

Trabajo de fin de grado.

*Grado en filosofía.*

2020-2021.

# DUPLICAR LA CONCIENCIA

*Una lectura de Chalmers y Dennett*

Alumna. Yamel Nayade Medina Martín.

Tutor. Antonio Manuel Liz Gutiérrez.

## ÍNDICE.

1. INTRODUCCIÓN.....	3
2. ANTECEDENTES. Historia de la problemática alma-cuerpo y mente-cuerpo.....	5
2.1 El problema del alma. La prehistoria y la Edad Medieval.....	5
2.2 El problema mente-cuerpo. La modernidad.....	7
2.3 Nuevas teorías de la mente. Desde la Ilustración hasta finales del siglo XX. ...	9
2.4 El problema de la conciencia. 1950-1990.....	10
3. ESTADO ACTUAL. Las teorías de la conciencia de Dennett y Chalmers .....	13
3.1 Daniel Dennett. <i>La conciencia explicada</i> . .....	14
3.2 David Chalmers. <i>La mente consciente</i> . .....	21
3.3 ¿Qué es la conciencia? .....	27
4. DISCUSIÓN Y POSICIONAMIENTO. Problemas de una teoría de la conciencia.28	
4.1 Los individuos mecánicos. Robots, ciborgs e Inteligencias Artificiales.....	29
4.2 La posibilidad de crear una Inteligencia Artificial.....	30
4.3 ¿Qué es el “yo”? .....	33
4.4 La inmortalidad virtual.....	35
5. CONCLUSIÓN Y VÍAS ABIERTAS. Desde el presente hacia el futuro.....	38
6. BIBLIOGRAFÍA CITADA. ....	40

## INTRODUCCIÓN.

Determinar qué es la conciencia, de qué cosas podemos decir que son conscientes, o cuáles son los límites que esta podría tener, son algunos de los objetivos actuales más importantes de la filosofía de la mente. Existen diversas maneras de abordar esta cuestión como, por ejemplo, recurrir al análisis del aprendizaje, al del lenguaje, o al de la propia sociología. Todas ellas parecen igual de válidas, pues una buena teoría de la conciencia debería estar basada en los resultados que puedan concluir distintas disciplinas. Dividiré este trabajo en cuatro apartados:

En el primero de ellos, “Antecedentes”, haremos un breve recorrido histórico acerca de la problemática de la conciencia, señalando sus inicios en la prehistoria. Estudiaremos las concepciones de distintos autores y sus épocas correspondientes, así como los diversos nombres que ha recibido lo que comprendemos como *yo*.

En segundo lugar, proseguiré con la presentación del “Estado actual” de la cuestión. Tendré como objetivo principal responder a la pregunta de qué es la conciencia. Para ello me apoyaré en las teorías de Dennett y Chalmers presentadas en sus obras, tituladas *La conciencia explicada* y *La mente consciente* respectivamente. Utilizando las conclusiones de ambos, desarrollaré una síntesis de sus postulados que me ayudará llegar a la tesis central de este proyecto: ¿es posible duplicar la conciencia?

Ese será el tema central de “Discusión y posicionamiento”. No obstante, al intentar responder a esta cuestión, ello me suscitará nuevas problemáticas tales como qué

podemos denominar “yo”, la creación de Inteligencias Artificiales y ciborgs, o cuestiones transhumanistas tales como la posibilidad de alcanzar la inmortalidad virtual.

Por último, en “Conclusiones y vías abiertas” haré un breve recorrido acerca de las cuestiones que habremos tratado y hablaré acerca del estado actual de la tecnología, concluyendo con una posible predicción acerca de qué pasará en el futuro.

Lo que me ha despertado el interés a tratar el tema de qué es la conciencia son tanto las distintas problemáticas epistemológicas, transhumanistas, sociológicas y acerca de la intencionalidad –entre otras– que se me han presentado a lo largo de la carrera. Se trata, a su vez, de un tema de gran actualidad que no se remonta únicamente a estudios académicos, sino también es presentada en la cultura popular bajo el formato de series de televisión, películas, libros y cómics. Entre ellos cabe mencionar *Black Mirror*, *Ghost in the shell*, *Cyberpunk 2077* o *The Stanley Parable*.

Entenderé la conciencia en términos de autoconciencia. Podemos verla, en sí, como la capacidad que tiene el sujeto de reflexionar tanto acerca de sí mismo como del resto de fenómenos y sujetos que se presentan en su mundo externo.

Para analizar este problema me remontaré concretamente al estudio de la conciencia desde las interpretaciones de Daniel Dennett y David Chalmers. Lo que me ha llevado a escogerlos fue el hecho de que tratan teorías de la conciencia actuales –fundamentadas tanto científica como filosóficamente– que parten del postulado de que esta puede ser analizada. Además de ello, fundamentan cada una de sus afirmaciones en investigaciones científicas actuales. La necesidad de encontrar una teoría apoyada tanto en la filosofía como en la ciencia –concretamente en la neurobiología– proviene a que, con el uso conjunto de ambas, es posible tanto elaborar una tesis como basarla en datos empíricos adquiridos por el método científico.

En síntesis, procederé a comenzar con el recorrido histórico de la cuestión, continuaré con una propia definición de lo que es la conciencia, siempre reconociendo que es posible realizar otras interpretaciones y, tras hacerlo, me meteré directamente con la cuestión principal de si la conciencia puede ser, de hecho, *duplicada*.

## **ANTECEDENTES.**

### **Historia de la problemática alma-cuerpo y mente-cuerpo.**

Si nos ceñimos a establecer los antecedentes del estudio de la *conciencia*, nos remitiríamos a, aproximadamente, un siglo atrás. Pero no se trata de un asunto nuevo que acabe de surgir, sino de una problemática filosófica cuyo nombre ha ido cambiando a lo largo de la historia. En un principio fue denominado el dualismo alma-cuerpo, aunque después pasó a denominarse cuerpo-mente, y, finalmente mente-cerebro. Citando al autor David Pineda, podemos establecer que “La reflexión filosófica sobre la mente es tan antigua como la propia filosofía”<sup>1</sup>, pues, si comenzamos a estudiarlo partiendo de la noción de *alma*, podemos remontarnos a la propia prehistoria. En este punto haremos un breve recorrido histórico acerca de cómo ha variado este asunto.

#### 2.1 El problema del alma. La prehistoria y la Edad Medieval.

En las últimas décadas hemos redescubierto muchos puntos interesantes sobre la prehistoria. Hemos aprendido que realizaban distintos ritos, tanto funerarios como para la caza, que dan a entender que creían en la existencia de un más allá. Espíritu Avila señaló que en estos pueblos existía la idea de que todo a su alrededor estaba animado. Les rendían culto a los seres de la naturaleza; "La idea es que nada muere en forma absoluta, sino que pasa a otro estado."<sup>2</sup> Tanto sus pintadas en cuevas como los restos arqueológicos que hemos rescatado son las mayores pruebas de ello. No obstante, situándonos concretamente en el inicio de las civilizaciones con lenguaje escrito, nos

---

<sup>1</sup> Pineda (2012), 11.

<sup>2</sup> Espíritu (2018), 49.

encontramos que ya había distintas concepciones del término alma. Algunas de ellas, siguiendo a Ferrater Mora, son el verlo como un aliento, una realidad aérea que vaga a través de los vivos manifestándose en fuerzas y acciones, e incluso como una sombra que desciende al seno de la tierra<sup>3</sup>. Si nos encontramos interesados en encontrar, más concretamente, a teorías de individuos concretos, deberemos viajar hasta la Antigua Grecia. En ella tenemos dos pensadores que hablaron acerca de la dualidad alma-cuerpo; Platón y a Aristóteles.

Platón, en el *Fedón*, señalará que el alma es tanto distinguible como separable del cuerpo. Lo importante para cada individuo es su alma, siendo que el cuerpo es algo que podemos dejar atrás. Esto es lo que asemeja a los hombres con los dioses, y siendo que el alma tiene un origen divino, su deseo será aspirar reencontrarse con ellos. Dividirá el alma –y, por tanto, la conciencia– en tres partes distintas: la sensitiva, la irascible y la inteligible. Las explicará a través de la metáfora del carro alado, que consiste en lo siguiente: un caballo negro y no obediente, representante de la parte sensitiva de nuestra alma, y otro caballo blanco y dócil, representante de la parte irascible, mueven al propio sujeto a actuar. Este, representando la parte inteligible, deberá tratar de guiarse por los mejores impulsos, ignorando los peores, si lo que desea es dejar el mundo físico y volver al mundo de las ideas, donde se encontrará con la verdadera *realidad*. Solo saldrá de este ciclo de reencarnaciones dependiendo del papel que desempeñe en su vida; si no vive una vida filosófica, volverá a renacer. Por otro lado, tenemos Aristóteles. Su teoría, también dualista, consistirá en ver al cuerpo como simple materia reemplazable, y al alma como la sustancia que es, en sí, “el *quid* esencial del cuerpo”<sup>4</sup>. Esta es la clave de movilidad del cuerpo, aquello que hace que esté vivo. El alma no es solamente es principio de movimiento, sino en términos de Ferrater Mora, sino también la propia *causa* de que el cuerpo sea lo que es. Aristóteles distinguirá tres tipos de alma, pero lo hará entre el alma animal, el alma vegetal y el alma propiamente humana: la racional.

La teoría de Platón será recuperada por la Iglesia Católica en la Edad Media; la cual se instaurará en occidente como constitutiva de un pensamiento hegemónico. Se hará una lectura del alma platónica siguiendo de los propios escritos de La Biblia que impedían avanzar en cualquier otro tipo de conocimiento que pudiese atentar contra ella. Hasta el fin de este periodo se mantuvo este ideal, por lo que no hubo un gran progreso.

---

<sup>3</sup> Cfr. Ferrater (2008), 22.

<sup>4</sup> Ferrater (2008), 23.

## 2.2 El problema mente-cuerpo. La modernidad.

Con el fin de la Edad Media las ramas de conocimiento comenzaron a prosperar de nuevo, como en la antigüedad, con la diferencia de que se habían perdido muchos conocimientos antiguos. A lo largo de esta época fueron recuperando más referentes y escribiendo más teorías acerca de nuestro mundo. En la filosofía esto también sucede, ya que surgieron grandes pensadores como René Descartes e Immanuel Kant. Ellos se encuentran justo en mitad de la transición entre el problema alma-cuerpo hacia el mente-cuerpo.

El gran autor de esta nueva dualidad es René Descartes (1596-1650), y lo podemos considerar como el punto de partida histórico del problema que nos atañe en este escrito. Vivió en mitad de la revolución científica que se produjo entre los siglos XVI y XVII, y eso lo marcó para elaborar sus propias teorías. Autores como Llinàs Begon al leer la teoría de Descartes concluyen que es una mezcla tanto de teorías científicas como en teorías metafísicas; a pesar de que sus teorías se apoyan en pruebas científicas, no descarta establecer principios universales para estas.

Descartes afirma que podemos explicar el cuerpo sin estudiar la propia alma, la cual es muy compleja. Considerará al ser humano como un ser “no exclusivamente mecánico”<sup>5</sup>, sino como un alma consciente y racional que se sitúa de un cuerpo mecánico. Su frase, *Cogito, ergo sum*, hace referencia al papel tan fuerte que tiene la propia alma para él, siendo que el propio hecho de ser autoconsciente es lo que te da garantía de incluso tu propia existencia. A pesar de que recurre a la introspección como garantía de su tesis, se trata de una actividad mental que actúa a niveles tan simples que sería complicado de negar.

Según su teoría cuerpo y el alma como dos cosas separadas que funcionan en conjunción. El cuerpo –res extensa– tiene como únicas cualidades su propia extensión física en el espacio, mientras que el alma –res cogitans– estará caracterizada por su capacidad racional; su propio pensamiento. Los cuerpos son solo autómatas, actúan como la “sede del alma”<sup>6</sup>, aunque son desechables. Establece que en la glándula pineal

---

<sup>5</sup> Llinàs (2017), 439.

<sup>6</sup> Aguilar (2010), 763.

se encuentra el alma humana, y aunque posteriormente en la historia se demostrará que dicha zona tiene una función tan distinta como controlar el equilibrio, resulta muy relevante que este autor trate de situar un centro de las operaciones cerebrales. Establecerá que lo único de lo que podemos estar seguros es tanto de nuestra propia existencia –y de nuestro pensamiento– como de la existencia de un ser divino. A pesar de que no se convirtió en una visión global en la época que vivió, sí que dejará marcado el inicio de este debate de forma de que todos los autores tendrán en cuenta su punto de vista o refutaciones/argumentaciones de este.

Otro autor de la Modernidad muy importante es Immanuel Kant (1724-1804). Estableció el concepto de giro copernicano en el prefacio de la segunda edición de *Crítica de la razón pura*. En este, partiendo del hecho de que las teorías de Copérnico dieron todo un vuelco a la ciencia de su época, señalará cómo la filosofía debería hacer algo muy parecido. Hasta este momento, todas las teorías de la conciencia más importantes tenían en cuenta, más que nada, lo que sucedía en el mundo exterior – Kant afirma que lo que deberá hacerse es dar un vuelco y mirar directamente no a los objetos físicos, sino al propio espectador. Su teoría se basará en una diferenciación clave: el fenómeno y el noumeno. El fenómeno hace referencia a los objetos tal y como las personas los percibimos, mientras que lo segundo hace alusión al objeto en sí. Cuando miramos a un objeto no lo vemos en todas sus dimensiones, y ante situaciones diferentes hasta podemos captar, en ellos, otra serie de colores, olores, texturas, formas, tamaños, etc. Cabe la posibilidad de plantearse que cada sujeto ve el mundo de forma distinta, y como conocemos el mundo únicamente a través de los fenómenos, los seres humanos no seremos nunca capaces de conocer el noumeno. A pesar de ello, los seres humanos están muy predispuestos a actuar en base a buscar respuestas y entender el mundo que les rodea; pero no podrán nunca llegar a conocer el objeto en sí. Se dirá que únicamente podemos conocer el mundo a través de nuestros sentidos y de nuestras experiencias: estas últimas serán “el producto resultante de aplicar la actividad del intelecto al material que ofrecen los sentidos”<sup>7</sup>, siguiendo a Vázquez Lobeiras, y serán el único recurso que tendrán los individuos para tratar de entender lo que les rodea.

---

<sup>7</sup> Vázquez (2004), 5.



### 2.3 Nuevas teorías de la mente. Desde la Ilustración hasta finales del siglo XX.

A finales del siglo XVIII se produjo, tras varias revoluciones científicas, el siglo de las luces: La Ilustración. Distintas disciplinas de pensamiento se independizaron en esta época, separándose de la rama de la ciencia. Con la creación de estas, junto con el giro copernicano de Kant, se dio más que nunca mucha importancia al papel de los sujetos a la hora de conocer el mundo que le rodeaba.

No es hasta el siglo XX donde surge una teoría en concreto que nos resulta muy relevante para el estudio del cuerpo-mente: el conductismo. Consistirá en el estudio de las capacidades humanas desde la dualidad estímulo-respuesta. Isaac Palvov, John Watson y Burrhus F. Skinner son tres grandes autores constructivistas.

En primer lugar, Palvov es el pensador del condicionamiento clásico. Uno de los experimentos que realizó dio grandes avances en el estudio de la conducta de los seres vivos: el experimento de la campana y el perro. En él trataba de condicionar a un perro utilizando comida y una campana. Cada vez que él tocaba la campana, le daba comida al perro. Con el tiempo este acababa estableciendo una relación entre el sonido y la comida, así que incluso era capaz de salivar tan solo con escuchar este tintineo.

En segundo lugar, tenemos a John Watson, quien realizará un experimento muy similar mezclando un peluche de una rata y un bebé humano. Al niño se le dará el juguete de la rata, mostrando que no presenta ningún tipo de rechazo en primer lugar a este ser. Watson tratará de condicionarlo haciendo un ruido desagradable cada vez que el niño tocaba el juguete. Con el tiempo el niño empezó a evitar jugar con la rata para que no se repitiese dicho sonido desagradable, e incluso llegó a tenerle miedo. Skinner establecerá una teoría de condicionamiento operante, que estudia el papel de los refuerzos positivos y negativos. Estudiará el comportamiento de una rata a la que encierra en una caja con un botón. Cada vez que la rata lo acciona, caerá comida, así que una vez que su conducta está condicionada, tocará todo el rato dicho botón. Al tiempo Skinner dejó de proporcionarle comida cuando accionase el botón, comprobando que al darse cuenta de que esto cambiaba, dejó de tocarlo.

Estas teorías demuestran que no reaccionamos de forma innata ante ciertos estímulos, sino casi todo es aprendido por medio de la cultura.

## 2.4 El problema de la conciencia. 1950-1990.

A partir de 1950 surgirá otra teoría de la conciencia muy relevante: el computacionalismo. Alan Turing es el autor de esta corriente que estudiaremos, pues es uno de los autores más relevantes para el tema del estudio de la conciencia. No cree que la pregunta correcta que podemos realizarnos sea si una máquina puede pensar, sino, más bien: “¿bajo qué circunstancias podría confundirse una máquina con una persona inteligente?”<sup>8</sup>. El computacionalismo, según Laureano, se basa en “la tesis de que la inteligencia humana es equivalente a un algoritmo”<sup>9</sup>. Cuando comenzaron a estudiar la capacidad de las máquinas de imitar el pensamiento humano, se empezó a hablar de las personas en términos computacionales. Turing, siguiendo estas teorías, establecerá un test que será capaz de determinar si una máquina es capaz de emular el pensamiento humano. Señalará, de esta manera que, si un sujeto es capaz de ver una conversación producida por una máquina y considerar que se trata de un ser humano real, entonces se puede considerar que es inteligente.

Una crítica a esta idea, según Pineda, se basa en la idea de que es muy fácil imitar un comportamiento de forma que parezca que es tuyo propio. Le hablará al lector pidiéndole que imagine que se encuentra jugando a dos partidas de ajedrez a la vez, sin saber jugarlo. En una partida podría jugar con las blancas, y en otro con las negras. Si quiere ganar contra un jugador experto, solo tendría que imitar los movimientos del otro en su tablero.

Laureano recoge otro tipo de crítica muy frecuente contra el computacionalismo. Asegura que la inteligencia humana parece capaz de funcionar de una manera tan compleja que un ordenador no sería capaz de realizar todo lo que nosotros somos capaces, por lo que no deberíamos hablar en estos términos.

Esta cuestión de si las máquinas son inteligentes deriva, décadas después, en: “¿podría pensar una máquina que manipulase símbolos obedeciendo a reglas sensibles a la

---

<sup>8</sup> Pineda (2012), 294.

<sup>9</sup> Laureano (2016), 3.

escritura?”<sup>10</sup>. Searle y los Churchland son dos de los autores que trataron de responder a esto.

Searle en “¿Es la mente un programa informático?” asegurará que si afirmamos que una máquina se trata de simplemente “un sistema material capaz de desempeñar ciertas funciones [...] [entonces] los humanos somos máquinas de una clase biológica especial”<sup>11</sup>. Se trata del autor creador de una muy conocida crítica a la Inteligencia Artificial (IA) fuerte, resumible en la idea de que “un ordenador digital programado adecuadamente y con la información de entrada y de salida apropiada tendría una mente exactamente en el mismo sentido que los seres humanos tienen una mente”<sup>12</sup>. A su vez, rechaza que el pensamiento se reduzca a la manipulación formal de símbolos. Para argumentar esto último recrea el experimento de la habitación China. En este se introduce a un sujeto, que no sabe chino, en una habitación con un mecanismo que es capaz de traducir –utilizando signos– su idioma a este. La particularidad de esta teoría se basa en que si un habitante de esta lengua leyese las frases que se generan en la habitación china, este superaría el *test* de Turing. Este argumento “refuta toda pretensión de IA fuerte para las nuevas tecnologías en paralelo, inspiradas y basadas en las redes neuronales”<sup>13</sup>. Se trata de un análisis de la conciencia en los términos más puramente lingüísticos en los que un sujeto que no tiene ningún tipo de contenido semántico de ese idioma en su mente es capaz de comunicarse con otro que cree que, de hecho, habla chino.

Searle afirma que hay que distinguir los conceptos de simulación y de duplicación: los cerebros tienen sus capacidades y hacen determinadas tareas, mientras los ordenadores *simulan* a los cerebros, pero no los duplican. Su conclusión es que los humanos son capaces de hacer cosas que una Inteligencia Artificial jamás podrá hacer. En clave cómica, enumera algunas de ellas: “enamorarse, tener sentido del humor, sentir la angustia de la sociedad postindustrial bajo el capitalismo tardío”<sup>14</sup>.

Mientras que Searle mantiene una posición realista respecto a la existencia de la conciencia y, a su vez, de la intencionalidad –siendo estos productos causales de nuestro

---

<sup>10</sup> Churchland (1990), 18.

<sup>11</sup> Searle (1990), 10.

<sup>12</sup> Dennett (1991), 538. Se encuentra citando a Searle; “Turing the Chinese Room” (1988).

<sup>13</sup> Searle (1990), 12.

<sup>14</sup> *Ibidem*, 18.

cerebro–, podemos señalar un extremo opuesto a esta visión: la de los autores Patricia y Paul Churchland. Defienden, en sus teorías, un eliminativismo neurofisiológico radical.

Establecen que la neurobiología es la forma *correcta* de estudiar la mente, y aceptan que los estados mentales con significados son constructos teóricos. No es que sostengan que los conceptos mentales sean reductibles a conceptos físicos, sino que directamente son falsos y deben ser eliminados de nuestras teorías. Se entenderá como teoría folk de la conciencia a la idea de que comprendemos al resto de personas gracias a las “relaciones legales que se dan entre circunstancias externas, estados internos y conductas.”<sup>15</sup>.

Con esto concluimos el recorrido histórico, colocándonos, concretamente a finales del siglo XX, en la década de 1990. En el siguiente punto explicaré las teorías de la conciencia que trataremos desde la pregunta de: ¿qué es la conciencia?

---

<sup>15</sup> Gomila (1990), 242.

## **ESTADO ACTUAL.**

### **Las teorías de la conciencia de Dennett y Chalmers.**

En el punto anterior vimos, a grandes rasgos, las teorías de la conciencia de Searle y de Patricia y Paul Churchland. Sus posturas son, en sí, representativas de dos polos opuestos en el espectro de teorías de la conciencia actuales, pero ambos pueden ser considerados muy *extremistas*. Con el propósito de ir más hacia *el centro* de la cuestión, procederé a centrarme, en sí, en dos teorías actuales que poseen tanto puntos en común como diferencias:

1. En primer lugar, la de Daniel Dennett. Se trata de un autor nacido en Estados Unidos en el año 1942, es el autor de *La Conciencia Explicada: una teoría interdisciplinar* (1991).
2. Y, a su vez, en la de David Chalmers. Nacido en 1966, en Australia, es el autor de *La mente consciente: en busca de una teoría fundamental* (1996).

Ambos son teóricos de la conciencia y realizan en sus respectivas obras un intento de establecer una última teoría acerca de esta. Mi propósito en este punto será tanto explicar sus teorías como ofrecer sus argumentos con el fin de responder a la pregunta de «¿qué es la conciencia?». Tanto Chalmers como Dennett al inicio de sus obras establecerán una serie de “normas” que consideran que hay que seguir para establecer una teoría de la conciencia. Estas son:

Por parte de Dennett:

1. “Nada de tejidos milagrosos”<sup>16</sup>; tratará de explicar cada punto de su teoría sin saltarse puntos ni salir con el comodín de «esto no podrá ser explicado nunca».
2. “Nada de anestésicos fingidos”<sup>17</sup>, si afirma que algo se trata de una ilusión, trabajará en argumentar que estas se den con pruebas físicas.
3. “Nada de regatear con los detalles empíricos”<sup>18</sup>, solo seguirá los preceptos de la ciencia tal y como se encuentran en su época actual.

Y, por otro lado, las de Chalmers:

1. Tomar en serio la conciencia; considera que no tiene sentido establecer una teoría de la conciencia si creemos que esta no existe o se trata únicamente de una *ilusión*.
2. Tomar en serio a la conciencia sin cuestionar teorías científicas actuales. Sus teorías deben seguir siendo “compatibles con la ciencia contemporánea”<sup>19</sup>.
3. Aceptar que la conciencia se trata de un “fenómeno natural sometido al dominio de las leyes naturales”<sup>20</sup>. Aunque esto no implica que *deba* ser explicada igual que cualquier otro fenómeno natural.

Habiendo establecido sus objetivos, procederé, directamente, a explicar de forma individual cada una de sus teorías.

### 3.1 Daniel Dennett. *La conciencia explicada*.

La conciencia, para Daniel Dennett, es el último de los grandes misterios. Señalará: “Cuando comprendamos la conciencia –cuando ya no haya misterio– ésta será diferente, pero seguirá habiendo belleza y más motivos que nunca para el asombro”<sup>21</sup>. Comenzar a abordar una teoría de la conciencia es muy complicado, puesto que no existen unos pilares establecidos como *la verdad* por los cuales partir. Parece que lo único que podemos hacer es andar a oscuras y escoger un punto cualquiera por el cual comenzar.

---

<sup>16</sup> Dennett (1991), 54.

<sup>17</sup> Ibidem, 55.

<sup>18</sup> Ibidem, 55.

<sup>19</sup> Chalmers (1996), 18.

<sup>20</sup> Ibidem.

<sup>21</sup> Dennett (1991), 36.

Dennett lo hará desde el análisis del cerebro, señalando que cuando contemplamos el cerebro lo que observamos es masa gris complicada, pero posible —o, al menos, aparentemente— de estudiar. A pesar de que no tengamos actualmente clara la función de cada parte del cerebro, no parece imposible llegar a establecerlo. Cuando nos imaginamos algo de forma consciente, o incluso cuando pensamos en la posibilidad de imaginar algo, en nuestro cerebro se da una serie de relaciones neuronales. Pero, aunque lo abramos y lo estudiemos a fondo, no vemos de ninguna forma físicamente qué clase de información se está pasando de un lado a otro, ni por qué es así. Dennett descarta cualquier teoría dualista con la cual tratemos de entender esto: si tratamos de encontrar en qué zona del cerebro se produce este tipo de situaciones, deberemos dejar de lado el ver al cerebro como algo separado del cuerpo físico. Por esto, partirá desde una definición importante:

*El fenómeno y el noúmeno. El problema de la memoria.*

La distinción entre fenómeno y noúmeno acontece al estudio de la fenomenología: el estudio descriptivo de la materia. En sí, el fenómeno es una mezcla de las experiencias del mundo obtenidas tanto de nuestro exterior, como de interior y de nuestras emociones. Las experiencias externas se obtienen mediante nuestros sentidos, pero esto no significa que tengamos por separado una experiencia del fenómeno distinta por cada uno de los cinco que tenemos —vista, oído, olfato, gusto y tacto—, sino que se dan a la vez. Las otras categorías, como nuestros pensamientos y emociones, se adquieren por introspección. Dennett estudiará las externas a través de los qualia: las experiencias cualitativas que tenemos de objetos, capaces de proporcionarnos sensaciones aparentemente imposibles de explicar tales como los colores o el dolor.

El estudio de los qualia suele acompañarse del experimento del Cuarto de Mary de Frank Jackson, que presenta un argumento en contra del fisicalismo. Expone el caso hipotético de una mujer, Mary, que viviría toda su vida en un laboratorio blanco y negro. En este entorno aprendería toda “la información física que se puede obtener”<sup>22</sup> acerca de los colores, convirtiéndose en una experta en este campo. Un día, cuando concluirían sus estudios, sería dejada en libertad. En este momento dejaría el laboratorio y observaría, por primera vez, cómo se ven los colores con sus propios ojos. La cuestión

---

<sup>22</sup> Dennett (1991), 410.

principal es. cuando Mary salga al mundo exterior, ¿aprenderá algo nuevo? Jackson determinará que sí, y que incluso tendrá que aprender los colores debido a que no podrá distinguir el rojo del azul. No obstante, según Dennett, si Mary había aprendido *todo* lo que se podía saber acerca de los colores, simplemente habría encontrado *otra forma* de experimentarlos. Ella ya sabría “qué impresión física exacta produciría un objeto amarillo o un objeto azul”<sup>23</sup>, puesto que, si ha estudiado *todo*, habrá comprendido que la sensación de color se encuentra no en los propios objetos, sino en su propio cerebro. Según Dennett los qualias no se encuentran en el mundo físico, sino del cerebro del propio sujeto. Con esto quiere decir que no es que los plátanos *sean* de color amarillo, es que nosotros mismos, los sujetos, los interpretamos de esta –y no de otra– manera. La sensación que tenemos de que se nos presenta un mundo con objetos de determinados colores se trata, únicamente, de una *ilusión*.

Esto no implica que por un lado estén las experiencias físicas y por otro las psicológicas, sino que con cada experiencia física que tengamos, esta estará acompañada de una emocional o introspectiva. De la misma manera que el conocimiento de las últimas está basado, a su vez, en las físicas, ya que se quedan almacenados en nuestra memoria: una fuente de acceso a experiencias pasadas.

El problema de la memoria se encuentra en que no hay garantía de que sean algo objetivo, pues cuando recordamos algo, no lo recordamos *tal y como se dio*, sino exactamente como *la última vez que lo recordamos*. Estudiará este fenómeno con «la mujer con sombrero»: tras una fiesta el domingo, el lunes hable con un amigo acerca de aquella persona; ella nunca existió, pero su amigo puede incluso llegar a creer que la llegó a ver. A este fenómeno lo analiza utilizando los conceptos de “revisión orwelliana” y “revisión estaliniana”. La primera de ellas, denominada así por *1984* de Orwell, hace referencia a “las contaminaciones postexperienciales o revisiones de la memoria”<sup>24</sup> que se producen cuando vuelves a recordar algo que viviste. Incluso puedes llegar a sostener que, como crees que sucedió así, entonces así se llevó a cabo, llegando a creer que experimentaste cosas que nunca sucedieron. Por otro lado, la estaliniana se basa en la creación y difusión de “falsos testimonios y confesiones, y completándolo todo con

---

<sup>23</sup> Dennett (1991), 411.

<sup>24</sup> *Ibidem*, 146.



falsas pruebas”<sup>25</sup>, lo que hace creer a la persona que realmente fue así porque hay “pruebas” que lo demuestran.

Tenemos la idea de que, en nuestra memoria, se encuentran ordenadas nuestras experiencias de forma cronológica. Esto no es así, puesto que podemos recordar algo que sucedió antes o después independientemente del orden en el que se dio en la vida real. Incluso podemos recordar el proceso por el que una pelota se deslizó por una rampa, comenzando por el punto donde esta se detuvo. Nuestros cerebros, según Dennett, son “máquinas de anticipación”<sup>26</sup> capaces de predecir el futuro. Los mejores son aquellos que son capaces de anticipar lo máximo posible.

No obstante, así es como funciona la memoria, y es innegable que es el gran pilar que poseemos para formular nuestro conocimiento. A pesar de que sea una reproducción de algo que se da, no por esto tiene menos valor. Cuando escuchamos la reproducción de una pieza musical, o cuando leemos una novela, somos capaces de emocionarnos con la misma intensidad que si lo viviésemos de primera mano. Entonces, ¿por qué le damos tanta importancia a las sensaciones de primera mano? Porque nos parece que son más vividas, más reales, y porque no podemos estar seguros de la veracidad de nuestros recuerdos. Pero ni siquiera estos parecen ser completamente fiables, pues nuestros sentidos nos engañan y los *qualia* –antes mencionados– no parecen ser compartidos por todas las personas. Nos es imposible describirle a un ciego de nacimiento lo que es el color rojo. Por más que lo intentemos, parece que nunca lo llegará a entender con la misma intensidad que nosotros. E, incluso, ni siquiera podemos estar seguros de que nuestra experiencia de rojo sea objetiva y compartida por todos, puesto que sabemos que no es así cuando comparamos nuestro *qualia* del rojo con el de un animal o un daltónico.

Pero, a pesar de que comprendamos el mundo de una forma sesgada, así es como la conocemos. pesar de todo, y aunque sea de forma sesgada esta es nuestra forma de conocer el mundo. No podemos desprendernos del papel que tienen en nosotros, pero podemos actuar teniendo en cuenta que sabemos que son engaños e *ilusiones*.

Nuestras memorias no se dan de forma lineal, y no conocemos el mundo físico.

---

<sup>25</sup> Dennett (1990), 148.

<sup>26</sup> Ibidem (1990), 220.

### *La heterofenomenología.*

*Cogito, ergo sum.* No parece que esta afirmación pueda generarnos ninguna duda. Si soy capaz de pensar, y sé que lo estoy haciendo, ¿qué duda puedo tener acerca de mi pensamiento? Incluso, a pesar de que no tengamos pruebas de ello, al relacionarnos con otros seres humanos damos por hecho que ellos también poseen pensamientos. E incluso conciencia. Dennett, por ello, mencionará un concepto filosófico surgido en esta época: la idea de los *zombis fenoménicos*. Se trata de individuos que, pese a tener una “conducta perfectamente natural”<sup>27</sup>, carecen de conciencia. Tacha a estos individuos de “una especie de autómeta”<sup>28</sup> ya que se mueven por el mundo, pero no poseen experiencias conscientes. Solo podemos estar seguros de que nosotros mismos tenemos conciencia gracias a un ejercicio de *introspección* en el que podemos llegar a la misma conclusión que Descartes, pero ¿qué fundamento hay para creer que el resto de los seres humanos sean así? ¿Cómo podemos estudiar su conciencia?

El autor presenta el término de la heterofenomenología: el acto de usar el método científico con el fin de conocer el estado mental interno de los sujetos, estudiándolo gracias a sus testimonios orales y las evidencias científicas de estos. Un investigador heterofenomenológico tendrá que no solamente escuchar a su estudiado, sino también que tomarse en serio al sujeto y reconocer que tiene conciencia.

### *La ilusión del Teatro Cartesiano.*

La teoría de Dennett está basada en la crítica de lo que denomina el Teatro Cartesiano; de René Descartes. Esta se basa en la idea de que cuando cada parte del cerebro opera en aquello de lo que está encargado, esta información será trasladada a un punto específico del cerebro. Allí se encuentra el Significador Central –el alma– del sujeto, el cual ejerce como un juez que, analizando toda la información que le llega, decide cómo el cuerpo va a actuar. El problema de esta teoría es que Descartes sitúa a dicho juez en la glándula pineal, y en la actualidad está demostrado que esta no es su función. Por ello,

---

<sup>27</sup> Dennett (1991), 93.

<sup>28</sup> Ibidem, 93.

y en contra de la idea de que exista una zona del cerebro “principal” donde todo esto se dé, Dennett elabora su teoría de las Versiones Múltiples (*the multiple drafts model*). Esta dice que no hay un centro en el cerebro, cada parte de este actúa de forma individual y en paralelo con el resto, trabajando en conjunto en sus diferentes tareas, interactuando entre sí y moviendo al cuerpo para que reaccione con eficacia ante cada estímulo externo percibido. Es decir, nuestros actos no se dan gracias a un único Significador Central, situado en nuestro cerebro, que dicte cómo vamos a reaccionar, sino por la función independiente de cada parte del cerebro. El Teatro Cartesiano se basa la ilusión que tenemos de que nuestra memoria, y nuestra conciencia, son lineales, inalterables y bien fundamentadas. Bebe mucho de “la manera en que a usted, real y objetivamente, le parece que son las cosas, incluso cuando a usted no le parece que le parezca que son así”<sup>29</sup>, como dirá Dennett citando a Smullyan.

Pero no es así; nuestro cerebro es plástico. Es adaptable y puede hacer varias tareas de forma simultánea sin que estas estén necesariamente relacionadas. No hay un centro del cerebro, y es incluso más complejo –aparentemente– que el del resto de seres vivos. No es solo que seamos capaces de señalar cada parte de nuestro cuerpo o saber cuánta hambre tenemos, sino que, también, podemos hacer operaciones matemáticas, nombrar y reconocer países o continentes, o jugar al ajedrez. A esto lo denomina *memes*: bloques de información y patrones de pensamiento socialmente relevantes<sup>30</sup>. Son traspasados, como si se trata de genes, por el proceso de socialización por el que pasan todos los individuos en sus primeros años de vida. Se nos presentan como necesarios para vivir en comunidades. Para Dennett los tres elementos que constituyen a la persona son su propia genética, la plasticidad de su cerebro y los *memes*. Su teoría de la conciencia se encuentra basada en el papel que ejercen, en el sujeto, cada uno estos tres elementos.

### *La teoría de la conciencia de Dennett.*

Pasaremos al último punto, la presentación de la teoría de la conciencia de Dennett. Siguiendo a von Neumann, afirma que esta, al igual que la memoria, no se da de forma lineal ni continua. Cuando dormimos, cuando nos despistamos y en momentos donde no nos encontramos en óptimas condiciones físicas, la autoconciencia desaparece. Cita a

---

<sup>29</sup> Dennett (1991), 165. Se encuentra citando a Smullyan; “An epistemological nightmare” (1981).

<sup>30</sup> Se trata de un concepto surgido de la sociobiología en la segunda mitad del siglo XX.

Minsky para explicar esto, diciendo: “nuestra continuidad proviene, no de una genuina perceptividad, sino de nuestra maravillosa insensibilidad ante la mayor parte de los cambios”<sup>31</sup>. Parece que la conciencia se trata de una especie de lectura del mundo que puede apagarse y encenderse a veces, y que, aún con ello, no somos conscientes de que esto sea así.

Dennett distingue lo que crees que sientes de lo que *realmente* sientes. Es distinto creer que algo es de una manera, a *saber*, conscientemente, que estás en una ilusión en la que algo *te parece* que es así, incluso si sabes que no lo es. Este es, para Dennett, justamente lo que es la conciencia. Esta “es dispersa y está llena de vacíos, y no contiene ni la mitad de lo que las personas creen que contiene.”<sup>32</sup>; creemos que es algo muy importante, que es *lo que somos*, por encima de nuestro cuerpo –que es reemplazable–. Pero simplemente se trata una ilusión en la que se fundamenta el Teatro Cartesiano. Las personas somos, en sí, el conjunto de los distintos procesos que se dan en nuestro cuerpo, sin distinguir la mente del corazón, por ejemplo. Todo actúa de manera simultánea y distinta. La conciencia es simplemente una ficción creada por nuestros propios cuerpos y que nos es útil para tratar de vivir en el mundo. Otros animales tienen –como mecanismo de defensa aportado por la evolución– garras, colmillos y venenos; pero nosotros lo que tenemos es nuestra sensación de conciencia. Pero, aún con todo, nuestros recuerdos no son objetivos, los qualia parecen depender de cada sujeto, y nuestra conciencia no es lineal.

Según Dennett por todo esto no podemos hablar de que haya nadie realmente *consciente*. La conciencia se trata tan solo de una ilusión de nuestro cerebro, por lo que todos somos *zombis fenoménicos*. Es tan solo el “resultado de una serie de ilusiones cognitivas”<sup>33</sup>, pero no un hecho físico en sí. Las personas no somos siempre conscientes de lo que nos rodea, e incluso parece que mientras nuestro cerebro puede captar todo lo que hay a nuestro alrededor, a “nosotros” –como la ilusión consciente del cerebro– solo nos llega una pequeña parte de la información que percibe nuestro cuerpo, y únicamente cuando prestamos atención a ello.

Por ejemplo, cuando una persona se encuentra conduciendo se cruza con numerosos coches e individuos, e incluso algunos animales, pero segundos después de haberlos

---

<sup>31</sup> Dennett (1991), 438. Se encuentra citando a Minsky; *The Society of Mind* (1985).

<sup>32</sup> Ibidem, 451.

<sup>33</sup> Ibidem, 532.

captado se olvida por completo de ello, a no ser que alguno le llame la atención. De la misma forma presenta, en este libro, un juego que hacía cuando era pequeño con sus amigos; en él escondían un dedal en cualquier parte de la habitación, pero tenía que estar a la vista, en un sitio donde cualquiera lo pudiese ver. El juego consistiría en que todos los niños entrarían al lugar y se pondrían a mirar toda la sala de arriba hacia abajo. Cuando lograbas encontrar el dedal, te sentabas en una esquina. Perdía el último niño que fuese capaz de encontrarlo. Dennett presenta este hecho como algo curioso, pues si hemos mirado la habitación en algún momento habremos visto el dedal, pues nuestros ojos han pasado por todos lados, pero podrían tardar largos minutos en encontrarlo a pesar de que estuviese a la vista. No obstante, parece que no fuimos *conscientes* de que este estaba ahí desde el principio. Y, la verdad es que, en sí, lo fuimos, pero *el constructo mental* que entendemos como lo que somos, no. Tu cuerpo ya había visto el dedal a pesar de que tú, no fueses consciente de ello. Según la teoría de las Versiones Múltiples, la idea de que nosotros somos como un Espectador Central de nuestra vida, racionalizadores de nuestras experiencias internas y externas, se trata tan solo de una fantasía; de una *ilusión* creada por nuestro cerebro.

### 3.2 David Chalmers. *La mente consciente*.

Chalmers –en la línea de Dennett– establece que la conciencia no es tan solo un gran misterio, sino “tal vez, el mayor obstáculo pendiente en nuestra búsqueda de una comprensión científica del universo”<sup>34</sup>. Es, a su vez, tanto lo más familiar como lo más misterioso. A lo largo de la historia, como hemos visto en *Antecedentes*, se ha tratado esta problemática como cualquier otra, pero no se ha llegado a ningún consenso. Puede darse incluso la ilusión de que no se ha avanzado en su investigación, pero cada vez parece que nos acercamos más a la respuesta. ¿Qué es la conciencia? Parece algo que *surge* del cerebro, pero no sabemos si de verdad es así, ni por qué lo sería o cómo se daría. Tratando de seguir sus tres normas de tomar en serio el estudio de la conciencia, seguir la ciencia actual y tomar su existencia como un fenómeno natural, llega a la conclusión de que la conciencia podría ser, quizá, “un problema científico que requiere

---

<sup>34</sup> Chalmers (1996), 15.

métodos filosóficos de comprensión”<sup>35</sup>. Este autor partirá su estudio desde el análisis de lo fenoménico y lo psicológico.

### *Lo fenoménico y lo psicológico.*

Chalmers contrapone estos dos conceptos. Ambos hacen referencia tanto a enfoques sobre cómo estudiar la teoría de la conciencia como a enfoques sobre la manera de tratar los términos mentales. Con respecto a la teoría de la conciencia, estudiarla en términos fenoménicos será centrándose en la propia experiencia autoconsciente, mientras los psicológicos en la idea de que nuestro cerebro se caracteriza por lo que es capaz de hacer. En cuanto a distinguirlos en términos mentales, esto hace referencia a la dualidad de conocer el mundo a través de nuestras experiencias fenoménicas físicas y de nuestros estados psicológicos; conocemos a través de una síntesis de ambas.

### *La conciencia y la percatación*

Podemos entender la conciencia de muchas maneras; como vigilia, introspección, por la capacidad de expresarnos acerca de contenidos de nuestro cerebro, por pensar sobre nosotros mismos, podemos prestar atención a las cosas, podemos hacer cosas de forma deliberada o porque podemos tener conocimiento acerca de algún hecho concreto<sup>36</sup>, pero cualquiera de estas visiones –según Chalmers– se quedan cortas ya que solo definen la conciencia en términos de su función.

No debemos confundir, asegura, entre conciencia y percatación. A pesar de que en nuestro lenguaje cotidiano no las dividimos, la percatación se trata de una acción que –aparentemente– todo ser vivo es capaz de cumplir: el comprender el mundo a través de tus propios sentidos, entendiendo lo que te rodea y actuando conforme a ello. Newell definirá percatación “la capacidad para hacer que su conducta dependa de algún conocimiento”. Por otro lado, la conciencia, que debe “estar siempre acompañada de percatación”<sup>37</sup>, hace referencia a lo que podamos reflexionar, verbal o internamente, sobre lo que percatamos. El autor Block también distingue entre *conciencia fenoménica*

---

<sup>35</sup> Chalmers (1996), 19.

<sup>36</sup> Cfr. Chalmers (1996), 53.

<sup>37</sup> Chalmers (1996), 55.

y *conciencia de acceso*, las cuales se refieren a los conceptos de conciencia y percatación respectivamente.

### *Sobrevenida*

Para comenzar a hablar más profundamente en su teoría, deberemos mover nuestra vista a un nuevo término: el término de sobrevenida<sup>38</sup>. Se trata de un concepto surgido a lo largo del siglo veinte en el cual se establece la ordenación de relaciones que se dan entre un grupo de propiedades (o hechos), A, y otro grupo, B, del mismo tipo. Su punto de partida es la teoría de que si nos encontramos en el mundo con cierta *cosa* (fenómeno o hecho) que posee determinadas características, “cualquier otra cosa exactamente como ella debe, bajo cualquier circunstancia, tenerlo en exactamente el mismo grado”<sup>39</sup> – citando a Moore–. En su artículo *Sobrevenida: Un Caso de Ingeniería Conceptual*, Manuel Liz señalará la anterior cita, proseguida por el ejemplo que da Hare utilizando al santo San Francisco. Él fue, en vida, muy buena persona. Así que por el uso del término sobrevenida es que podemos decir que un hombre que haya vivido de la misma forma que San Francisco en cada aspecto, se comportará de la misma manera que este. Chalmers utilizará este concepto para establecer que los hechos biológicos y psicológicos sobrevienen a los hechos físicos, que considera los más básicos de nuestro universo. Muy en la línea de Davidson, a quien secunda. Los hechos psicológicos sobrevendrán, concretamente, si se da una situación en la que ambas cosas sean biológicamente iguales: por ejemplo, dos mesas que son físicamente idénticas.

Siguiendo con el concepto de sobrevenida, afirmará que dos mundos posibles completamente idénticos en sus propiedades físicas deben ser también idénticos respecto a cualquier otra propiedad.

Con todo esto, debemos distinguir, para esto, la diferencia entre lo *naturalmente* posible y lo *lógicamente* posible. Cabe la posibilidad de imaginar un mundo donde haya teléfonos voladores, aunque esto no se de en la vida real, pero no podemos siquiera imaginar la posibilidad de que Dios crease una zorra macho; pues es una contradicción. Lo naturalmente posible hace referencia a los hechos que se dan en el mundo, mientras lo lógicamente posible a aquello que, aunque no se de en el mundo real, podría darse sin

---

<sup>38</sup> También traducido como *superveniencia*.

<sup>39</sup> Liz (2008), 246. Citando a Moore en “The Conception of Intrinsic Value”.

alterar para nada el mundo en el que vivimos. Las propiedades físicas que se comparten en todos los mundos posibles son propiedades positivas; pero cuando nos referimos a aquello que puede darse en uno y no en otro, esto será una propiedad negativa. Las leyes de la física son un ejemplo de propiedad positiva, pero el tener un hijo es negativa en tanto en cuanto que no es necesario que tu gemelo en un mundo posible deba tenerlo. Por todo ello Chalmers establecerá que la sobrevenida lógica global se basará en el hecho de que: “para cualquier mundo lógicamente posible  $W$  que es A-indiscernible de nuestro mundo–, los hechos B verdaderos en nuestro mundo son también verdaderos en  $W$ .”<sup>40</sup>.

### *Crítica al materialismo*

Tras hablar sobre cómo conocemos el mundo y establecer que existen propiedades físicas positivas que se dan en todos los mundos posibles, pasaremos a hablar ahora de los propios fenómenos concretos. Se ha dado la idea, a lo largo de la historia, de que todo hecho parece explicable usando la ciencia. Se ha optado incluso por las explicaciones reductivas, que son aquellas donde se divide y separa cada punto de cada explicación hasta que se llega a una en la que se ha reducido el hecho a términos más simples. El problema de este tipo de explicaciones es que ayudan a entender el problema conceptualmente, pero se quedan cortas. Según Chalmers, una explicación de este tipo “elimina el misterio”<sup>41</sup>, pero no proporciona una explicación profunda del fenómeno, sino crea una abstracción que ayuda a comprenderlo mejor.

Las explicaciones reductivas del materialismo se basan en el supuesto de que todo es sobreviniente a lo físico. Aunque en un principio parezca que esto es así, Chalmers señalará distintos argumentos con los que podemos que no. He aquí dos experimentos mentales que señala:

El primero hace referencia a los zombis filosóficos (1). Esto ya lo hemos visto con anterioridad; individuos capaces de captar el mundo en que viven, pudiendo *percibir* todo aquello que está a su alrededor, pero que carecen de conciencia. Chalmers los llamará zombis fenoménicos para distinguirlo de los psicológicos, que son aquellos que carecen de mente. Nos resulta complicado imaginar este tipo de individuos puesto que

---

<sup>40</sup> Chalmers (1996), 69.

<sup>41</sup> Ibidem, 79.



se da la idea de que, si no tienen mente, no se comportarán como otras personas, pero no por carecer de experiencia consciente se deben comportar distinto al resto. Según la teoría de Chalmers, estos seres son lógicamente posibles, y cabe la posibilidad de que existan, puesto que no hemos analizado la mente de otras personas. Para hablar en profundidad de esto, nos presentará el experimento mental de Block de crear una simulación de la conciencia con la población de China. Cada individuo se encargaría de hacer la función de una neurona del cerebro, y se crearía, de esta manera, un sistema físicamente distinto pero que fuese capaz de realizar estas funciones. La afirmación de Chalmers será que mi gemelo zombi, tenga la estructura cerebral que sea, puede seguir pasando el *test* de Turing –aplicado a zombis– de forma que a mí me parezca un ser humano tan racional y coherente que yo. Incluso si en vez de neuronas posee chips de silicio.

Podemos hablar también sobre el argumento del espectro invertido (2). En este se nos presenta un individuo, también físicamente indistinguible de usted mismo, pero con una configuración cerebral distinta que hace que sus qualia sean diferentes a los suyos. Ante una misma experiencia de rojo, su gemelo con el espectro invertido podría captar azul. Y no es que esté percibiendo lo que conocemos como *azul*, sino se encuentra captando un rojo que su propia mente acoge como *azul*. Es más, los colores nos remiten a sensaciones y recuerdos; por lo que este individuo, al ver el color rojo que usted capta, no solamente captaría el azul, sino lo entenderá como un color frío. Cabe la posibilidad lógica de imaginar que nuestros qualia son, en sí, distintos al de los demás individuos. Ya sabemos que esto sucede con los animales; los murciélagos, por ejemplo, ven el mundo no a través de su visión, sino de ultrasonidos. Al igual que Mary puede ser una experta en el color rojo y conocer sobre este sin haberlo experimentado, se le presentaría el mismo problema a un experto en murciélagos que nunca haya vivido lo que, en sí, es ser uno.

Junto con estos dos, Chalmers presentará un hecho que debemos tener en cuenta. Y es que, en lo que respecta a nuestro avance del conocimiento del mundo físico, hemos avanzado muchísimo a lo largo de la historia, y cada vez estos avances se producen más rápido. No obstante, y como hemos dicho con anterioridad, con respecto a nuestro conocimiento de nuestra conciencia, pareciera que no nos hemos movido. A esto se le conoce como asimetría epistémica (3), ya que conocemos mucho sobre el mundo físico, pero casi nada sobre nosotros mismos. La única herramienta conocida para movernos en

el ámbito de la conciencia es el acto introspectivo, pero este no es objetivo. Según Chalmers: “Mi conocimiento de la conciencia proviene, en primer lugar, de mi propia persona, no de alguna otra observación externa.”<sup>42</sup>, y este es el gran problema de por qué no avanzamos. Por esto concluye que esto sucede porque la conciencia no puede ser explicada, de ninguna forma, ni en términos físicos.

### *El dualismo naturalista*

Todo esto, según Chalmers, nos lleva a una conclusión muy importante: si nuestro mundo carece de experiencias conscientes, si cabe pensar en un mundo lógicamente posible donde nuestra conciencia no se dé, y si los hechos de nuestra conciencia son ulteriores con respecto a los hechos fenomenológicos de nuestro mundo, “Si la conciencia no sobreviene lógicamente a lo físico”, entonces el materialismo es falso.

No obstante, desmarcarte del materialismo no es necesariamente positivo, puesto nos hace caer en una dualidad: nos encontramos en un mundo con características físicas y no físicas. Aunque sea complicado defender este tipo de teorías, al estudiar el cuerpo utilizando el método científico y las propiedades físicas nos damos cuenta de que la conciencia no parece hallarse en ningún lado. No está en la glándula pineal, referenciando a Descartes. ¿Y si la conciencia se trata de algo que surge del cerebro?, aparentemente esto es lo que se da; Chalmers afirmará, entonces, “El hecho de que la mente necesita surgir del cerebro indica que ocurre algo ulterior, más allá de los hechos físicos”<sup>43</sup>.

El dualismo naturalista que este autor defiende se basa la idea de que el mundo no está formado únicamente por propiedades físicas, sino también por propiedades fenoménicas que surgen de las físicas: propiedades *protofenoménicas*.

Los elementos con propiedades protofenoménicas, como la conciencia, surge directamente de elementos físicos. Si, como hemos establecido con anterioridad, determinada estructura física *X* es capaz de hacer surgir experiencias fenoménicas *Y*, esto se dará en cualquier situación posible; si la vida de San Francisco (la estructura física *X*) ha llevado a que sea tan buena persona que fuese convertida en un santo (*Y*), entonces cualquier otra persona que haya vivido su vida será igual.

---

<sup>42</sup> Chalmers (1996), 141.

<sup>43</sup> Ibidem, 169.

Teniendo todos estos hechos, Chalmers concluirá lo siguiente. La conciencia se trata de un elemento protofenoménico que surge del mecanismo del cerebro. Si este mecanismo llegase a ser copiado con exactamente los mismos elementos, la conciencia tendría, por consiguiente, que darse a su vez. Si construimos un individuo que en vez de neuronas tenga chips de silicio, pero posee una estructura cerebral con exactamente los mismos patrones que el cerebro humano, entonces la conciencia también surgirá de esta. De la misma manera que si esto se da de forma parecida –como con la habitación china de Searle o con la simulación cerebral de Block con los habitantes de China– de forma que es lo suficientemente compleja y parecida al cerebro humano, una conciencia también surgiría de esta.

### 3.3 ¿Qué es la conciencia?

Tanto Daniel Dennett como David Chalmers tratan el problema de la conciencia con el mismo objetivo: elaborar una nueva y mejor teoría. Ambos pasan por argumentos muy similares, puesto que son lectores y herederos de los antecedentes de su problemática tales como Descartes o Searle. Poseen puntos en común, así como también difieren en otros. Para comenzar, Dennett rechaza el basarse en una teoría dualista, mientras Chalmers critica que la teoría de las Versiones Múltiples se base en establecer ilusiones. No obstante, ambos autores llegan a la conclusión de que la conciencia –ya sea algo que surge o una ilusión– es el producto de las interconexiones que se producen dentro del mecanismo cerebral. No sabemos si es algo que realmente está ahí físicamente, pero somos conscientes de nosotros mismos y de nuestra actividad cerebral. Podemos cuestionar cómo surge esta, pero lo que me resulta muy relevante de ambas teorías es que podemos sacar la siguiente conclusión:

La conciencia es un producto de nuestro cerebro que se da gracias al papel que cada parte de este rinde. Por tanto, si el cerebro pudiese ser *simulado*, una conciencia nueva podría ser *creada*.

## **DISCUSIÓN Y POSICIONAMIENTO.**

### **Problemas de una teoría de la conciencia.**

En el capítulo anterior se mostraba bien cómo, si bien las teorías de Dennett y Chalmers, examinadas en profundidad, pueden ser capaces de contradecir a la otra, no ofrecen visiones completamente opuestas, pues tienen sus puntos en común. Como he señalado en la introducción a este trabajo, mi punto central será tratar de responder a la pregunta de si la mente puede ser duplicada. Sostengo que la conjunción de las teorías de Chalmers y Dennett en este sentido pueden servir como fundamento para defender que, efectivamente, puede crearse una conciencia de forma artificial. Recapitulando: Dennett ofrece la visión de que la mente se trata, tan solo, de una mera ilusión que tenemos los seres humanos sobre nosotros mismos. Un constructo que crea nuestro propio cerebro por motivos evolutivos. Chalmers, por su parte, analiza las estructuras donde la conciencia se produce, barajando la posibilidad de establecer simulaciones de conciencia que actúen de la misma manera que haría un cerebro real. Es decir, con la conjunción de estas dos conclusiones de sus propias teorías existiría la posibilidad dar un paso más allá en la creación de una teoría de la conciencia, puesto que la conjunción de ambas teorías puede ser capaz de definir qué es la conciencia.

No obstante, cabe reflexionar acerca del estudio del cerebro. Parece difícil de imaginar que podamos llegar a hacer grandes progresos en el estudio de la conciencia si no conseguimos llegar a estudiar el propio cerebro en profundidad. La tesis que defiende sobre este punto se basa en que para llegar a establecer una teoría sobre la conciencia necesitamos utilizar tanto a la neurobiología –como ciencia que aporte datos– como la

filosofía –como encargada de crear una teoría a través de estas investigaciones–. Para estudiar la conciencia deberemos, en primer lugar, estudiar el cerebro; y para estudiar el cerebro puede resultar, a su vez, muy positivo ir estableciendo desde antes teorías de la conciencia que ayuden a imaginar cómo podríamos empezar a analizarlo.

Ya hemos concluido la primera pregunta de qué es la conciencia, así que ahora procederé a pasar a responder a la siguiente: ¿puede ser duplicada la conciencia? Según la teoría de Chalmers y Dennett esto puede ser así si conseguimos crear una máquina que sea capaz de simular un cerebro, pero ¿cómo podría lograrse esto? En los puntos siguientes de este apartado voy a hablar acerca de las Inteligencias Artificiales, del concepto del *yo* y, por último, de la inmortalidad virtual.

#### 4.1 Los individuos mecánicos. Robots, ciborgs e Inteligencias Artificiales.

Si nos cuestionamos cómo podría ser posible duplicar a un individuo, uno de los términos a los que enseguida nos remite esta idea es el concepto de clonación. Es posible hacer un clon de una persona, en el experimento de la oveja Dolly esto ya se logró. No obstante, se trata tan solo de un clon *biológico*, de sus propiedades físicas, pero no de sus cualidades mentales. Con duplicar, por el contrario, nos referimos a establecer una copia exacta del cerebro del individuo: de su personalidad, vivencias, traumas y de cada neurona de este. Una de las formas en las que esto puede conseguirse sería con la elaboración de una máquina que sea capaz de funcionar exactamente igual que nuestro cerebro (que podría, perfectamente, tratarse de un ciborg o de un robot completamente), mientras otra sería en términos de *software*. Subir a la red nuestra propia conciencia.

Antes de proseguir, definiremos estos tres términos para que no haya confusión entre a lo que me refiero con cada uno. Un robot consiste en una serie de “creaciones tecnológicas que están destinados a realizar tareas repetitivas, riesgosas, o de altísima precisión en las que pueden superar las limitaciones biológicas del ser humano”<sup>44</sup>, puede ser humanoide como bien puede no serlo. Un ciborg, por otro lado, se trata de

---

<sup>44</sup> Rabadán (2019), 398.

una “criatura compuesta por elementos orgánicos y dispositivos cibernéticos”<sup>45</sup>. Por último, con Inteligencia Artificial entendemos “programas” –que pueden tener o no un cuerpo físico, un *hardware*– que, en términos de Rabadán, realiza una “simulación de algunas actividades del sistema nervioso humano mediante máquinas”<sup>46</sup>. En otros términos, las Inteligencias Artificiales hacen referencia a estos programas que son capaces de simular una o todas las partes del cerebro, los robots son criaturas mecánicas –no biológicas– que realizan tareas peligrosas y repetitivas y los ciborgs son una mezcla biológica entre robots –máquinas– y seres vivos. Un ciborg puede ser tanto un robot al que hayan añadido partes biológicas de un ser vivo no mecánico como exactamente lo contrario: un humano con un cuerpo con piezas mecánicas. El límite de considerar a alguien humano o un ciborg se encuentra todavía muy difuso, pues queda abierta la cuestión de hasta qué punto puedes tener partes de tu cuerpo mecanizadas y seguir siendo humano. En este escrito lo que más nos interesan son las Inteligencias Artificiales, pues son programas que podrían ser capaces de crear conciencias artificiales. No obstante, volveré sobre el tema de los robots y ciborgs más adelante.

#### 4.2 La posibilidad de crear una Inteligencia Artificial.

Chalmers señaló en *La mente consciente*, a dos grandes grupos de pensadores acerca de las Inteligencias Artificiales: los que defienden su posibilidad, y aquellos que se posicionan en contra de esta. Sus defensores, como señaló este autor, sostienen que eventualmente se empezarán a crear máquinas que pueden llegar a ser consideradas como conscientes. No obstante, aquellos que están en contra de ello lo hacen por una serie de argumentos que Chalmers divide en dos: quienes atacan a la teoría de forma externa y los que lo hacen de forma interna. No les da mucha relevancia a los primeros, pues sostiene que quienes consideran que un sistema computacional jamás será capaz de tener conciencia, debido a que se trata de una cualidad intrínsecamente humana, se equivocan. Recordemos que su teoría se basa en la idea de que una simulación del cerebro puede ser capaz de generar una conciencia.

---

<sup>45</sup> Rabadán (2019), 399.

<sup>46</sup> Ibidem, 397.

Los autores que defienden las críticas externas son Dreyfus, Lucas y Penrose; mientras el primero defiende que un ordenador no puede ser capaz de ser creativo, los otros defienden que no serían capaces de poseer una comprensión matemática del mismo nivel que los seres humanos. Pero ante estas ideas podemos establecer la crítica de que no hay un fundamento para ello. Incluso las Inteligencias Artificiales que ya existen aprenden de su alrededor y recogen información que almacenan con el fin de utilizarla más adelante. Exactamente igual que un cerebro humano. Debemos abandonar la imagen de una conciencia artificial que sea parecida al funcionamiento de una calculadora; ya no se trata de máquinas que no pueden salirse de lo que sus propios programadores deseaban, sino son capaces de aprender de su entorno y adaptarse.

Por otro lado, tenemos las críticas internas, que son las grandes críticas a las Inteligencias Artificiales. En estas se argumenta que es lógicamente posible que una máquina sea capaz de simular la conducta humana, e incluso pasar el *test* de Turing, pero no se produce ninguna garantía de que dicha máquina en sí posea vida. De que haga este tipo de actos por su propia cuenta y de forma consciente. También se ofrece un segundo argumento en base a la idea de que la conciencia se trata de una “característica que nosotros tenemos pero que los ordenadores no poseen”<sup>47</sup>, y no solo los ordenadores. Cabe pensar la posibilidad de que gran parte de los animales tampoco la tengan. Pero podemos objetar estas teorías también. La primera podemos refutarla, como la propia habitación china de Searle con la ayuda de la teoría de la conciencia de Chalmers. Ambas teorías se basan concretamente en el enfoque reduccionista, y como bien el autor criticó, la mente no puede ser explicada en estos términos.

Al igual que Chalmers, sostengo que el surgimiento de Inteligencias Artificiales es bastante probable que se dé en algún momento si seguimos avanzando en nuestros conocimientos de ingeniería y robótica. No parece que haya ningún impedimento para que estas puedan ser creadas. No obstante, ¿podemos considerar que son verdaderamente conscientes? ¿En qué punto podemos establecer que una máquina lo es? ¿Cuál es la línea que divide la conciencia de la inconsciencia en una máquina? Incluso, podemos entrar más concretamente a esta pregunta: ¿podemos considerar consciente una máquina?

---

<sup>47</sup> Chalmers (1996), 396.

Dennett, defensor del modelo de las Versiones Múltiples (en las que cada parte del cerebro hace su papel de forma semiindependiente), y de que el yo se trata simplemente de una abstracción afirmará: “más de una vez has imaginado un robot consciente. El problema no es que no *puedas* imaginar un robot consciente, sino que no puedes imaginar *cómo* puede ser consciente un robot”<sup>48</sup>. No hay ninguna complicación en imaginar que un objeto animado tenga conciencia y pueda llegar a establecer conversaciones con otros seres humanos. Los niños juegan utilizando este recurso. Y tampoco se da la situación de que nunca hayamos podido ver referencias de un robot consciente, puesto que, si bien no existe, la ciencia ficción está llena de ellos. Este autor considerará que podemos aceptar la idea de que un ordenador sea capaz de poseer conciencia si empezamos con la idea computacionalista de “pensar en el cerebro como si fuera una especie de ordenador”<sup>49</sup>. Él sostiene que utilizando la introspección podemos llegar a establecer una buena teoría de la conciencia si conseguimos eliminar aquellas cosas que nos confunden tales como la existencia del Teatro Cartesiano o de los qualia. Incluso sostiene la posibilidad de que puedan ser capaces de establecer puentes entre el mundo físico y el suyo propio como máquina, “capaz de establecer los vínculos necesarios con el mundo del contenido, con los mundos de la (hetero-)fenomenología.”<sup>50</sup>, tal y como podemos hacer el resto de las personas.

La verdadera cuestión se encuentra en si los individuos humanos sabríamos cómo posicionarnos al lado de una Inteligencia Artificial, pues no tenemos del todo claro si estas puedan llegar a ser conscientes por sí mismas. ¿Al hablar con una estaríamos ante alguien que te comprende, o ante una habitación china donde solo se da esa apariencia? Pero como hemos visto con la teoría de Chalmers y Dennett, si conseguimos crear una máquina más o menos consciente, ¿por qué no surgiría de esta una conciencia tal y como nosotros la conocemos? Quizá, porque estas aparentemente no podrían ser consideradas *un yo*.

---

<sup>48</sup> Dennett (1991), 533.

<sup>49</sup> Ibidem, 535.

<sup>50</sup> Ibidem, 536.



### 4.3 ¿Qué es el “yo”?

Si seguimos el planteamiento de Dennett y Chalmers, no habría ningún atisbo de dudas con que, de hecho, la conciencia puede ser duplicada. El problema de la duplicación de la conciencia pertenece a la rama de la filosofía transhumanista. Víctor Kerbs la define como “un movimiento cultural o una filosofía, para la cual el futuro de nuestra especie consiste en su liberación del cuerpo y superación de la muerte”<sup>51</sup>.

Estamos utilizando en este trabajo la idea de la conciencia como autoconciencia. No obstante, existen otras teorías de la conciencia diversas, de la misma manera que hay diferentes concepciones de lo que es un ser humano. En este trabajo sostengo la idea de que un ser humano es, en sí, la información que posee en su cerebro. Cuando esta queda dañada incluso puede darse la idea de que *no se trata de la misma persona*, a pesar de que físicamente lo siga siendo. Pero es posible concebir a tu persona no solamente como un cúmulo de información, sino también como tu propio cuerpo físico. ¿Podrías seguir considerándote tú mismo si añades partes mecánicas a tu cuerpo? ¿Sucede algo si nos implantamos un chip? ¿Y si cambiamos una extremidad? ¿Hasta qué punto podríamos seguir hablando de que se trata de la misma persona? Es exactamente el mismo problema del barco de Teseo aplicado al cuerpo humano. Este, planteado por Plutarco, nos narra acerca de un barco con el cual se hacía, a modo de ritual, un recorrido desde Atenas hasta Creta. En el barco cada vez que se producía un daño de cualquier tipo se reemplazaba la pieza por otra nueva. Si esto continúa hasta el punto en el que el barco ya no conserve ni una sola pieza de las que tenía al principio, ¿podría seguir considerándose el barco de Teseo? Es posible que esa cuestión se nos plantee en un futuro a las propias personas en un posible futuro –y no *tan* futuro– mundo donde seamos ciborgs.

Existen personas dispuestas a defender que lo que te constituye a ti mismo no es simplemente esa información que posees en tu cerebro, después de todo esta es reemplazable y está en constante cambio. Vamos olvidando cosas y aprendemos otras nuevas, dejamos de lado aparcados algunos conocimientos mientras vamos cambiando otros. ¿No se podría decir que incluso podría darse el dilema del barco de Teseo en ti mismo? ¿Sigues siendo, acaso, la misma persona que fuiste hace una semana? ¿Y hace diez años? ¿Serás el mismo dentro de otros diez?

---

<sup>51</sup> Kerbs (2020), 14.

No hay fundamento para sostener que sí ni que no, puesto que depende completamente de lo que cada individuo comprenda como *él mismo*. Si usted defiende una postura donde el barco de Teseo siga siendo el mismo barco, pues funciona para el mismo propósito, y es solo una *evolución* de este, diferirá completamente de quienes consideran que el barco de Teseo se encuentra en las antiguas tablas que fueron sustituidas. Incluso es posible establecer posiciones intermedias donde se considere que el sujeto es tanto el *nuevo* “barco” como el *viejo*. Personalmente me decanto por esto último, puesto que, aunque no seas exactamente igual que hace un cierto número de años, hay algo que se mantiene en ti, *algo* que hace que sigas siendo tú mismo: tu continuidad física y psicológica. Víctor Krebs afirmará, en línea de esto: “Lo que somos no es solo esa parte de nosotros que ya ha sucedido sino, más intrínsecamente, la vitalidad espontánea que permanece detrás de lo real, en lo más profundo de la materia”<sup>52</sup>.

El problema de la duplicación de la mente, a su vez, trae otro dilema que es digno de mención. Y es que, si conseguimos establecer una máquina capaz de duplicar tu propia conciencia, ¿esta máquina serías tú mismo, o se trataría de un *clon* con experiencias iguales o similares? Si se trata de un clon, ¿no podría darse, con el tiempo, que el “clon” difiera de su original? De nuevo, la respuesta dependerá del propio sujeto, pero personalmente sostengo que se trataría de un ser distinto a ti mismo, puesto que incluso si tiene tus mismos recuerdos y experiencias pasadas, una vez que esté “vivo”, comenzará a tener experiencias del mundo físico que, con el tiempo, hará que seáis capaces hasta de diferir completamente.

De todas maneras, resulta complicado imaginar la posibilidad de crear una máquina capaz de duplicar conciencias. Incluso si las teorías de Chalmers y Dennett son capaces de explicar *cómo* podría darse, carecemos del material necesario como para llevarlo a cabo. Esto no significa que en un futuro no pueda suceder, pero mirándolo desde nuestra actualidad, se crea el debate de si es posible sacar la conciencia de su propio espacio físico, su cerebro. Espíritu Avila se encuentra completamente de acuerdo con ello, estableciendo que “es demasiado arrogante para este momento de la historia de la humanidad negar la posibilidad de que la conciencia pueda extraerse del cerebro

---

<sup>52</sup> Krebs (2020), 20.

físico”<sup>53</sup>. No obstante, personalmente no considero que su punto sea cierto, puesto que a pesar de que no parece imposible, la investigación científica no ha avanzado tanto como para estar cerca de que esto suceda.

Si sostenemos la teoría de que lo que te constituye a ti como lo que eres se trata tanto de tu cerebro como de tu conciencia, hasta el punto de que todo el resto de tu cuerpo puede ser sustituido, si todo lo que te constituye como *tú* es esta información, entonces es perfectamente posible argumentar que esta conciencia, duplicada, seas también tú. Si tu cuerpo físico fallece nada más traspasar tu conciencia, directamente podría sostenerse que ha sido uno de estos momentos descritos por Dennett donde tu conciencia se apaga y se vuelve a encender un poco más tarde. Además, esto genera una gran ventaja al sujeto, puesto que si tu cuerpo, caduco, es reemplazado por uno que puede ser sustituido, dejando intacta tu conciencia, ¿no estaríamos ante un tipo de *inmortalidad*?

#### 4.4 La inmortalidad virtual.

La posibilidad de duplicar la conciencia no solamente resulta un gran avance en lo que respecta aprender más sobre nosotros mismos, o una forma de superar la inteligencia que un ser humano puede alcanzar en vida, sino también si se hace un traspaso de conciencia perfecto en el cual no se pierda ninguna parte de esta, si la máquina donde esto sucede puede sustentarse de forma infinita, estaríamos ante una forma de alcanzar la inmortalidad virtual.

En términos de Justo Domínguez, con la duplicación de la conciencia una de las cosas que se pretenden es “alcanzar la mayor longevidad posible, evitando el envejecimiento y prolongando un estado físico de juventud”<sup>54</sup>. En un mundo donde todas las personas sean simple información guardada en la nube, o “un USB de millones de terabytes.”<sup>55</sup> ya no tendríamos que hablar en términos de salud, enfermedad o muerte, pues todos estos problemas podrían ser suplidos.

Como afirmó Víctor Krebs, si se crease este mundo virtual de conciencias subidas a un ordenador, esto implicaría que deberíamos replantearnos las diferencias que se producen

---

<sup>53</sup> Espíritu (2018), 46.

<sup>54</sup> Justo (2019), 234.

<sup>55</sup> Espíritu (2018), 48.

entre lo *virtual* y lo *real*. Pero este propio replanteamiento se está dando ya, hoy en día, debido a que el uso de las distintas tecnologías se ha vuelto una parte completamente importante en nuestra vida. Hace tan solo unas décadas se hablaba diferenciando lo que sucedía «en internet» contra «en el mundo real», pero ahora esta línea ha sido difuminada hasta el punto donde nos preguntamos si deberíamos dejar de distinguir ambos “mundos” y constituirlos como uno solo.

En el artículo de Espíritu Avila titulado “¿Es posible o no la inmortalidad cibernética?” hablará acerca del proyecto ruso Iniciativa 2045 de Dmitry Itskov. En él el multimillonario promete que para dicho año conseguirá hacer posible “transferir las funciones humanas de un individuo a una entidad artificial para de esa forma darle inmortalidad, es decir, hacer que la conciencia sea transferida a un objeto no orgánico”<sup>56</sup>. Actualmente se encuentra trabajando con un grupo de investigadores que poseen este objetivo, estableciendo como meta que entre los años 2015 y 2020 se podría crear un cuerpo artificial que pueda ser sustentador de una inteligencia artificial; esta última no será creada hasta los años 2030-2035. Aunque se trata de una especulación un tanto arriesgada, debido a las fechas tan tempranas que establece, y no se ha cumplido su primer objetivo, se está investigando desde hace décadas en la creación de robots artificiales humanoides. Estos serían capaces de hacer funciones que un ser humano natural no puede o prefiere no hacer. Recordemos que los robots son máquinas creadas para hacer trabajos que son peligrosos para los seres humanos, o para hacer el trabajo de varios en menos tiempo. No veo posible que este proyecto pueda llevarse a cabo en tan poco tiempo por los escasos avances en la robótica que hemos poseído, pero no descarto que sea posible crear esa máquina de la que hablan tanto Dennett como Chalmers; una que sea capaz de imitar a un cerebro. Chalmers señalará que “la gran plausibilidad de la tesis de la IA se debe precisamente a que la clase de sistemas computacionales es tan amplia”<sup>57</sup>, como existe la posibilidad lógica de que pueda ser creado un sistema que simule el cerebro, independientemente del material con el que esté hecho, entonces no se da ningún problema con ello.

Justo Domínguez señalará en “Vida mortal y eternidad” un hecho que, también, resulta digno de mención, y es el problema de para qué podría quererse ser inmortal. Qué podría implicar para los seres humanos que esto se diese. ¿Acabarían siendo

---

<sup>56</sup> Espíritu (2018), 42.

<sup>57</sup> Chalmers (1996), 418.

considerados de menor categoría los humanos antes de las conciencias duplicadas? ¿Seguirían naciendo biológicamente humanos? ¿O solo artificialmente? ¿Dónde quedaría la humanidad ahí? Este tipo de teorías me parecen muy cercanas a las de la Antigua Grecia, pues valoran más la propia conciencia, la mente, el alma de la persona, antes que la materia. Dejan de lado el mundo físico, como si fuese algo de segundo orden, para dar pie a un mundo virtual más perfecto. En sí esto podría incluso parecer el planteamiento de una utopía: algo deseable, pero que en la práctica dejase de serlo. Y es que estamos hablando, con esto, de teorías tales como las de los cerebros en tarros de Hilary Putnam. Esta se basa en la teoría mental, a modo de supuesto hipotético, de que todos seamos cerebros en cubetas que están conectados a un ordenador que simula el mundo *como si* lo estuviésemos experimentando en primera persona. El problema de esto se basa en la idea de que, si somos cerebros en cubetas, nunca lo sabríamos, pues no tendríamos conciencia de que el mundo físico que experimentamos sea *el real*.

En un mundo donde la conciencia pueda ser duplicada y subida a la red, uno de los destinos posibles de este sería crear todo un sistema de cerebros en cubetas, o información en una nube/USB. ¿Es esto a lo que la humanidad aspira? ¿O nos detendremos mucho antes de llegar a este nivel de investigación? Los recursos de la Tierra son finitos, y parece que estamos llegando a estar cerca del límite de estos. Quizá sea posible crear una máquina que simule la conciencia, pero nunca seremos capaces de simplemente comenzar a crearla por la falta de materiales. No obstante, incluso si no podremos ser capaces nunca de hacer este tipo de proyectos en el mundo físico real, resulta muy interesante plantearse estas preguntas.

En un mundo donde gran parte de la realidad se dé también de forma virtual, ya sea porque nuestros cuerpos físicos sean descartados y nuestras consciencias pasen a estar en un gran programa artificial, o ya sea porque con el uso de la Realidad Virtual podríamos tener una vida más allá de lo físico, la humanidad sería replanteada. Quizá este mundo nunca se dé, pero creo que solamente podría no darse si frenamos los avances científicos. Al menos las investigaciones tecnológicas están yendo, directamente a este futuro. Si bien es probable que nunca podamos comprender en este mundo, si conseguimos teorizar qué es la conciencia, ¿no sería posible, aunque sea, establecer los *planos* de este mundo virtual?

## **CONCLUSIÓN Y VÍAS ABIERTAS.**

### **Desde el presente hacia el futuro.**

Tras realizar el recorrido histórico de los análisis de la conciencia, vimos que a lo largo de la historia este problema ha sido presentado como el problema alma-cuerpo, el problema mente-cuerpo, el problema mente-mente y, por último, el problema de la conciencia. Estudiamos, a su vez, las teorías de la conciencia de Daniel Dennett y David Chalmers con el fin de establecer una unificación entre ambas. Concluimos que esta se trata tan solo de una mera abstracción de nuestro cerebro que tiene un fin práctico. Así como que, ante una duplicación exacta de los procesos cerebrales de una persona, podríamos crear una nueva conciencia, pues esta es una propiedad emergente surgida de ellos. De esta forma podría ser posible crear una conciencia artificial. Por último, nos planteamos la posibilidad de duplicar y simular la conciencia humana y de, incluso, alcanzar la inmortalidad virtual.

A pesar de que no se ha llegado a ningún consenso sobre lo que es la conciencia en sí, y esta y cualquier otra tesis pueden recibir refutaciones, cada teoría de la conciencia que se hace resulta un avance para la investigación del mundo físico. Con frecuencia se busca darle gran importancia al estudio de la naturaleza y se deja de lado a las cuestiones del propio sujeto. E incluso si distintas disciplinas científicas han avanzado estudiando el cuerpo humano y la realidad física, el cerebro sigue siendo el órgano de nuestro cuerpo que menos conocemos.

En sí me gustaría concluir que, si bien no parece del todo *naturalmente posible* duplicar la conciencia –pues sería difícil construir una máquina de dicha magnitud–, sí que es *lógicamente posible*, y más en el momento presente. Por ello es necesario plantearse de forma seria, pues puede que con los avances tecnológicos pueda volverse una realidad. Cada día, la tecnología llena más y más nuestro día a día; vivimos en redes sociales, al lado de nuestros teléfonos, tabletas y ordenadores. Habitamos entre cables, viendo la contaminación lumínica como algo normal, y siendo completamente dependientes de todos estos aparatos. Además, a pesar de que las Inteligencias Artificiales y los ciborgs parecen temas muy lejanos a lo que nos acontece hoy en día, los robots llevan ya numerosas décadas entre nosotros. Incluso se prevé que es posible que en un futuro cercano más puestos de trabajos sean eliminados por robots, y que nos los repartiremos equitativamente para tan solo dentro de cuatro años<sup>58</sup>. Las funciones que pueden desempeñar estas máquinas son innumerables; nos cuesta establecer límites para algo que se encuentra, aún, en desarrollo. Ya hay empresas y modelos de robots que se venden libremente a las personas, como Boston Dynamics. Esta es una de las más conocidas, sobre todo por su perro robot llamado Spot, con funciones que van actualizándose periódicamente y con una libertad que permite al propio usuario hacer lo que prefiera con ellos. De esta forma el mismo modelo ha sido utilizado tanto para ayudar en tiempos de pandemia a familiares a comunicarse por videollamada con alguien ingresado en el hospital, como para ayudar a personas en trabajos peligrosos, o incluso para ser utilizado en movimientos belicistas por el gobierno de Francia<sup>59</sup>. Además de esto tenemos otras compañías tales como Misty Robotics, que fabrica robots que pueden ayudar a las personas mayores –y a todo el que lo desee– en sus tareas y aportarle compañía.

No podemos afirmar si algún día llegaremos a comprender lo que es la conciencia, ni si podremos llegar a duplicarla, pero con los avances que está teniendo la ingeniería y la robótica, no parece descabellado pensar que nos encontramos a tan solo unas décadas de que esto se haga realidad. Quizá, en menos de cien años, tengamos un mundo lleno de ciborgs en el que habiten con Inteligencias Artificiales capaces de superar el test de Turing.

---

<sup>58</sup> Pérez (2020).

<sup>59</sup> CNN (2021), Heriot Watt University (2021) y Durden, T. (2021) respectivamente.

## **BIBLIOGRAFÍA CITADA.**

### 5.1 Bibliografía.

Chalmers, D. J. (1999). *La mente consciente: En busca de una teoría fundamental*. Barcelona: Gedisa editorial.

Churchland, P. M. y Churchland, P.S. (1990). ¿Podría pensar una máquina? *Investigación y ciencia*, 162, marzo, 18-24.

Dennett, D. (1991). *La conciencia explicada: Una teoría interdisciplinar*. Barcelona: Paidós.

Ferrater, J. (2008). *Diccionario de filosofía abreviado*. Barcelona: Libros de Sísifo, Edhasa.

Flanagan, O. (1991). *The science of the mind*. (2ª ed.). Estados Unidos: Instituto de Tecnología de Massachusetts.

Gellatly A. y Zarate, O. (1994). *Introducing mind & brain: A graphic guide*. Londres: Totem Books.

Gomila, A. (1990). El materialismo eliminativista de los Churchland. *Contextos*, VIII (15-16), 241-259.

Laureano, L. C. (2016). Tres argumentos ordinales contra el computacionalismo. *Thémata, Revista de Filosofía*, 55, octubre, 327-340. DOI: 10.12795



Liz, A. M. (2001-a). *Perspectivas actuales en filosofía de la mente*. Tenerife: Colección textos universitarios.

Liz, A. M. (2001-b). *La vida mental de algunos trozos de materia: Teorías de la sobrevenida*. Colección tropos 2.

Liz, A. M. (2008). Sobrevenida: Un caso de ingeniería conceptual. *Logos. Anales del seminario de Metafísica*. 42, 243-260.

Pineda, D. (2012). *La mente humana: Introducción a la filosofía de la psicología*. Madrid: Cátedra.

Searle, J. R. (1990). ¿Es la mente un programa informático? *Investigación y ciencia*, 162, marzo, 10-16.

## 5.2 Webgrafía.

Aguilar, M. T. (2010). Descartes y el cuerpo máquina. *Pensamiento. Revista de investigación e información filosófica*, 66(249), 755-770. Recuperado de: <https://revistas.comillas.edu/index.php/pensamiento/article/view/2491>

CNN. (23 de marzo de 2021). *Mira este perro robótico interactuar con pacientes en un hospital*. CNN. Recuperado de: <https://edition.cnn.com/videos/spanish/2021/03/23/perro-robotico-hospital-boston-dynamics-dr-spot-perro-robot-ayuda-medicos-pacientes-klv-cnn-dinero.cnn>

Cotrina, J. (22 de febrero del 2021). Robots que hacen compañía. *El periódico*. <https://www.elperiodico.com/es/barcelona/20210222/robots-compania-11536432>

Durden, T. (14 de abril de 2021). Boston Dynamics' Robot Dog Tested In War-Training Exercise. *Zero Hedge*. Recuperado de: <https://www.zerohedge.com/political/boston-dynamics-spot-robot-tested-french-army-field-training-exercise>

Espíritu, A. R. (2018). ¿Es posible o no la inmortalidad cibernética? *Tierra Nuestra*, 12, 41-51. Recuperado de: <https://revistas.lamolina.edu.pe/index.php/tnu/article/view/1268/1318>

Heriot Watt University. (2 de junio de 2021). New robot set to save lives, cut CO2 and support construction. Heriot Watt University. Recuperado de: <https://www.hw.ac.uk/news/articles/2021/new-robot-set-to-save-lives-cut-co2-and.htm>

Justo, E. J. (2019). Vida inmortal y eternidad Sobre el proyecto transhumanista de inmortalidad. *Scientia et Fides*, 7(2), 233-246. Recuperado de <https://dadun.unav.edu/bitstream/10171/58302/1/4%20Vida%20inmortal%20y%20eternidad.pdf>

Krebs, V. J. (2020). Cuerpo virtual. Avatares de la digitalidad. *LÓGOI Revista de filosofía*, 37, 13-26. ISSN: 1316-693X. Recuperado de: <http://revistasenlinea.saber.ucab.edu.ve/temas/index.php/logoi/article/view/4543>

Leiva, C. (2005). Conductismo, cognitivismo y aprendizaje. *Tecnología en Marcha*, 18, (1), 66-73. Recuperado de: <http://www.scielo.org.co/pdf/rcps/v22n2/v22n2a12.pdf>

Llinàs, J. L. (2017). El cuerpo como máquina: la aportación del dualismo cartesiano al desarrollo de la ciencia moderna. *Daimon Revista Internacional de Filosofía*, 437-443. Recuperado de: <https://doi.org/10.6018/daimon/268851>

Pérez, G. R. (22 de octubre de 2020). Robots y humanos se repartirán por igual los trabajos en 2025. *El País*. Recuperado de: <https://elpais.com/economia/2020-10-21/robots-y-humanos-se-repartiran-por-igual-los-trabajos-en-2025.html>

Rabadán, A. T. (2019). Horizonte de la inteligencia artificial y neurociencias. Acerca de robots, androides y cyborgs. *Medicina (Buenos Aires)*, 79, 397-400. Recuperado de: <https://www.medicinabuenosaires.com/revistas/vol79-19/n5/Rabadan.pdf>

Stangroom, J. (Consultado el 3 de junio de 2021). Staying Alive - The Scenarios. Recuperado de: <https://www.philosophyexperiments.com/stayingalive/Default.aspx>

The Engineer Jobs. (2 de junio de 2021). Telexistence technology to assess robot in hazardous tasks. *The Engineer Jobs*. Recuperado de: <https://www.theengineer.co.uk/telexistence-technology-robot-hazardous-tasks/>

Vázquez, M. J. (2004). Immanuel Kant: El giro copernicano como ontología de la experiencia. *ÉXODA: Series filosóficas*, 18, 69-93. Recuperado de: <http://revistas.uned.es/index.php/endoxa/article/view/5082/4901>

