realmente existen experiencias conscientes, porque los intentos de reducción suelen presuponer un elemento causalmente cerrado que hace peligrar otros elementos relativos a la acción humana y los intentos no-reductivos suelen presuponer una excesiva separación entre aquello mental y aquello físico.

En relación al problema de la existencia de los ‘zombies’, parece ser que estudios científicos como los relacionados con la visión ciega permiten negar incluso su posibilidad lógica, porque siempre existiría a un nivel sub-personal un elemento que permitiría afirmar la presencia de una experiencia consciente, a pesar de que a nivel personal el individuo no sea consciente de dicha experiencia.⁷ Por tanto, parece coherente decir que una explicación

⁶ Este método descriptivo sería la heterofenomenología como fue descrita en Dennett (1991), como una mezcla de descripción personal de aquello que cuenta como experiencia consciente y una reconstrucción de la historia cognitiva de cada uno, pero teniendo en cuenta que es el contexto el que, en última instancia, cuenta y que es un acceso en tercera persona el único que podemos tener sobre la propia historia cognitiva (véase sección 5.2.2.2.).

⁷ Una crítica que podría hacerse a esta afirmación es que, en un momento determinado se está sosteniendo la posibilidad de la existencia de elementos en la conciencia que no son accesibles a un nivel personal, por lo que abrimos las puertas a la presencia de elementos inconscientes. Una solución de este tipo aparece en los escritos de John Searle, que considera la existencia de elementos inconscientes que, a pesar de ello, tienen un papel causal en nuestra conducta, como por ejemplo en los casos en los que conducimos mecánicamente como guiados por un piloto

adecuada de la conciencia debe, primero, poder responder a los problemas planteados por la intuición sensible y, segundo, también debe poder rendir cuentas de los datos proporcionados por los avances en las ciencias específicas.

En el plano epistemológico del argumento del conocimiento parece que afirmar la existencia de elementos pre-lingüísticos no-conceptuales capaces de explicar el conocimiento de qué sea estar en un determinado estado consciente es una ventaja porque permite establecer una teoría que contenga la descripción de una cierta arquitectura funcional capaz de dar cuenta de la presencia de los *qualia* antes de toda experiencia. Pero el problema aquí estibaría en el modo en que podemos entender esa construcción pre-conceptual. Una posible respuesta pasaría por tener en cuenta que los individuos tenemos una capacidad previa de acceso a la realidad, en definitiva que la identificación es previa a toda descripción, en tanto que disponemos de ciertas herramientas cognitivas que nos aportan información. Pero también poseemos otro tipo de capacidades conceptualmente participativas, por lo que la misión aquí reside ahora en cómo explicar la presencia de dicho contenido conceptual y no-conceptual previo a toda experiencia consciente.

6.2.1.2. El fisicalismo es verdadero aunque Mary aprenda algo nuevo

Pensemos por un momento en el caso concreto de Mary, nuestra neuro-científica confinada desde su nacimiento a un encarcelamiento cromático. Mary conoce todo aquello que puede llegar a conocerse acerca de la física, la química y la neurofisiología (es decir, *todo* aquello cognoscible desde un plano puramente físico) sobre la percepción del color. El problema principal, como nos indica Jackson (1982 y 1986), es que Mary ha tenido que aprender desde su encierro en una habitación en blanco y negro. La idea principal del argumento del conocimiento, tal como lo plantea Jackson, es que si Mary es liberada de su confinamiento mono-cromático y se le presenta una experiencia de color (digamos, por ejemplo, se le proporciona la percepción de un tomate maduro, algo que podría experimentar como una percepción de rojo), entonces Mary deberá tener una experiencia totalmente diferente de las que, hasta el momento, se le han presentado (deberá entenderse como una experiencia totalmente nueva), por lo que su conocimiento crecerá. En otras palabras, Mary aprenderá algo nuevo. A saber, Mary aprenderá cómo es experimentar una percepción de rojo, algo que hasta el momento no sabía. Porque aunque fuera la mayor experta a nivel mundial sobre los aspectos físico-químico-neurológicos acerca de la percepción del color, nunca había experimentado

automático (Searle 1990: 633), como ya hemos mostrado en el capítulo 2. En este sentido, podríamos responder apelando a la clásica diferenciación entre conciencia por acceso y conciencia fenoménica establecida por Block (1995). Así, tendríamos a un nivel sub-personal un correlato neuronal y nervioso capaz de proporcionar una respuesta semi-automática a cuestiones habituales y/o límite (como por ejemplo, la conducción automática o el llevarse la mano a la mejilla y emitir un grito ante un dolor de muelas) dependientes de nuestra propia *vivencia* (de nuestra propia conciencia fenoménica), mientras que la posibilidad de tener acceso epistémico a dicho contenido fenoménico requeriría de elementos-extra como la atención y/o la memoria, como cuando por un momento nos olvidamos de lo tristes que estamos ante la pérdida reciente de un ser querido al recibir la enhorabuena por un trabajo bien hecho). Cf. Block (1998) para una ampliación de estas nociones. Por otro lado tendríamos posiciones como la defendida por Hatzomoyisis (2007), que defiende que la existencia de experiencias conscientes no requiere del conocimiento de las causas de la misma.

ninguna experiencia real de color ‘en sus propias carnes’. Si esto fuese así, nos indica Jackson, entonces deberemos presuponer la falsedad del fisicalismo. Resumamos el argumento de Jackson.

La Mary de Jackson (el Argumento del Conocimiento):

a) Mary, a pesar de haber sido confinada a una vida mono-cromática, ha llegado a saber todo aquello que desde el punto de vista físico puede conocerse sobre el color.

b) Cuando a Mary se le presenta una experiencia de rojo por primera vez, aprende algo nuevo: cómo es experimentar una sensación de rojo (desde el punto de vista cualitativo subjetivo).

c) Si Mary aprende algo nuevo cuando experimenta el rojo por primera vez, entonces cómo es experimentar una sensación de rojo no puede ser considerado un hecho físico.

d) Por tanto, el fisicalismo es falso.

Pero, como hemos comentado en las secciones anteriores, parece posible considerar que el fisicalismo puede ser consistente con la idea de que Mary aprenda algo nuevo cuando experimenta por primera vez la percepción de rojo ante, por ejemplo, la presencia de un tomate maduro. Simplemente tenemos que considerar la posibilidad de que Mary dispone de ciertos elementos pre-conceptuales capaces de identificar una sensación concreta como una sensación de *algo* que, sin embargo, son previos a toda descripción.

Es decir, Mary (como sujeto perceptor normal y en situaciones de percepción normales) dispone de una serie de elementos (podemos denominarlos habilidades o poderes) de deducción capaces de permitirle concebir la experiencia de dicha percepción como una sensación de rojo. Porque, no lo olvidemos, Mary conoce todo aquello que desde el plano físico puede conocerse acerca del color. Por dicho motivo, ante una determinada experiencia nueva, una experiencia de la que (presuponemos) Mary está siendo consciente, ésta tan sólo tiene que acudir a sus manuales (o a su propio conocimiento previo) para identificar esa nueva experiencia como la experiencia de rojo, en este caso.

Pensemos en un ejemplo similar. Mary, que conoce desde el punto de vista físico todo aquello que puede saberse sobre la física, la química y la neuro-fisiología del sentido del gusto (aunque ha permanecido siempre encerrada en un mundo donde nunca ha podido saborear o degustar nada a través de sus papilas gustativas), ha sido invitada a un helado de chocolate. Mary, como sujeto perceptor normal en una situación normal, es capaz de degustar perfectamente el helado. Tal vez, podamos conceder que conoce algo nuevo: cómo sabe el helado de chocolate (desde el plano subjetivo y cualitativo de poseer una nueva experiencia gustativa). Pero, no lo olvidemos, Mary conoce todo aquello que puede saberse sobre la física del gusto, por lo que tiene (como sujeto perceptor normal) la capacidad de deducir que aquello que experimenta desde el punto de vista subjetivo es lo mismo que sus manuales describían desde un punto de vista objetivo. Todo esto parece indicarnos que lo único que Mary podría llegar a hacer es, digamos, coger su cuaderno de notas y, justo en el margen donde tiene anotada

la descripción física de lo que es tener una experiencia de saborear un helado de chocolate, indicar ‘y es maravilloso experimentar dicha sensación cuando lo saboreo’. Nada cambiaría, entonces. El fisicalismo, por ende, puede ser verdadero porque permite describir perfectamente cómo sería tener dicha experiencia concreta.⁸

Una dificultad añadida que podemos encontrarnos a la hora de sostener dicho argumento nos viene de la mano del propio Jackson. En un texto posterior,⁹ Jackson insiste en la posibilidad de que para que ciertos hechos físicos de nivel superior, de nivel personal, puedan ser deducidos de los hechos físicos de nivel inferior, de un nivel sub-doxástico, debe darse un requisito: dicha deducción debería entenderse siempre como una deducción *a priori*. Si consideramos que dicha sensación es capaz de ser identificada a partir de una serie de elementos pre-conceptuales como una determinada sensación, ello no quiere decir que dicha deducción deba realizarse *a priori*, porque la conceptualización de dicha sensación puede llevarse a cabo de manera totalmente *a posteriori*. Es decir, en este caso, lo primeramente importante es que Mary sea capaz de poder identificar la sensación concreta como una percepción determinada, y que pueda evaluar dicha sensación como placentera, por ejemplo. De qué sea dicha sensación es algo que puede ser descrito *a posteriori* a partir de la propia historia cognitiva de Mary o a partir de todos aquellos recursos conceptuales que Mary ha llegado a adquirir a lo largo de los años.

Como venimos considerando, podemos afirmar que Mary posee una serie de elementos que le permiten realizar deducciones acertadas. Es decir, Mary puede determinar desde sus hechos físicos primarios concretos determinados hechos a un nivel personal que le permitan indicar qué clase de sensación tiene en un momento determinado. Pero ello no quiere decir que esté capacitada para poder deducir *todos* los hechos respecto de la percepción del color de los hechos físicos fundamentales. A menudo, Mary necesitará apelar a sus experiencias previas (aunque sea a aquellas experiencias que ha tenido tan sólo en blanco y negro) para poder saber que está teniendo una experiencia visual de color y no, por ejemplo, una experiencia gustativa de (sabor a) chocolate. A partir de dicha experimentación, entonces, es desde la que Mary estará capacitada para deducir qué clase de sensación es la que está teniendo en dicho momento particular. Pero, como estamos viendo, no depende de elementos *a priori* la posibilidad de deducción. Pero que podamos apelar a otras experiencias previas no quiere decir que sea un requisito imprescindible la presencia de conceptos fenoménicos.

Como hemos comentado anteriormente (véase la nota 3 de la sección 6.2.1.1., cuando atendimos a la polémica entre Balog y Stoljar respecto de la estrategia conceptual) existe la posibilidad de que alguien esté tentado a insistir en la necesaria presencia de elementos

⁸ Tal vez el principal problema aquí, cuando afirmamos la verdad del fisicalismo, sea encontrar un cierto equilibrio a la hora de determinar las propiedades relevantes que algo debería tener para que podamos entenderlo como físico. Es decir, y teniendo presentes las ideas de Liz (1999), deberíamos atender a una serie de condiciones (pp. 194-195) a la hora de establecer una definición de realidad física (existenciales, de continuidad, de contraste, de alternativas posibles y de criterios de satisfacción) que impidan, como parece haber sido el caso a lo largo de los años, que el fisicalismo aparezca (p. 193) como literalmente falso (por ser excesivamente restrictivo a la hora de caracterizar lo físico) o como vacuamente verdadero (por tener una concepción tan amplia de lo físico que puede ser aplicado a cualquier cosa). No voy a detenerme en esta idea aquí (véase nota 113).

⁹ Cf. Jackson (1993). Véase también Chalmers and Jackson (2001).

conceptuales que permitan determinar el tipo concreto de sensación que alguien está teniendo en un momento concreto (Cf. Chalmers 2007a y 2007b). Visto de este modo, sólo si alguien tiene ya presente un determinado concepto previo de, digamos, lo que es experimentar una sensación de rojo, podemos decir que sabe que aquella sensación que en este preciso momento está experimentando es una sensación de rojo. En este sentido, tendríamos que, por un lado, Mary necesariamente debería aprender algo nuevo cuando por primera vez tiene una experiencia de rojo y, por otro lado, ello nos obligaría a afirmar la existencia de un hiato entre aquellos elementos puramente mentales que un sujeto perceptor es capaz de tener y los elementos físicos que permiten a dicho sujeto tener dichas percepciones. Es decir, como indica Chalmers, aquello que parece querer decirnos el defensor de la estrategia conceptual es que para que un sujeto sea capaz de asociar la sensación concreta que se tiene (pongamos por caso, el demostrativo *esto*) con la experiencia concreta de, digamos, rojo dicho sujeto debe poseer ya el concepto de experimentar una sensación de rojo.¹⁰ Pero, vamos a insistir, éste no tiene por qué ser el caso.

Pensemos de nuevo en Mary. Mary sabe todo aquello que desde el plano físico puede saberse sobre la química y la neuro-fisiología del olfato. Pero Mary ha estado siempre privada de una sensación olfativa. Es decir, Mary no tiene un concepto fenoménico que le permita decir cómo es tener una sensación olfativa. Pensemos que por primera vez a Mary se le ofrece la posibilidad de oler. Se le da a oler una rosa. Mary sabe que lo que está haciendo es oler una rosa. Es decir, puede deducir que aquella sensación concreta que está teniendo ahora es una sensación olfativa porque puede diferenciarla (aunque sea por eliminación) del resto de las sensaciones que es capaz de tener. Y a pesar de que pueda ser cierto que aprende algo nuevo al presentársele por primera vez los aspectos cualitativos y subjetivos de la sensación de oler una rosa, Mary sabe perfectamente que es la sensación de una rosa porque, como le indicaban sus manuales, aquella parte del cerebro que tiene activada es la parte del cerebro que debería activarse cuando alguien (en condiciones normales) huele una rosa (Nida-Rümelin 1996) y se dan el resto de condiciones físicas que permiten determinar su experiencia concreta con la experiencia determinada de oler una rosa. No requiere de ningún tipo de concepto fenoménico previo para poder experimentar una sensación olfativa concreta como una experiencia de oler una rosa. Del mismo modo, no requiere ningún tipo de concepto fenoménico previo para saber que la sensación que está teniendo es una experiencia de rojo (cuando se le presenta un tomate maduro).

En este sentido, lo que pretendemos defender es la posibilidad de que, a pesar de que Mary pueda aprender algo nuevo cuando experimenta una sensación de rojo por primera vez, el fisicalismo sea verdadero. Porque, por un lado, los sujetos perceptores disponen de una serie determinada de habilidades de percepción e identificación que son previas a toda percepción y a

¹⁰ Una conclusión similar podemos extraer de Bigelow and Pargetter (1990 y 2006). Aunque estos autores consideran que Jackson se equivoca a la hora de defender la falsedad del fisicalismo, consideran que tiene razón a la hora de llamar la atención acerca de algunos errores que éste comente. El principal problema que plantearía una teoría fisicalista, según estos autores, sería el abandono deliberado, al menos, en algunos de sus casos, de la posibilidad de la existencia de propiedades subjetivas (dependientes de diferentes modos de *familiaridad* sujetos a la perspectiva en primera persona del sujeto perceptor) que permitirían determinar el carácter fenoménico concreto de determinados estados mentales. Por supuesto, como muestro a continuación, no estoy de acuerdo con ellos.

toda conceptualización. Pero que, por otro lado, también disponen de toda una serie de capacidades de deducción que, a partir de elementos mínimos,¹¹ son capaces de identificar experiencias concretas como experiencias de rojo, por ejemplo. Y ello, en ningún caso, demuestra que la propiedad concreta de un sujeto perceptor de tener una determinada sensación visual (pongamos por caso, una experiencia de rojo) sea una propiedad no-física (Montero 2007). Bien al contrario, lo que vendría a indicarnos es que, en última instancia, aquello que cuenta como una determinada sensación visual, por ejemplo, vendrá determinado por completo por las propiedades físicas fundamentales del sujeto perceptor que permiten describir su experiencia como una experiencia visual y por aquellas capacidades discriminatorias que permiten realizar las deducciones pertinentes que le permitirían identificarla como una experiencia concreta, digamos, de rojo, aunque sea de un modo *a posteriori*.

Todavía podríamos estar tentados a decir, como haría por ejemplo Prinz (forthcoming), que nuestra estrategia es similar a aquella defendida por los defensores de los modos de presentación fregeanos. Es decir, podríamos sostener que estamos defendiendo que Mary realmente no experimenta nada nuevo desde un punto de vista metafísico, sino que lo que está adquiriendo es un nuevo modo de presentación (desde el punto de vista epistémico) de los mismos hechos concretos que ya conocía (Cf. Travis 2002). De todos modos, éste no es el caso. Principalmente por dos motivos. En primer lugar, considerar la posibilidad de la existencia a este nivel de diferentes modos de presentación presupondría, por un lado, una dificultad considerable al impedir la posibilidad de que dos modos diferentes de presentación, aunque co-referenciales, pudieran referirse a la misma propiedad física (recordemos que esta es la objeción que Max Black, tal y como la identificó Smart (1959), realizó a la aproximación fregeana del pensamiento) y, por otro lado, dificultaría la posibilidad de la existencia de las capacidades cognoscitivas que hemos supuesto que poseería todo sujeto perceptor normal, en tanto que requería presuponer, también, la existencia previa de un nivel conceptual que no estamos dispuestos a aceptar. En segundo lugar, como defiende Montero (2007: 187-189), la posibilidad que estamos negando (esto es, la posibilidad de aceptar que Mary posee un acceso epistémico nuevo a partir de un nuevo modo de presentación a un hecho viejo que ya conocía) presupone también la posibilidad de que las capacidades de deducción de los sujetos perceptores normales sean entendidas de un modo *a priori*, lo que presupondría la existencia del mismo nivel conceptual que estamos negando. La existencia de un nivel conceptual no es problemática si se considera existente de un modo posterior a la misma experiencia. La dificultad reside en el hecho, defendido por el internismo (por ejemplo, de Chalmers y Jackson, pero también de Stalnaker, Brandom, McDowell o Boghossian), de que dicho nivel conceptual deba presuponerse *a priori*, como previo a toda experiencia.

¹¹ Desde este punto de vista, se podría decir, incluso, que el conocimiento neurológico de Mary le proporciona un campo de posibles cualidades (de sensaciones cualitativas). Si suponemos esto, entonces, al oler la rosa podríamos decir que Mary ni siquiera conocería nada nuevo. Simplemente se actualiza uno de esos posibles hechos neurológicos que ya conoce como siendo capaz de constituir un peculiar campo de contrastes cualitativos.

6.2.1.3. El curioso caso de la heladería mágica

Imaginemos por un momento la siguiente escena. Los mismos científicos que mantuvieron confinada en su exilio cromático a Mary, trazaron también un plan paralelo. Decidieron clonar otro ser, idéntico molécula por molécula a Mary. Es decir, gestaron a Shally. La única diferencia entre Mary y Shally es que, a diferencia de la primera, Shally tiene invertidos los canales neurológicos que le permitirían experimentar el sabor del helado de chocolate y el sabor del helado de fresa. Es decir, fue concebida como un ser espectralmente invertido respecto del chocolate y la fresa en referencia a la experiencia gustativa normal de su gemela genética Mary. Pensemos también que ambas han sido educadas hasta llegar a ser las mayores expertas respecto de la biología y la neurofisiología del sabor, pero con la peculiaridad de que a ninguna de ellas se le ha permitido nunca tener una experiencia gustativa. Podría objetarse que Shally, al ser una experta en la neurofisiología del cerebro respecto de las experiencias gustativas, sabría detectar esa diferencia espectralmente invertida en sus respectivos canales neurológicos. Pero, por mor del argumento, supondremos también que se los científicos que la confinan han manipulado su conocimiento para evitar en todo momento que Shally supiera de dicha inversión neurofisiológica.

Un buen día, sus carceleros deciden dejarlas salir de paseo. Llevadas por un ataque goloso, deciden visitar una heladería cercana al centro de investigación del cual partieron. Mary elige un helado de fresa, mientras que Shally se decanta por un helado de chocolate. Mary disfruta de su helado de fresa, experimenta por primera vez lo que es tener una sensación de sabor a fresa, pero también determina que es algo que había deducido hace ya mucho tiempo, aunque le descubra una determinada serie de particularidades cualitativas y subjetivas que no había experimentado de manera previa. Shally disfruta de su helado de chocolate. Experimenta por primera vez lo que es tener una sensación de sabor a chocolate. Pero, espera un momento, piensa. Algo no encaja en lo que tenía entendido que debería ser una experiencia de sabor a chocolate. Necesita realizar una comprobación. Le pide, por favor, a Mary una cucharada de su helado de fresa. Vaya. Ahora sí que está realmente confundida. El helado de fresa de Mary tampoco sabe como debería saber un helado de fresa. Es curioso. Consulta su manual, para verificar que sus sospechas no se deben a un error suyo, que no se deba a una confusión en sus propias investigaciones. Comprueba que realmente algo extraño ocurre en su propia experiencia personal de saborear los helados de fresa y chocolate y no en los informes que conserva después de años de investigación. Realmente, llega a determinar, lo que ocurre es que cuando yo, se dice Shally, experimento el sabor del chocolate, realmente se me aparece como una sensación de sabor a fresa, y a la inversa. Entonces toma su cuaderno de notas y, justo en el margen de las descripciones de lo que sería tener una experiencia de (helado con) sabor a chocolate y tener una experiencia de (helado con) sabor a fresa, escribe: ‘y yo tengo los canales neurológicos de los que dependen dichas experiencias invertidos’. Y añade: ‘aunque ello no impide que pueda decir que es maravilloso tener dichas experiencias de sabor’.

Como parece demostrar el caso concreto que hemos visto, presuponer una inversión de espectro tampoco parecería ser una dificultad para una solución fisicalista de la conciencia. Una

persona como Shally, que experiencialmente está invertida respecto del sabor del chocolate y de la fresa, podría perfectamente deducir tan sólo a partir de los propios hechos físicos cómo debería ser una experiencia concreta de sabor. De este modo, además de plantear una respuesta viable a la problemática del espectro invertido, esta solución permite mantener también una perspectiva materialista respecto de la conciencia.

6.2.1.4. ¿Podemos plantear una solución materialista al problema de la unidad de la experiencia?

En el capítulo introductorio (véase sección 1.4.14.), introdujimos el problema de la unidad de lo consciente (*binding problem*). Este problema se refiere realmente a una dificultad explicativa derivada de una constatación que, de una manera habitual, cualquier ser consciente es capaz de realizar. Cualquier ser consciente que, por ejemplo, ve acercarse a una velocidad más o menos constante un objeto, digamos un coche, con una cierta forma, color, movimiento, velocidad (y un largo etcétera de características particulares concretas), lo concibe como una unidad (lo que le permite, entre otras cosas, poder esquivarlo, por ejemplo). Imaginemos por un solo momento la catástrofe que sería para un ser concreto si no fuese capaz de concebir dicho objeto como una unidad, sino como un color, una forma, un movimiento, etc., concretos pero sin unidad. Moriría atropellado, porque la confusión se apoderaría de él al no ser capaz de distinguirlo de otras formas, colores, movimientos, etc., y concebirlo como una unidad. Precisamente por ello estamos prácticamente obligados a admitir que la conciencia (o, al menos, la experiencia consciente) debe estar necesariamente unificada.

Distintos autores, desde diferentes posiciones, han planteado una explicación para dicho fenómeno. En el capítulo cuatro, atendimos a la solución proporcionada por el representacionalismo de Tye. Para Tye, la unidad perceptiva es una cuestión de que, en una experiencia concreta, exista una serie de cualidades perceptivas simultáneamente experimentadas que compartan el mismo contenido (Tye 2003). Pero, como también mostramos, parece que esta respuesta es un tanto forzada porque no parece respetar la exigencia de transparencia que el mismo Tye supone a la conciencia. Por un lado, parece que la propuesta de Tye no consigue respetar la multiplicidad de la experiencia sensible, en tanto que presupone que dos experiencias concretas del mismo objeto (pongamos, una visual y la otra táctil) deberán tener el mismo contenido representacional (Cf. O’Dea ms.). Pero nada nos asegura que ello sea el caso, porque una experiencia táctil de dicho objeto podría representarse con un contenido diferente de una experiencia visual (pensemos tan sólo en la posibilidad de que el sujeto perceptor tenga visión borrosa). Por otro lado, mucho más preocupante es que la opción representacional de Tye no consiga diferenciar entre casos de unidad inter- e intra-modales. Es decir, no consiga diferenciar entre casos de unidad referidas a los mismos tipos de sensación y casos de unidad referida a tipos diferentes de sensación (Bayne 2005). Parece ser el caso que existe una fuerte conexión entre los contenidos de una experiencia y el modo en que dicho contenido se hace presente en la experiencia. Si el representacionalismo no consigue diferenciar entre casos donde la experiencia se muestra diferente, por tanto, deberíamos concluir, no

consigue mantener su exigencia de transparencia, por lo que deberíamos dejarla de lado (por ser demasiado estricta) y abandonar la perspectiva representacional.

La importancia de la unidad de la conciencia, como hemos comentado al principio de la sección, se muestra especialmente en los desórdenes que impiden la unificación de la experiencia consciente. Un caso paradigmático lo encontramos en los casos de cerebro escindido (véase sección 1.4.13.). Por definición, un paciente que padece el denominado *split-brain syndrome* es alguien que, normalmente a partir de una intervención quirúrgica, tiene separados los dos hemisferios de su cerebro. Consecuentemente, no es capaz de interconectar las sensaciones que, aunque referidas al mismo objeto, son dependientes de hemisferios diferentes (Bayne 2008). En este sentido, no es capaz de tener una imagen unificada del objeto en cuestión. ¡Ni siquiera es capaz de tener una imagen unificada de sí mismo! Desde este punto de vista, parece mucho más coherente afirmar que la organización unificadora de la conducta sensitiva depende de una cierta oscilación sincronizada de determinadas células nerviosas situadas en el córtex cerebral (Watanabe 1996), lo que explicaría la falta de unidad consciente en pacientes que, precisamente por tener el cerebro escindido, han perdido la conexión con dicha parte de nuestro cerebro (Singer 1993 y 1995). Visto de este modo, sólo la existencia de determinadas condiciones físicas permite una explicación coherente de la unidad de la conciencia, por lo que el fisicalismo sale reforzado a la hora de situarse como la única explicación consistente de la conciencia.

6.2.2. El problema difícil, revisado

En esta sección intentamos aportar razones para solucionar el *hard problem* identificado por Chalmers, a la luz de las respuestas que hemos dado a los anteriores problemas débiles de la conciencia.

6.2.2.1. ¿Existe realmente el hiato explicativo?

Por mucho que lo pueda parecer, la pregunta que encabeza esta sección no es una pregunta retórica. No pretende una respuesta concreta ni fácil. Alguien puede estar tentado a responder con un rotundo sí, por supuesto que existe un hiato entre los hechos físicos y los hechos mentales. Esta idea se derivaría de nuestra autoridad de primera persona respecto de nuestros propios estados mentales. Cómo podríamos estar todos tan equivocados a la hora de vivir nuestras propias mentes. Todos podemos dar cuenta del modo en cómo nos aparece individualmente el mundo. Cómo es el mundo para nosotros. Por supuesto, podríamos concluir, debe existir una diferencia clara entre cómo sea el mundo y cómo lo vivimos. Esto es, todos tenemos la sensación, derivada del sentido común, de que el mundo es como en realidad se nos aparece, independientemente de cómo sea el mundo. En otras palabras, que existe un modo metafísicamente diferente de existencia de los hechos físicos que lo distingue claramente de los hechos mentales.

Como hemos comentado anteriormente, dicha caracterización pretende ser realista. Es decir, persigue mostrar fielmente cómo es el mundo. Por supuesto, si el mundo es tal y como se

nos aparece a los individuos perceptores, por tanto, no puede ser que el mundo tenga unas características diferentes a aquellas que nosotros percibimos. Por tanto, debemos concluir, el mundo debe tener las características que nosotros le presuponemos. Pero la visión común sobre el mundo entra en conflicto con la explicación que del mismo ofrecen nuestras actuales ciencias físicas. Desde el punto de vista de la ciencia actual, el mundo tiene unas características concretas que, de forma constitutiva, lo definen (seamos capaces o no de llegar a captarlas). Visto de este modo, el sentido común parte de una idea irrealista del mundo.

Nosotros pretendemos defender una imagen realista de la conciencia, una imagen fisicalista, en tanto que ésta puede ser descrita en términos físicos y en tanto que permite ofrecer una imagen concreta del mundo.

Pensemos por un momento en un ejemplo (extremadamente ridiculizado) de la anterior perspectiva derivada del sentido común: la carrera en comité que describe Lewis Carroll en *Alicia en el país de las maravillas*. Tras haber caído en su propio mar de lágrimas y haber coincidido con el Dodo, el Ratón y los demás animales, Alicia llega (junto con todos) a tierra firme. Todos ellos están empapados, por lo que el Dodo decide proponer una carrera en comité (consistente en correr en círculos) para secarse. El *quid* de la cuestión estriba, precisamente, en que lo relevante de esta historia no reside en el hecho de que deban correr para secarse sino en que, mientras realizan la carrera, todos ellos deben pensar en cosas calientes, y esto es lo que les permitirá secarse. Por supuesto, podemos decir, que el que lleguen a secarse depende del hecho (perfectamente describable desde la física actual) de que al correr se genera una pequeña corriente que, sumada al calor que desprende el cuerpo, permite evaporar la humedad, y otros hechos físicos similares. Precisamente esto nos indica que los estados fenoménicos no tienen efecto causal, pero que pueden estar ahí y ser importantes para nosotros, aunque no permitan una explicación causal de los hechos físicos a los que, a menudo, parecen estar asociados. Creer que por tan sólo pensar en cosas calientes podemos llegar a secarnos indica ciertamente un desconocimiento total de cómo funciona el mundo y de las leyes físicas que lo explican. En pocas palabras, estamos presuponiendo una visión irrealista del mundo al creer que el mundo es como creemos que es (o se nos aparece) de modo subjetivo. Algo similar pasaría cuando se defiende la posibilidad de que existan, de manera separada, una mente como desconectada del cuerpo o una serie de elementos fenoménicos caracterizados como causalmente eficaces (como muchos defensores del dualismo sostendrían).¹²

¹² Una idea similar, aunque no idéntica, sería la defendida actualmente por todos aquellos filósofos que mantienen una posición aislacionista, tal y como la define Davies (2000: 64): ésta es “una posición que mantiene que la filosofía de la mente y del lenguaje ofrecen una metodología distintiva –el análisis conceptual, laxamente entendido– con la que investigar una clase de cuestiones sustantivas sobre las nociones que aparecen en nuestro pensamiento cotidiano sobre cuestiones lingüísticas y psicológicas. Aquellos que ocupan dicha posición están dispuestos a admitir que las ciencias cognitivas son muy interesantes por sí mismas, pero que poco o nada tienen que decir al proyecto filosófico de delinear los contornos de nuestro esquema conceptual”. Muchos de los actuales intentos de aportar una explicación similar al problema de la conciencia presuponen este aislacionismo. Posiblemente los más influyentes, actualmente, sobre todo por su número, son aquellos que se posicionan en supuestos wittgensteinianos. Normalmente, a partir de algún dato *a priori* (como la normatividad de nuestros actos y estados mentales) niegan aquellos aspectos relevantes desde el punto de vista físico que permitirían reforzar dicha noción. Lo que de este modo consiguen, precisamente, es un efecto contrario: al negar la relevancia de las consideraciones que los elementos físicos tendrían a la hora de

Como estamos diciendo, desde un punto de vista epistemológico, el sentido común defendería que los seres conscientes sienten la realidad como se presenta a sus conciencias. Aunque intuitivamente parece cierta (¿cómo si no se va a sentir la realidad?, podríamos preguntar), debemos ser cautos con esta afirmación, porque puede hacernos recaer en una temida tesis metafísica (identificada como idealista). Ésta afirmaría que todo aquello que existe es producto por completo de nuestra actividad mental. Ante esta idea, queremos contraponer una visión fisicalista del mundo desligada de las soluciones aportadas a los anteriormente tratados problemas de la conciencia.

La suposición de la existencia de una brecha explicativa entre lo físico y lo mental supone concebir la existencia separada de dos sustancias. Hemos defendido a lo largo de todo este trabajo la existencia de una base física para la presencia de todo estado mental, por lo que no vamos a tener más remedio que concluir negando la existencia de dicho hiato explicativo presupuesto desde posiciones internistas y dualistas. Actualmente existe un intento de transformar este hiato en un problema semántico, más que metafísico o epistémico. La idea subyacente a dicho intento reside en el hecho de que, afirman, la metafísica es en última instancia una forma de análisis conceptual, por lo que (en definitiva) la disputa acerca del presunto hiato entre lo mental y lo físico apunta a una lucha entre diferentes categorías estructurales, por lo que un análisis semántico permitiría una explicación satisfactoria de dichas nociones. Cf. D'Oro 2007. Pero, como hemos dicho anteriormente (véase la nota 123), estas estrategias tan sólo evitan los problemas al aislarse de las verdaderas cuestiones.

Los argumentos a favor de dicho hiato suelen moverse en dos afirmaciones básicas. En primer lugar, consideran que una aproximación dualista de las experiencias fenoménicas ofrecería una explicación del hecho de que el conocimiento fenoménico (y de la existencia del acceso privilegiado en primera persona) es infalible. Las características metafísicas de la experiencia consciente permitirían explicar las propiedades cognoscitivas de lo mental, especialmente de lo fenoménico, en tanto que aportarían una explicación representacional de los modos en los que los estados mentales fenoménicamente conscientes se presentan a los sujetos.¹³ En segundo lugar, y como presupone la primera afirmación, consideran que todo

determinar el conocimiento innato de dicha normatividad, lo que hacen es mostrar cómo, después de todo (aunque lo parezca *a priori*), no existe dicha normatividad intrínseca en los estados fenoménicos. Cuando se les acusa de esto, acaban por escudarse en una simple constatación: la noción de *qualia* defendida por un fisicalismo como el que estamos sosteniendo aquí acaba por ser tan extremadamente débil que aboca a los problemas que hemos identificado en el estudio de la conciencia por tener un criterio *a posteriori* de identificación entre *qualia* y elementos físicos y por suponer una irrelevancia causal que lo aproximaría peligrosamente al epifenomenalismo. Desde mi punto de vista, lo que realmente esconde el empleo de una noción normativamente robusta de *qualia* (como la que defiende este aislacionismo wittgensteiniano) es, por un lado, un error sustancial al pretender determinar mediante argumentos *a priori* algo que tan sólo podemos realizar mediante investigación empírica (y que dejaría abierta la amenaza que hemos identificado anteriormente al hablar de la estrategia de presuponer la existencia de un contenido conceptual *a priori* en los estados mentales con carácter fenoménico) y, por otro lado, un miedo patente a enfrentarse a los problemas débiles de la conciencia. El problema principal, creemos, es que rehuir estos problemas no los hace desaparecer, por lo que tarde o temprano deberán acabar enfrentándose a ellos.

¹³ Un ejemplo claro de una aproximación de este tipo lo encontraríamos en la teoría auto-representacional defendida actualmente por Kriegel. Básicamente, el argumento de Kriegel pretende salvaguardar la infalibilidad del acceso de primera persona a los propios estados mentales a partir del supuesto de la existencia de un sentido interno periférico que permite atender a todos aquellos elementos fenoménicos presentes en el sujeto. Cf. Horgan and Kriegel 2007.

estado fenoménico es transparente. Es decir, consideran que la única teoría adecuada para dar cuenta de lo fenoménico es aquella que consigue establecer una adecuada relación entre los estados físicos y los estados fenoménicos (Levin 2007).

Pero estas dos afirmaciones parecen insostenibles. Por un lado, no tenemos más remedio que negar la tesis de la transparencia por ser un criterio demasiado exigente para la conciencia. Aunque la mayoría de las veces somos conscientes de aquello que nos pasa, es decir, somos conscientes de estar teniendo una experiencia determinada, no siempre es este el caso. A menudo, tenemos una sensación de algo que no somos capaces de identificar (por ejemplo, una molestia en uno de nuestros hombros o en una de nuestras rodillas, pero que no conseguimos identificar como un dolor, un picor o algo similar en concreto) o nos encontramos en un estado del cual no somos enteramente conscientes (como cuando inadvertidamente entramos en un estado de ansiedad por la presión del trabajo acumulado). Visto que no siempre podemos defender la presencia de un criterio de transparencia en la identificación de los estados mentales con carácter fenoménico, no tenemos más remedio que negar este requisito porque, en caso contrario, siempre sería necesario a la hora de rendir cuentas de nuestras intuiciones de sentido común respecto de los contenidos conscientes (respecto del acceso privilegiado en primera persona) aceptar la existencia de una relación de co-referencia *de jure*¹⁴ entre los estados fenoménicos y los pensamientos relativos a los mismos (Schroeter 2007), algo que no parece posible, como hemos visto, al analizar la imposibilidad de la existencia de contenido previos establecidos metafísicamente *a priori*. Pero tampoco sería defendible un criterio de transparencia más débil que hemos identificado también a lo largo de este trabajo: la posibilidad de que el sujeto tenga acceso al mundo y a sus propiedades. Del mismo modo que antes, este criterio parece demasiado exigente en tanto existen esferas de la realidad que nos están vetadas y a las cuales no tenemos ningún tipo de acceso consciente.

Por otro lado, precisamente atender a las propiedades metafísicas de los estados fenoménicos (como hemos mostrado en la sección 6.1.3.) nos indica la presencia de elementos objetivos (más allá del descubrimiento particular de los correlatos neuronales de la conciencia o de la posibilidad de la existencia de un nivel cuántico. Cf. Kuttner and Rosenblum 2006) que permiten identificar (o, al menos, explicar) los estados mentales con carácter cualitativo con propiedades físicas y con respuestas perfectamente identificables desde el plano de la investigación empírica. Una serie de elementos objetivos que, precisamente, permitirían a

Véase también Jehle and Kriegel 2006 para un enfrentamiento entre las teorías disposicionales de nivel superior (como la de Carruthers) y las auto-representacionales. Para una introducción general a las teorías auto-representacionales, véase Kriegel and Williford 2009.

¹⁴ Esta co-referencia *de jure* sería aquella que se podría establecer por principio, de modo esencial, entre todo estado fenoménico y el nivel de conocimiento consciente que de dicho estado debería necesariamente tener el sujeto perceptor, en tanto que lo que defendería la tesis de la transparencia es, precisamente, que los estados fenoménicos siempre son accesibles de modo directo al sujeto consciente. Lo que hemos pretendido en las últimas páginas es cuestionar esta posibilidad.

cualquier sujeto ‘ponerse en la piel’ de los otros y, aunque no siempre, saber qué tipo de estado está presente actualmente en un sujeto concreto en un determinado momento.¹⁵

En vistas a intentar aportar algo de claridad al tema tratado, pasamos a presentar un breve esbozo de cómo, en principio, creemos que es la base física de las experiencias cualitativas. Desde nuestro punto de vista, la información recibida del mundo externo es internalizada a partir de los mecanismos de transmisión de información cognitiva dispuestos a lo largo de nuestro sistema receptor y dicha información es procesada y codificada por determinados sistemas de procesamiento mediante intercambios físico-químicos realizados en el cerebro y en el sistema nervioso central. Estos estados cerebrales actuarían como correlatos neuronales de *qualia* sensitivamente captados. Deberían entenderse como los correlatos cerebrales de aquellos estados que nos aparecen como mentales. Cada uno de estos *qualia* por separado puede ser identificado como una serie de señales químicas y físicas que asignan cualidades diferentes, dependientes del sujeto particular que los experimente, a fenómenos numéricamente distintos, pero en todo sujeto tendría un alto valor para la supervivencia del organismo. En este sentido, la información del entorno es internalizada a partir de elementos biológicamente evolucionados que han ido refinando su capacidad de reconocimiento de un modo filogenético, pero que es elaborada por sistemas de procesamiento capaces de generar un particular modo de experimentación a partir de su peculiar desarrollo ontogenético, dependiente de la particular historia cognitiva de cada sujeto. En este último caso, la información relativa al mundo exterior se internaliza en cada individuo a partir de una serie de ‘apaños’ realizados por los sistemas de procesamiento de la información, en tanto que, primero, permitirían captar de un modo directo aquella información del exterior (y del interior) en tanto que experiencialmente relevante y, segundo, mediante una larga cadena de combinaciones y procesos empíricos (en los que, a menudo, intervendrían procesos de control racional y conceptualizaciones debidas a periodos previos de aprendizaje) en los que cada uno de los individuos estaría en disposición de confeccionar su propia imagen del mundo a partir de las propiedades externas, lo que le permitiría disponer de una imagen más o menos fidedigna de su propia conciencia.

Desde este punto de vista, las experiencias cualitativamente relevantes, los *qualia* (como las percepciones, los estados corporales o los estados de ánimo), no serían más que internalizaciones biológicamente determinadas de estímulos consistentes en señales bioquímicas que son contingentemente plasmadas como fenoménicamente relevantes. Pero que, en un sentido pragmático, tienen un carácter de alerta (hablando en términos de supervivencia). Sólo debemos recordar el hecho de la presencia de la sensación de extremo calor cuando cogemos un

¹⁵ A este respecto, dos son actualmente las posiciones que intentan aportar una explicación a la posibilidad de intentar conocer aquello que están sintiendo los demás (o que intentan explicar la presencia en los demás de dichos elementos objetivos, el conocido como el problema de las otras mentes, y que deliberadamente hemos dejado al margen en este trabajo). Por un lado tendríamos la teoría-teoría, el punto de vista de que nuestra capacidad para comprender a los otros es una capacidad de teorización acerca del comportamiento de los demás. Por otro lado, tendríamos la teoría de la simulación, la posición que defiende que nosotros atribuimos estados mentales a los demás mediante la simulación de su contenido mental. Para una defensa de la primera perspectiva, véase especialmente Carruthers and Smith 1996. Para una defensa de la segunda, véase Ravenscroft 1998. Ojeda 2006 ofrece una aproximación crítica a ambas perspectivas.

vaso de agua hirviendo, o nuestra retirada instintiva del brazo cuando notamos un exceso de calor por aproximarnos distraídamente a una caldera. Claramente, la experiencia fenoménica que extraemos en estos casos permite una mayor supervivencia de los individuos que la padecen. Más difícil de explicar es la relevancia de otro tipo de estados con carácter fenoménico. Visto de este modo, que un determinado objeto aparezca como rojo o que, en un momento determinado, mi estado de ánimo definible como de abatimiento coloree el resto de mis estados mentales tiene una explicación menos plausible, sobre todo por la dificultad a la hora de poder atribuirle un poder causal. De todos modos, como hemos comentado en la sección 6.1.3., ello no debería impedirnos afirmar su presencia.

En casos concretos como los anteriormente tenidos en cuenta, tal vez, el papel evolutivamente relevante sea mucho más laxo o, simplemente, su falta de eficacia causal se deba al hecho de que, en la actualidad, los seres humanos (aquellos seres con una conciencia mucho más compleja) hayan llegado a un extremo en el que han conseguido establecer un medio externo de socialización que permite economizar la carga genética individual mediante la institucionalización de la mayor parte de los recursos de supervivencia. Pero, lo que sí es cierto, es que la alternativa física permite salvaguardar dicho poder causal. Cuando un correlato neuronal como los tenidos en cuenta aquí se incorpora a los procesos conscientes, no sólo permite tener presente su impacto sobre la realidad de un modo normal, sino que también permite explicar determinados estados disfuncionales (como las alucinaciones). Pensemos tan sólo en una experiencia perceptiva alterada de rojo. Al observar un plátano, por un momento, se nos aparece como rojo. En dicho momento concreto, se ha hecho consciente un determinado estado fenoménico, aunque erróneo, que podría explicarse desde el plano cerebral, en tanto que (accidental o erróneamente; tal vez debido a una disfunción momentánea o una alteración genética) una explicación física permitiría dar perfectamente cuenta de la posibilidad de la aparición de este tipo de nuevas entidades físicas y, lo que sería más importante para nuestro propósito final, permitiría explicar (mediante su disolución) el temido problema de la causación mental (evitando así la acusación de epifenomenalismo).

Por supuesto, nosotros no estamos ofreciendo una detallada descripción de los procesos cerebrales que permitirían explicar de un modo completo todo hecho causalmente relevante. De momento, tampoco disponemos de una respuesta física total. Pero sí disponemos de un cierto número de explicaciones más o menos detalladas del funcionamiento de determinados procesos de tratamiento de la información cognitiva (Cf. Churchland and Sejnowski 1992; Churchland, Ramachandran and Sejnowski 1994; Koch and Davis 1994; Abbot and Sejnowski 1999; LeDoux 2002; Llinás 2003). En dichos desarrollos se muestra claramente cómo parece más o menos cierto que la información cognitivamente relevante se codifica a partir de los circuitos neuronales presentes en el córtex, pero no existen indicios de que dichas individuaciones tengan una predisposición a ser 'archivadas' o procesadas por determinados centros concretos. Según esto, parece que la decodificación de las señales particulares relativas a experiencias concretas no depende apriorísticamente de ningún sistema cerebral en concreto, sino que aparece relegado a una capacidad (digamos) interpretativa propia de cada sujeto que se conforma a partir de su

propia plasticidad ontogenética y de su propio desarrollo histórico cognitivo individual. En este sentido, los mecanismos de integración de dichos correlatos neuronales en procesos conscientes parecen depender por completo de los propios mecanismos individuales de control sensitivo que cada individuo, por separado, desarrolla de modo particular a partir de las comunes estructuras neuronales y nerviosas.

En este sentido, sólo tenemos que atender a la evolución de sistemas de percepción concreta, como el ojo. Parece demostrarse que el procesamiento de la información visual comienza en la retina (Cf. Zeki 1993; Humphreys 2003; Llinás 2003: 109-128). Allí, diferentes células nerviosas captan diferentes elementos lumínicos que permiten, tras su transporte y procesamiento en sistemas situados en el córtex occipital, establecer diferentes tipos de sensaciones visuales. Desde el córtex occipital, los flujos de información se redistribuyen por los sistemas de procesamiento del córtex parietal posterior y de las áreas temporales inferiores, permitiendo unificar de este modo las diferentes sensaciones visuales para conformar, mediante el procesamiento conjunto de los diferentes correlatos neuronales, los diferentes elementos retinales. Esto viene a indicarnos que, en realidad, aunque la retina capta de modo directo la realidad a partir de la absorción de diferentes células fotorreceptoras, la percepción visual es un proceso creativo, en tanto que presupone el trabajo de diferentes sistemas de procesamiento que permiten elaborar la información dependiente de diferentes correlatos neuronales que, en última instancia, permitirán componer una metáfora del mundo. (Tan sólo debemos recordar que los elementos relativos al color y las formas se procesan de manera simultánea, mientras que la percepción del movimiento puede durar hasta 50 milisegundos. Cf. Kinsbourne 2000).

Si tenemos en cuenta estas ideas, parece que estamos sosteniendo una reformulación representacional de la conciencia perceptiva, pero éste no es el caso aquí. Pensemos por un momento en el cerebro. El cerebro por sí sólo no tiene salidas ni entradas sensoriales directas. Básicamente, el cerebro es una máquina virtual capaz de emular la realidad. Somos básicamente máquinas dispuestas a soñar, por lo que la actividad cerebral es una metáfora para toda actividad posterior. Desde el punto de vista evolutivo, el único objetivo del cerebro es conseguir la supervivencia del organismo mediante la anticipación a todo acontecimiento futuro, por lo que necesita recopilar la mayor cantidad posible de información del mundo físico exterior e interior al propio organismo que le sea posible para, así, poder establecer determinados planes de acción mediante la disposición de conductas y poder prever posibles respuestas a problemas concretos. (Todo esto lo podría lograr con tan sólo un kilo y medio de peso y un consumo limitado a 14 vatios. Cf. Erecinska and Silver 1994). Visto así, entonces, en tanto que el cerebro es un sistema inherentemente cerrado, la evolución no tuvo más remedio que inventar los órganos capaces de captar la realidad externa, los órganos sensitivos, en tanto que sistemas especializados de asimilación (e internalización y conducción cerebral) de las propiedades físicas externas.

Es conveniente recordar aquí que la selección natural actúa sin objetivos y sin intención final, por lo que carece de toda teleología. Entonces, el desarrollo de todo órgano no tiende hacia la consecución de un fin concreto, sino que en un momento determinado un órgano se instaaura como un modo posible de realizar determinadas funciones, como le ocurre al ojo. Pero

lo que precisamente nos indica el caso de la percepción visual, como podemos comprobar a partir de los numerosos y diferentes diseños existentes de ojos, es que la evolución crea una variedad de diferentes posibles soluciones al problema de captación de los datos externos, permitiendo la múltiple realizabilidad de dicha tarea. Lo mismo ocurriría en casos como la percepción de otros tipos de experiencias fenoménicas: la evolución desarrollaría diferentes modos de captación de las mismas experiencias a partir de la potenciación de aquellos diseños que, funcionalmente hablando, permiten una mayor supervivencia a los individuos que los poseen.

Si una conclusión podemos extraer teniendo presente todo lo anteriormente defendido es que, en última instancia, la intencionalidad presente en los estados mentales no es intrínseca (como podría afirmar, por ejemplo, una posición cercana al naturalismo biológico como la defendida por Searle), sino relacional (como insinúa una posición como la defendida por Dennett). Es decir, la asociación existente entre diferentes experiencias fenoménicas (capaz de hacer aparecer como unitarios los objetos y las experiencias de los mismos) se lleva a cabo a partir del procesamiento por separado, aunque trabajando conjuntamente, de diferentes mecanismos que manipulan los datos informacionales relativos a las experiencias internas o externas al propio organismo. En otras palabras, la intencionalidad no es más que la relación existente entre los correlatos neuronales de las diferentes experiencias y estados mentales que componen el aparato cerebral del sujeto perceptor concreto (Cf. Musacchio 2005: 450; Pestana 2006). Ello nos llevaría a concluir que no existe una intrinsicidad propia en ningún ser concreto que pueda indicarnos una naturaleza puramente esencial de la intencionalidad presente en los estados mentales.

Del mismo modo, queremos afirmar también que la conciencia fenoménica no es más que un sistema de navegación que permite a los sujetos perceptores detectar determinadas propiedades físicas existentes en el mundo (tanto externo como interno), lo que le permite al sujeto darse cuenta de aquello que acontece tanto dentro como fuera de su cuerpo y que, de algún u otro modo, le afecta directamente. Tal vez, como hemos dicho anteriormente, los *qualia* actualmente no tienen un extremado valor de supervivencia ni un extremado papel causal a la hora de explicar la conducta de los seres conscientes. Pero sin duda, están presentes a lo largo de toda nuestra vida consciente.

6.2.2.2. Entonces, ¿por qué los *qualia* parecen no ser físicos?

La naturaleza de los conceptos que empleamos para hablar de la experiencia tiene un extremado interés por sí misma. Pero ello no significa que el empleo de dichos conceptos fenoménicos deba ser estrictamente entendido en un sentido intrínseco (Crane 2005: 162). Muy probablemente, como hemos comentado a lo largo de este trabajo, la idea de la existencia de una vinculación apriórica entre concepto y fenómeno sea falsa. Si existe algún tipo de relación, ésta debe ser entendida en un sentido trivial, en un sentido relacional, establecido tan sólo de modo empírico y *a posteriori* (sobre todo si tenemos en cuenta que la existencia de la estrategia conceptual anteriormente expuesta, nos parece, no permite una explicación satisfactoria del

modo en que vivimos nuestra propia experiencia consciente). Así concebida, no tenemos más remedio que intentar explicar cómo puede ser que, si la experiencia fenoménica depende de estados físicos, nos aparezca como subjetiva y dependiente de elementos intrínsecos a la mente (lo que nos obliga, de nuevo, a enfrentarnos al *hard problem*).

Como hemos dicho antes, los *qualia* están presentes de un modo directo en todo ser consciente. Todos nosotros podemos saborear, podemos oler, podemos sentir la suavidad del tejido del que se componen nuestras sábanas, podemos sentir calor y frío, podemos tener miedo, sed, hambre, sentirnos abatidos, cansados, y un largísimo etcétera de ejemplos similares. Todas estas sensaciones están presentes en nosotros en forma de experiencia, pero responden causalmente a los correlatos neuronales que los provocan.¹⁶

Precisamente porque la única forma que los seres conscientes tenemos de sentir dichas experiencias es a partir de nuestros mecanismos individuales de percepción (a los que tenemos un acceso directo) y no podemos sentir de modo directo las sensaciones que padecen los demás (podemos tener un acceso por analogía, en tercera persona, presuponiendo que los demás son capaces de sentir lo mismo que nosotros cuando parece que tenemos determinadas sensaciones a partir de determinados comportamientos), es por lo que, creemos, nos aparecen de un modo cualitativamente subjetivo, en tanto que tan sólo somos capaces de captar el modo en cómo se nos aparecen a nosotros.¹⁷

Precisamente esta situación es la que, a través de generaciones, ha permitido alimentar y mantener vivo un mito: la existencia de dos tipos diferentes de sustancias (la mental, por un lado, y la material, por otro). Pero, como ha sido nuestra intención a lo largo de todo este trabajo, nosotros nos decantamos más por aceptar la idea de que la única sustancia propiamente dicha es la sustancia material, de la cual se compone nuestro cuerpo y nuestro cerebro y que, de un modo múltiplemente realizable, como hemos visto, permite la existencia de sensaciones en tanto que correlatos de determinadas distribuciones neuronales y nerviosas. Si esto es así, entonces, todo ser capaz de tener (aunque sea funcionalmente hablando) correlatos físicos similares puede, suponemos, tener las mismas sensaciones que nosotros cuando experimentamos. Tener una conciencia, por tanto, es tener una determinada serie de habilidades y capacidades cognitivas pero, en sí misma, la conciencia (al igual que la mente) no puede ser una sustancia. Para que algo pueda ser considerado una sustancia debe poder demostrar ciertas propiedades causales (Crane 2003). Esto es, el modo más claro de demostrar la existencia de una sustancia es poder aportar pruebas o razones que indiquen su autonomía, en tanto que por un lado sea capaz de demostrar un cierto nivel de causalidad o, al menos, la posesión de una serie determinada de poderes causales propios (es decir, que es capaz de tener ciertos efectos) y por otro lado que es independiente de cualquier otro substrato o sustancia. En el caso de la

¹⁶ Como hemos expuesto en el capítulo 2, Searle defiende precisamente que la mente es un producto causal del cerebro. Como ya dejamos constancia allí, es esta tesis la que permite afirmar una posición naturalista respecto de la mente. Podemos indicar de nuevo, como ya lo hicimos en el capítulo 2, que esta misma tesis es la parte aceptable de la teoría de Searle.

¹⁷ Muchos son los autores que defienden una posición similar para explicar el carácter subjetivo de los estados mentales con contenido fenoménico. Tal vez el más representativo es McGinn 1982.

materia, de aquello físico, es claramente identificable como sustancia por sí misma porque, primero, es capaz de demostrar toda una serie de leyes causales que permiten la consecución de toda una serie definida de efectos y, segundo, además, dichos poderes causales lo tiene de manera esencial, es decir, por sí misma, demostrándose así independiente de cualquier otro tipo de sustancia, siendo suficiente para la consecución de los efectos deseados que se dé uno de sus principios causales (aunque tenga que ser el adecuado). Pero, como hemos visto, no parece ocurrir lo mismo con la mente, en general, y con la conciencia, en particular. Aunque hemos estado postulando la posibilidad de que ambas tengan ciertas características causales (ciertas propiedades causales capaces de causar algún efecto), también hemos establecido la imposibilidad de que dichas propiedades causales se den de modo autónomo e independiente de la presencia de una sustancia material, física. Es decir, a diferencia de la sustancia material, que demuestra su causalidad y su independencia a la hora de poder conseguir ciertos efectos por la acción independiente de sus propios poderes causales, la conciencia, como hemos comentado, tan sólo consigue tener determinadas propiedades causales si la consideramos el correlato de ciertas propiedades físicas dependientes de las estructuras neuronales y nerviosas de los sujetos en las que acontece, demostrando así su dependencia de la existencia de una sustancia material y negando así el requisito de autonomía que una entidad requiere para poder ser considerada sustancia.

CONCLUSIONES GENERALES

Tal vez el pensamiento no sea más que un capricho de la Naturaleza, que se lo ha concedido a una única especie de igual modo que concedió a los mamíferos raros o extintos que vemos en nuestros museos sus absurdas astas: armas u ornamentos tan curiosamente desplegados, enlazados, retorcidos o ramificados que son más peligrosas que inútiles para el animal que los lleva.

Después de todo, ¿por qué no? Nuestras cabezas están atestadas de especulaciones e ideas que se quedan atrapadas en los laberintos del bosque de los hechos incontestables y nos ponen en un dilema, orgullosos de nuestras excrescencias y condenados a berrear, como ciervos en celo, nuestros poemas y nuestras teorías, grandilocuentes y mortificados.
(Paul Valéry, *Estudios filosóficos*)

*Siéntate ante los hechos como un niño pequeño,
prepárate a renunciar a cualquier noción preconcebida.*
(T. H. Huxley, *Science and education*)

CONCLUSIONES GENERALES

Como propusimos al comienzo de este trabajo, nuestro objetivo principal ha sido realizar una revisión crítica de los actuales intentos representacionistas respecto del estudio de la conciencia para intentar, por un lado, mostrar los elementos que una teoría de la conciencia debería necesariamente poder explicar en la actualidad y, por otro lado, establecer las limitaciones en las que las actuales aproximaciones a la noción de conciencia parecen acabar. Pero también, como indicamos desde el inicio, no sólo pretendimos realizar una labor negativa. También era nuestra intención intentar aportar alguna solución a los (o al menos, alguna perspectiva nueva respecto de ciertos) problemas que, hoy en día, el análisis analítico de la Filosofía de la Mente, en general, y de la Filosofía de la Conciencia, en particular, no pueden dejar de afrontar.

Lo que realizaremos en las siguientes páginas es un resumen de las ideas y conclusiones más relevantes que pueden ser extraídas de este trabajo.

* * *

A la hora de mostrar la influencia que la noción de conciencia ha tenido a lo largo de los siglos en el estudio filosófico de la mente, comenzamos nuestra exposición con una breve contextualización de la misma. Afirmamos allí, en el prólogo, que en la Antigüedad la noción de conciencia no era entendida de un modo separado de la noción completa de alma. La conciencia, aunque no entendida en el sentido en que hoy en día es concebida, no era diferenciada de la posesión general de un alma, de un espíritu, que animaba los cuerpos y que insuflaba el hálito vital que todo ser vivo parecía poseer. Aunque en un sentido restringido, el alma contenía una serie de características racionales que le permitían, por una parte, y de manera intuitiva, aprehender la realidad circundante mediante un ejercicio de contemplación racional pero también, por otra parte, se le concedía una serie de características causales que, vinculadas de modo estrecho con la posesión de un cuerpo, le permitían actuar en el mundo material, aunque el precio a pagar fuera la limitación, precisamente, de dicha actividad racional a aquello que le aportaban sus sentidos imperfectos.

Esa noción genérica de alma dominó la perspectiva filosófica durante los años de la Antigüedad, en sus dos vertientes: la platónica y la aristotélica. La escuela platónica entendía el alma como aquella entidad metafísicamente independiente destinada a vagar por el mundo de las Ideas, capaz de llegar a alcanzar el Conocimiento, y aunque habitara limitadamente en un cuerpo sensible estaba capacitada para trascender lo mundano y alcanzar dicha certeza. La escuela aristotélica, en cambio, aunque también consideraba el alma como el aparato racional, la vinculaba directamente con la posesión de un cuerpo, en su sentido físico, y aunque también estaba capacitada para alcanzar el conocimiento cierto de la realidad, en ningún momento podría ser entendida como separada de la limitación sensible, dado que el único modo de captar el mundo era, precisamente, a través de los sentidos imperfectos del cuerpo al que irremediamente estaba unida. Pero dicha noción de alma también será heredada por la filosofía medieval, donde será cristianizada y se convertirá en el instrumento perfecto para

demostrar las propiedades divinas del dios cristiano. Como comentamos al comienzo, de nuevo, la noción de alma adquiere una mayor preponderancia frente a la noción de materia, en tanto que la materia adquiere peyorativamente una serie de características limitativas que, en vistas a la consecución de la verdad y la salvación (fundadas en la divinidad), imposibilitarían la contemplación de la bondad divina. Tan sólo a las características superiores del alma (por su racionalidad e intuitividad puras) le correspondería la misión de captar los atributos divinos, en tanto que reflejados en la creación y en la naturaleza, pero a partir (por supuesto) de un acto de abstracción trascendental que permitiera liberar al sujeto de toda limitación terrenal.

En este momento crucial, cuando la visión religiosa del cristianismo inunda el panorama filosófico, cobra crucial importancia la noción de acción. En tanto que la bondad divina debía permitir la existencia en los humanos (al menos) de la posibilidad real de elegir, esto es, de la posesión por parte de las criaturas racionales creadas por dios de un libre albedrío que permitiera a los sujetos individuales actuar según sus propios razonamientos, la noción de conciencia comienza a cobrar relevancia. Debido a la posibilidad de actuar de manera libre, los sujetos serían capaces de decidir por sí mismos cómo actuar: si según los designios de la Razón Divina (presta a conceder la salvación) o según los criterios de la terrenalidad.

Con la llegada de la concepción racionalista (y a la vez escéptica), iniciada en Descartes y Montaigne, esta perspectiva metafísica es abandonada paulatinamente en beneficio de una nueva concepción cada vez más epistemológica. Como vimos, a partir de los estudios cartesianos, la prioridad ya no es tanto la salvación del alma inmortal como la correcta aprehensión de la realidad. La noción metafísica de alma es sustituida por la noción epistemológica de razón, una razón capaz de captar mediante intuición racional la realidad de lo que nos rodea mediante la superación de las limitaciones del conocimiento sensible.

Contra el solipsismo metodológico de Descartes, el empirismo inglés concibe una nueva noción de razón, mucho más vinculada a la realidad material, donde la verdadera imagen del mundo aparece vinculada estrechamente con la noción de los datos de los sentidos. A diferencia de la propuesta racionalista, obsesionada por desligarse de toda sensibilidad, Locke y Hume apuestan por la existencia de una razón vinculada directamente con aquella información extraída a partir de los sentidos materialmente vinculados que poseen los sujetos. La realidad no sería ni más ni menos que el conjunto organizado de aquellas sensaciones que, mediante la construcción realizada *a posteriori* por la razón, los sujetos materiales (esto es, un alma unida a un cuerpo) son capaces de captar.

Llegados al siglo XX, y como mostramos, la noción de conciencia adquiere una nueva dimensión, una mayor relevancia. Si en los anteriores intentos la cuestión principal era determinar la relación existente entre un alma racional capaz de determinar por sí misma mediante la pura intuición la realidad interna y un cuerpo material capaz de captar mediante sensaciones la realidad externa (el tradicionalmente denominado problema Mente-Cuerpo), a comienzos del siglo XX, con la aparición de nuevas teorías científicas que presentan datos nuevos sobre la realidad física, se apuesta por una mayor visión científica del problema. Justo es

en este momento donde aparece nuestra noción en su actual sentido: la conciencia entendida como carácter fenoménico.

Como dejamos constancia en el apartado 1.3., parece que a lo largo del siglo XX la preocupación respecto de la relación existente entre el cuerpo y el alma se reformula en términos, de nuevo, epistémicos. Si durante los anteriores siglos la relación era entendida en un sentido metafísico (la existencia unida o separada de dos entidades, dos sustancias, diferentes), a partir de los avances científicos realizados a comienzos de dicho siglo se obtiene una visión diferente de dicha relación. Aún así, como también llamamos la atención, parece una idea comúnmente aceptada la existencia de un único problema: la relación entre un cuerpo material y una mente inmaterial.

Aunque cierto es que, a grandes rasgos, actualmente la problemática puede seguir reduciéndose a un único problema (¿cómo encaja la supuesta existencia de una entidad como la mente en un mundo físico?), la nueva concepción de la física permite incluir dicha existencia dentro de la materialidad. A la luz de dichos descubrimientos, no es de extrañar que la anterior concepción de alma como entidad trascendental sea sustituida progresivamente por una noción no tan metafísicamente anclada a la irrealidad material, sino que cada vez adquiere mayor relevancia la posibilidad de una relación de causalidad entre lo material y lo espiritual mediante la paulatina asimilación de dicha entidad como una mente (entidad también difícil de describir materialmente) con implicaciones causales sobre el cuerpo. Y la mente entendida en un sentido abstracto es cada vez más entendida como cierto tipo de estados materialmente dependientes que, sin embargo, tienen implicaciones que pueden ir más allá de la materialidad. Dado este contexto, aparece la noción de conciencia tal y como se entiende actualmente. En tanto que estados mentales que, de un modo u otro, están presentes en el sujeto y que, a pesar de depender de ciertos elementos físicos, no pueden ser reducidos a éstos porque aparecen de un modo subjetivo al sujeto de conciencia.

En este sentido, como indicamos, parece plausible suponer con Chalmers la existencia diferenciada de un problema ‘difícil’ de la conciencia y de una serie de problemas ‘fáciles’. Los segundos se refieren a problemas objetivos respecto del estudio de la base cerebral de la conciencia y que se refieren a la posibilidad de que los estímulos ambientales aporten información que el cerebro integra y sintetiza para producir estados, sensaciones, internos al sujeto. El primero se refiere a la posibilidad de que dicho procesamiento de información llevado a cabo por una base material, el cerebro, cause la existencia de una vida mental separada de toda materialidad, en tanto que dicha vida mental mostraría una realidad interna, subjetiva, que permitiría a todo individuo sentir de manera diferente las mismas fuerzas físicas que interactúan con nosotros.

Diferenciamos dos posiciones básicas a la hora de afrontar este segundo problema. Por un lado, desarrollamos los intentos de aquellos que, como Nagel y Jackson, defienden la importancia de dichas características subjetivas últimas, una defensa que, en última instancia, insinuaría la existencia de dos entidades: una mental, los elementos subjetivos anteriormente mencionados, que permitiría explicar la presencia de sensaciones subjetivamente relevantes

pero que no sería reducible a la segunda entidad, la material, la base física del cerebro y del sistema nervioso. Visto de este modo, y en palabras de Levine, deberíamos presuponer la existencia de un hiato explicativo entre lo mental y lo físico, en tanto que lo mental dependería causalmente de la base cerebral pero no sería reductivamente explicable por lo físico. Por otro lado, llamamos la atención sobre diferentes posiciones naturalistas, que aunque no niegan la existencia de elementos mentales consideran oportuno apostar por la reducción de éstos a elementos puramente físicos y, algunos, apostarían por la eliminación de la explicación en términos mentales de los primeros en beneficio de su explicación en términos físicos.

De todas maneras, y como dejamos constancia desde el principio, el propósito de dicho recorrido histórico por las concepciones mantenidas a lo largo de los años, desde la Antigüedad hasta nuestros días, respecto de la relación entre el cuerpo y la mente era llamar la atención sobre lo problemática que resulta la apelación a la presencia de elementos mentales con carácter fenoménico, a la presencia de la conciencia entendida como estados cualitativos subjetivamente dependientes. Y, sobre todo, a esa serie de problemas ‘fáciles’ que nos interesaría resaltar por ser extremadamente importantes a la hora de poder explicar la relación existente entre dichos estados y la realidad física en la que se sitúan.

Teniendo en cuenta dicha dificultad, dedicamos el apartado 1.4. a elencar una serie de problemas que toda teoría de la conciencia debería ser capaz de, por lo menos, explicar. A partir de un anterior intento de ofrecer una lista de problemas, el que realiza Michael Tye en su *Ten Problems of Consciousness*, decidimos ampliar su lista a 15 problemas básicos que podríamos identificar en el estudio de la conciencia, matizando algunos de los ya consideramos en la mencionada obra y ampliando las perspectivas ontológicas que aparecen en su elección a una serie de dificultades epistemológicas no tenidas en cuenta de manera explícita en la mencionada obra aunque sí aparecen implícitas en ella (y cuya elaboración a partir de la misma la realiza en obras posteriores). Lo que nos interesaba destacar con dicha lista de problemas era, por supuesto, la dificultad que la mayoría de teorías actuales sobre la conciencia, especialmente aquellas de corte fisicalista, encuentran a la hora de tener que describir el funcionamiento de la conciencia dentro de un mundo puramente físico, dado que ésta parece ser puramente subjetiva y situada en una perspectiva en primera persona. Llamamos la atención sobre el clásico problema del conocimiento reformulado por Jackson. Personificado en Mary, una neurocientífica criada en un ambiente monocromático que ha llegado a convertirse en la mayor experta mundial respecto de la física del color. Lo que este problema pretende resaltar en la relevancia de la perspectiva personal y subjetiva de los estados cualitativos, en tanto que parece razonable suponer que una explicación fisicalista de dichos elementos subjetivos dejaría por explicar, precisamente, la riqueza de dichas sensaciones cualitativas, accesibles tan sólo en primera persona al propio sujeto, que todo ser consciente parece experimentar ante una determinada sensación. Lo mismo nos sucedería ante el experimento mental ofrecido por Nagel. Aunque consigamos describir el modo en que un murciélago es capaz de sentir (saber-cómo), todavía nos faltaría el sentir como lo hace un murciélago (saber-qué).

Estrechamente relacionado con los dos anteriores problemas, presentamos la posibilidad tenida en cuenta por aquellos filósofos antireduccionistas de la existencia de un hiato explicativo, esto es, la posibilidad de que una explicación puramente física no pueda rendir cuentas de la riqueza de las experiencias conscientes. Aunque también reivindicamos la posibilidad (y acabaremos defendiendo precisamente una posición que, aunque no es reduccionista, consigue superar el mencionado hiato) de poder cruzar dicha separación a partir de explicaciones fisicalistas, en tanto que la posibilidad de la existencia de dicho hiato dificultaría en demasía el papel causal que lo fenoménico parece tener (nuestro cuarto problema).

Nos detenemos algo más en el problema de los duplicados físicos, por ser un argumento especialmente importante, sobre todo por la gran cantidad de argumentaciones que se han sucedido a lo largo de los años tanto a su favor como en su contra. La posibilidad de la existencia de seres aparentemente conscientes, molécula por molécula idénticos a un ser con conciencia, pero que sin embargo carecen del contenido cualitativo que todo ser consciente posee ante la experiencia de una determinada sensación plantea un especial problema a aquellas teorías fisicalistas que sostienen la posibilidad de identificar lo mental con elementos físicos. Aunque es un argumento más lógico que real, su implicación no decrece. Dado que aunque tan sólo sea de manera lógica, la posible existencia de *zombies* (como son denominados en la literatura) parece reivindicar la existencia de un elemento esencial que los seres conscientes poseen, a diferencia de los seres no-conscientes, por lo que aquellos que consideran que la posesión de determinados estados físicos es suficiente para poseer elementos conscientes (y mentales) estarían en un error. Nos resistimos, por supuesto, a aceptar dicha conclusión y, como mostramos en el capítulo final, aunque consideramos plausible (no sólo de modo lógico sino, incluso, materialmente) la existencia de *zombies*, nos parece profundamente equivocada la conclusión a la que llegan sus defensores.

Un capítulo aparte hubieran merecido los tres problemas que aquí resumimos. El espectro invertido, la variación estándar y la visión borrosa se refieren, todos, a la posibilidad de que ante la misma sensación dos sujetos conscientes normales puedan percibir experiencias subjetivamente diferentes. Los defensores de dicha tesis consideran dicha posibilidad suficiente para desmentir que la conciencia pueda ser explicada de manera física, optando por definir la conciencia como un elemento subjetivo y diferente de todo elemento físico. Siguiendo con nuestra posición fisicalista, aportamos alguna pista que nos llevará (en el capítulo final) a desmentir dicha posibilidad y a asimilar dichas dificultades a determinadas anomalías perfectamente explicables desde el plano puramente físico.

Los problemas de la transparencia y de la luminosidad llaman la atención especialmente sobre la dificultad de la posición epistemológica del sujeto respecto de sus propios estados conscientes. Es decir, a diferencia de aquellos que consideran que la experiencia es accesible de modo inmediato al sujeto, de modo subjetivo, es probable que exista un problema a la hora de determinar, primero, la posibilidad de que seamos conscientes de todos nuestros estados fenoménicos (transparencia) y, segundo, además, de poder determinar si dichos estados son

realmente conscientes (luminosidad). Una aproximación fisicalista a los mismos permitiría, al menos, mitigar dicha dificultad por permitir dejar en manos de la contrastación empírica la posibilidad de que realmente lleguemos a conocer la existencia de nuestros propios estados conscientes.

Al presentar el problema de la localización del contenido fenoménico advertimos la dificultad presente a la hora de determinar en qué lugar podríamos situarlo. ¿Dónde estaría dicho contenido? Si consideramos que el contenido fenoménico es un objeto mental, entonces debería estar situado en nuestra mente. Pero entonces deberíamos lograr especificar en qué sentido deberíamos entender dicha mente para poder contener elementos causales. Si, como realmente defendemos, entendemos dicho contenido como dependiente de elementos físicos, entonces el problema virtualmente desaparece al poder explicar la presencia de contenido cualitativo en los sujetos conscientes a partir de la presencia de elementos de bajo nivel, subdóxicos, que nos presentarían de modo objetivo las características de nuestras experiencias. Pero dicha problemática se vería acrecentada por la existencia de anomalías, como la presencia de miembros-fantasma y de percepciones sin percepción. Desde un plano puramente subjetivo, la sensación proporcionada por una pierna amputada (una posibilidad ya tenida en cuenta en sus *Meditaciones* por el propio Descartes) o la percepción sin conciencia de obstáculos en el camino de un paciente de *blindsight* es difícilmente explicada sin apelar a una entidad separada de la física. Desde una explicación fisicalista, apelar a lesiones cerebrales o anomalías en el campo sensitivo tampoco mitiga la dificultad que acompaña a dichas posibilidades, pero permite una explicación de la misma a partir de elementos físicos que no requerirían multiplicar las entidades metafísicamente relevantes. Del mismo modo, la posibilidad de que exista un eje físico de procesamiento de la información, el cerebro, permitiría también aportar una explicación física a la unidad de la experiencia consciente, sin necesidad de apelar a entidades inmateriales, algo que garantizaría también el soporte idóneo para generar una imagen personal suficientemente robusta para soportar el peso metafísico de la unidad personal, la agencia y la responsabilidad de nuestras acciones.

* * *

Al comienzo del capítulo 1, realizamos una diferenciación crucial que presentaba, precisamente, las peculiares características del contenido de los estados conscientes. Dijimos que los estados mentales pueden definirse a partir de su carácter relacional, esto es, por la relación existente entre un agente psicológicamente dotado y sus contenidos mentales, entendidos como proposicionales, y que podríamos entender los estados conscientes también como relacionales. Pero, a diferencia de otros estados mentales, los estados conscientes se caracterizarían por ser la relación entre un agente capaz de darse cuenta de sus propiedades psicológicas y el contenido de dichos estados, un contenido que no parece poder reducirse a mera proposicionalidad, sino que debería entenderse en un sentido fenomenológico, esto es, inmediato al sujeto. Y como todo estado mental, también los estados mentales con carácter fenoménico estarían dotados de una cierta normatividad. La dificultad contemplada en ese momento era, precisamente, estipular el modo correcto de entender dicha relación, dado que a

diferencia de estados mentales como las actitudes proposicionales, los estados mentales con carácter fenoménico no parecen responder de modo proposicional. Si pudiera entenderse dicha relación como proposicional, entonces parecería perderse la gran riqueza experiencial que los individuos conscientes tenemos. Por dicho motivo propusimos la posibilidad de que dicha relación debiera entenderse de otro modo.

Propusimos también la revisión de las dos formas tradicionales en las que la literatura científica trata dicha relación. Por un lado, contemplamos la posibilidad de que el contenido de los estados con carácter fenoménico fuese entendido como conceptual, esto es, dependiente de una categorización objetiva externa al propio sujeto. Contrapuesta a esta posibilidad, por otro lado, también analizamos los intentos de aquellos que entienden dicho contenido como no-conceptual, esto es, ya no dependiente de categorizaciones externas sino dependiente de la distribución subdoxástica interna a los propios individuos conscientes.

Llamamos especialmente la atención sobre un principio básico, el Principio de Autonomía (PA), que Bermúdez atribuye al contenido de los estados conscientes. Dicho PA estipula la posibilidad de que una criatura no requiera de concepto alguno para tener estados mentales conscientes, lo que relegaría el contenido de dichos estados a una relación con elementos previos a toda categorización. Contrapusimos a la concepción de Bermúdez la posición débil de Peacocke, una posición que defiende la presencia (la mayoría de las veces) de algún tipo de elemento conceptual que permitiría afirmar en el contenido de los estados mentales la presencia de elementos no-conceptuales, con lo que se establecería una dependencia del contenido no-conceptual sobre el conceptual, negando así PA. La dificultad sobre la que llamamos la atención en dicha polémica es que, si por un lado, sostenemos una posición fuerte como la de Bermúdez, el contenido no-conceptual quedaría desconectado por completo del contenido conceptual, por lo que se perdería la posibilidad de que, en un momento determinado, pudiéramos apelar a elementos conceptuales para fundamentar la presencia de elementos no-conceptuales, elementos de bajo nivel, subdoxásticos, que permitieran afirmar la posesión de determinados estados mentales por parte de sujetos conscientes. Pero si, por otro lado, defendiéramos una posición cercana a la de Peacocke, entonces, aunque pudiéramos establecer esa relación (al menos, en principio, deseable) entre contenido conceptual y no-conceptual, lo que nos encontraríamos es con una nueva entidad metafísicamente relevante que (como ocurre con los datos de los sentidos) dificultaría e incrementaría innesariamente nuestros elementos reales. Concluimos, entonces, en dicho momento, la necesidad de rechazar dichas concepciones en beneficio de dos alternativas viables. La primera, la defensa de una posición representacional intermedia (que también desechamos a lo largo de la sección 2.1.3.2.) o bien, la segunda, un progresivo abandono de las posiciones representacionales (algo que intentamos realizar a lo largo de todo este trabajo) en beneficio de una posición fisicalista que sea capaz de rendir cuentas del contenido de los estados mentales con carácter fenoménico.

La primera opción, una posición representacional respecto del contenido fenoménico, como dijimos, tiene por objetivo básico evitar en la medida de lo posible extender la cantidad de entidades metafísicamente relevante en beneficio de un nivel puramente representacional que

permita servir de base para el contenido de las experiencias fenoménicas. Destacamos dos estrategias básicas. En primer lugar, analizamos la propuesta de Fodor. Concluimos que su posición, la posibilidad de la existencia de un contenido conceptual, era excesivamente restrictiva, en tanto que la posesión del mismo concepto por parte de dos sujetos es lo que permitiría determinar que se encuentran en el mismo estado. Esto es, que el contenido fenoménico sería reducido a elementos puramente conceptuales, en tanto que lo que determinaría el estado mental en el que se encontraría un sujeto en un momento concreto dependería del papel funcional del concepto que lo implementara, siendo la naturaleza de dichos conceptos pública. La segunda estrategia persigue el mismo fin, explicar el contenido fenoménico de las experiencias conscientes a partir también de elementos representacionales, pero sin fundamentarlo en elementos conceptuales, sino fundándolo en elementos no-conceptuales. Evitando cualquier tipo de mención al contenido conceptual, lo que permite esta estrategia es identificar el contenido fenoménico con elementos subdóxicos, siendo relevante tan sólo para determinar el contenido de una experiencia la existencia de mecanismos universales de interacción entre sujeto y entorno.

La segunda opción, la que pretendo defender, como consecuencia directa de nuestro rechazo de PA, apuesta por la interdependencia entre el contenido conceptual y el contenido no-conceptual de los estados mentales con carácter fenoménico. Dicha concepción afirmaría que para que pudiéramos decir que un sujeto se encuentra en un determinado estado con carácter fenoménico sería suficiente la presencia de elementos relaciones de nivel primario, subpersonal, de elementos cognitivos capaces de captar, sintetizar y manipular la información aportada por el entorno. Pero que, para que dicho estado fuera realmente consciente y pudiera ser tenido como un estado de un tipo concreto, para que el sujeto adquiriera un conocimiento respecto del mismo y pudiera catalogarse dicho estado, requeriríamos de un concepto que permitiera categorizar dicha información dentro de nuestra imagen de la realidad. Lo que estaríamos proponiendo es una explicación del contenido de los estados con carácter fenoménico como la confluencia de dos clases diferentes de elementos. Por un lado, la existencia de una serie de estados informacionales (sin contenido previo) sintetizados y procesados por elementos neurofisiológicos a partir de los datos captados por una serie de sistemas sensitivos que permitiría, por otro lado, el desarrollo histórico y cognitivo particular de cada individuo en su entorno concreto a partir de la historia de entrenamientos cognitivos que haya recibido a lo largo de su desarrollo vital. El error sería presuponer que, al considerar su posible interdependencia, estamos afirmando que podrían identificarse. Pero nada más lejos. Lo que pretendemos con dicha declaración es encontrar una explicación plausible, a la vez, a la intuición cotidiana de que los seres conscientes compartimos experiencias y permitir, al mismo tiempo, salvaguardar la riqueza individual que para cada uno de nosotros supone tener dichas experiencias.

* * *

También realizamos una diferenciación entre dos tipos distintos de conciencia. Siguiendo las definiciones marcadas inicialmente por Block, distinguimos entre una conciencia fenomenológica y una conciencia representacional. La diferenciación original de Block

pretendía llamar la atención sobre dos aspectos relevantes, pero distintos, de la conciencia: las vivencias y el contenido de las mismas. Para Block, la conciencia fenomenológica es el conjunto de estados mediante el que los seres conscientes nos damos cuenta de nuestras propias vivencias, siendo por tanto indistinguible de las experiencias mismas. La conciencia representacional, en cambio, es el conjunto de estados que nos proporcionaría acceso a los contenidos informaciones de dichas experiencias. A diferencia de la primera, que sería marcadamente causal, la conciencia por acceso, por su contenido representacional, tiene un papel funcional al permitir nuestra acción a partir de la imagen que de nuestro mundo realiza.

Aunque aceptamos la diferenciación propuesta por Block, dado que consideramos que las dos nociones de conciencia resaltan aspectos diferentes, también defendimos que sería conveniente mantener una posición integracionista respecto de la conciencia. Esto es, defendimos la posibilidad de que las dos clases de conciencia diferenciadas por Block tan sólo sean dos aspectos diferentes de una única conciencia. Así evitábamos la tendencia segregacionista actual y evitábamos también las dificultades que se desligan de tener que relacionar estas dos nociones diferentes. Al apostar por esta tendencia de integración, entonces, también tuvimos la necesidad de matizar ciertas propiedades que parecen tener ambos aspectos. A diferencia de Block, nosotros defendemos la posibilidad de que la conciencia tenga un aspecto de acceso que nos informe de qué tipo de experiencia estamos teniendo en un momento determinado, pero también un aspecto fenomenológico que nos permita determinar la experiencia misma. Obviamente, el acceso a la información permite al sujeto consciente poder actuar de modo racional, dado que dicho aspecto de la conciencia le proporciona, funcionalmente hablando, toda aquella información requerida para poder forjar sus propias estrategias. Pero también llamamos la atención sobre el aspecto fenomenológico de la conciencia por dicho motivo, dado que no está del todo claro que la conciencia fenomenológica no pueda tener un papel funcional en la acción. Dados los datos cotidianos, parecería desprenderse lo contrario, en tanto que la percepción de determinadas características cualitativas puede ser relevante para un sujeto en un momento determinado y puede permitir la formación de nuevas estrategias de acción.

Si esto fuera así, como defendemos, entonces se vuelve necesario diferenciar dos aspectos importantes dentro de la misma experiencia. Por un lado, el carácter cualitativo de la misma y, por otro, su carácter subjetivo. El carácter cualitativo (los *qualia* concretos de los que se trata) es el modo concreto en que una experiencia es capaz de representar el mundo para un sujeto concreto. El carácter subjetivo aporta el componente atencional, en tanto que si la experiencia se presenta a un sujeto, el sujeto debe ser capaz de darse cuenta de dicha experiencia. Por tanto, este hecho parecería demostrar la existencia de una estrecha relación entre ambos aspectos de la experiencia.

La dificultad sobre la que llamamos la atención cuando defendimos esta relación era mostrar en qué sentido podemos defender que un estado mental pueda estar incluido en su propia vivencia. Esto es, que aparte de la vivencia de una experiencia, también podamos diferenciar un aspecto individual de relevancia para un sujeto en dicha experiencia. Sobre todo

si, como pretendemos, defendemos una posición fiscalista de dicha relación. Para intentar solucionar dicha problemática, decidimos realizar un pequeño rodeo y presentar los intentos que ciertas teorías representacionistas y no-representacionistas respecto de la conciencia han realizado en los últimos años para intentar superarla con la intención de analizar los beneficios y los inconvenientes de sus conclusiones.

* * *

En el apartado 2.1.2. atendimos brevemente a ciertos planteamientos no-representacionales de explicación de la conciencia. Como ya dijimos allí, todos ellos tienen en común una tendencia naturalizadora que pretende aportar una explicación en términos físicos a la presencia de determinado contenido fenoménico. El punto básico de dichos intentos, lo que precisamente marca su carácter no-representacional, es la idea de que se puede ofrecer una explicación del carácter subjetivo y cualitativo de las experiencias a partir de la descripción de la relación existente entre un sujeto concreto y sus propias vivencias, en tanto que dicha relación sería directa, cognitiva, porque la presencia de un determinado estado mental con un cierto contenido fenoménico parecería ser suficiente para que el sujeto pueda ser consciente del mismo (siempre que dicho sujeto disponga de los mecanismos suficientes para captar dicho contenido).

Dos son las estrategias básicas en las que nos detuvimos. Por un lado, la posición conceptualista de Gennaro y, por otro, las tesis no-conceptualistas ofrecidas por ciertos funcionalistas y ciertos cognitivistas. Para Gennaro, experiencia y vivencia irían de la mano, nunca podrían ser considerados como estados separados. Son dos componentes del mismo estado. Pero vimos que esta teoría adolece de varias dificultades. En primer lugar, Gennaro no consigue demostrar que el carácter subjetivo de la conciencia sea conceptual. Segundo, no consigue describir adecuadamente dicha interdependencia entre vivencia y experiencia. Y tercero, que sea realmente así no quiere decir que un sujeto concreto sea capaz de saberlo, por lo que de poco serviría poder determinarlo de modo objetivo, dejando de nuevo abierta la puerta de la existencia de un hiato explicativo.

La segunda estrategia, la sostenida por funcionalistas y cognitivistas, define la relación entre vivencia y experiencia en términos no-conceptuales. Para los primeros, los estados mentales pueden ser individualizados según el papel funcional que desempeñan en la vida mental de los sujetos. De este modo, es conveniente separar entre papel funcional y contenido, en tanto que el papel funcional sería una abstracción de las causas y los efectos posibles que un estado podría provocar y el contenido es el estado neurofisiológico del sistema nervioso central capaz de provocar dicho estado. Entonces, el estado neurofisiológico es capaz de realizar el estado mental concreto en tanto que es capaz de ocupar el papel funcional adecuado para implementar dicho estado. La dificultad principal que denunciamos en dichas aproximaciones es el modo en que podemos describir cómo dicha interdependencia entre papel funcional y contenido permite mostrar la relación existente entre vivencia y experiencia, entre fenomenología y carácter subjetivo. Una respuesta es la aportada por Carruthers, que reivindica una relación semántica entre el papel funcional y el contenido neurofisiológico, una relación que permitiría generar pensamientos de segundo orden (el carácter subjetivo) dependientes de

pensamientos de primer orden (experiencia directa) capaces de hacer consciente al sujeto de sus propias sensaciones. El inconveniente de dicha posición (como también de las posiciones representacionales de nivel superior, como la de Rosenthal, que analizamos también en este punto) es que, defendiendo esta diferenciación entre niveles, estaríamos presuponiendo la existencia de todo un nivel periférico de estados primitivos que, potencialmente, permitirían que el sujeto fuera capaz de darse cuenta de todo aquello que le ocurre. Pero no tenemos más remedio que afirmar su incoherencia, al ser excesivamente amplio el espectro de posibilidades de conciencia que abre dicha suposición y por multiplicar demasiado *in extenso* las sensaciones que seríamos capaces de captar al mismo tiempo. Intentando salvar esta objeción, encontraríamos la propuesta funcionalista de Shoemaker. Al evitar diferenciar entre la experiencia y la vivencia, propone que la determinación del estado neurofisiológico capaz de ocupar el papel funcional adecuado que permitiría implementar el estado mental concreto tan sólo dependería de una simple constatación empírica. Cierto es que al evitar la existencia de niveles semánticos, Shoemaker consigue mitigar la dificultad anterior. Pero ello no le exime de inconvenientes. Si consideramos que tiene razón, entonces, deberíamos considerar la posibilidad de que el papel funcional fuera capaz de implementar un contenido intencional, pero no (como sería su intención) que constituya el contenido intencional mismo. Si esto fuera posible, entonces deberíamos presuponer una tercera instancia que permitiera hacer de puente entre el estado mental concreto y el contenido neurofisiológico del mismo, por lo que violaríamos una de las leyes que aceptamos desde el principio (la simplificación de las entidades metafísicamente relevantes) y tornaríamos inconsistente la misma tesis funcionalista. Pero, además, si esto fuera así, deberíamos presuponer que el papel funcional estaría dotado de características disposicionales, pero no intencionales, en tanto que la separación existente entre contenido y papel funcional sería establecida tan sólo a partir de una relación de contingencia que permitiría en un momento determinado que el sujeto estuviera capacitado para captar determinadas sensaciones (tener ciertas experiencias) pero no nos indicaría nada respecto a la posibilidad de que dicho sujeto pudiera ser capaz de darse cuenta de la presencia de vivencias respecto de dicha experiencia.

Para los segundos, los cognitivistas, sería posible encontrar en la naturaleza el carácter intencional de la experiencia misma. Ésta es, por ejemplo, la propuesta de Baars. Para él, sería posible presentar una modelización cognitiva de la dinámica causal que estaría implicada en los procesos cognitivos capaces de explicar las causas del comportamiento de los agentes, en tanto que este tipo de construcción teórica sería fácilmente inferible a partir de la evidencia misma. La conciencia sería, entonces, una noción operativa. Si esto es así, entonces, como aceptamos, la probabilidad de que un suceso mental concreto pueda llegar a ser consciente depende de su valor informativo, en tanto que sería capaz de proporcionar los datos necesarios para que un sujeto concreto pudiera, por un lado, tener experiencias (a partir de la existencia de estados neurofisiológicos que le permitirían directamente captar dicha información sensible) y, por otro lado, llegar a saber qué tipo de experiencia está teniendo en cada momento (a partir de la manipulación y sintetización de la información relevante captada por los sistemas cognitivos).

La dificultad principal, como ya observamos, que podríamos imputar a la teoría cognitiva, en general, y a la de Baars, en particular, es que parece dejar fuera de su análisis a la conciencia fenomenológica. El modelo cognitivo, como hemos visto, permite explicar el acceso del sujeto a la información necesaria, pero no es capaz de explicar satisfactoriamente por qué dichos procesos cognitivos deben siempre estar acompañados de una experiencia consciente, perdiendo la posibilidad de que un sujeto pueda tener vivencias al mismo tiempo.

* * *

A lo largo del apartado 2.2. expusimos las diferentes aproximaciones representacionistas que actualmente dominan el panorama filosófico en cuanto al estudio de la conciencia. Una teoría representacional cree posible individuar los estados mentales en términos de su contenido representacional. En este sentido, diferencian entre el objeto representado y el vehículo mediante el cual se realiza dicha representación. Normalmente, asocian la presencia de determinados estados neurofisiológicos del sistema nervioso central (el vehículo transmisor) con la posibilidad de captación de determinadas características físicas del entorno (aquello que acabará siendo representado), estableciendo entre ambos, entre vehículo y contenido, una relación física. Pero, aunque toda explicación representacional parece defender de alguna forma esta relación física, dependiendo del modo en que sea entendida la misma las conclusiones que las diferentes posiciones mantendrán varían.

Al analizar las teorías informacionales (de las cuales la de Dretske es la más relevante) advertimos que lo que realmente persiguen es una naturalización de la mente y la conciencia que permita demostrar cómo el carácter cualitativo de las experiencias estaría determinado, primero, por la información recibida por los sistemas perceptivos (y por sus posteriores elaboraciones) y, segundo, por el modo en que dicha información se presenta a dichos sistemas perceptivos. Si esto es así, entonces, aquello que marcaría las diferencias entre las experiencias sería la existencia de diferentes procesos cognitivos particularmente desarrollados según su propia historia cognitiva capaces de elaborar y procesar de manera diferente la información recibida del exterior. La dificultad que destacamos en este tipo de posición se deriva, precisamente, de entender esta relación entre contenido y estado mental en un sentido funcionalista, en tanto que exigiría una excesiva determinación y una especialización demasiado restrictiva a los diferentes sistemas de procesamiento, impidiendo la posibilidad de que funcionen de manera adecuada en determinados casos donde no está definida del todo la clase de información aportada.

También destacamos la existencia de teorías teleológicas. Éstas, a diferencia de las anteriores, defienden que aquello que permitirá determinar la relación existente entre un contenido y un estado representacional son los procesos evolutivos que han permitido su desarrollo, consiguiendo de este modo establecer las funciones propias mediante el proceso llevado a cabo por la selección evolutiva, lo que determinaría su normatividad. Este es el caso de Millikan. Por su parte, Papineau reconoce la existencia de este tipo de funcionalidad, aunque cree posible que en determinados momentos puedan existir en la vida individual de cada sujeto una serie de modificaciones que establezcan nuevas disposiciones conductuales, mejoras

estructurales que permitirían un mejor desarrollo de los mecanismos de toma de decisión y de razonamiento para la consecución de objetivos, lo que permitiría una mejora sustancial en las respuestas de los organismos. Este tipo de explicación adolece de la misma dificultad que la anterior, aunque su funcionalidad sea entendida de un modo más biológico que evolutivo. La exigencia de determinación de la relación entre contenido y estado obligaría a presuponer una relación demasiado estrecha en los diferentes procesos de generación funcional, por lo que la función propia de cada sistema sería excesivamente restrictiva, siendo imposible que actúe en casos en los que la información no esté definida claramente.

El primitivismo considera que los objetos disponen de características que pueden aparecer en nuestra experiencia, pero que no son disposicionales. Es decir, pensando en el ejemplo de contemplar un tomate rojo, presuponen de un modo realista que el tomate tiene la propiedad de ser rojo, pero que dicha propiedad no es una disposición para que captemos el tomate como rojo, sino que dicha propiedad superviene sobre otras propiedades físicas, lo que permitiría que emergiera en los sujetos la experiencia de rojo al contemplar un tomate. La relación entre contenido y estado, aquí, entonces, aparece como una relación de dependencia, siendo necesario considerar, primero, la existencia de propiedades reales capaces de supervenir sobre otras propiedades y, segundo, dejando en manos del sujeto la posibilidad de que emerjan dichas experiencias siempre que su sistema perceptivo sea capaz de captar dichas propiedades. Esto último, precisamente, es una dificultad añadida, sobre todo si tenemos en cuenta que en la percepción de determinadas características es irrelevante la función concreta que dichos sistemas puedan tener.

Por su parte, las posiciones intencionalistas consideran que el carácter fenoménico de ciertos estados mentales es lo mismo que tener cierta clase de contenido intencional. Así, el intencionalismo defiende que el contenido proposicional de las experiencias perceptivas determinaría su propio carácter fenoménico, siendo el contenido representacional de las experiencias el que determine el tipo de experiencia que se está teniendo en un momento determinado. Si esto es así, entonces existe una asimilación entre contenido y estado, dado que depende del contenido de dicho estado el que el sujeto se encuentre en un estado o en otro. Dependiendo del papel funcional que se le otorgue a cada contenido (en palabras de Lycan y Prinz) o del tipo de intención que determine el contenido representacional de un estado determinado (según Tye y Byrne), estaríamos hablando de un tipo diferente de estado que se implementa en un sujeto. La dificultad subyacente a esta teoría es, precisamente, poder determinar de manera concreta cómo todo estado fenoménico puede ser intencional, dado que dicha relación dependería de la presencia de elementos puramente no-conceptuales y, como venimos defendiendo, a menudo la mera apelación a elementos no-conceptuales no es suficiente para dicho propósito. Deberíamos suponer la posibilidad de la existencia de elementos conceptuales que, al menos en determinados momentos, pudieran ayudarnos a determinar el tipo de estado en el que nos encontramos dada una cierta experiencia.

Para simplificar la visión representacional del contenido fenoménico realizamos una diferenciación entre las diferentes perspectivas según su concepción del carácter cualitativo de

las experiencias. Aunque toda posición representacionista considera que un *quale* son las propiedades de los objetos representadas en la experiencia y dados al sujeto de un modo concreto, existe una ambigüedad a la hora de emplear dicho término. Como dijimos en su momento (apartado 2.2.3.), ciertos posicionamientos emplean indistintamente el término *qualia* para referirse tanto al carácter subjetivo como al carácter cualitativo de las experiencias. Pero, como ya avanzamos, aunque ambos son aspectos del mismo estado, deberíamos poder diferenciarlos en tanto que el primero se refiere al modo particular en que la experiencia se presenta a un sujeto y el segundo indicaría el contenido particular representado por una determinada experiencia. Esto es, deberíamos poder distinguir entre el modo en que un sujeto es capaz de experimentar una experiencia concreta (el aspecto subjetivo) y las propiedades cualitativamente relevantes de una experiencia concreta (las propiedades del *quale* particular percibido, que serían diferentes de la experiencia misma). La razón principal por la cual consideramos necesario diferenciar entre el carácter cualitativo y el carácter subjetivo de los *qualia* es porque el segundo (el *quale* concreto) podría ser perfectamente explicable desde un lenguaje público, objetivo, en tanto que no requiere de la posibilidad de que un sujeto sea consciente del mismo para que podamos decir que se encuentra en dicho estado concreto, mientras que el peculiar carácter cualitativo de una experiencia determinada viene asociado directamente con la presencia de un sujeto para el cual dicho carácter se presenta, siendo inefable dicho carácter. Si estamos en lo cierto, entonces, las teorías representacionales del contenido de las experiencias fenoménicas son capaces de explicar tan sólo el segundo aspecto, el carácter cualitativo de las experiencias, en tanto que una aproximación representacional del mismo, al menos en algunos casos, permite salvaguardar la presencia de un sujeto perceptor y de una serie de características físicas externas que son relevantes para dicho sujeto en un momento determinado. Pero cuando pretenden asimilar dicho aspecto cualitativo con el aspecto subjetivo, entonces parecen incurrir en una inconsistencia, dado que es plausible (al menos en principio) que pueda darse un determinado estado experiencial sin necesidad de que el sujeto se dé cuenta de que lo está teniendo, demostrándose entonces que no existiría el nivel representacional que reivindican estas teorías.

Tan sólo tenemos que pensar en un ejemplo concreto: nuestros casos de visión ciega. Recordemos en qué consisten. Varios experimentos con personas que han sufrido algún tipo de lesión cerebral que ha afectado a su campo visual (es decir, algún tipo de anomalía cerebral que ha causado la pérdida de visión) pero que, sin embargo, tienen intacto su sistema visual (no tienen ninguna disfunción en los ojos y en los nervios ópticos) parecen darnos alguna pista sobre la posibilidad real de separar los dos aspectos de la experiencia sobre los venimos llamando la atención. Un caso concreto lo encontramos en las pruebas realizadas con un paciente de visión ciega al cual se instó a pasar de una habitación a otra a través de un pequeño pasillo donde estratégicamente se habían colocado varios obstáculos. Aunque el sujeto en cuestión no era capaz de 'ver' dichos objetos (esto es, no era consciente de la presencia de los obstáculos interpuestos en su camino), sin embargo fue capaz de atravesar el pasillo sorteando todos los obstáculos. Incluso una vez realizado el experimento e interrogado sobre la posible

existencia de obstáculos en su camino, el sujeto seguía sin ser capaz de expresar ningún tipo de conciencia respecto de la presencia de obstáculos en el pasillo que acababa de recorrer. Otro experimento relevante con otro paciente de visión ciega consistía en presentarle algún tipo de objeto (un vaso en la mayoría de los casos) en la parte ciega de su campo visual. El paciente en ningún caso mostraba la más mínima señal de darse cuenta de lo que se le pedía, dado que aseguraba que no podía ver ningún vaso en la zona en la que se le aseguraba estaría situado dicho objeto. Sin embargo, cuando se le instaba a asirlo, el paciente era capaz de cogerlo sin dificultad, incluso sorprendiéndose de la presencia de un objeto que era capaz de coger pero sobre el cual no tenía ningún tipo de experiencia visual. Pero, ¿qué relevancia puede tener estos casos para nuestro propósito?

Como hemos indicado anteriormente, consideramos necesaria la diferenciación de dos aspectos de la experiencia. Por un lado, un aspecto subjetivo y, por otro, un aspecto cualitativo. Precisamente lo que nos aportan los casos de visión ciega es la prueba empírica de que debe existir dicha distinción y de que, en última instancia, el aspecto subjetivo no parece ser representacional, como insistirían las posiciones que hemos analizado. Porque, como nos dicen los casos de visión ciega, aunque el sujeto es capaz de interactuar con los objetos externos sobre los que no tienen ningún tipo de percepción, objetos que no pueden experimentar, no es capaz de darse cuenta de su presencia. Lo que parecen probar estos experimentos es la existencia de un nivel previo, no-representacional, que permitiría al sujeto tener un acceso subjetivo puramente informacional a los objetos sin necesidad de que dicho acceso trascienda a su nivel personal, a su nivel consciente. Ello parece indicarnos que, a grandes rasgos, existiría cierto tipo de elementos no-representacionales que permitirían determinar el contenido de determinadas experiencias perceptivas independientemente de que el sujeto tuviera un acceso consciente a las mismas e independientemente de que el sujeto pudiera interactuar con los objetos de su entorno sobre los cuales no tiene una experiencia consciente.

* * *

Dedicamos el capítulo 3 por completo a analizar la posición representacional de John R. Searle. Éste es uno de los filósofos que con mayor fuerza retoma la lucha y la defensa por situar a la conciencia de nuevo dentro de la vida mental de los sujetos, en tanto que pretende reivindicar el papel central que tiene dentro de una explicación naturalizada de la mente. Sostendrá Searle la imposibilidad de poder encontrar una mente sin conciencia, lo que demuestra su importancia dentro de cualquier teoría de la mente que pretenda ser completa.

En la teoría de Searle, la mente y la conciencia son analizadas desde una perspectiva biológica. Su principal pretensión es ver en qué sentido podemos naturalizarlas. Para Searle, la conciencia no es más que un fenómeno biológico producido a partir de la ontología y la peculiar estructura del sistema nervioso central y, principalmente, del cerebro humano. La conciencia es entendida como un producto emergente de los procesos realizados a un nivel inferior en el cerebro y que es capaz de emerger al nivel personal, el nivel de las actitudes proposicionales, de los sujetos. Así, el cerebro segregaría conciencia del mismo modo que el páncreas segrega la bilis. La conciencia, entonces, es una propiedad causalmente explicable a partir de los elementos

constituyentes del cerebro, pero no sería propiedad de ninguno de ellos en particular, sino que tan sólo es definible a partir de considerar todas sus propiedades en conjunto. Y el único modo plausible de abordar su estudio sería a través de una investigación científicamente dirigida a dilucidar la ontología profunda subyacente a la estructura cerebral y de las redes neuronales que permiten su emergencia. Vista la dificultad de la ciencia actual para poder aportar una explicación física de la ontología del aparato cerebral causante de la conciencia, no queda más remedio que renovar dicha perspectiva a partir de los datos intuitivamente accesibles desde el sentido común. De este modo, la conciencia es concebida como el conjunto de elementos fenoméricamente relevantes que componen toda la vida mental de los sujetos (humanos) y que tan sólo es accesible desde la perspectiva en primera persona. Dada la imposibilidad de afrontar su estudio de un modo científico, físico, entonces no quedaría más remedio que explicarla a partir de la información que nos aportan nuestras propias experiencias fenoméricas subjetivamente dependientes.

Una dificultad añadida que deberá salvar toda aproximación a la conciencia, desde el punto de vista de Searle, es el vocabulario dualista que está inserto en la tradición, y que heredamos directamente de la metafísica cartesiana. Según su análisis, todavía en la ciencia moderna podemos encontrar científicos, como Crick y Edelman (como analizamos en los apartados 3.1.1. y 3.1.2.) que pretenden describir el sustrato físico subyacente a la conciencia y que, al mismo tiempo, acaban por defender la posibilidad de un elemento no-físico capaz de sustentar el conjunto de nuestras vivencias conscientes. Al hablar de emergencia, lejos de considerar la existencia de dos entidades separadas, conseguiríamos entender la conciencia como un rasgo esencial del cerebro.

En el estudio de la conciencia, el principal problema a tratar sería la existencia de *qualia*. Para Searle, la conciencia no es más que el conjunto de estados cualitativos, de vivencias subjetivamente dependientes, que son causadas a partir de la base física del cerebro. Cuando surge la necesidad de tener que explicar científicamente cómo el sistema biológico que compone el cerebro y sus redes neuronales es capaz de causar el conjunto de nuestra experiencias subjetivas, esto es, cuando tenemos que explicar cómo la química, física y biología de nuestro sistema nervioso central es capaz de causar la conciencia, aparece una nueva dificultad, en tanto que tenemos que poder explicar cómo de una realidad objetiva puede emerger una serie de cualidades subjetivas. Los intentos de aquellos que, como Penrose, apuestan por un análisis del nivel micro de la física del cerebro acaban por malentender el problema al multiplicar los ya difíciles problemas que aparecen al tener que dar cuenta del carácter subjetivo de la conciencia a partir de la base objetiva del cerebro con nuevos problemas desligados de la ontología de nivel cuántico identificada de modo esencial en la física del cerebro.

Pero todavía peor, para Searle, se plantea el estudio de la conciencia cuando, como Dennett y Chalmers, se acepta como válida la metáfora computacional de la mente. Al considerar al cerebro como el *hardware* donde se implementa, como un *software*, la conciencia no hacemos más que complicar la difícil situación de tener que explicar la aparición de las

cualidades subjetivas a partir de la base objetiva del cerebro, dado que damos por supuesta la aparición de la semántica a partir de la sintaxis. Contra esta posibilidad, Searle postula su famoso ejemplo de la habitación china, un argumento que demostraría, al menos según los términos searleanos, la imposibilidad de que un mero sistema manipulador de símbolos fuera capaz de generar significado.

Podríamos resumir la posición inicial de Searle sobre la conciencia, a la espera de atender a su naturalismo biológico, del siguiente modo. Para Searle, que la conciencia sea una cuestión de experiencia subjetiva y personal es tan sólo una simple cuestión empírica. Aunque dicha noción de conciencia también se encuentra mediada por la acción social, lo que permitiría explicar la presencia en todos los seres humanos de determinadas vivencias fenoménicas involucradas en el carácter cualitativo de ciertas sensaciones, si la conciencia es una mera cuestión empírica, entonces puede ser explicada a partir de las ciencias físicas. Pero que ello sea posible no quiere decir que la conciencia sea una noción reducible tan sólo a elementos físicos. Aunque es cierto que la conciencia, según Searle, tiene profundos niveles funcionales dependientes causalmente de la estructura biológica del sistema nervioso central, la actividad consciente aparece a un nivel subjetivo, por lo que debe representar un nivel ontológico profundo que esconda la verdadera naturaleza de los fenómenos conscientes. En su afán por descubrir la naturaleza de la conciencia, Searle pretende, por una parte, afirmar la existencia de un nivel ontológicamente profundo, subjetivo y cualitativo, que sin embargo, por otra parte, es dependiente de la fuerte concepción científica del mundo que poseemos. Pero al mismo tiempo, también afirma Searle que la primera no es reducible a la segunda, por lo que a la hora de intentar su aproximación conciliadora, compatibilista (como la hemos definido), no puede dejar de ser parcial al considerar que tan sólo dos grandes orientaciones persiguen la explicación de la conciencia, precisamente aquellas dos orientaciones que o bien acaban por postular un dualismo de propiedades (Crick, Edelman y Penrose) o bien intentan aportar una visión computacional de la conciencia (Dennett y Chalmers, aunque con finalidades diferentes. Eliminativistas, el primero, y dualistas, el segundo). La gran dificultad que, como explicamos, aparece a lo largo del desarrollo de los argumentos de Searle, sobre todo, se refiere a la contradicción que aparece en su descripción de la conciencia, dado que con la intención de mostrar la irreducibilidad de la conciencia a partir de la noción de emergencia, Searle acaba por defender, también, una concepción dualista de la misma que presupone, sin pretenderlo, el mismo vocabulario mentalista cartesiano que tanto afán pone en denunciar.

Naturalismo biológico es el nombre que Searle adopta para definir su proyecto de naturalización de la mente. Los dos rasgos definitorios de la mente, para Searle, son la intencionalidad y la conciencia. Por tanto, en el proyecto searleano, ambas nociones deben ser entendidas como productos biológicos causalmente derivados de la estructura física cerebral, explicables desde el marco objetivo de las ciencias físicas.

Con la noción de intencionalidad, Searle se refiere a aquella característica que presentan la mayoría de nuestros estados mentales de estar dirigidos a objetos y estados del mundo diferentes de ellos mismos. En este nivel es donde podemos descubrir el representacionalismo

de Searle. Si un estado mental determinado tiene un contenido capaz de representar en el mundo interno del mismo sujeto el objeto externo al cual se dirige, entonces el contenido de dicho estado mental debe ser representacional y dependiente del modo en que al sujeto se le presenta dicho objeto. Pero, entonces, surge la pregunta crucial. Si la capacidad representacional de nuestra mente depende por completo de la estructura física cerebral, ¿cómo dicha estructura, objetiva por definición, es capaz de representar de manera subjetiva un determinado rasgo del mundo? El modo de aclarar esta dificultad, en el programa de Searle, pasa por describir el modo en que los fenómenos intencionales son causados por procesos biológicos y cómo éstos permiten la existencia en la mente de ciertas características: la existencia tanto de un modo psicológico representacional como de un contenido (diferente del objeto mismo) en los estados intencionales, la presencia de una dirección de ajuste entre el modo psicológico de los estados mentales y el mundo y la existencia de una serie de condiciones de satisfacción para los estados mentales intencionales.

Como vemos, la posición de Searle, aunque afirma la relación causal entre la biología del cerebro y la presencia de determinados estados mentales, parece defender en realidad un marcado internismo respecto de lo mental, en tanto que la determinación de los estados mentales en los que se encuentra un sujeto puede establecerse a partir de su propia teoría perceptiva. Si, como defiende Searle a partir de su concepción realista y directa de la percepción, la percepción de los objetos materiales que observamos es debida a la existencia de vivencias perceptivas de nuestras propias experiencias en tanto que somos capaces de aprehender la relación causal existente en las que está involucrado dicho acto perceptivo, entonces recaemos en una seria dificultad al pretender que también de los objetos internos de nuestra mente (dependientes de los elementos neurofisiológicos capaces de explicar causalmente su aparición) tenemos un claro acceso consciente a partir de las vivencias que permiten demostrar su relación con nuestras experiencias, en tanto que parece difícil establecer empíricamente dicha relación. La estrategia de Searle, como hemos defendido, pasa tan sólo por volver a repetir los mismos argumentos sin aportar nunca una explicación más allá de su mera plausibilidad. Pero ante todas las incógnitas abiertas, Searle todavía nos debe una respuesta razonable que permita establecer claramente dicha relación entre nuestras experiencias y los objetos capaces de causarlas.

Searle concibe la mente de un modo holista. Esto es, la mente no es la composición atómica de estados mentales aislados, sino que es un todo (unificado mediante redes conceptuales) que permite confrontar la naturaleza interna de los diferentes estados mentales entre sí y con el contexto externo hacia el cual los mismos se dirigen. Dichas redes conceptuales configuran el entorno y están compuestas de un número indeterminado de estados intencionales que permiten individuar cada estado mental que cada sujeto es capaz de tener y dotarlo de sentido. El conjunto de redes conceptuales tejen aquello que Searle denomina el Trasfondo (*Background*): el conjunto de emisiones y estados intencionales que permiten dotar de significado no sólo a nuestro lenguaje sino también a toda nuestra vida mental. El problema que advertimos con esta definición de Trasfondo, entonces, es precisamente su presunto carácter

pre-intencional. Si consideramos que el Trasfondo es algo previo a toda intención que está ya inserto en el contexto, entonces parece difícil indicar el modo en que entramos en contacto con él. Searle intenta mitigar esta dificultad afirmando que debemos entender el Trasfondo en un sentido de saber-cómo: es decir, en tanto que el Trasfondo sería un conjunto de capacidades mentales que permitirían englobar todas las representaciones que sobre el mundo podemos llegar a tener, en tanto que nos indicaría cómo es el mundo y cómo debemos comportarnos en y con él, siendo entonces un modo primario que engloba nuestras reacciones habituales ante las diferentes sensaciones y experiencias que sobre los objetos externos tenemos, un saber previo a cualquier conceptualización. Éste es, como dijimos, el punto en el que podríamos aceptar la teoría de Searle. Pero sólo en este punto, en tanto que parece que Searle también presupone un sentido metafísico en su defensa del Trasfondo. Dado que Searle afirma que el Trasfondo está en nuestras cabezas, no podemos más que distanciarnos de sus conclusiones porque, lejos de permitir (como sugiere Searle) que compartamos el trasfondo concreto que cada cual posee a través de las redes conceptuales implicadas, suponer la existencia interna de dicho nivel primario supondría instaurar un nivel psicológico privado que correría el riesgo de dejarnos inmersos en un solipsismo irreconciliable con nuestras intuiciones externalistas.

Ante esta suposición, Searle parece redefinir su noción de Trasfondo a partir de una sutil diferenciación entre aquella serie de elementos que afloran a un nivel consciente, personal, y todo el conjunto de elementos subdóxicos, a nivel neurofisiológico, que permiten explicar (pero no reducir) la noción de intencionalidad. Así, lo que parece estar proponiendo Searle es una noción de intencionalidad que permita, por un lado, la existencia de un nivel inconsciente (un nivel neurofisiológico, marcado por su Principio de Conexión, como mostramos) capaz de causar, por otro lado, un nivel consciente, pero separando conceptualmente entre la perspectiva subjetivista de primera persona que todos parecemos tener a la hora de acceder a nuestras propias vivencias y sensaciones y la perspectiva objetivista que nos brinda el conocimiento en tercera persona. De nuevo, como denunciábamos, parece irreconciliable esta propuesta, en tanto que suponer la existencia separada de dos niveles de reconocimiento supondría, cuanto menos, la existencia de un dualismo conceptual que, lejos de permitir la reconciliación entre lo objetivo y lo subjetivo, tan sólo agravaría nuestro problema presuponiendo la existencia separada de una sustancia subjetiva, la mental, y una sustancia objetiva, la física.

Cuando atiende al segundo rasgo definitorio de lo mental, la conciencia, Searle se refiere al conjunto de aquellos estados con carácter fenoménico de los cuales somos capaces (o podemos llegar a ser capaces) de darnos cuenta. La existencia de un nivel consciente, de una entidad biológica denominada conciencia, sería, por definición, propia de todo ser con una mente, en tanto que es lo que permitiría la existencia de ciertos contenidos proposicionales. Pero, como mostramos, no todos los estados mentales son proposicionales. Existen ciertos estados mentales que podríamos caracterizar como fenoménicos, precisamente, por carecer de contenido proposicional pero que, sin embargo, serían mentalmente relevantes para los sujetos. Pensemos tan sólo en la sensación de miedo o de ansiedad. Ninguno de ellos parece ser

proposicional (ni tan siquiera intencional, al no estar del todo claro a que objeto se refiere), pero sin embargo son relevantes para la vida mental del sujeto que los tiene.

En su teoría, Searle insiste en que la conciencia tiene dos rasgos definitorios. El primero, por un lado, es la presencia en el mundo mental de una serie de estados asociados a nuestras capacidades cognitivas caracterizadas por su carácter fenomenológico y subjetivo. El segundo, por otro lado, es la posibilidad objetiva de que dichos estados puedan ser explicados de modo científico. Como ya hemos comentado, al igual que la intencionalidad, también la conciencia debe considerarse como un elemento biológico. En este sentido, los aspectos mentales no deben ser reducidos a elementos físicos, dado que éstos ya serían, en sí mismos, un subtipo de propiedades físicas. Varios de dichos estados se caracterizarían por su carácter cualitativo y subjetivo. Existen estados, como el dolor o la ansiedad, que aunque pueden no tener un objeto determinado, se caracterizarían por ser la vivencia concreta que presentaría de cierta manera una serie de cualidades a un sujeto, de tal modo que el sujeto al cual se presentan dichas cualidades tan sólo puede captarlas de modo subjetivo.

La concepción que Searle tiene de la conciencia presenta a todo ser consciente como plenamente reflexivo. Esto es, los seres conscientes no sólo tienen vivencias experimentadas subjetivamente, sino que además dichos seres son capaces de prestar atención a dichas vivencias para llegar a darse cuenta de que las tienen. Pero que esto sea así no significa que siempre requiramos de la introspección para llegar a ser conscientes. La mayoría de las veces somos conscientes porque la mente, en sí misma, ya es reflexiva. Pero entonces, ¿cómo se convierte una experiencia en consciente? Es en este preciso momento cuando Searle apela a su Principio de Conexión. Para Searle, existe una necesaria conexión entre la conciencia y la intencionalidad. Toda experiencia antes de ser consciente debe ser inconsciente. Es decir, todo estado intencional, antes de llegar a emerger a un nivel personal, ha debido permanecer en un nivel inconsciente, subdóxico. De este modo, un estado intencional inconsciente se demuestra con estrechas relaciones con estados mentales conscientes.

Básicamente, el Principio de Conexión de Searle diferencia entre una conciencia intrínseca (aquella propia de los estados mentales y que se encuentra directamente en la naturaleza) y una conciencia derivada (propia del lenguaje, que sólo se adquiere a partir de la primera), establece la existencia de rasgos aspectuales en los estados mentales que determinan sus condiciones de satisfacción (es decir, una serie de aspectos relevantes en aquello representado por dichos estados que se cumplirá sólo si existe una correlación con el objeto representado) y la existencia de estados mentales inconscientes (aunque intencionales) que son potencialmente conscientes en tanto dependientes de los rasgos objetivos del cerebro.

Como mostramos en la sección 3.2.2.3., las dificultades que el Principio de Conexión tiene a la hora de tener que explicar los problemas derivados de los casos de visión ciega y de la posible existencia de duplicados físicos sin conciencia parecen apostar por un abandono del mismo. En el primero de los casos, la visión ciega, como vimos, el comportamiento de los pacientes parece causado por un estado intencional cuyo contenido no es accesible a la conciencia. Si, con Searle, afirmamos que los estados perceptivos de dichos pacientes son

inconscientes, deberíamos presuponer que son potencialmente conscientes. Pero lo que principalmente nos muestran los casos de visión ciega es que, aunque en ningún momento puedan llegar a ser conscientes para el sujeto, las percepciones que parece tener se dan a un nivel no-consciente y que el sujeto puede actuar causalmente a partir de ellas, por lo que la conexión entre conciencia e inconsciente parece no sostenerse. Ante el segundo caso, el problema *zombie*, Searle afirma, apelando al argumento de la indeterminación de la traducción de Quine, que al no existir hechos objetivos que permitan especificar en los duplicados los rasgos aspectuales de sus estados mentales, entonces parece que es imposible decir que dichos duplicados tendrían realmente estados mentales aunque carentes de conciencia. El problema con esta posible defensa lo encontramos en sus mismos términos, dado que podríamos decir que, entonces, si esto fuera así, nunca existirían hechos objetivos para determinar los rasgos aspectuales de un estado mental, dado que dependen de estados meramente subjetivos, por lo que en realidad ningún estado mental podría llegar a ser considerado intencional y, por tanto, ninguno de ellos podría llegar a ser nunca consciente.

De todos modos, el inconveniente principal con la teoría de Searle lo encontramos en la base de su teoría misma, en su defensa del naturalismo biológico. Recordemos que para Searle los estados conscientes son el producto de ciertos procesos neurobiológicos de bajo nivel, de nivel micro, que acontecen en el cerebro y que permiten la emergencia de determinados estados mentales a un nivel macro, personal, sobre los cuales los sujetos tienen un acceso consciente de un modo subjetivo. En este sentido, Searle está defendiendo una teoría representacional en tanto que a un nivel consciente los estados y las sensaciones que tiene un sujeto representan a un nivel personal aquello que ocurre en el nivel microfísico de su cerebro. Pero aunque esto es así, los estados mentales conscientes no son físicamente reducibles a los estados cerebrales que lo permiten, aunque sí son causalmente dependientes de ellos, demostrándose la conciencia como un producto emergente de los procesos biológicos que acontecen en el sistema nervioso central.

A pesar de los esfuerzos de Searle, creemos reconocer bajo estas afirmaciones la sombra del dualismo de propiedades porque en sus argumentaciones resuenan, todavía, las viejas pretensiones mentalistas cartesianas. En primer lugar, Searle parte de la suposición de que existe en la naturaleza un nivel intencional básico que nada que no sea un cerebro humano es capaz de tener. Si esto es así, entonces, sólo un cerebro humano es capaz de causar una mente y de causar el significado que acabará siendo representado en los estados de la misma. Pero, como defendimos, esta posibilidad es demasiado restrictiva, porque nada hace pensar que una mente pueda, *a priori*, ser causada a partir de otros modos (pusimos el ejemplo de un termostato, capaz de representar la temperatura de una habitación). Que realmente la base física del cerebro cause la conciencia es plausible para la existencia de la conciencia a un nivel humano, pero nada impide la existencia de un nivel representacional formado a partir de otros tipos de cerebro (es plausible suponer que los cerebros y sistemas nerviosos de los murciélagos permiten que éstos se representen la realidad a su modo) o de otro tipo de sistemas (también es plausible suponer que un velocímetro es capaz de representar la velocidad del objeto al que funcionalmente se

adhiera). Ello supondría establecer la conciencia como algo esencial y exclusivo de los seres humanos, mitificándola en exceso y convirtiendo la posición searleana en chauvinista.

Más preocupante, si cabe, nos parece la posición emergentista de Searle. Si consideramos que la conciencia emerge de los estados biológicos del cerebro, entonces parecería plausible la opción de reducir causalmente la primera a los segundos, en tanto que es perfectamente explicable desde el punto de vista de la reducción la existencia de un nivel consciente que dé cuenta de las sensaciones subyacentes a nivel neurobiológico. Pretender rescatar una irreductibilidad ontológica, como hace Searle, sólo consigue establecer una noción autónoma de conciencia que puede llegar a ser interpretada como independiente y desvinculada de aquellos procesos biológicos y neurofisiológicos que la ocasionan, recayendo en el inconveniente del hiato explicativo. Aunque Searle apuesta por una solución causal para solucionar dicha aparente desconexión entre lo mental y lo físico, la apelación a un mecanismo de relación que permita conectar ambas entidades, ello no sería suficiente para librarse de la acusación de dualismo. Y ello por dos motivos. Por un lado, porque parece posible afirmar la existencia de un criterio epistémico que indica que si disponemos de una explicación completa en términos físicos de un fenómeno concreto, entonces no requerimos de la apelación a ningún otro fenómeno para poder explicar su aparición. Por otro lado, además, porque aunque quiere explicar causalmente dicha relación, todavía respetaría la dualidad ontológica que tanto quiere eliminar, lo que le aproxima peligrosamente a posicionamientos misterianistas.

* * *

El capítulo 4 está dedicado a revisar el planteamiento de Michael Tye. Como dijimos allí, la posición representacionista de Tye ha ido radicalizándose hasta llegar a defender un intencionalismo que afirma la existencia necesaria de una representación de las características más relevantes para los sujetos de los objetos externos hacia las cuales se dirigirían todas sus experiencias conscientes. Tal vez aquello que mejor caracterizaría la posición de Tye es su teoría PANIC. Dicha teoría define al contenido de los estados conscientes como intencional, abstracto, no-conceptual y preparado para despertar otros estados mentales en el sujeto. Pero seguramente la noción sobre la que pivota dicha posición sea la noción de transparencia. Según Tye, toda experiencia fenoménica estaría presente de modo inmediato en las vivencias que de la misma los sujetos tienen y dichas vivencias se mostrarían de modo subjetivo a los perceptores (por lo que no cabría el error a la hora de individualizar e identificar una determinada sensación como una sensación de algo en particular) de la experiencia mostrando un peculiar carácter sensitivo (su particular modo de darse para dicho sujeto) y un marcado carácter cualitativo (el modo concreto mediante el cual el sujeto evaluaría dichas vivencias).

Lo primero que podemos decir de la teoría de Tye es que defiende que aquello que un sujeto siente al tener una vivencia son las características públicamente observables de las cosas y que cualquiera puede sentir, que dichas características pertenecen al objeto y no a las propias vivencias y que dichas vivencias no son respecto de la propia experiencia sino respecto del objeto externo representado por las mismas en tanto poseyendo ciertas cualidades públicamente observables que encajan en los sistemas perceptivos de los sujetos sensibles. Visto así, el

contenido representacional de las vivencias es intencional en tanto que se relaciona directamente con las características que los objetos externos poseen y que son representadas en la experiencia de los mismos. Cuando el sujeto experimenta una determinada sensación lo hace porque posee una serie de estados neurofisiológicos que lo capacitan para captar dichas características de los objetos externos, unos estados previos a cualquier otra significación o información y que son capaces de captar y elaborar la información recibida a partir de los sistemas cognitivos sin necesidad de recurrir a ninguna categorización previa. Dichas representaciones tendrán un carácter abstracto, en tanto que tan sólo presentarían al sujeto las cualidades físicas públicamente observables de los objetos externos pero sin ser ellas mismas un objeto. Dicho contenido permitiría, una vez elaborado, encontrar las salidas cognitivas adecuadas a cada situación perceptiva concreta, unas salidas capaces de responder adecuadamente a los estímulos experienciales de los sujetos.

Si, como hemos dicho, todo contenido fenoménico de una experiencia consciente es una representación preparada, abstracta, no-conceptual e intencionalmente dirigida hacia las cualidades más relevantes para los sujetos perceptores de los objetos que pueblan el mundo, pero sin llegar a ser en sí misma un objeto, entonces ¿cómo explicar la presencia de experiencias fenoménicas que habitualmente podríamos considerar puramente introspectivas y subjetivamente dependientes si no parecen referirse a objetos externos que las causen? Analizamos varios casos. En el caso de las cualidades secundarias (como el color, el sabor o el olor), aunque las experiencias parecen ser enteramente subjetivas, lo que realmente contaría como tener dicha experiencia es que el sujeto perceptor sea capaz de representarse adecuadamente dicha experiencia a partir de las cualidades físicas que las provocan. Así, el poder tener una experiencia adecuada de un color dependería de que el sujeto perceptor tuviera un sistema visual adecuado para captar de manera correcta la onda lumínica concreta capaz de causar dicho color. Del mismo modo ocurriría con el sabor y el olor. Que un sujeto sea capaz de captar un olor o un sabor concretos dependerá de que su sistema olfativo y su sistema gustativo sean capaces de captar adecuadamente las condiciones físicas que causan el olor y el sabor concretos. Pero, entonces, ¿qué ocurre en situaciones como las desarrolladas por los casos de espectro invertido, la variación estándar o la visión borrosa? Dichas situaciones, para Tye, serían perfectamente explicables desde su teoría PANIC, en tanto que ésta prevé que, en determinados casos, dichos sistemas son capaces de adaptarse a las circunstancias en tanto que la capacidad cognitiva de los sistemas perceptivos de los sujetos no sólo depende del desarrollo filogenético de especie, sino que también depende del particular desarrollo ontogenético que cada individuo realiza a lo largo de su vida. Es decir, que además de la existencia a un nivel de especie de un proceso de desarrollo evolutivo de los diferentes sistemas perceptivos y de representación, también existe un peculiar proceso de adaptación progresiva de cada sistema concreto dependiente del particular proceso de aprendizaje cognitivo que dichos sistemas han sufrido a lo largo de la vida de cada individuo concreto. Así, como vimos en el apartado 2.2., los casos de variación estándar podrían ser perfectamente explicables si apelamos a que dos sujetos, ante una misma experiencia, pueden captar el mismo color pero representárselo de manera

diferente, en tanto que a pesar de que comparten el mismo sistema perceptor y los mismos sistemas de elaboración de la información perceptiva, dichos sistemas han sufrido desarrollos ontogenéticos distintos dependientes de los aprendizajes seguidos durante la vida particular de cada uno de dichos individuos que posibilitan que cada cual se represente de manera distinta las mismas características físicas de los objetos externos. Del mismo modo, en los casos de espectro invertido, ocurriría lo mismo. A pesar de que los sistemas cognitivos de cada sujeto perceptor son idénticos filogenéticamente hablando, aunque fenoménicamente invertidos, la historia particular de ambos sujetos ha permitido a sus respectivos sistemas perceptivos y de elaboración de información adaptarse adecuadamente al mundo en el que viven a partir del aprendizaje realizado al captar sucesivamente los objetos externos que pueblan su entorno y permitirles representarse a partir de ellos de modo adecuado sus respectivas experiencias perceptivas.

El problema principal en esta teoría, como llamamos la atención, lo encontramos precisamente en el nivel representacional que Tye reconoce en el contenido de las experiencias fenoménicas. Porque para Tye, si aquello que realmente cuenta como una experiencia fenoménica es la existencia de una representación creada por los sistemas cognitivos y los sistemas procesadores de la información aportada por los mismos a partir de las características relevantes de los objetos externos, entonces parecería suponer la existencia de un cierre cognitivo en tanto que presupondría que aquello que contaría como una experiencia fenoménica tan sólo sería aquello que fuera capaz de encajar representacionalmente en la subestructura capaz de generar la imagen global de nuestro entorno y que es capaz de ocasionar las salidas comportamentales adecuadas. Pero, aún peor, si con Tye afirmáramos que el contenido fenoménico de una experiencia es idéntico a su contenido representacional, entonces estaríamos suponiendo la identidad entre el contenido de una experiencia y una proposición, en tanto que dicho contenido representacional ha sido identificado con una proposición respecto de las características relevantes para un sujeto de los objetos externos hacia el cual se dirigen las experiencias (al ser entendida como *de dicto*). Si esto es así, entonces la teoría de Tye parece caer en una inconsistencia al no poder demostrar que el contenido de una experiencia sea no-conceptual. Si entendemos el carácter subjetivo de una experiencia como proposicional, entonces ¿cómo entender que pueda ser identificado con su carácter cualitativo, en tanto que el primero sería un objeto abstracto (una proposición) y el segundo sería un término simple (la cualidad concreta de tener dicha experiencia)? Y aunque, a menudo parece insinuar Tye, podamos entender su posición representacional como afirmando que para que una experiencia tenga un cierto carácter fenoménico sólo debe poseer un determinado carácter intencional (negando la posibilidad de identificar carácter con contenido pero permitiendo que el hecho concreto de tener una experiencia sea idéntico a poseer un determinado contenido), de nuevo caeríamos en una inconsistencia al pretender que el contenido sea entendido, otra vez, como proposicional, volviendo a contradecir su propia afirmación de que el contenido de una experiencia fenoménica debe ser entendido como no-conceptual. Para marcar las distancias con la teoría representacional de Tye, nosotros preferimos, como dejamos constancia en el apartado 5.1., interpretar dicho contenido pre-conceptual presupuesto en las experiencias fenoménicas

desde una perspectiva personal que suscriba la teoría dennettiana que afirma la existencia de una actitud intencional, que tenga presente la posibilidad de que todo aquello que sentimos puede ser interpretado desde el punto de vista de las respuestas habituales que los seres conscientes tenemos ante ciertos estímulos. Como dijimos anteriormente, afirmar esto no presupone considerar que todo contenido fenoménico es conceptual. Todo lo contrario. La mayoría de nuestro contenido fenoménico es no-conceptual, previo a toda categorización, en tanto que se refiere a la carga informacional que nuestros sistemas perceptivos reciben cuando son estimulados por ciertas sensaciones externas y al proceso de elaboración que nuestros sistemas cognitivos son capaces de gestionar a partir de nuestra distribución biológica evolucionada para permitir la aparición de un nivel comportamental que permita nuestra adaptación a nuestro medio circundante y una correcta comprensión de nuestro entorno. Pero ello no quiere decir que en un momento dado dicho contenido no pueda ser conceptualizado (es decir, que el contenido fenoménico siempre será conceptualizable), en tanto que puede ser ajustado de tal forma que cierta información pueda ser entendida como una experiencia concreta.

Para explicar esto más adecuadamente, tomemos como ejemplo los casos de visión borrosa. Como dijimos, en un caso de visión borrosa, el sujeto perceptor es capaz de observar cualquier objeto situado en su entorno, pero éste siempre le aparecerá como borroso. Según la teoría de Tye, la experiencia concreta siempre representará el objeto de nuestra visión y sus cualidades relevantes para el sujeto, pero nunca nos dirá nada respecto de aquellas cualidades internas al sujeto que permiten dicha representación (que siempre dependerán de la particular distribución de los sistemas perceptivos del sujeto). Esto es, para Tye, parece que ante casos como los mencionados de visión borrosa el sujeto perceptor es un mero paciente de una experiencia concreta, una experiencia que sólo puede ser representada como borrosa, en tanto que dependerá del modo en como su sistema perceptivo está captando en un momento determinado las cualidades físicas relevantes de los objetos externos. Pero, pensamos, en determinados casos un individuo puede llegar a ser consciente de algo más que del modo particular en que se representa los objetos que aparecen en su experiencia. Es decir, lo que los casos de visión borrosa (y casos afines) precisamente nos dicen es que, a menudo, cuando a alguien se le presenta una determinada experiencia es que, simplemente, en un momento concreto, las experiencias están informando al sujeto que en dichos casos el sujeto percibe el mundo de manera borrosa. Esto es, que la visión borrosa puede formar parte del carácter intencional de la misma experiencia perceptiva, sin necesidad de recurrir a una explicación representacional de la misma, por lo que podríamos decir que nos informa del modo particular en que un sujeto es consciente del mundo, dependiendo tan sólo de la propia agudeza de su sistema perceptivo la determinación del carácter fenoménico de una experiencia concreta (independientemente de su contenido). Simplemente, dicho individuo parece ser capaz de captar el mundo de manera diferente a todos aquellos que contemplaríamos el mundo de manera no-borrosa, por lo que además de inferir un mero elemento paciente en la percepción, también estaríamos reivindicando un elemento activo en la misma, en tanto que el modo concreto en el que el sujeto estaría captando el mundo aportaría la información requerida para que éste sea

capaz de forjarse las estrategias comportamentales necesarias para su correcta adecuación al mundo que lo rodea.

El que Tye apele a la introspección y al requisito de transparencia parece insuficiente para poder explicar la intencionalidad presente en las experiencias. El que algo pueda ser captado de un cierto modo por un sujeto tan sólo puede ser, como hemos dicho, el modo concreto con el que un sujeto puede estar familiarizado al interactuar con el mundo. Que un individuo pueda equivocarse (como en los casos de variación estándar) o captar de manera anómala la realidad (como en los casos de visión borrosa) no quiere decir que el individuo esté representando de manera diferente el mundo, sino que tan sólo parece decirnos que así es el modo habitual en como su conciencia funciona cuando se le presentan experiencias fenoménicas. Si estamos en lo cierto, entonces sería conveniente comenzar por analizar la relación que existe entre el modo subjetivo y el modo cualitativo de las experiencias fenoménicas, como anteriormente recomendábamos, en tanto que seríamos capaces de salvaguardar la intencionalidad presente en las mismas sin necesidad de apelar a un elemento representacional y permitiendo una explicación satisfactoria de la unidad de nuestra experiencia consciente a partir de la noción de sujeto de experiencia experiencialmente situado. Entonces, en lugar de insistir como hace Tye en que todo contenido fenoménico tiene un único aspecto evaluativo (puramente pasivo), podríamos decir que dicho contenido puede presentarse de modos muy diferentes, muchas veces afectivo (lo que explicaría su carácter activo), capaz de englobar distintas experiencias y diferentes estados mentales hasta el punto de inundar toda la vida mental de un sujeto perceptivo concreto y guiar sus futuras estrategias comportamentales.

Cuando Tye pasa a analizar las sensaciones corporales emplea la misma estrategia representacional. En los casos concretos de dolor, por ejemplo (aunque sería ampliable a otras sensaciones corporales como el picor), Tye insinúa que nuestro cerebro establece una imagen de nuestro cuerpo a partir de la observación del mismo y representa también aquella parte del mismo que en dicho momento determinado nos está doliendo. Del mismo modo que antes, Tye afirma que cuando sentimos dolor lo que ocurre es que el sujeto representa una cierta experiencia como de dolor en tanto que representa las cualidades del cuerpo que en dicha situación son relevantes para el sujeto. La diferencia aquí la encontramos a la hora de localizar el lugar exacto en el que ocurre el dolor, en tanto que sería suficiente pensar que el sujeto tan sólo se representa una cierta sensación como de dolor y que ese dolor se experimenta de un modo concreto. Bastaría con saber que dichas experiencias representan cualidades diferentes que pueden referirse a partes concretas del cuerpo (el dolor) o simplemente a la representación de las mismas (aunque carezcamos de dichas partes, como en los casos de los miembros-fantasma). Pero de nuevo, creemos posible que se den dichas experiencias de dolor sin necesidad de apelar a un nivel representacional. La existencia de sistemas perceptivos internos que son capaces de captar la información interna del cuerpo y de sistemas que permiten elaborar dicha información, a menudo, nos basta para ser capaces de informarnos de la existencia de anomalías en nuestro cuerpo (pequeños dolores o simplemente situaciones de malestar) que se darían a un nivel previo, a un nivel primario, sin necesidad de recurrir a una representación de

las mismas. A menudo, sentimos un dolor sin necesidad de tener que localizarlo. Nos basta con saber que sentimos dolor. Y esa sensación es suficiente para inundar nuestro mundo experiencial.

Del mismo modo, cuando atendimos a los estados emocionales y a los estados de ánimo, vimos como Tye concibe las emociones como experiencias dirigidas hacia objetos que estarían situados fuera del cuerpo del sujeto perceptor, quién sería capaz de representarlos como teniendo tanto sensaciones sensitivas como evaluativas. Dichas sensaciones estarían representando ciertas disfunciones o anomalías que ciertos objetos externos provocarían en el cuerpo de los sujetos. La vivencia fenoménica de una experiencia emocional será, entonces, fijada por su contenido representacional, en tanto que representaría de determinado modo para el sujeto la situación externa capaz de provocar dicha sensación. Visto así, entonces, como Tye insiste, lo que nos permitirían las emociones y los estados de ánimo sería ofrecernos una imagen de las situaciones externas que pueden interferir con nuestro desarrollo y mostrarnos el impacto que dichas situaciones pueden llegar a tener sobre nuestro cuerpo. De este modo, las experiencias emocionales serían estados disposicionales dependientes de un contenido no-conceptual que pueden llegar a ser racionalmente justificados. Pero si, como nosotros lo entendemos, las experiencias emocionales son estados disposicionales con un valor de justificación, entonces nos vemos de nuevo involucrados en un nivel proposicional, violando nuevamente sus propias premisas no-conceptuales. Además, si como Tye insiste, las sensaciones emocionales deben ser transparentes, entonces dichas experiencias deben representar de modo adecuado determinadas cualidades físicas externas e internas al sujeto, unas cualidades representacionales que el sujeto estaría siempre en condiciones de captar. Pero, como hemos demostrado a partir de nuestros ejemplos del miedo sentido por un sujeto en el parking a raíz de una mala representación de una situación o del miedo visceral sentido por Bart hacia los payasos, lo que dichas experiencias emocionales vienen a decirnos, en contra de la tesis de transparencia, es que existen casos posibles en donde las sensaciones emocionales pueden ocurrir sin necesidad de que éstas se dirijan a ningún objeto concreto, a ninguna situación representada más allá del simple hecho de sentir miedo, y que a menudo en dichas experiencias de miedo (pero también podría ampliarse a otras situaciones de experiencia corporal, emocional o de estado de ánimo) no son realmente accesibles al sujeto de una manera consciente. Si esto es así, entonces, parece negarse la existencia de un nivel representacional que siempre permitiría la existencia de contenido experiencial. Nosotros optamos, de nuevo, por defender la existencia de un nivel previo, de un nivel intencional primario, que no tiene por qué estar involucrado con representaciones (aunque ello no quiere decir que en algún momento concreto sí pueda estarlo), y que permite decirnos que (al menos en el momento determinado en que ocurre la experiencia concreta) el mundo del sujeto perceptor se transforma, de tal modo que toda su vida mental se ve condicionada por dicha sensación, haciendo que su modo de contemplar y entender el mundo se vea modificada. Que en dichas situaciones, el sujeto perceptor ve de manera diferente el mundo porque las experiencias concretas que pasan a un

primer plano (como el dolor, el miedo o la sensación de desasosiego por la pérdida de un ser querido) tiñen, colorean, la entera vida mental del sujeto.

Si hemos entendido bien las argumentaciones de Tye, entonces parece claro que su teoría se encuentra con serias dificultades a la hora de poder explicar el modo en que un sujeto es capaz de experimentar sensaciones y, lo que es todavía peor, a la hora de establecer las condiciones adecuadas que permiten distinguir entre aquellos estados mentales con carácter fenoménico de aquellos estados mentales que carecerían del mismo. El peso de la prueba para poder realizar dicha diferenciación, según Tye, recaería sobre la noción de preparación del contenido fenoménico. Tye entiende dicha condición de preparación de dos maneras diferentes. Para Tye, primero, sería posible que en el momento en que un estado mental con un determinado contenido juega un papel funcional concreto en la vida mental de un sujeto, dicho contenido es necesariamente capaz de preparar la instanciación de ciertas actitudes proposicionales o bien puede llevar al sujeto a un estado diferente capaz de instanciar actitudes proposicionales. Pero, si esto fuera así, el representacionalismo de Tye se volvería inconsistente, en tanto que dicho estado podría interpretarse como teniendo o no un determinado contenido representacional según la conveniencia de la situación lo requiriera, algo que iría contra los principios básicos del intencionalismo defendido por Tye. Pero, segundo, tampoco la noción de preparación parece ser suficiente para permitir la diferenciación entre estados fenoménicos y no-fenoménicos. Pensemos, de nuevo, en el caso concreto de la visión ciega. Un sujeto con visión ciega puede ser capaz de evitar obstáculos en su camino aunque no puede tener acceso consciente a la percepción de los mismos. Dentro del sistema de Tye, entonces, las percepciones realizadas por el sujeto de visión ciega no pueden ser consideradas como fenoménicas, pero sin embargo cumplen todos los requisitos que Tye propone para que esto sea así, menos su condición de preparación. Pero vemos que ello no impide al sujeto de visión ciega poder realizar un uso correcto y adecuado de aquellas percepciones intencionales pero sin acceso consciente, dado que le permiten desenvolverse adecuadamente en su entorno. Por tanto, a pesar de que estas cuasi-experiencias realizadas a partir de los sistemas perceptivos de bajo nivel que se encuentran en su cerebro (recordemos que en la actualidad ya no parece plausible referirse tan sólo a 5 sentidos básicos, sino que los neurocientíficos proponen una disgregación de los sentidos a lo largo y ancho de nuestro sistema nervioso central que permite establecer diferentes capacidades perceptivas y sistemas de elaboración de información que llegan aproximadamente a multiplicar hasta 26 o 28 los diferentes sentidos) no aparecen a un nivel personal, debemos presuponer que estas experiencias se manifiestan a través del comportamiento del sujeto, lo que nos indicaría que la condición de preparación tampoco sería suficiente para determinar si un estado mental concreto es o no es fenoménico. Al evitar la adopción de esta noción de preparación (al separarnos de su propuesta representacional), nuestra propuesta es capaz de recoger este tipo de casos como casos de percepción intencional, una percepción que aunque no es capaz de florecer a un nivel verbal, creencial o consciente, sin embargo, es capaz de influir intencionalmente sobre la conducta del sujeto perceptor.

Aunque aquí estamos insistiendo en nuestras diferencias con la posición de Tye, consideramos que su intencionalismo también tiene varias ventajas. Por ejemplo, cuando Tye analiza más recientemente (en su *Consciousness Revisited*, editado en febrero de 2009) los casos de la gallina de Guinea, un clásico ejemplo dentro de las teorías acerca de la determinación del contenido perceptivo, que se remonta hasta Chisholm y Ayer y su controvertida disputa acerca de la plausibilidad o implausibilidad de los datos de los sentidos, como mostramos en el apartado 4.3.3.2., pero también en otros muchos casos donde la atención adquiere un papel prominente a la hora de determinar el contenido de un acto perceptivo, no podemos más que estar de acuerdo con su teoría. En este tipo de casos, como ya vimos, lo que Tye estaría defendiendo, más que un representacionalismo fuerte (en tanto que en estos trabajos el mismo Tye se cuestiona la plausibilidad de que toda experiencia fenoménica pueda ser considerada como transparente), es una especie de teoría demostrativa de las experiencias fenoménicas, en donde aquello que contaría como contenido de las mismas es la representación más o menos adecuada de aquellas partes del mundo, las cualidades físicas de los objetos externos, que son relevantes para el sujeto perceptor a partir de un acto atencional sobre las mismas. Este tipo de casos vendrían a demostrar, contra Block y otros, que la atención adquiere a menudo en las experiencias fenoménicas el papel determinante a la hora de diferenciar entre aquello que contaría como una experiencia fenoménica y aquello que no podría contar como tal. Detengámonos un momento en el ejemplo de Austin (extraído de *Sense and Sensibilia*) de dos individuos ante una experiencia fenoménica aparentemente idéntica: la percepción de un limón. Uno de ellos, nuestro sujeto A, percibe un limón real. El otro, nuestro sujeto B, percibe en realidad una pastilla de jabón que es idéntica fenoménicamente a un limón real. Para Tye, que los dos sujetos puedan tener experiencias indistinguibles (aquella que tienen sobre un limón) es plausible desde esta nueva formulación de su teoría. Pero que ambas experiencias tengan el mismo carácter fenoménico no quiere decir que compartan el mismo contenido representacional, dado que la primera experiencia se refiere a un limón real, mientras que la segunda se refiere a una pastilla de jabón que se parece a un limón real. Desde este punto de vista, parece plausible determinar el error en la representación del sujeto B, en tanto que representa erróneamente como un limón real lo que tan sólo es una pastilla de jabón. El nivel atencional es lo que aquí realmente determinaría la existencia de una experiencia fenoménica, en tanto que el carácter fenoménico vendría determinado por una posición (podríamos decir) adverbial del sujeto frente a la experiencia concreta del objeto. Que exista una diferencia a un nivel representacional, al menos en este momento, no nos preocupa en absoluto, dado que lo que querríamos resaltar aquí es la necesidad de que la atención se centre en un determinado punto de la realidad en un momento concreto para permitir determinar que contaría como una experiencia perceptiva y determinar aquello que contaría como contenido fenoménico. El nivel representacional, como también defendimos a lo largo del apartado 4.3., sería prescindible, pero no voy a ahondar aquí de nuevo en mis argumentos.

También estoy de acuerdo con Tye en la falacia de considerar plausible la crítica dualista contra el fisicalismo ofrecida por los argumentos del conocimiento y de los *zombies*,

como dejo constancia en las secciones 4.3.5.1. y 4.3.5.2. Pero me voy a detener brevemente en la solución que Tye propone para el hiato explicativo. Según hemos mostrado a lo largo de este capítulo, Tye está firmemente convencido de la verdad del fisicalismo. Aunque defiende la existencia de elementos fenoménicos en las experiencias conscientes, Tye cree posible que dichos elementos pueden ser explicados enteramente a partir de los elementos físicos que son capaces de causarlos. Al considerar que los estados fenoménicos están presentes en el sujeto en perspectiva, la teoría de Tye consigue explicar el elemento subjetivo que toda experiencia fenoménica presenta al sujeto perceptor. Y al considerar que el contenido fenoménico representa las cualidades físicas de los objetos externos que las ocasionan, Tye consigue poder explicar la presencia en la experiencia de elementos cualitativamente relevantes para el sujeto. Aunque no estamos de acuerdo con el nivel representacional que Tye defiende para el contenido de las experiencias con carácter fenoménico, compartimos su entusiasmo, como hemos dejado constancia, en la plausibilidad de una explicación fisicalista para los problemas de la conciencia.

* * *

Realizamos un recorrido crítico por la teoría ficcionalista de Daniel C. Dennett a lo largo del capítulo 5. Dennett parte de una visión científica de la filosofía, heredada directamente de sus mentores Ryle y Quine, que le permita construir una estructura física en tercera persona capaz de desentrañar los enrevesados problemas de la mente y la conciencia. Como nos dice, pretende construir una teoría del contenido para después poder construir, a partir de dicha noción, una correcta teoría de la conciencia. Precisamente por ello debemos tener presente en su concepción de la mente dos partes interconectadas: primero, una teoría filogenética de la conciencia y el pensamiento que establece un desarrollo de la mente a partir de la teoría darwiniana de la evolución y, segundo, un desarrollo ontogenético que permita determinar el modo concreto en que cada individuo es capaz de construir su propia conciencia individual.

Desde el punto de vista filogenético, de especie, Dennett considera que los seres humanos no somos más que mecanismos dotados de una intencionalidad derivada configurada por la evolución de las especies. En realidad, defiende que todo presunto significado o intencionalidad presente en la naturaleza en realidad es derivada, creada a partir de las leyes de la teoría evolutiva. Esto es, que todo ser aparentemente consciente no es más que un conjunto de pequeños sistemas intencionales organizados según los ciegos designios de la evolución. A partir de las teorías de R. Dawkins, que afirma que la evolución actuaría como un relojero ciego, Dennett pretende demostrar que toda intencionalidad presente en la mente humana en realidad se debe a la evolución a partir de una serie de escalas de re-representación (la torre de la generación y la prueba) que permitiría a los seres humanos auparse hasta el nivel consciente desde el desarrollo de sus primeros estadios como seres meramente conductuales genéticamente cerrados para la acción (las criaturas darwinianas y skinnerianas) y seres capaces de pre-seleccionar cierto tipo de comportamientos (las criaturas popperianas) hasta llegar a convertirse en los seres aparentemente intencionales y libres de decisión que son (las criaturas gregorianas) que, a partir de sus genes egoístas, han conseguido liberarse de su determinación genética hasta el punto de poder crear nuevas herramientas conceptuales (las palabras) mediante las que poder

aprehender el mundo y los objetos que lo pueblan para poder forjar nuevas estrategias de acción mediante la anticipación a los futuros acontecimientos. La idea principal de Dennett es, frente a otros como Burge, Dretske o Fodor, que defenderían la existencia de una intencionalidad intrínseca en la naturaleza (algo que otorgaría un significado primordial y único, en tanto que natural, a aquellos órganos, funciones y estados físicos) que tan sólo existe de un modo derivado adquirido a través del largo proceso adaptativo que posibilita la evolución.

Entonces, la teoría dennettiana estaría proponiendo que la gran serie de conductas organizadas que los seres conscientes parecen tener a la hora de mostrar su comportamiento requiere de un elevado número de estructuras de control interdependientes derivadas de la organización existente a un nivel subdóxico en el sistema nervioso central de los individuos, unas estructuras que han sido creadas a partir de la fuerte presión selectiva que obliga a los organismos a reorganizar continuamente sus estructuras fenotípicas de control de tal modo que puedan llegar a ser capaces de conformar nuevas estructuras capaces de llevarlos a la expansión y a la supervivencia. Dicha presión selectiva obligaría a los individuos a representarse aquellas situaciones externas (los objetos del mundo) que en un momento determinado pueden ser relevantes y a volverse a representar aquellas situaciones anteriores que no han funcionado de una manera adecuada, permitiendo así la proliferación de representaciones (entendidas en un sentido meramente utilitarista) que permitirán una paulatina sofisticación de los significados, eliminando de raíz ciertas representaciones equivocadas y matizando los significados de tal manera que los individuos sean capaces de llegar a conformar conceptos que permitan ahorrar tiempo a la hora de forjarse sus próximas estrategias de acción y de maximizar las posibilidades de éxito y supervivencia a partir de la mayor rapidez de manipulación de información. Si esto es cierto, entonces los seres conscientes (especialmente los humanos) son capaces de etiquetar aquellos objetos del mundo externo que son relevantes para los sujetos de acción, lo que les permitiría una mayor probabilidad de supervivencia al permitir aumentar el éxito de sus conductas y minimizar los errores. Esta dimensión interna tiene un mero sentido representacional en tanto que de un modo regular es capaz de actualizar continuamente los continuos cambios que se producen en la dimensión externa. Esos cambios externos obligan continuamente a los individuos a reorganizar su estructura interna para no perecer. Y dicho entorno externo sólo tiene sentido en tanto que se ven representados por los recursos internos que permiten su identificación al considerarlos relevantes. De este modo, los objetos externos obligan continuamente a las estructuras internas a modificarse a partir de nuevas y continuas re-representaciones de sus aspectos relevantes, por lo que dichas representaciones consiguen, así, convertirse en instrumentos (entendidos como objetos) mentales.

Como hemos matizado a lo largo del apartado 5.1., el concepto de 'derivado' en tanto que aplicado a la intencionalidad, al significado y a la conciencia empleado por Dennett debe entenderse en un sentido externista. Esto es, en tanto que nuestros estados mentales, donde adquieren especial relevancia aquellos conscientes, dependen del significado lingüístico que les hemos proporcionado a partir de nuestras estructuras internas evolutivamente desarrolladas. También hemos defendido la idea de que la teoría de Dennett, al pretender restringirse a un

ámbito científicamente cerrado, a un materialismo estricto conclusivo, podría ser interpretado con pretensiones de cierre cognitivo.

Como afirmamos en la sección 5.3.2.1., si Dennett pretende considerar la intencionalidad como derivada (una tesis que nosotros también suscribimos) ésta debe ser explicada de manera completa y necesaria por la evolución. En Dennett, la existencia de ciertos rasgos que permiten postular la aparición de intencionalidad en los organismos es explicada de una manera adaptativa. El producto resultante de dicha adaptación sería el proceso que, de un modo funcional, mejor responda a dicha adaptación. Pero si lo concebimos de este modo, como vimos, entonces recaemos en una petición de principio al pretender que es la misma selección natural la que permite explicar el proceso adaptativo mismo. Si queremos defender una posición evolucionista, como insistimos, deberíamos poder diferenciar entre la propia selección adaptativa y sus objetivos. De este modo, podríamos suponer que si un organismo posee una determinada característica se debe a la existencia de ciertos procesos selectivos que han permitido su potenciación y no, como Dennett pretende, a la existencia de procesos selectivos que permitirían la aparición de dicha característica. La noción de selección natural parece un elemento necesario para poder explicar la aparición de la intencionalidad en la mente humana, pero parece no ser suficiente apelar a elementos adaptativos para explicarla porque, entonces, debería permitirnos indicar el modo en que dicha característica aparece en la mente, algo que no parece plausible desde la posición de Dennett. Apelar a criterios funcionales, morfológicos y nomológicos de individuación de dichas características parece mucho más plausible que apelar simplemente a elementos adaptativos, incluso nos permitiría salvaguardar mejor nuestras pretensiones externistas.

Desde el punto de vista ontogénico, la estrategia dennettiana consiste en un intento de construir una teoría capaz de describir el modo en que nuestra base biológica sea capaz de producir nuestras mentes y conciencias particulares.

Dennett parte de la irrealidad de un teatro cartesiano en la conciencia. Dicho teatro haría de la conciencia un escenario en el que se presentarían los datos de nuestros sentidos. La idea general es sustituir dicha explicación por una teoría naturalizada capaz de describir todos nuestros estados fenoménicos como un producto de nuestra biología evolucionada, a partir de la eficacia causal de nuestro cerebro y los sistemas de procesamiento de la información sensible que componen nuestro sistema nervioso central. El modelo de los borradores múltiples es la alternativa propuesta por Dennett. Lejos de considerar la existencia de un núcleo de la conciencia (como lo concebía el Teatro Cartesiano), Dennett considera que toda nuestra actividad mental se realiza en el cerebro a partir de una serie de procesos de distribución y manipulación de la información sensible captada por nuestros sistemas cognitivos y perceptivos capaces de procesarla incesantemente y siempre siendo revisables sus resultados. Dicha operación de reelaboración permanente de la información sensible permite la existencia al unísono de varios borradores (entendidos como interpretaciones diferentes de los datos recibidos desde los sistemas cognitivos) que los sistemas de procesamiento elaboran para que los sujetos puedan forjar sus estrategias de acción. De este modo, según el modo particular en

que en cada individuo se imponga uno de dichos borradores frente a los otros dará lugar a la aparición de su conciencia particular. Pero que uno de los borradores se imponga no impide que en un momento concreto otro de esos borradores pueda alcanzar el nivel consciente. El conjunto de representaciones que acabe por imponerse lo hará de manera aleatoria, siendo imprevisible cual de todos pueda llegar al nivel consciente.

El modo en que un borrador se crea depende del tipo de revisión de la información que lo configure. El sujeto recibe información procedente del exterior a través de sus sistemas perceptivos. Dicha información se procesa en los sistemas adecuados y sufre un proceso de discriminación guiado por los propios intereses particulares del sujeto receptor, pasando así el filtro de la heterofenomenología (una especie de cuenta-historias personal que permite al sujeto concreto describir para sí mismo el modo en que los datos deben ser entendidos y permitiéndole configurar un borrador de cómo puede ser entendida la experiencia concreta que está teniendo en un determinado momento). Este intento de naturalizar la experiencia fenoménica (lo que Dennett denominará una estrategia de quineación de los *qualia*) permite hacer depender al contenido fenoménico de nuestros estados mentales de un nivel físico de explicación de la experiencia fenoménica vinculado a los diferentes sistemas cognitivos de captación y elaboración de la información sensible y de un nivel puramente funcional que permitiría la existencia de una competencia interpretativa a los agentes que les llevaría a producir diferentes conjuntos de representaciones de lo que serían sus propias experiencias, permitiendo así fijar el contenido fenoménico a partir de multitud de microjuicios que les permitiría forjar sus posibles respuestas comportamentales futuras. Pero, entonces, si con Dennett afirmáramos que ambos niveles se encuentran a un nivel subpersonal, entonces tanto el proceso de captación y elaboración de la información sensible como el proceso interpretativo dependerían de un nivel meramente funcional, con lo que las experiencias fenoménicas quedarían relegadas a un mero conjunto de ficciones útiles capaces de permitir el correcto desenvolvimiento del sujeto en su entorno, pudiendo incluso negar la realidad de las mismas experiencias, algo que parece contraintuitivo. Si lo que Dennett está afirmando es que el acto de eliminación de las vivencias en tanto que acceso subjetivo al carácter cualitativo de los propios estados fenoménicos es plausible, entonces el único modo en que parece entender la conciencia es al mero hecho de que los sujetos tuvieran estados con un cierto nivel representacional (en tanto que ficcional pero útil, lo que permitiría interpretar la teoría dennettiana como una teoría de nivel superior) cuyo contenido tan sólo concierne a sus propios estados cognitivos y funcionales, por lo que la única noción relevante de conciencia para Dennett parece ser aquella que se podría estipular como una 'conciencia de', meramente reflexiva y dirigida a esos mismos estados funcionales, como perfectamente diferenciable de la conciencia por acceso que tanto empeño estamos poniendo en defender y que poseería un contenido inferencialmente disponible que posibilitaría el control racional de la acción de los sujetos. Si nosotros estamos en lo cierto, entonces sería conveniente postular, al menos, la existencia de estados mentales que, aunque tan sólo fuera a un nivel mínimo, todavía dispusieran de un contenido fenoménico que garantizara la posibilidad de que los sujetos perceptores fueran capaces de sentir sus propias vivencias, en tanto que los *qualia*

sería aquello que haría que la vida mereciera la pena. Si esto fuera así, entonces bastaría para afirmar que el mero hecho de tener una experiencia ya acredita al sujeto para saber que la tiene, sin necesidad de apelar a ningún nivel conceptualmente superior que permita diferenciar dichas experiencias del resto de estados mentales.

* * *

En el capítulo 6 desarrollamos algunas ideas de aquello que podríamos denominar posición intencionalista no sólo representacional de la conciencia. Según dicha concepción, la nuestra propia, la conciencia debería entenderse como un mecanismo de anticipación que permitiría al sujeto prever los acontecimientos futuros a partir de la percepción directa de los elementos que conforman el mundo externo. La conciencia sería, entonces, un ingenio tecnológico configurado por la evolución del cual podemos beneficiarnos aquellos seres capaces de darnos cuenta de los elementos relevantes de nuestro entorno en tanto que maximizarían las posibilidades de nuestra supervivencia y permitirían un desarrollo óptimo de nuestras capacidades cognitivas. Para realizar dicha defensa, partimos de la suposición materialista que afirma que, al menos, muchos de nuestros estados mentales pueden identificarse con ciertos estados físicos del cerebro. Por ello mismo, pretendemos mantener una posición fisicalista que permita, por un lado, ser realistas respecto del modo en como los seres conscientes captamos el mundo y los hechos que acontecen en él y, por otro lado, ser coherentes con la imagen que la ciencia actual conforma del mundo físico.

Frente a tesis fisicalistas reduccionistas, como las de Kim, reivindicamos un elemento no-reductivo que permita salvaguardar nuestra noción ordinaria de *qualia*. Nuestra reivindicación se dirige especialmente hacia una noción de *qualia* que no exija una separación tajante entre aquello que podríamos considerar elementos mentales y aquello que son elementos físicos y que tenga presente a la evolución de nuestros sistemas perceptivos y de reconocimiento y procesamiento de la información a la hora de elaborar una explicación plausible del trasfondo biológico que permite un crecimiento progresivo de nuestro conocimiento del mundo a partir de nuestras acciones y reacciones naturales a la presencia de determinados estados con cierto contenido cualitativo.

Para evitar una acusación de epifenomenismo hacia nuestra concepción de la conciencia, debimos hacer frente a una dificultad: debimos ofrecer una explicación del modo en como el contenido fenoménico es capaz de ser causalmente eficaz. Cierto es que no siempre encontramos elementos conscientes involucrados en una acción. Pero no menos cierto es que siempre parece estar presente en los sujetos algún tipo de sensación, incluso muchas veces coexisten varias sensaciones al mismo tiempo. Como dijimos anteriormente, y defendemos en el apartado 6.1.3., muchas veces nuestras sensaciones o las vivencias que de ellas tenemos son capaces de causar nuestras conductas o de hacer aflorar determinados estados mentales (también explicables físicamente) que permitirían explicar nuestro comportamiento. Aunque también es cierto que, a menudo, dichas sensaciones no pasan de ser simplemente una mera anécdota en nuestra vida mental. Pero de todas maneras, dichas vivencias nunca pasan inadvertidas para nosotros: las vivencias de las experiencias son relevantes para un sujeto en tanto que las percibe

de modo subjetivo y de una manera cualitativamente relevante. Y aunque no podamos decir de todas ellas que son causantes de nuestra acción, sí podemos decir al menos que son las causantes de que, muchas veces, toda nuestra vida mental se vea coloreada por su presencia. En definitiva, aunque de muchas de nuestras experiencias podamos decir que tienen un pobre poder causal sobre nuestra conducta, no podemos decir que sean irrelevantes. La presencia de elementos de conciencia con carácter cualitativo tiñe todo nuestro mundo y es lo que hace que valga la pena vivir.

Como también hemos mostrado, una estrategia explicativa de corte fisicalista de la conciencia permitiría explicar adecuadamente el modo en que dicho teñido tiene lugar en los seres conscientes. Frente a todos aquellos que consideran plausible las conclusiones de la corazonada *zombie*, nosotros afirmamos la posibilidad de la inexistencia de dichos seres comportamentalmente idéntico a un individuo consciente pero carente de elementos cualitativos. Las razones que aportamos se dirigen principalmente a datos extraídos de los experimentos realizados con pacientes con visión ciega. Dichos experimentos indican un nivel previo a la afloración a nivel personal de todo elemento consciente, un nivel subdoxástico, que compartiríamos todos los seres conscientes (al menos en un sentido metafórico.). A partir de dichos elementos pre-conscientes, el individuo (que no fuera consciente de la presencia de elementos cualitativos en su nivel personal) podría interactuar con su entorno de modo tal que no parecería importar mucho si es consciente o no de la existencia de elementos fenoménicos en su mente. Si esto fuera cierto, como defendemos, la existencia de *zombies* tampoco resultaría problemática, dado que cualquiera de nosotros en un momento determinado podría actuar como uno de estos seres carentes de conciencia. Pero gracias a la existencia de este nivel previo no importaría mucho que careciéramos de elementos conscientes en dicho momento concreto, porque sabemos que el teñido de nuestro mundo mental puede existir a un nivel subdoxástico, por lo que los elementos que permiten explicar nuestro comportamiento pueden salvarse a partir de una explicación física.

Del mismo modo, cuando atendimos al problema del argumento del conocimiento, afirmamos la verdad del fisicalismo independientemente de que Mary pudiera aprender algo nuevo cuando tuviera por primera vez una experiencia de rojo. Nuestra idea al defender esta conclusión fue demostrar en qué sentido alguien que sabe todo aquello que desde el plano físico puede saberse sobre la percepción del color no puede nunca haber pasado por alto los elementos cualitativos relevantes que a un nivel consciente tendría un ser que percibe el color rojo en tanto, hemos defendido, los elementos mentales serían totalmente de los elementos físicos, aunque puedan ser entendidos como causales. Si Mary conoce todo aquello que desde el plano físico se puede conocer sobre la física de la percepción del color, entonces Mary conoce perfectamente qué tipo de sensación tendrá cuando perciba por primera vez la sensación de rojo, aunque sólo será de un modo *a posteriori*, después de tenerla, el modo en que será capaz de determinar el *quale* concreto que supone percibir el color rojo. Del mismo modo, defendimos, podríamos solucionar el problema del espectro invertido.

Si estamos en lo cierto, entonces, no debe preocuparnos la posibilidad de un hiato explicativo a la hora de rendir cuentas de la conciencia, porque éste no existiría. La explicación fisicalista sería perfectamente posible en tanto que, como hemos pretendido mostrar, la única posible sustancia existente es una sustancia material y física. Y aunque a menudo podamos decir que parece existir una entidad separada de ella, ésta tan sólo es una ilusión creada a partir de la acción de dichos elementos físicos que componen nuestro sistema nervioso central. La presencia de elementos mentales, entonces, como explicamos, serían correlatos de la actividad física que se da en nuestro cerebro y en el sistema nervioso, en tanto que la posibilidad de la existencia de causalidad sólo sería posible si se demuestra la autonomía de dicha sustancia. Aunque hemos estado postulando la posibilidad de que ambas (la material y la mental) tengan ciertas características causales (ciertas propiedades causales capaces de causar algún efecto), también hemos establecido la imposibilidad de que dichas propiedades causales se den de modo autónomo e independiente de la presencia de una sustancia material, física. Sólo la sustancia material demostraría su causalidad y su independencia a la hora de poder conseguir ciertos efectos por la acción independiente de sus propios poderes causales. La conciencia, en cambio, tan sólo conseguiría tener determinadas propiedades causales si es considerada como dependiente de ciertas propiedades físicas a partir de las estructuras neuronales y nerviosas de los sujetos en las que acontece. Ello demostraría su dependencia de la existencia de una sustancia material y negaría la autonomía de la conciencia, por lo que ya no podría ser considerada sustancia.

Nuestra principal conclusión, por tanto, es que todo ser capaz de tener sensaciones subjetivas cualitativamente relevantes, aunque no siempre pueda llegar a darse cuenta de que las tiene (esto es, aunque a menudo dichas sensaciones queden ancladas en un nivel diferente del consciente), se ve influido por las vivencias que de dichas experiencias tiene. Tal vez dichas sensaciones sean sumamente pobres a la hora de poder explicar sus comportamientos futuros y, aunque algunas de ellas puedan ser capaces de hacerlo, a menudo sería difícil poder establecer con precisión el modo en que éstas consiguen determinar nuestras acciones. Lo que sí parece claro es que todas estas sensaciones nos permiten un aprendizaje vital para nuestra supervivencia. Nos aportan una valiosa información acerca de cómo son las cosas en el mundo en que vivimos, cómo se nos aparecen, y nos aportan una serie de herramientas más o menos eficaces que nos permiten poder interactuar adecuadamente (aunque siempre existe la posibilidad de error) con ellas y con nuestro entorno. Todas estas sensaciones forjan nuestra mente consciente, aquello que somos, aquello que hacemos y aquello que sentimos, en tanto que el más leve cambio en dichas sensaciones puede ser extremadamente crucial para nuestro devenir futuro. No hace falta recurrir a ejemplos clásicos extraídos de la literatura, como el caso del monstruo de Frankenstein (cuyo contacto con la única persona capaz de interactuar con él, un pobre ciego abandonado en su cabaña, le proporciona el lenguaje y la posibilidad de abandonar su ira para sustituirla por venganza, a la que paulatinamente también irá abandonando a cambio de una búsqueda incesante de respuestas racionales que acabará en devoción filial) y al más reciente ejemplo de Arte (el androide programado por el filósofo

William que B. Beckett nos presenta en *Génesis*, un ser capaz de ser consciente y de aprender a partir su interacción con otros seres conscientes, especialmente Adán, su cautivo compañero, y que al final demostrará la importancia que la vida consciente tiene a la hora de explicar nuestras estrategias de supervivencia y nuestra capacidad de aprendizaje y acción). Tal vez nos basta con apelar a la propia historia personal de cada uno de nosotros para demostrar en qué sentido son sumamente importantes estos detalles cualitativamente relevantes en nuestras vidas. Cómo una caricia del ser amado puede alegrarnos el día, cómo el dolor provocado por la picadura de una avispa puede amargarnos la soleada mañana en la piscina que habíamos planeado, cómo el confort que sentimos al sentarnos en la butaca tras una dura jornada de bicicleta puede relajarnos y hacernos olvidar momentáneamente nuestra sensación de fatiga, cómo el sabor intenso de un café puede invadirnos de placer... Tal vez el papel causal de dichos elementos sea especialmente difícil de determinar, pero de lo que sí podemos estar seguros es de su tremenda importancia en nuestra propia vida consciente.

* * *

Para concluir, y de un modo un tanto más esquemático, presentamos aquí las que, creemos, serían nuestras principales conclusiones.

1. Sería preciso diferenciar entre dos aspectos de la experiencia consciente, su aspecto cualitativo y su aspecto subjetivo. Los problemas sugeridos por cada uno de estos aspectos son distintos y asimilar ambos aspectos da origen a muchas confusiones.

2. La conciencia es intencional en un sentido no sólo representacional y no sólo conceptual. La intencionalidad de la conciencia estaría estrechamente vinculada a la acción y a la previsión del futuro en un nivel personal. Los casos de “blind sight” y de “visión borrosa” serían especialmente relevantes aquí, dado que nos aportarían datos suficientes para demostrar la existencia de niveles subpersonales de conciencia que influyen en nuestro comportamiento.

3. Es posible mantener una concepción fisicalista de la conciencia capaz de salvar el ‘hiato explicativo’ y el ‘problema difícil’. Por tanto, se puede contrarrestar la ‘corazonada zombie’, el ‘argumento del conocimiento’ y el ‘espectro invertido’.

4. La conciencia podría tener un cierto papel causal relacionado con el ‘coloreado’ cualitativo de toda nuestra vida mental. Quedaría abierto el problema de desarrollar una teoría sustantiva (tanto filosófica como empírica) sobre este sugerente papel causal relativo a la operación de ‘teñido’, aunque los pasos iniciales ya han sido mostrados.

5. Respecto a los planteamientos de los tres autores que han sido analizados con mayor detalle (Searle, Tye y Dennett), la posición adoptada ha sido:

- a) Frente al naturalismo de Searle: la defensa de un fisicalismo no misterianista.
- b) Frente al representacionalismo de Tye: la defensa de un intencionalismo no sólo representacional y no sólo conceptual.
- c) Frente al ficcionalismo de Dennett: la defensa de un realismo mínimo respecto a los aspectos cualitativos y subjetivos de la experiencia consciente.

REFERENCIAS BIBLIOGRÁFICAS

- ABBOTT, L. and SEJNOWSKI, T. J. (1999); *Neural codes and distributed representations. Foundations of neural computation*. Cambridge (Ma.): MIT Press.
- ACERO, J. J. (1997): “Mentes, máquinas e intencionalidad original: el argumento de Dennett”, en J. E. Corbí y C. J. Moya (eds.), *Ensayos sobre libertad y necesidad*. Valencia: Pretextos, pp. 187-206.
- (2001). “El lenguaje y el origen de la intencionalidad”, en M. C. Paredes Martín (ed.); *Mente, conciencia y conocimiento*. Salamanca: Ediciones Universidad de Salamanca, pp. 29-54.
- , FRÁPOLLI, M. J. y ROMERO, E. (1997); “El significado y las actitudes proposicionales, I: Los problemas y las respuestas clásicas”, *Ágora. Papeles de Filosofía*, Vol. 16, No. 2, pp. 5-37.
- (1998); “El significado y las actitudes proposicionales, II: Mundos posibles, proposiciones y estados psicológicos”, *Ágora. Papeles de Filosofía*, Vol. 17, No. 1, pp. 5-30.
- AKINS, K. (1993); “What is it like to be boring and myopic”, en E. Dahlbom (ed.); *Dennett and his critics: demystifying mind*. Oxford: Blackwell, pp. 124-160.
- ANDERSON, S. R. y LIGHTFOOT, D. W. (2002); *The language organ: linguistics as cognitive physiology*. Cambridge: Cambridge University Press.
- ANSCOMBE, G. E. M. (1957); *Intention*. Oxford: Blackwell (trad. esp. *Intención*. Barcelona: Paidós, 1991).
- ANTONY, M. V. (2006); “Papineau on the vagueness of phenomenal concepts”, *Dialectica*, Vol. 60, No. 4, pp. 475-483.
- ARMSTRONG, D. M. (1969); *A materialist theory of mind*. London: Routledge.
- (1981); “The causal theory of the mind”, en *The nature of the mind and other essays*. Ithaca: Cornell University Press (reimpreso en Borst, C. V. (ed.); *The mind-brain identity theory*. London: MacMillan, 1970, pp. 67-82).
- (1991); “Intentionality, perception, and causality: reflections on Searle’s *Intentionality*”, en Lepore and Gulick (1991), pp. 149-158.
- (1996); “Qualia ain’t in the head: Review to Tye’s *Ten problems of consciousness*”, *Psyche*, Vol. 2, No. 31.
- (1999); *The mind-body problem: an opinated introduction*. Boulder: Westview Press.
- and MALCOM, N. (1984); *Consciousness and causality*. Oxford: Blackwell.
- AUSTIN, J. L. (1962); *Sense and sensibilia*. Oxford: Clarendon Press.
- AYDEDE, M. (2005); “Cognitive Architecture, Concepts, and Introspection: An Information-Theoretic Solution to the Problem of Phenomenal Consciousness”, *Noûs*, Vol. 39, No. 2, pp. 197-255.
- (ed.) (2006); *Pain: new essays on its nature and the methodology of its study*. Cambridge (Ma.): MIT Press.
- (2006); “The main difficulty with pain: commentary on Tye”, en Aydede (ed.), pp. 123-136.

- AYER, A. J. (1940); *The foundations of empirical knowledge*. London: Macmillan, 1998.
- BAARS, B. J. (1988); *A cognitive theory of consciousness*. Cambridge: Cambridge University Press.
- (1997); *In the theater of consciousness: the workspace of the mind*. Oxford: Oxford University Press.
- BACH, K. (1997); "Engineering the mind. Review of *Naturalizing the mind* by Fred Dretske", *Philosophy and Phenomenological Research*, Vol. 57, No. 2, pp. 459-468.
- BALOG, K. (1999); "Conceivability, possibility, and the Mind-Body Problem", *Philosophical Review*, Vol. 108, No. 4, pp. 497-528.
- (2000); "Conceivability arguments or the revenge of the zombies", en <http://philpapers.org/browse/zombies-and-the-conceivability-argument>.
- (forthcoming); "In defense of the Phenomenal Concept Strategy", *Philosophy and Phenomenological Research*, en prensa.
- BAYNE, T. (2005); "Divided brains and unified phenomenology: a review essay on M. Tye's *Consciousness and persons*", *Philosophical Psychology*, Vol. 18, No. 4, pp. 495-512.
- (2008); "The unity of consciousness and the split-brain syndrome", *The Journal of Philosophy*, Vol. 105, No. 6, pp. 277-300.
- and CHALMERS, D. (2003); "Why is the unity of consciousness?", en A. Cleeremans (ed.); *The unity of consciousness*. Oxford: Oxford University Press, pp. 23-58.
- BEISECKER, D. (2006); "Dennett's overlooked originality", *Minds and Machines*, Vol. 16, No. 1, pp. 43-55.
- BERMÚDEZ, J. L. (1994); "Peacocke's argument against the autonomy of non-conceptual content", *Mind and Language*, Vol. 9, No. 3, pp. 403-418.
- (1998); *The paradox of self-consciousness*. Cambridge (Ma.): MIT Press.
- (2003); *Thinking without words*. Oxford: Oxford University Press.
- BERNSTEIN, R. J. (1968); "The challenge of scientific materialism", *International Philosophical Quarterly*, Vol. VIII, No. 2, pp. 252-275 (reimpreso en Rosenthal, D. M. (ed.); *The nature of mind*. Oxford: Oxford University Press, 1991, pp. 200-222).
- BIGELOW, J. and PARGETTER, R. (1990); "Acquaintance with qualia", *Theoria*, Vol. 61, No. 1, pp. 129-147.
- (2006); "Re-acquaintance with qualia", *Australasian Journal of Philosophy*, Vol. 84, No. 3, pp. 353-378.
- BLOCK, N. (1978); "Troubles with functionalism", en N. Block (ed.) (1980); *Readings in the Philosophy of Psychology, Vol. 1*. Cambridge (Ma.): Harvard University Press.
- (1986); "Advertisement for a semantics for psychology", *Midwest Studies in Philosophy*, Vol. 10, pp. 615-677.
- (1990); "Inverted earth", en J. Tomberlin (ed.) (1990), pp. 53-79.
- (1993); "Review of Dennett's *Consciousness Explained*", *The Journal of Philosophy*, Vol. XC, pp. 181-193.

- (1995): "On a confusion about a Function of Mind", *Behavioral and Brain Sciences*, No. 18, pp. 227-247.
- (1998); "How to find the neural correlate of consciousness", en A. O'Hear (ed.); *Current issues in Philosophy of Mind*, Supplement of *Philosophy* no. 43, pp. 23-34.
- (1999); "Sexism, racism, ageism and the nature of consciousness", *Philosophical Topics (The Philosophy of Sydney Shoemaker)*, Vol. 26, No. 1-2.
- (2001); "How not to find the neural correlate of consciousness", en J. Branquino (ed.); *The foundations of Cognitive Science*. Oxford: Oxford University Press.
- (2006); "Bodily sensations as an obstacle for representationalism", en Aydede (ed.), pp. 137-142.
- (2007a); "Consciousness, accessibility and the mesh between psychology and neuroscience", *Behavioral and Brain Science*, Vol. 30, No. 4, pp. 481-499.
- (2007b); "Overflow, access, and attention", *Behavioral and Brain Science*, Vol. 39, No. 4, pp. 530-548.
- and STALNAKER (1999); "Conceptual analysis, dualism, and the explanatory gap", *Philosophical Review*, Vol. 108, No. 1, pp. 1-46.
- BOGHOSSIAN, P (1992); "The status of content", *Philosophical Review*, Vol. 99, pp. 157-184.
- and VELLEMAN, D. (1989); "Colour as a secondary quality", *Mind*, Vol. 98, No. 1, pp. 81-103.
- BORST, C. V. (ed.) (1970; *The mind-brain identity theory*. London: MacMillan.
- BOTTERILL, G. and CARRUTHERS, P. (1999); *The philosophy of psychology*. Cambridge: Cambridge University Press.
- BRADDON-MITCHELL, D. y JACKSON, F. (1996); *Philosophy of mind and cognition*. Oxford: Blackwell.
- BRENTANO, F. (1874); *Psychology from empirical standpoint*. London: Routledge, 1973.
- BROACKES, J. (2007); "Black and white and the Inverted Spectrum", *The Philosophical Quarterly*, Vol. 57, No. 227, pp. 161-175.
- BRONCANO, F. (2006); "Sujeto y subjetividad en la mente extensa", *Revista de Filosofía*, Vol. 31, No. 2, pp. 109-133.
- BURGE, T. (1979); "Individualism and the mental", *Midwest Studies in Philosophy*, Vol. 4, pp. 73-121.
- (1985); "Individualism and psychology", *Philosophical Review*, Vol. XCV, No. 1, pp. 3-46.
- (1989); "Individuation and Causation in Psychology", *Pacific Philosophical Quarterly*, Vol. 70, No. 2, pp. 303-322.
- BURTON, R. G. (1995); "Searle on rediscovering the mind", *Man and World*, Vol. 28, pp. 163-174.
- BYRNE, A. (2001); "Intentionalism defended", *Philosophical Review*, Vol. 110, No. 2, pp. 199-240.

--- (2002a); "Don't PANIC: Tye's intentionalist theory of consciousness", *A field guide to Philosophy of Mind. Symposium on Tye's Consciousness, color, and content*, <http://uniroma3.it/progetti/kant/field/tyesymp.pdf>, 3-27.

--- (2002b); "Tye on color and the explanatory gap", *Pacific APA*, <http://mit.edu/abyrne/www/tyeonColorandtheGap.pdf>.

--- (2006); "Color and the mind-body problem", *Dialectica*, Vol. 60, No. 3, pp.223-244.

--- and TYE, M. (2006); "Qualia ain't in the head", *Nous*, Vol. 40, No. 2, pp. 241-255.

--- and HILBERT, David. R. (2007); "Truest blue", *Analysis*, Vol. 67, No. 1, pp. 87-92.

--- and LOGUE, H. (2009); *Disjunctivism: Contemporary Readings*. Cambridge (Ma.): MIT Press.

CAMPBELL, K. (1970); *Body and mind*. Oxford: Oxford University Press.

CARRUTHERS, P. (1996a); *Language, thought, and consciousness*. Cambridge: Cambridge University Press.

--- (1996b); "Simulation and self-knowledge: a defence of theory-theory", en P. Carruthers and P. K. Smith (eds.); *Theories of theories of mind*. Cambridge: Cambridge University Press, pp. 22-38.

--- (2000a); "The evolution of consciousness", en Carruthers and Chamberlain (2000), pp. 254-275.

--- (2000b); *Phenomenal Consciousness*. Cambridge: Cambridge University Press.

--- (2001); "Consciousness: explaining the phenomena", en D. M. Walsh (ed.); *Naturalism, evolution, and mind. Philosophy suppl. No. 49*, pp. 61-85.

--- (2004); *The nature of mind*. London: Routledge.

--- and SMITH, P. K. (eds.) (1996); *Theories of theories of mind*. Cambridge: Cambridge University Press.

--- and CHAMBERLAIN, A. (eds.) (2000); *Evolution and the human mind*. Cambridge: Cambridge University Press.

CASACUBERTA, D. (2000); *Qué es una emoción*. Madrid: Crítica.

CAVANAGH, P. (1999); "Attentional resolution: the grain and the resolution of visual awareness", en C. Taddei-Ferretti and C. Musio (eds.); *Neuronal basis and psychological aspects of consciousness*. New York: World Scientific.

CHALMERS, D. (1996); *The conscious mind*, Oxford, Oxford University Press (trad. esp. *La mente consciente*. Barcelona: Gedisa, pp. 1999).

--- (ed.) (2002); *Philosophy of mind: classical and contemporary readings*. Oxford: Oxford University Press.

--- (2007a); "Phenomenal Concepts and the Explanatory Gap", en T. Alter and S. Walter (eds.); *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Oxford: Oxford University Press.

--- (2007b); "Phenomenal Concepts and the Knowledge Argument", en P. Ludlow, Y. Nagasawa and D. Stoljar (2004), pp. 269-298.

--- and CLARK, A. (1998); "The extended mind", *Analysis*, Vol. 58, No. 1, pp. 7-19.

--- and JACKSON, F. (2001); "Conceptual analysis and reductive explanation", *Philosophical Review*, Vol. 110, No. 2, pp. 315-360.

CHACÓN, P. y RODRÍGUEZ, R. (eds.) (2000); *Pensando la mente. Perspectivas en filosofía y psicología*. Madrid: Biblioteca Nueva.

CHISHOLM, R. M. (1942); "The problem of the speckled hen", *Mind*, Vol. 51, No. 4, pp. 368-373.

--- (1989); *Theory of knowledge*. Englewood (Nj.): Prentice.

CHURCHLAND, Patricia (1986); *Neurophilosophy*. Cambridge (Ma.): MIT Press.

--- (2002); *Brain-wise*. Cambridge (Ma.): MIT Press.

--- and SEJNOWSKI, T. J. (1992); *The computational brain*. Cambridge (Ma.): MIT Press.

---, RAMACHANDRAN, V. S., and SEJNOWSKI, T. J. (1994); "A critique of pure vision", en C. Koch and J. L. Davis (eds.); *Large scale neuronal theories of the brain*. Cambridge (Ma.): MIT Press, pp. 132-157.

CHURCHLAND, Paul (1979); *Scientific realism and the plasticity of mind*. Cambridge: Cambridge University Press.

--- (1984); *Materia y conciencia*. Barcelona: Gedisa, 1999.

--- (1989); *A neurocomputational perspective*. Cambridge (Ma.): MIT Press.

--- (2005); "Quimerical colors: some phenomenological predictions from cognitive neuroscience", *Philosophical Psychology*, Vol. 18, No. 5, pp. 527-560.

CLARK, A. (1993); *Associative engines: connectionism, concepts and representational change*. Cambridge (Ma): MIT Press.

--- (2003); *Natural-born cyborgs. Minds, Technologies, and the future of human intelligence*. Oxford: Oxford University Press.

CODE, A. (1991); "Aristotle, Searle, and the Mind-Body Problem", en Lepore and Gulick (1991), pp. 105-113.

COHEN, J. (2004); "Color Properties and Color Ascriptions: a Relationalist Manifesto", *Philosophical Review*, Vol. 113, pp. 37-43.

--- HARDIN, C. L.; and McLAUGHLIN, B. P. (2006); "True colours", *Analysis*, Vol. 66, No. 4, pp. 335-340.

COLOMINA, J. J. (2006); "Sobre libertad y necesidad", en *Actas del V Congreso de la Sociedad de Lógica, Metodología y Filosofía de la Ciencia en España*. Granada: Universidad de Granada-SLMFCE, pp. 467-470.

--- (2007); "Austin sobre percepción. Por una eliminación de los datos de los sentidos", *Eikasía*, No. 11, pp. 61-70.

--- (2008a); "¿Es subjetiva la experiencia del color? Sobre el argumento representacionista de Tye acerca de la Variación Estándar", en Vázquez, M. E. y Alonso, A. (eds.); *Periferias. El extremo como término medio*. Madrid: Verbum, pp. 93-103.

--- (2008b); "Tomando conciencia de nuestro cuerpo: las explicaciones de Tye sobre las sensaciones corporales", *Dilema*, Vol. XII, No. 1, pp. 103-123.

--- (2008c); “Acerca de la implausibilidad de la existencia de un órgano del lenguaje”, *Ludus Vitalis*, Vol. XVI, No. 29, 143-161.

--- (2008d); “Identificación y representación. Un intento dualista de explicar la percepción”, *Laguna*, No. 22, pp. 122-124.

--- (por aparecer); “El contextualismo de los actos de habla. Lo que la filosofía lingüística puede aportar al estudio de la percepción”, en M. Liz (ed.); *Los velos de la realidad*. Madrid: Laertes.

CORCORAN, K. (2001); “The trouble with Searle’s biological naturalism”, *Erkenntnis*, Vol 55, No. 2., pp. 307-324.

CRANE, T. (1992); “The nonconceptual content of experience”, en T. Crane (ed.); *The contents of experience*. Cambridge: Cambridge University Press.

--- (1995); *The mechanical mind*. London: Routledge.

--- (1998); “Intentionality as the mark of the mental”, en A. O’Hear (ed.); *Current issues in Philosophy of Mind*. Suppl. of *Philosophy*, No. 43, pp. 229-251.

--- (2003); “Mental substances”, en A. O’Hear (ed.); *Minds and persons*. *Philosophy suppl.*, No. 53, pp. 229-250.

--- (2005); “Papineau on phenomenal concepts”, *Philosophy and Phenomenological Research*, Vol. 71, No. 1, pp. 155-162.

--- (2007); “Intentionality”, en A. Beckermann and B. McLaughlin (eds.); *Oxford Handbook to the Philosophy of Mind*. Oxford: Oxford University Press (trad. esp. “Intencionalidad”, *Laguna*, No. 19, 2006, pp. 9-28).

CRICK, F. (1989); *The Astonishing hypothesis*. New York: Scribner.

--- and KOCH, (1990); “Towards a neurobiological theory of consciousness”, *Seminars in the Neuroscience*, Vol. 2, pp. 263-275.

CUMMINS, R. (1975); “Functional analysis”, *The Journal of Philosophy*, Vol. 72, 741-765.

--- (2002); “Neo-teleology”, en A. Ariew, R. Cummins, and M. Perlman (eds.); *Functions*. Oxford: Oxford University Press, pp. 157-173).

CUSSINS, A. (1990); “The connectionist construction of concepts”, en M. Boden (ed.); *The philosophy of artificial intelligence*. Oxford: Oxford University Press, pp. 368-440.

DAMASIO, A. (1994); *Descartes’s error*. New York: Putnam’s Sons (trad. esp. El error de Descartes. Madrid: Crítica, 2006).

DAVIDSON, D. (1987); “Knowing one’s own mind”, en D. Davidson (2001); *Subjective, Intersubjective, Objective*, Oxford, Oxford University Press, pp. 41-71.

DAVIES, M. (1989); “Tacit knowledge and subdoxastic states”, en A. George (ed.); *Reflections on Chomsky*. Oxford: Blackwell.

--- (1992); “Perceptual content and local supervenience”, *Proceedings of the Aristotelian Society*, Vol. 92, pp. 21-45.

--- (1996); “Externalism and experience”, en N. Block, G. Güzeldere, and O. Flanagan (eds.); *The nature of consciousness*. Cambridge (Ma.): MIT Press, 1997.

--- (1998); “Language, thought and the language of thought (Aunty’s Own argument revised)”, en P. Carruthers and J. Boucher (eds.); *Language and thought: interdisciplinary themes*. Cambridge: Cambridge University Press.

--- (2000); “Persons and their underpinnings”, en J. L. Bermúdez and M. Elton (eds.); *Personal and subpersonal: Essays on Psychological Explanation. Philosophical Explorations*, Vol. 3, pp. 63-82.

DAWKINS, Richard (1986); *The blind watchmaker*. London: Penguin Books.

--- (1989); *The selfish gene*. Oxford: Oxford University Press (trad esp. *El gen egoísta*. Barcelona: Salvat, 2002).

DEHAENE, S. and NACCACHE, L. (2001); “Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework”, *Cognition*, Vol. 79, pp. 1-37.

DENNETT, Daniel C. (1969); *Content and consciousness*. London: Routledge.

--- (1978a); *Brainstorms: philosophical essays on mind and psychology*, Brighton: Harvester Press.

--- (1978b); “Current issues in the philosophy of mind”, *American Philosophical Quarterly*, Vol. 15, No. 4, pp. 249-261.

--- (1982); “How to study consciousness empirically: or nothing comes to mind,” *Synthese*, Vol. 53, pp. 159-80.

--- (1984a); *Elbow room: the varieties of free will worth wanting*. Cambridge (Ma.): MIT Press.

--- (1984b); “I could not have done otherwise –so what?”, *The Journal of Philosophy*, Vol. 81, No. 2, pp. 553-567.

--- (1986a); “The moral first Aid Manual”, en S. McMurrin (ed.); *The Tanner Lectures on Human Values*, Vol. 7. University of Utah Press, pp. 121-147.

--- (1986b); “Cognitive wheels: the frame problem of AI”, en Ch. Hookway (ed.); *Minds, machines and evolution*. Cambridge: Cambridge University Press, pp. 129-151.

--- (1987a); *The intentional stance*. Cambridge (Ma.): MIT Press.

--- (1987b); “The logical geography of computational approaches: a view from the East Pole”, en M. Brand & M. Harnish (eds.); *Problems in the representation of knowledge*. Tucson: University of Arizona Press, pp. 59-79.

--- (1988a); “Quining qualia”, en W. Lycan (ed.) (1990); *Mind and Cognition*. Oxford: Blackwell, pp. 519-547. Reimpreso en Chalmers (eds.), 2002 (por donde se cita)

--- (1988b); “Review of J. Fodor, Psychosemantics”, *The Journal of Philosophy*, Vol. LXXXV, pp. 384-389.

--- (1989); “The origins of Selves”, *Cogito*, No. 3, pp. 163-173.

--- (1990); “The interpretation of texts, people and other artifacts”, *Philosophy and Phenomenological Research*, Vol. 50, supplement, pp. 177-194.

--- (1991); *Consciousness explained*. Boston: Little Brown (trad. esp. *La conciencia explicada*. Barcelona: Paidós, 1995).

--- (1992a); "Temporal anomalies of Consciousness: implications of the uncentered brain", en Y. Christen and P. S. Churchland (eds.); *Neurophilosophy and Alzheimer's Disease*. Berlin: Springer, pp. 5-17.

--- (1992b) "The Self as a Center of Narrative Gravity", en F. Kessel, P. Cole and D. Johnson (eds.); *Self and Consciousness: Multiple Perspectives*. Hillsdale: Erlbaum, pp. 299-325.

--- (1993a); "Précis of *Consciousness Explained*", *Philosophy and Phenomenological Research*, Vol. LIII, No. 4, pp. 889-891.

--- (1993b); "The message is: There is no *Medium*", *Philosophy and Phenomenological Research*, Vol. LIII, No. 4, pp. 919-931.

--- (1993c); "Evolution, Teleology, Intentionality (Reply to Ringen and Bennett)", *Behavioral and Brain Sciences*, Vol. 16, No. 2, pp. 289-291.

--- (1993d); "Review of Searle's *The rediscovery of the mind*", *The Journal of Philosophy*, 60, pp. 193-205.

--- (1993e); "Learning and labeling", *Mind and Language*, vol. 8, pp. 540-547.

--- (1994a); "Get real", *Philosophical Topics*, vol. 22, pp. 505-568.

--- (1994b); "The role of language in intelligence", en J. Khalifa (ed.), *What is intelligence?*. Cambridge: Cambridge University Press.

--- (1995a); *Darwin's dangerous idea: evolution and the meaning of life*. New York: Simon and Schuster.

--- (1995b); "The unimagined preposterousness of zombies", *Journal of Consciousness Studies*, Vol. 2, No. 4, pp. 322-336.

--- (1995c); "Get real", *Philosophical Topics*, Vol. 22, No. 1/2, pp. 505-568.

--- (1996a); *Kinds of minds: toward an understanding of consciousness*. New York: Basic Books.

--- (1996b); "Cow-sharks, magnets, and swampman", *Mind and Language*, Vol. 11, No. 1, pp. 76-77.

--- (1997); "How to do other things with words", en J. Preston (ed.) *Philosophy*, Vol. 42, supplement, pp. 219-235.

--- (1998); *Brainchildren. Essays on disenging minds*. Cambridge (Ma): MIT Press.

--- (2001); "The zombie hunch: extinction of an intuition?", en A. O'Hear (ed.); *Philosophy at the new millenium*, suppl. No. 48 of the Royal Institute of Philosophy. Cambridge: Cambridge University Press, pp. 27-43.

--- (2003a); *Freedom evolves*, New York: Viking Penguin.

--- (2003b); "On failures of freedom and the fear of science", *Daedalus*, No. 9, pp. 126-130.

--- (2003c); "Making ourselves at home in our machines", *Journal of Mathematical Psychology*, Vol. 47, pp. 101-104.

--- (2005a); *Sweet dreams*, Cambridge (Ma.), MIT Press (trad. esp. *Dulces sueños*. Buenos Aires: Katz, 2006).

--- (2005b); "Natural freedom(Reply to A. Mele, J. M. Fischer and T. O'Connor)", *Metaphilosophy*, Vol. 36, No. 4, pp.449-459.

- (2006); *Breaking the spell*. New York: Penguin Books.
- (2008); "Descartes argument from design", *The Journal of Philosophy*, Vol. CV, No. 7, pp. 333-345.
- and HAUGELAND, J. (1987a); "Consciousness", en R. L. Gregory (ed.), *The Oxford Companion to the mind*. Oxford: Oxford University Press, pp. 160-164.
- and HAUGELAND, J. (1987b); "Intentionality", en R. L. Gregory (ed.), *The Oxford Companion to the mind*. Oxford: Oxford University Press, pp. 383-386.
- DERKSEN, A. A. (2005); "Dennett's rhetorical strategies in *Consciousness Explained*", *Journal for General Philosophy of Science*, Vol. 36, No. 1, pp. 29-48.
- DEVITT, M. (1998); "A narrow representation theory of the mind", en S. Silvers (ed.); *Re-representations. Readings in the Philosophy of Mental Representations*. Dordrecht: Kluwer Academic.
- DÍAZ-LEÓN, E. (2008); "We are living in a material worlds (and I am a material girl)", *Teorema*, Vol. XXVII, No. 3, pp. 85-102.
- DORBOLO, J. (2006); "Intuition pumps", *Minds and Machines*, Vol. 16, No. 1, pp. 81-86.
- D'ORO, G. (2007); "The gap is semantic, not epistemological", *Ratio*, Vol. 20, No. 2, 168-178.
- DRETSKE, Fred (1969); *Seeing and Knowing*. Chicago: The University of Chicago Press.
- (1981); *Knowledge and the flow of information*. Cambridge (Ma.): MIT Press.
- (1986); "Misrepresentation", en A. Goldman (1993), pp. 297-314.
- (1988); *Explaining Behavior: Reasons in a World of Causes*. Cambridge (Ma): MIT Press.
- (1990); "Reply to reviewers", *Philosophy and Phenomenological Research*, Vol. 50, pp. 819-839.
- (1993); "Conscious experience", *Mind*, Vol. 102, No. 406, pp. 263-283.
- (1995); *Naturalizing the mind*. Cambridge (Ma.): MIT Press.
- (1996); "Phenomenal externalism", en Villanueva (1996).
- (2001); "Norms, history and the mental", en D. M. Walsh (ed.); *Naturalism, evolution and mind. Philosophy suppl.*, No. 49, pp. 87-104.
- (2003); "Experience as representation", *Philosophical Issues, No. 13. Philosophy of mind*, pp. 67-82.
- (2007); "What change blindness teaches about consciousness", *Philosophical Perspectives*, Vol. 21, No. 2, pp. 216-230.
- EDELMAN, G. (1987); *Neural darwinism*. New York: Basic Books.
- (1989); *Remembered present: a biological theory of consciousness*. New York: Basic Books.
- ERECINSKA, M. and SILVER, I. (1994); "Ions and energy in the mammalian brain", *Programme in Neurobiology*, Vol. 43, No. 1, pp. 37-71.
- EVANS, G. (1982); *The varieties of reference*. Oxford: Oxford University Press.
- FARKAS, K. (2008); *The subject's point of view*. Oxford: Oxford University Press.

FEILG, H. (1959); "The 'mental' and the 'physical'", *Minnesota Studies in the Philosophy of Science*, Vol. II.

--- (1960); "Mind-body, not a pseudo-problem", en Hook, S. (ed.); *Dimensions of mind*. New York: New York University Press (reimpreso en Borst (ed.), pp. 33-41).

FEYERABEND, P. (1963a); "Mental events and the brain", *The Journal of Philosophy*, Vol. LX, No. 11, pp. 295-296 (reimpreso en Borst (ed.), pp. 140-141).

--- (1963b); "Materialism and the mind-body problem", *The Review of Metaphysics*, Vol. XVII (reimpreso en Borst (ed.), pp. 142-156).

FIELD, H. (1977); "Logic, meaning, and conceptual role", *The Journal of Philosophy*, Vol. 74, No. 2, pp. 379-409.

FISCHER and RAVIZZA (1998); *Responsibility and control: a theory of moral responsibility*. Cambridge: Cambridge University Press.

FODOR, J. (1975); *Solipsismo metodológico*

--- (1980); *The language of thought*. Cambridge (Ma.): Harvard University Press.

--- (1987); *Psychosemantics*. Cambridge (Ma.): MIT Press.

--- (1990); *A theory of content*. Cambridge (Ma.): MIT Press.

--- (1998); *Concepts*. Oxford: Oxford University Press.

--- (2000); *The mind doesn't work that way*. Cambridge (Ma.): MIT Press.

--- and LEPORE, E. (1994); "What is the Connection Principle?", *Philosophy and Phenomenological Research*, Vol. 54, No. 4, pp. 837-845.

--- and PYLYSHYN (1988); "Connectionism and cognitive architecture: a critical analysis", *Cognition*, Vol. 28, pp. 3-71.

FOTION, N. (2001); *John Searle*. New Jersey (NJ): Princeton University Press.

FRANKFURT (1982); *On Bullshit*. Princeton: Princeton University Press.

FREEMAN, W. J. and SKARDA, C. A. (1991); "Mind/Brain Science: Neuroscience on Philosophy of Mind", en Lepore and Gulick (1991), pp. 115-127.

GARCÍA-CARPINTERO, M. (1996); "Superveniencia y determinación del contenido amplio", *Revista de Filosofía*, Vol. IX, No. 16, pp. 57-92.

--- (1998); "La naturalización de las ciencias cognitivas", en N. Català (ed.); *El lenguaje y la mente humana*. Madrid: Ariel, pp. 181-200.

--- (1999); "Searle on perception", *Teorema*, Vol. XVIII, No. 1, pp. 19-41.

--- (2000); "Las razones para el dualismo", en Chacón y Rodríguez (2000), pp. 27-120.

--- (2001); "El carácter externo de la conciencia", en M. C. Paredes (ed.); *Mente, Conciencia y Conocimiento*. Salamanca: Ediciones de la Universidad de Salamanca, pp. 55-90.

--- and MACIÀ, J. (eds.) (2006); *Two-dimensional semantics*. Oxford: Oxford University Press.

GARRETT, B. J. (1995); "Non-reductionism and John Searle's *The rediscovery of mind*", *Philosophy and Phenomenological Research*, Vol. 55, No. 1, pp. 209-215.

GENNARO, R. J. (1996); *Consciousness and Self-consciousness*. Amsterdam: John Benjamin.

GOLDMAN, A. (1993); "Consciousness, Folk Psychology, and Cognitive Science", *Consciousness and Cognition*, Vol. 2, pp. 364-388.

--- (ed.) (1993); *Readings in Philosophy and Cognitive Science*. Cambridge (Ma.): MIT Press.

GONZÁLEZ-CASTÁN, Ó. L. (1992); "Intencionalidad sin conciencia: Brentano, Searle y las ciencias cognitivas", *Revista de Filosofía*, No.7, pp. 99-118.

--- (1999); "The connection principle and the classificatory scheme of reality", *Teorema*, Vol. XVIII, No. 1, pp. 85-98.

GOULD, S. J. (1980); *The panda's thumb*. New York: Norton.

--- (1991); *Bully for Brontosaurus*. New York: Norton.

--- and LEWONTIN, R. (1979); "The sprandels of San Marco and the Panglossian paradigm", *Proceedings of the Royal Society*, B205, pp. 581-598.

GUERRERO, J. A. (2001a); "El naturalismo biológico", en P. Chacón (ed.); *Filosofía de la Psicología*. Madrid: Biblioteca Nueva, pp. 153-168.

--- (2001b); "Perspectivas actuales sobre la conciencia", en P. Chacón (ed.); *Filosofía de la Psicología*. Madrid: Biblioteca Nueva, pp. 171-204.

--- (2001c); "Problemas epistemológicos subyacentes a la teoría de la mente de Searle", *Logos. Anales del Seminario de Metafísica*, Vol. 34, pp. 297-316.

GULICK, R. van (1995); "Why the connection argument doesn't work", *Philosophy and Phenomenological Research*, Vol. 55, No. 1, 201-207.

GÜZELDERE, G. (1997); "The many faces of consciousness", en N. Block, O. Flanagan and G. Güzelde (eds.); *The nature of consciousness*. Cambridge (Ma.): MIT Press, pp. 1-67.

HARDIN, C. L. (1988); *Color for philosophers: unweaving the rainbow*. Indianapolis (In.): Hackett Books.

--- (2003); "A Spectral Reflectance doth not a Color make", *The Journal of Philosophy*, Vol. 100, pp. 192-202.

HARMAN, G. (1973); *Thought*. Princeton: Princeton University Press.

--- (1990); "The intrinsic quality of experience", en Tomberlin (1990), pp. 31-52.

HATZIMOYSIS, A. (2007); "The case against unconscious emotions", *Analysis*, Vol 67, No. 4, pp. 292-299.

HAUGELAND, J. (1981); *Mind design*. Cambridge (Ma.): MIT Press.

--- (1985); *Artificial Intelligence: The Very Idea*. Cambridge (Ma.): MIT Press.

HECK, R. G. (2000); "Nonconceptual content and the 'space of reasons'", *Philosophical Review*, Vol. 109, No. 4, pp. 483-523.

HERMOSO, J. y CHACÓN, P. (2000); "Sobre el carácter irreductible de la intencionalidad: la ontología del inconsciente y los dos conceptos de Trasfondo en Searle", en Chacón y Rodríguez (eds.), pp. 167-194.

HILL, C. (2005); "Ow! The paradox of pain", en Aydede (2005), pp. 75-98.

HIRSTEIN, W. (2001); *On Searle*. Belmont (Ca.): Wadsworth.

HONDERICH, T. (1995); "Consciousness, Neural functionalism, real subjectivity", *American Philosophical Quarterly*, Vol. 32, No. 4, pp. 369-381.

--- (2002);

HORGAN, T. and TIENSON, T. (2002); "The Intentionality of Phenomenology and the Phenomenology of Intentionality", en D. Chalmers (2002), pp. 520-33.

HORGAN, T. and KRIEGEL, U. (2007); "Phenomenal epistemology: what is consciousness that we know it so well?", en E. Sosa and E. Villanueva (eds.); *The Metaphysics of Epistemology. Philosophical Issues*, Vol. 17, pp. 123-144.

HUME, D. (1739); *Tratado sobre la naturaleza humana*. Madrid: Akal, 2000.

--- (1748); *Investigación sobre el entendimiento humano*. Madrid: Istmo, 2005.

HUMPHREYS, G. W. (2003); "Conscious visual representations built from multiple binding processes: evidence from neuropsychology", *Progress in Brain Research*, Vol. 142, no. 3, pp. 243-255.

HURLEY, S. L. (1998); *Consciousness in action*. Cambridge (Ma.): Harvard University Press.

--- (2000); "Clarifications: responses to Kobes and Kinsbourne", *Mind and Language*, Vol. 15, No. 5, pp. 556-561.

HUXLEY, T. H. (1866); *Lessons on elementary physiology*. London: Macmillan.

JACKENDOFF, R. (1987); *Consciousness and computational mind*. Cambridge (Ma.): MIT Press.

JACKSON, F. (1977); *Perception*. Cambridge: Cambridge University Press.

--- (1982); "Epiphenomenal qualia", *The Philosophical Quarterly*, Vol. 32, pp. 127-136.

--- (1986); "What Mary didn't know", *The Journal of Philosophy*, Vol. 83, pp. 291-295.

--- (1993); "Armchair Metaphysics", en J. O. Hawthorne and M. Michaelis (eds.); *Philosophy in Mind, Philosophical Studies*, vol. 60, pp. 23-42.

--- (2000). "Representation, scepticism, and the a priori", en P. Boghossian and C. Peacocke (eds.) (2000). *New essays on the a priori*. Oxford: Oxford University Press, pp. 320-332.

--- (2001). "Locke-ing onto content", en D. M. Walsh (ed.) (2001). *Naturalism, evolution and mind*. Royal Institute of Philosophy, *Philosophy* suppl., No. 49, pp. 127- 143.

--- (2003); "Representation and narrow belief", *Philosophical Issues*, No. 13, *Philosophy of mind*, pp. 99-113.

--- and PETTIT, P. (1990); "Causation in the Philosophy of Mind", *Philosophy and Phenomenological Research*, Vol. 50, suppl., pp. 195-214.

JACOB, P. (1995); "Consciousness, intentionality and function. What is the right order of explanation?", *Philosophy and Phenomenological Research*, Vol. 55, No. 1, 195-200.

JAMES, W. (1884); "What is an emotion?", en *Essays in Philosophy*. Cambridge (Ma.): Harvard University Press, 1990.

--- (1890); *The principles of Psychology*. New York: Dover, 1999.

JEHLE, D. and KRIEGEL, U. (2006); "An argument against dispositionalist HOT Theory", *Philosophical Psychology*, Vol. 19, No. 4, pp. 463-476.

- KANE (1996); *The significance on free will*. Oxford: Oxford University Press.
- KAPLAN, D. (1989); “Demonstratives” and “Afterthoughts”, en J. Almog (ed.); *Themes from Kaplan*. Oxford: Oxford University Press.
- KIM, J. (1984); “Concepts of supervenience”, *Philosophy and Phenomenological Research*, Vol. 45, No. 1, pp. 153-176.
- (1995); “Mental Causation in Searle’s ‘Biological Naturalism’”, *Philosophy and Phenomenological Research*, Vol. 55, pp. 189-194.
- (1998); *Philosophy of mind*. Oxford: Westview Press.
- (2005); “Mental Causation, Reduction, and Consciousness: Where We Now Stand with the Mind-Body Problem”, en O. Kang (ed.), *Contemporary British and American Philosophy*.
- KINSBOURNE, M. (2000); “Consciousness in action: antecedents and origins”, *Mind and Language*, Vol. 15, No. 5, pp. 545-555.
- KIRK, R. (1974a); “Sentience and Behaviour”, *Mind*, Vol. 83, No.1, pp. 43-60.
- (1974b); “Zombies vs. Materialists”, *Aristotelian Society Proceedings*, suppl., Vol. 48, pp. 135-152.
- (1994); *Raw feeling*. Oxford: Oxford University Press.
- (2005); *Zombies and consciousness*; Oxford: Oxford University Press.
- (2008); “Physicalism and phenomenal consciousness”, *Teorema*, Vol. XXVII, No. 3, pp. 71-84.
- KOBES, B. W. (2000); “Unity of Consciousness and bi-level externalism”, *Mind and Language*, Vol. 15, No. 5, pp. 528-544.
- KOCH, C. and Davis, J. L. (eds.) (1994); *Large-Scale neuronal theories of the brain*. Cambridge (Ma.): MIT Press.
- KRIEGEL, U. (2002); “PANIC theory and the prospects for a representational theory of phenomenal consciousness”, *Philosophical Psychology*, Vol. 15, No. 1, pp. 55-64.
- (2006); “Naturalizing subjective character”, *Philosophy and Phenomenological Research*, Vol. LXXI, No. 1, pp. 23-57.
- and WILLIFORD, K. (eds.) (2009); *Self-Representational approaches to Consciousness*. New York: Bradford Books.
- KRIPKE, S. (1971/1980); “Identity and necessity” (trad. esp. “Identidad y necesidad”, en L. M. Valdés Villanueva (ed.), *La búsqueda del significado*. Madrid: Tecnos, 1999.
- (1972/1981); “Naming and necessity”, en D. Davidson and G. Harman (eds.); *Semantics of natural languages*. Dordrecht: Riedel.
- KUPFERMANN, I. (1991); “Genetic determinants of behaviour”, en E. Kandel and T. Jessell (eds.); *Principles of neural science*. New York: Elsevier.
- KUTTNER, F. and ROSENBLUM, B. (2006); “The only objective evidence for Consciousness”, *The Journal of Mind and Behavior*, Vol. 27, N. 1, pp. 43-56.
- LAPIEDRA, R. (2005); *Els dèficits de la realitat i la creació del món*. València: Servei de Publicacions de la Universitat de València.
- LEDOUX, J. (2002); *Synaptic self*. New York: Viking Books.

- LEIBNIZ, W. (1840); *Monadología*. Oviedo: Pentalfa, 1981.
- LEPORE, E. and GULICK, R. van (1991); *Searle and his critics*. Oxford: Blackwell.
- LETTVIN *et alri* (1959): "What the frog's eye tells the frog's brain", *Proceedings of the IRE*, pp. 1940-1951.
- LEVINE, J. (1983); "Materialism and qualia: the explanatory gap", *Pacific Philosophical Quarterly*, Vol. 64, No. 3, pp. 354-361. También en Lycan (1990).
- (2001); *Purple haze: the puzzle of consciousness*. Oxford: Oxford University Press.
- (2007); "Antimaterialist arguments and influential replies", en Velmans y Scheneider (2007), capítulo 29.
- LEVIN, J. (2007); "Nagel vs. Nagel on the nature of phenomenal concepts", *Ratio*, Vol. 20, No. 3, pp. 293-307.
- LEWIS, D. (1978); "Mad pain and martian pain", en D. Lewis (1983); *Philosophical Papers (I)*. Oxford: Oxford University Press, pp 99-109.
- (1990); "What experience teaches", en Lycan (1990), pp. 499-519.
- LIGHTFOOT, D. W. (1999); *The development of language: Acquisition, change and evolution*. Oxford: Blackwell.
- LIZ, M. (1995); "Causalidad y contenido mental", en F. Broncano (ed.); *La mente humana*. Madrid: Trotta, pp. 207-243.
- (1999); "Las fronteras de lo físico", en *Proceedings of the Congress The Analytic Philosophy at the Turn of Millenium*, pp. 193-201.
- (2000); "Razones y otras causas perdidas", en Chacón y Rodríguez (2000), pp. 121-166.
- (2001a); *Perspectivas actuales en Filosofía de la Mente*. Tenerife: Gobierno de Canarias.
- (2001b); *La vida mental de algunos trozos de materia. Teorías de la sobreveniencia*. Salamanca: Factótum.
- (2003); *Justificar y explicar*. La Laguna: Servicio de Publicaciones de la Universidad de La Laguna.
- (2007); "El mundo físico y el mundo. Otros fisicalismos son posibles", en D. Pérez Chico y M. Barroso (eds.); *Pluralidad de la filosofía analítica*. Madrid: Plaza y Valdés, pp. 95-128.
- LLINÁS, R. (2003); *El cerebro y el mito del yo*. Madrid: Belacqva.
- LOAR, B. (1997); "Phenomenal states", en N. Block, O. Flanagan and G. Guzeldere (eds.); *The Nature of Consciousness*. Cambridge (Ma.): MIT Press.
- LOCKE, J. (1689); *An essay concerning human understanding*. Hertforshire: Worsworth Editions, 1998.
- LOWE (1993); "The causal autonomy of mental", *Mind*, No. 102, pp. 629-644.
- LUDLOW, P., NAGASAWA, Y. and STOLJAR, D. (eds.) (2004); *There's Something About Mary*. Cambridge (Ma): MIT Press.
- LYCAN, W. (1973); "Inverted spectrum", *Ratio*, Vol. 15, No. 3 pp. 315-319.
- (1986); "Two Concepts of Reduction: Modal Realism at Risk", *The Journal of Philosophy*, Vol. 83, No. 4, pp. 693-694.
- (ed.) (1990); *Mind and Cognition*. Oxford: Blackwell.

- (1996); *Consciousness and experience*. Cambridge (Ma.): MIT Press.
- (2000); "Representational theories of consciousness", *Stanford Encyclopedia of Philosophy*.
- (2002); "The case for phenomenal externalism", en J. E. Tomberlin (ed.); *Philosophical Perspectives, Vol. 16*. Atascadero (Ca.): Ridgeview.
- (2003); "Vs. A new a priorist argument for dualism", *Philosophical Issues, No. 13, Philosophy of mind*, pp. 130-147.
- (2006); "Pautz vs. Byrne and Tye about Externalism Intentionalism", *Online Philosophy Conference 2006*,
<http://garnet.acns.fsu.edu/~tan02/OPC%20Week%20Four>.
- MAERZ, A. and PAUL, M. (1950). *A dictionary of color*. New York: McGraw Hill.
- MALCOCK, G., KAY, P. and WEBSTER, M. A. (2005); "Variations in normal color vision. IV. Binary hues and hue scaling", *Journal of the Optical Society of America, Vol. 22*, pp. 2154–2168.
- MANRIQUE, F. M. (1999); "Contenido conceptual y no-conceptual: ¿Autonomía o interdependencia?", *Proceedings of the Congress The Analytic Philosophy at the Turn of Millenium*, pp. 633-646.
- MARGOLIS, E. and LAURENCE, S. (2007); "The ontology of concepts-Abstract objects or mental representations?", *Noûs, Vol. 41, No. 4*, pp. 561-593.
- MARKS, C (1980); *Commissurotomy, consciousness, and the unity of mind*. Cambridge (Ma.): MIT Press.
- MARR, D. (1982); *Vision*. San Francisco (Ca.): Freeman.
- MARTÍNEZ, M. (2008); "La P de PANIC. Representacionalismo y fenomenología del dolor", *Teorema, Vol. XXVII, No. 3*.
- MAUND, B. (2006); "Michael Tye on pain and representational content", en Aydede (ed.) (2006), pp. 143-150.
- McDOWELL, J. (1994); *Mind and world*. Cambridge (Ma.): Harvard University Press.
- McGINN, C. (1982); *The character of mind*. Oxford: Oxford University Press.
- (1983); *The Subjective View: Secondary Qualities and Indexical Thoughts* Oxford: Oxford University Press.
- (1989); *Mental Content*. Oxford: Blackwell.
- (1991); *The problem of consciousness*. Oxford: Blackwell.
- (1996); "Another look at color", *The Journal of Philosophy, Vol. 93, No. 3*, pp. 537-553.
- (1999); *The mysterious flame*. Cambridge (Ma.): MIT Press.
- (2004); *Consciousness and its objects*. Oxford: Clarendon Press.
- McLAUGHLIN, B. P. (2003); "Color, consciousness, and color consciousness", en Q. Smith (ed.); *Consciousness: New Philosophical Perspectives*. New York: Oxford University Press.
- and TYE, M. (1998); "Is content-externalism compatible with privileged acces?", *Philosophical Review, Vol. 107, No. 3*, pp. 349-380.

MEGILL, J. (2007); "Naturalism, physicalism and epiphenomenalism", *Philosophical Psychology*, Vol. 20, No. 6, pp. 681-686.

MESTRE D. R. *et al.* (1992); "Perception of optical flow in cortical blindness: a case report", *Neuropsychologica*, Vol. 30, pp. 783-795.

MIGUENS, S. (2002); "Qualia or non epistemic perception: Dennett's and Dretske's representational theories of consciousness", *Ágora*, Vol 21, No. 2, pp. 193-208.

MILLIKAN, R. G. (1984); *Language, thought, and other biological categories*. Cambridge (Ma.): MIT Press.

--- (1993); *White Queen Psychology and other essays for Alice*. Cambridge (Ma.): MIT Press.

--- (2001); "What has natural information to do with intentional representation?", en D. M. Walsh (ed.); *Naturalism, evolution, and mind*. Supplement of *Philosophy*, No. 49, pp. 105-125.

--- (2003); "Teleological theories of mental content", en *Encyclopedia of Cognitive Science*, Vol. 11. London: MacMillan, pp. 1138-1141.

--- (2005); *Language: a biological model*. New York: Oxford University Press.

MONTERO, B. (2007); "Physicalism could be true even if Mary learns something new", *The Philosophical Quarterly*, Vol. 57, No. 227, pp. 176-189.

MOORE, G. E. (1922); "The refutation of idealism", en su *Philosophical Studies*. London: Routledge.

MUSACCHIO, J. M. (2005); "Why do qualia and the mind seem nonphysical?", *Synthese*, Vol. 147, No. 3, pp. 425-460.

NAGEL, E. and NEWMAN, J. R. (2007); *El teorema de Gödel*. Madrid: Tecnos.

NAGEL, T. (1971); "Brain bisection and the unity of consciousness", *Synthese*, Vol. 22, pp. 396-413.

--- (1972); "Dennett's *Content and Consciousness*", *The Journal of Philosophy*, Vol 69 (trad. esp. en *Otras mentes*. Barcelona: Gedisa, pp. 107-116).

--- (1974); "What is it like to be a bat?", *Philosophical Review*, Vol. 83, pp. 435-450.

NANAY, B. (2006); "Symmetry between the intentionality of minds and machines? The biological plausibility of Dennett's position", *Minds and Machines*, Vol. 16, No. 1, pp. 73-80.

NEANDER, K. (1991); "Functions as selected effects", *Philosophy of Science*, Vol. 58, pp. 168-184.

--- (1995); "Misrepresentation and malfunctioning", *Philosophical Studies*, Vol. 79, No. 1, pp. 109-141.

NEMIROW, L. (1990); "Physicalism and the cognitive role of acquaintance", en Lycan (1990), pp. 478-489.

NEWMAN, John von (1958); *The computer and the brain*. New Haven (Co.): Yale University Press.

NEWMAN, A. (2004); "The good, the bad, and the irrational: three views about mental content", *Philosophical Psychology*, Vol 17., No. 1, pp. 95-106.

- NIDA-RÜMELIN, M. (1996); "What Mary couldn't know", en T. Metzinger (ed.); *Conscious Experience*. Paderborn: Schöningh Books, pp. 219-241.
- NÖE, A. (2004); *Action in perception*. Cambridge (Ma.): MIT Press.
- (2009); *Out of our heads. Why you are not your brain, and other lessons from the biology of consciousness*. New York: Hill and Wang.
- NOORDHOF, P. (2003); "Review of *Consciousness, color and content*", *Mind and Language*, Vol. 18, No. 5, pp. 538-545.
- (2006); "In a state of pain", en Aydede (ed.) (2006), pp. 151-162.
- NORTOFF, G. and MUSHOLT, K. (2006); "How can Searle avoid Property Dualism? Epistemic-Ontological Inference and Autoepistemic Limitation", *Philosophical Psychology*, Vol. 19, No. 5, pp. 589-605.
- O'DEA, J. (ms.); "Tye on the sense and the unity of experience", en <http://jodea.net/papers/Tye-on-phenomenal-unity.pdf>.
- OJEDA, T. (2006); *Psicología popular: simulación versus teoría de la teoría*. La Laguna: Servicio de Publicaciones de la Universidad de La Laguna.
- OLAFSON, F. A. (1994); "Brain dualism", *Inquiry*, Vol. 37, pp. 253-265.
- PALMER, D. E. (1998); "Searle on consciousness: or how not to be a physicalist", *Ratio*, Vol. XI, No. 2, pp. 159-169.
- PAPINEAU, D. (1993a); *Philosophical naturalism*. Oxford: Blackwell.
- (1993b); "Physicalism, consciousness, and the antipathetic fallacy", *Australasian Journal of Philosophy*, Vol. 79, No. 2, pp. 169-183.
- (1998a); "Mind the gap", *Philosophical Perspectives*, Vol. 12. *Language, mind, and ontology*, pp. 373-388.
- (1998b); "Teleosemantics, and indeterminacy", *Australasian Journal of Philosophy*, Vol. 76, No. 1, pp. 1-14.
- (2000); *Introducing consciousness*. London: Icon Books.
- (2001); "The evolution of means-end reasoning", en D. Walsh (ed.); *Naturalism, evolution and mind. Philosophy suppl*, No. 51.
- (2002); *Thinking about consciousness*. Oxford: Oxford University Press.
- (2003a); "Could there be a science of consciousness?", *Philosophical Issues*, Vol. 13, *Philosophy of mind*, pp. 205-220.
- (2003b); "Human minds", en A. O'Hear (ed.); *Minds and Persons. Philosophy suppl.*, No. 53, pp. 159-183.
- (2005); "The cultural origins of cognitive adaptations", en A. O'Hear (ed.); *Philosophy, biology, and life. Philosophy suppl.*, No. 56.
- PARFIT, D. (1987); "Divided minds and the nature of persons", en C. Blakemore and S. Greenfield (eds.); *Mindwaves*. Oxford: Blackwell, pp. 19-26.
- PAUTZ, A. (2006a); "Sensory Awareness is not a Wide Physical Relation: an Empirical Argument against Externalist Intentionalism", *Noûs*, Vol. 40, No. 2, pp. 205-240.

--- (2006b); "Can Intentionalists be Reductivists?", *Online Philosophy Conference 2006*, <http://garnet.acns.fsu.edu/~tan02/OPC%20Week%20Four>.

--- (2008); "The interdependence of phenomenology and intentionality", *The Monist*, Vol. 91, No. 2, pp. 255-277.

--- (2009); "The problem of color", en T. Bayne, A. Cleeremans and P. Weelkens (eds.); *Oxford Companion to Consciousness*. Oxford: Oxford University Press.

--- (forthcoming a); "Tracking Intentionalism and Optimal Conditions: a reply to Byrne and Tye", *Nous* (también disponible en la web del autor, <http://garnet.acns.fsu.edu/~tan02/OPC%20Week%20Four/Pautz.pdf>).

--- (forthcoming b); "Variation in Normal Color Vision and the Nature of Consciousness", disponible en la web del autor,

<https://webpace.utexas.edu/arp424/www/website.variation.pdf>

--- (forthcoming c); "The problem of Standard Variation is a problem about Sensory Representation: Tye on the Puzzle about True Blue", <https://webpace.utexas.edu/arp424/www/WEBSITE%20STANDARD%20VARIATION.pdf>.

PEACOCKE, C. (1983); *Sense and content. Experience, thought, and their relation*. Oxford: Clarendon Press.

--- (1992); *A study of concepts*. Cambridge (Ma.): MIT Press.

--- (1994); "Non-conceptual content: kinds, rationales and relations", *Mind and Language*, Vol. 9, No. 3, pp. 419-430.

--- (1998); "Nonconceptual content defended", *Philosophy and Phenomenological Research*, Vol. 58, pp. 381-388.

--- (1999); *Being known*. Oxford: Oxford University Press.

--- (2001). "Does perception have a nonconceptual content?". *The Journal of Philosophy*, 98, pp. 239-264.

--- (2008); "Sensational properties: Theses to accept and theses to reject", *Revue Internationale de Philosophie*, Vol. 62, No. 243, pp. 7-24.

PENROSE, R. (1989); *La nueva mente del emperador*. Barcelona: Crítica, 2007.

--- (1994); *Las sombras de la mente*. Barcelona: Crítica, 1996.

PESTANA, M. S. (2006); "Association mechanisms and the intentionality of the mental", *The Journal of Mind and Behavior*, Vol. 27, No. 2, pp. 91-120.

PINEDA, D. (1999); "Searle y el problema de la exclusión causal", *Teorema*, Vol. XVIII, No. 1, pp. 155-170.

--- (2008); "Conciencia y dualismo", *Teorema*, Vol. XXVII, No. 3, pp. 131-148.

PINKER, S. (1997); *How the mind works*. New York: Norton.

PLACE, U. T. (1956); "Is consciousness a brain process?", *British Journal of Psychology*, Vol. XLVII, pp. 44-50 (reimpreso en Borst (ed.), pp. 42-51).

PLATCHIAS, D. (2008); "Experiencing a Hard Problem?", *Teorema*, Vol. XXVII, No. 3, pp. 115-130.

- POLGER, T. (2000); "Zombies explained", en D. Ross, A. Brook and D. Thompson (eds.); *Dennett's Philosophy*. Cambridge (Ma): MIT Press, pp. 259-286.
- (2001); "Review of Dennett's *Brainchildren*", *Philosophical Psychology*, Vol. 14, No. 2, pp. 245-250.
- (2004); *Natural minds*. Cambridge (Ma): MIT Press.
- POPPER, K. And ECCLES, J. (1978); *The self and its brain* (trad. esp. *El yo y su cerebro*. Madrid: Labor, 1980).
- PRINZ, J. (2000); "A neurofunctional theory of visual consciousness", *Consciousness and cognition*, Vol. 9, No. 2, pp. 243-259.
- (2002); *Furnishing the mind. Concepts and their perceptual basis*. Cambridge (Ma.): MIT Press.
- (2004); *Gut reactions. A perceptual theory of emotion*. Oxford: Oxford University Press.
- (2005); "A neurofunctional theory of consciousness", en A. Brook and K. Akins (eds.); *Cognition and the brain. The philosophy and neuroscience movement*. Cambridge: Cambridge University Press, pp. 381-396.
- (2008a); "Précis of *Gut reactions*", *Philosophy and Phenomenological Research*, Vol. LXXVI, No. 3, pp. 707-711.
- (2008b); "Response to D'Arms and Hills", *Philosophy and Phenomenological Research*, Vol. LXXVI, No. 3, pp. 729-732.
- (forthcoming); "Mental maintenance: a response to the knowledge argument", en *The conscious brain*. Oxford: Oxford University Press (original de 2005 disponible en <http://www.nyu.edu/gsas/dept/philo/courses/consciousness05/PrinzKnowledgeArgument.pdf>).
- PUC CETI, R. (1972); "Multiple identity", *The Personalist*, Vol. 54, pp. 203-215.
- PUTNAM, H. (1975); "El significado de 'significado'", *Teorema*, Vol. XIV, No. 3-4, pp. 345-405.
- QUINE, W. (1960); *Palabra y objeto*. Barcelona: Labor, 1968.
- RAMACHANDRAN, V. S. and HIRSTEIN, W. (1997); "Three laws of qualia: What neurology tell us about the biological functions of consciousness", *Journal of Consciousness Studies*, No. 4, pp. 429-457.
- and BLAKESLEE, S. (1998); *Phantoms in the brain*. New York: W. Morrow.
- RAVENS CROFT, I. (1998); "What is it like to be someone else? Simulation and empathy", *Ratio*, Vol. 9, No. 2, pp. 170-185.
- REBER, A. S. (1997); "Caterpillars and consciousness", *Philosophical Psychology*, Vol. 10, No. 4, pp. 437-449.
- REID, T. (1785); *Essays on the intellectual powers of man*. Cambridge (Ma.): MIT Press, 1969.
- REY, G. (1998); "A narrow representationalist account of cualitative experience", en Tomberlin (1998), pp. 430-450.
- (2008); "(Even Higer-Order) Intentionality without consciousness", *Revue Internationale de Philosophie*, Vol. 62, No. 243, pp. 51-78.

RORTY, R. (1965); "Mind-body identity, privacy and categories", *The Review of Metaphysics*, Vol. XIX (reimpreso en Borst (ed.), pp. 187-213).

--- (1970a); "In defense of eliminative materialism", *The Review of Metaphysics*, Vol. XXIV, No. 1, pp. 112-121 (reimpreso en Rosenthal (ed.), pp. 223-231).

--- (1970b); "Incorrigibility and the mark of the mental", *The Journal of Philosophy*, Vol. LXVII, pp. 399-424.

--- (1979); *La filosofía y el espejo de la naturaleza*. Madrid: Cátedra, 1989.

--- (1982); *Consecuencias del pragmatismo*. Madrid: Tecnos, 1996.

ROSENTHAL, D. M. (1986); "Two concepts of consciousness", *Philosophical Studies*, Vol. 49, pp. 329-359.

--- (2000); "Content, interpretation, and consciousness", en D. Ross, A. Brook and D. Thompson (eds.); *Dennett's Philosophy*. Cambridge (Ma): MIT Press, pp. 287-308.

ROSEN, D. (2001); "Abstract objects", *Stanford Encyclopedia of Philosophy*.

RUZ, M., TUDELA, P. y ACERO, J. J. (2002); "La conciencia explicada por Dennett: una revisión crítica desde la neurociencia cognitiva", *Theoria*, Vol. 17, No. 1, pp. 81-112.

SABATÉS, M. H. (1999); "Consciousness, emergence and naturalism", *Teorema*, Vol. XVIII, No. 1, pp. 139-153.

SACKS, O. (1984); *Con una sola pierna*. Barcelona: Anagrama, 1998.

--- (1987); *El hombre que confundió a su mujer con un sombrero*. Barcelona, Anagrama, 2002.

SAMUELS, R. (2000); "Massively modular minds: evolutionary psychology and cognitive architecture", en Carruthers and Chamberlain (2000), pp. 13-46.

SANFÉLIX, V. (2003); *Mente y conocimiento*. Madrid: Biblioteca Nueva.

SCHIER, E. (2007); "The represented object of color experience", *Philosophical Psychology*, Vol. 20, No. 1, pp. 1-27.

SCHROETER, L. (2007); "The illusion of transparency", *Australasian Journal of Philosophy*, Vol. 85, No. 4, pp. 597-618.

SEAGER, W. (2002); "Some awkwardness in poised content?", *A field guide to Philosophy of mind*. Symposium on Michael Tye's *Consciousness, color, and content*, en <http://host.uniroma3.it/progetti/kant/field/tyesymp.pdf>, pp. 54-64.

SEARLE, J. (1979a); "What is an intentional state?", *Mind*, Vol. 88, No. 349.

--- (1979b); "Literal meaning", *Erkenntnis*, Vol. 1, No. 2, pp. 207-224.

--- (1980a); "Minds, brains and programs", *Behavioral and Brain Science*, Vol. 3, pp. 417-423, disponible en T. O'Connor and D. Robb (eds.); *Philosophy of mind*. London: Routledge, 2003, pp. 332-352.

--- (1980b); "The background of meaning", en J. R. Searle, F. Keifer and M. Bierwisch (eds.); *Speech Acts Theory and Pragmatics*. Dordrecht: Reidel.

--- (1981); "Intentionality and method", *The Journal of Philosophy*, Vol. 78, pp. 720-733.

--- (1982); "Proper names and intentionality", *Pacific Philosophical Quarterly*, Vol. 63, No.

3.

- (1983); *Intentionality. An essay in the philosophy of mind*. Cambridge: Cambridge University Press (trad esp. *Intencionalidad*. Barcelona: Paidós, 1992).
- (1984); "Intentionality and its place in Nature", *Synthese*, Vol. 61, pp. 3-16.
- (1985); *Minds, brains and science*. Boston: Harvard University Press (trad. esp. *Mentes, cerebros y ciencia*. Madrid: Cátedra, 1994).
- (1989); "Consciousness, unconsciousness, and intentionality", *Philosophical Topics*, Vol. XVII, No. 2, pp. 193-209.
- (1990a); "Who is computing with the brain?", *Behavioral and Brain Science*, Vol. 13, No. 2, pp. 632-639.
- (1990b); "Consciousness, Explanatory Inversion and Cognitive Science", *Behavioral and Brain Science*, Vol. 13, pp. 585-642.
- (1991a); "Response: Meaning, intentionality, and speech acts", en Lepore and Gulick (1991), pp. 81-102.
- (1991b); "Response: The mind-body problem", en Lepore and Gulick (1991), pp. 141-146.
- (1991c); "Response: perception and the satisfactions of intentionality", en Lepore and Gulick (1991), pp. 181-192.
- (1991d); "Response: the Background of intentionality and action", en Lepore and Gulick (1991), pp. 290-299.
- (1992); *The rediscovery of mind*. Cambridge (Ma.): MIT Press (trad esp. *El redescubrimiento de la mente*. Madrid: Crítica, 1996).
- (1994); "The connection Principle and the ontology of the unconscious: a reply to Fodor and Lepore", *Philosophy and Phenomenological Research*, Vol. 54, No. 4, pp. 847-855.
- (1995a); *The construction of social reality*. New York: The free Press.
- (1995b); "Consciousness, the brain and the Connection Principle: a reply", *Philosophy and Phenomenological Research*, Vol. 55, No. 1, 217-232.
- (1997); *The mystery of consciousness*. London: Granta (trad esp. *El misterio de la conciencia*. Barcelona: Paidós, 2000).
- (1998); "The explanation of cognition", en J. Preston (ed.); *Thought and language*. Suppl. *Philosophy*, No. 42, Royal Institute of Philosophy, pp. 103-126.
- (1999); "El trasfondo de la intencionalidad", *Teorema*, Vol. XVIII, No. 1, pp. 7-18.
- (2000); *Razones para actuar. Una teoría del libre albedrío*. Oviedo: Nobel.
- (2002a); *Consciousness and language*. Cambridge: Cambridge University Press.
- (2002b); "Why I am not a property dualist", *Journal of Consciousness Studies*, Vol. 9, No. 12, pp. 57-64.
- (2004); *Mind. A brief introduction*. Oxford: Oxford University Press.
- (2007); "Biological naturalism", en Velmans and Sneider (2006), cap. 25.
- SELLARS, W. (1963); "Empiricism and the philosophy of mind", en *Science, perception and reality*. Atascadero (Ca.): Ridgeview, pp. 127-196.

- SHOEMAKER, S. (1982); "The inverted spectrum", *The Journal of Philosophy*, Vol. 79, pp. 357-381.
- (1984); *Identity, causes, and mind*. Cambridge: Cambridge University Press.
- (1990); "Qualities and qualia: What's in the mind", *Philosophy and Phenomenological Research*, Vol. 50, suppl., pp. 109-131.
- (1991); "Qualia and consciousness", *Mind*, Vol. 100, pp. 507-524.
- (1994); "Self-knowledge and 'inner sense'. Lecture II: the broad perceptual model", *Philosophy and Phenomenological Research*, Vol. 54, No. 2, pp. 271-290.
- (1997); "Phenomenal character", en A. Byrne and D. Hilbert (eds.); *Readings on color, volume 1*. Cambridge (Ma.): MIT Press.
- SINGER, W. (1993); "Synchronisation of cortical activity and its putative role in information processing and learning", *Annual Review of Physiology*, Vol. 55, No. 2, 349-374.
- (1995); "Development and plasticity of cortical processing architectures", *Science*, No. 270, pp. 758-764.
- SKIDELSKY, L. (2007); "La distinción doxástico-subdoxástico", *Crítica*, Vol. 39, No. 115, pp. 31-60.
- SMART, J. J. C. (1959); "Sensations and brain processes", *The Philosophical Review*, Vol. 68, No. 2, pp. 141-156 (reimpreso en Borst (ed.), pp. 52-66).
- (1962); "Brain processes and incorrigibility", *The Australasian Journal of Philosophy*, Vol. XL (reimpreso en Borst (ed.), pp. 107-109).
- SOBER, E. (1995); "Natural selection and distributive explanation", *British Journal for the Philosophy of Science*, Vol. 16, pp. 384-397.
- SPERLING, G. (1960); "The information available in brief visual presentation", *Psychological Monographs*, Vol. 74, No. 11, pp. 1-29.
- SPERRY, (1968); "Hemisphere disconnection and unity in consciousness awareness", *American Psychologist*, Vol. 23, No. 5, pp. 723-733.
- STALNAKER, R. (1984); *Inquiry*. Cambridge (Ma.): MIT Press.
- (1989); "On what is in the head", en Tomberlin (1989), pp. 287-316.
- (1996); "On a defence of the hegemony of the representation", en E. Villanueva (ed.); *Perception. Philosophical Issues*, No. 7. Atascadero (Ca.): Ridgeview, pp. 101-108.
- (1998a); "What might nonconceptual content be?", en E. Villanueva (ed.); *Philosophical Issues*, No. 9. Atascadero (Ca.): Ridgeview.
- (1998b); "Replies to comments", en E. Villanueva (ed.); *Philosophical Issues*, No. 9. Atascadero (Ca.): Ridgeview.
- STICH, S. (1979); "Autonomous Psychology and the belief-desire thesis", en A. Goldman (1993), pp. 699-718.
- (1983); *From Folk Psychology to Cognitive Science*. Cambridge (Ma.): MIT Press.
- (1990); *The fragmentation on reason*. Cambridge (Ma.): MIT Press.
- STIGOL, N. (1999); "Acerca del carácter consciente de los estados mentales", *Teorema*, Vol. XVIII, No. 1, pp. 99- 108.

- STOLJAR, D. (2005); "Physicalism and phenomenal concepts", *Mind and Language*, Vol. 20, No. 5, pp. 469-494.
- and GOLD, I. (1998); "On biological and Cognitive Neuroscience", *Mind and Language*, Vol. 13, No. 1, pp. 110-131.
- STRAWSON, G. (1986); *Freedom and belief*. Oxford: Oxford University Press.
- (1994); *Mental reality*. Cambridge (Ma.): MIT Press.
- THAU, M. (2002); *Consciousness and cognition*. Oxford: Oxford University Press.
- TOMASELLO, M. and CALL, J. (1997); *Primate cognition*. Oxford: Oxford University Press,
- TOMBERLIN, J. E. (ed.) (1990); *Philosophical Perspectives, Vol. 4. Philosophy of mind and Action Theory*. Atascadero (Ca.): Ridgeview.
- (ed.) (1998); *Philosophical Perspectives, Vol. 12. Language, Mind, and Ontology*. Atascadero (Ca.): Ridgeview.
- TRAVIS, Ch. (2002); "Frege's target", A. O'Hear (ed.); *Logic, Thought and Language. Philosophy suppl.*, No. 51, pp. 305-343.
- (2004). "The silence of senses". *Mind*, Vol. 113, No. 449, pp. 57-94.
- TURING, A. (1937); "On computable numbrers, with and application to the *Entscheidungsproblem*", *Proceedings of the London Mathematical Society*, Vol. XLII, pp. 230-265.
- TYE, M. (1991); "Visual qualia and visual content", en T. Crane (ed.); *The contents of experience*. Cambridge: Cambridge University Press, pp. 158-176.
- (1995); *Ten problems of consciousness*. Cambridge (Ma.): MIT Press.
- (1999); "Phenomenal consciousness: the explanatory gap as a cognitive illusion", *Mind*, Vol. 108, No. 4, pp. 705-725.
- (2000); *Consciousness, color, and content*. Cambridge (Ma.): MIT Press.
- (2002a); "Précis to *Consciousness, color, and content*", *A field guide to Philosophy of mind*. Symposium on Michael Tye's *Consciousness, color, and content*, en <http://host.uniroma3.it/progetti/kant/field/tyesymp.pdf>, pp. 1-2.
- (2002b); "To PANIC or not to PANIC. Reply to Byrne", *A field guide to Philosophy of mind*. Symposium on Michael Tye's *Consciousness, color, and content*, en <http://host.uniroma3.it/progetti/kant/field/tyesymp.pdf>, pp. 28-33.
- (2002c); "On the virtue of being poised. Reply to Seager", *A field guide to Philosophy of mind*. Symposium on Michael Tye's *Consciousness, color, and content*, en <http://host.uniroma3.it/progetti/kant/field/tyesymp.pdf>, pp.65-69.
- (2003a); "A theory of phenomenal concepts", en A. O'Hear (ed.) (2003); *Minds and persons*. Royal Institute of Philosophy, *Philosophy suppl.*, No. 53, pp. 91-106.
- (2003b); *Consciousness and persons*. Cambridge (Ma.): MIT Press.
- (2004); "The nature of nonconceptual content", en M. Reicher and J. Marek (eds.); *Experience and Analysis*. Proceedings 2004 Kirchberg Philosophy Conference, pp. 221-239.

- (2005); "Qualia", *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/qualia>.
- (2006a); "Another look at representationalism about pain", en Aydede (ed.) (2006), pp. 99-120.
- (2006b); "In defense of representationalism: reply to commentaries", en Aydede (ed.) (2006), pp. 163-176.
- (2006c); "Absent Qualia and the Mind-Body Problem", *Philosophical Review*, Vol. 115, No. 2, pp. 139-168.
- (2006d); "The puzzle of true blue", *Analysis*, Vol. 66, No. 3, pp. 173-178.
- (2006e); "The true about true blue", *Analysis*, Vol. 66, No. 4, pp. 340-344.
- (2006f); "Nonconceptual content, richness, and fineness of grain", en T. Gendler and J. Hawthorne (eds.); *Perceptual Experience*. Oxford: Oxford University Press.
- (2007a); "Representationalist theories of consciousness", en McLaughlin and Beckermann (eds.), cap. 14.
- (2007b); "Philosophical problems of consciousness", en M. Velmans and S. Schneider (eds.); *The Blackwell companion to consciousness*. Oxford: Blackwell, pp. 23-35.
- (2007c); "Intentionalism and the argument from no common content", en *Philosophical Perspectives*, No. 21. *Philosophy of Mind*, pp. 589-613.
- (2007d); "True blue redux", *Analysis*, Vol. 67, No. 1, pp. 92-93.
- (2007e); "New troubles for the qualia freak", en J. Cohen and B. McLaughlin (eds.); *Contemporary debates in the Philosophy of Mind*. Oxford: Blackwell, capítulo 17.
- (2008); "The experience of emotions: an intentionalist theory", *Revue Internationale de Philosophie*, Vol. 62, No. 243, pp. 25-50.
- (2009); *Consciousness revisited. Materialism without phenomenal concepts*. Cambridge (Ma.): MIT Press.
- (forthcoming a); "A new look at the speckled hen", *Analysis*.
- (forthcoming b); "The admissible contents of visual experience", *Philosophical Quarterly*.
- VARELA, F. (1990); *Conocer*. Barcelona: Gedisa.
- VELMANS, M. y SCHNEIDER, S. (eds.) (2007); *The Blackwell companion to consciousness*. Oxford: Blackwell.
- VILLANUEVA, E. (ed.) (1996); *Philosophical issues, 7: Perception*. Atascadero (Ca.): Ridgeview.
- WARFIELD, T. A. (1999); "Searle's causal powers", *Analysis*, Vol. 59, No. 1, pp. 29-32.
- WATANABE, M. (1996); "Reward expectancy in primate prefrontal neurons", *Nature*, No. 382, pp. 629-632.
- WATSON (1987); "Responsibility and the limits of evil: variations on a Strawsonian theme", en F. Schoeman (ed.); *Responsibility, character, and the emotions: new essays in moral psychology*. Cambridge, Cambridge University Press.

WEISKRANTZ, L. (1986); *Blindsight: a case study and its implications*. Oxford: Oxford University Press.

WILLIAMS, B. (1993); "Postscript", en D. Statman; *Moral Luck*. Albany: State University of New York Press, 1993.

WILLIAMSON, T. (1998); "The broadness of the mental: some logical considerations", en Tomberlin (1998), pp. 389-410.

--- (2000); *Knowledge and its limits*. Oxford: Oxford University Press.

WITTGENSTEIN, L. (1953); *Philosophical Investigations* (trad. esp. *Investigaciones Filosóficas*. Madrid: Crítica, 1989).

WOLF, S. (1990); *Freedom Within Reason*. Oxford: Oxford University Press.

WRIGHT, E. (ed.) (2008); *The case for qualia*. Cambridge (Ma.): MIT Press.

WYNN, T. (2000); "Symmetry and the evolution of the modular linguistic mind", en Carruthers and Chamberlain (2000), pp. 113-139.

ZEKI, S. (1990); "Color vision and functional specialisation in the brain cortex", *Neuroscience*, No. 6.

--- (1992); "The visual image in mind and brain", *Scientific American*, Vol. 267, No. 3, pp. 68-77.

--- (1993); *A vision of the brain*. Oxford: Blackwell.