

**MEDIDAS BASADAS EN EL CURRÍCULO PARA LA  
IDENTIFICACIÓN TEMPRANA DE DIFICULTADES DE  
APRENDIZAJE EN MATEMÁTICAS: UNA REVISIÓN  
SISTEMÁTICA**

**Trabajo Fin de Grado Psicología**

**Ainoa Campos Martín**

**Judit Falcón Rodríguez**

**Tutorizado por: Juan Eugenio Jiménez González**

**Curso académico 2021-22**

## Resumen

El principal objetivo de este trabajo ha sido realizar una revisión sistemática basada en el método PRISMA sobre las medidas de evaluación basada en el currículo (CBM, en sus siglas en inglés) para la identificación temprana de las dificultades de aprendizaje en matemáticas (DAM). La detección temprana de las DAM es de vital importancia para proporcionar al alumnado una correcta instrucción que garantice su correcto desarrollo y progreso tanto a nivel tanto educativo, como a nivel social y emocional. En los últimos años, las medidas basadas en el currículo se han establecido como una de las herramientas más beneficiosas para la evaluación de los niños/as en el ámbito escolar, dadas sus ventajas y su forma de aplicación. Se han revisado aquellos estudios que se han centrado en niños/as de edades comprendidas entre 3 y 5 años, donde se han analizado las propiedades de las CBM. Se concluye que, a pesar de contar con ciertas limitaciones, la CBM, sin embargo, se trata de una herramienta eficaz y fiable para la detección temprana de las DAM.

**Palabras clave:** medidas basadas en el currículo, preescolar, matemáticas, estudiantes en riesgo.

## **Abstract**

The main objective of this study has been to carry out a systematic review based on curriculum-based assessment (CBM) measures for early identification of learning difficulties in mathematics (LDM) using the PRISMA method. The early detection of LDM is of vital importance to provide students correct instruction that ensures their correct development and progress both at the educational level, as well as at the social and emotional level. In recent years, measures based on the curriculum have been established as one of the most beneficial tools for the assessment of boys and girls in the school environment, given their advantages and their manner of application. Studies focusing on children between the ages of 3 and 5 have been reviewed and the properties of CBMs have been analysed. It is concluded that, despite certain limitations, CBM, however, is an effective and reliable tool for the early detection of LDM.

**Keywords:** curriculum based measures, kindergarten, mathematics, at-risk students.

## Introducción

Como es sabido, hoy en día vivimos en una sociedad bastante evolucionada, pues gracias a las nuevas tecnologías y los avances científicos se ha conseguido mejorar en gran medida las condiciones de vida de la población, incluyendo la de los/as niños/as; y no solo mejorar la calidad de vida, sino prevenir la mortalidad. Es por ello por lo que en la actualidad la esperanza de vida está en torno a los 72 años, algo que era impensable años atrás. Más concretamente, en décadas anteriores la mortalidad infantil era mucho más alta que hoy en día, al igual que la tecnología era mucho más precaria, lo cual está relacionado, ya que la falta de recursos tecnológicos en medicina podía propiciar dicha mortalidad infantil al no tener los recursos suficientes y/o necesarios para paliarla.

Actualmente, ya no contamos con ese tipo de desventajas, y se podría decir que tenemos muchos más beneficios. Esto puede ser al mismo tiempo, aunque contradictorio, un inconveniente tal y como afirma Mateos (2009). Con esto se hace referencia a que sí, ha disminuido la mortalidad infantil, pero a su vez ha ido aumentando la morbilidad. Por eso se observa como algo contradictorio: el aumento de la tecnología tiene como consecuencias positivas que no haya tantas muertes infantiles de manera “natural”, pero a su vez se fomenta que dichos niños/as enfermen debido a esto.

Al decir que esto promueve que los niños/as enfermen, se hace alusión a todos aquellos “trastornos” que tuvieran que ver con el aprendizaje, ya que como se comentó anteriormente, dichos avances hacen que se generen muchos más problemas neuropsicológicos, lo que podría derivar a su vez en alteraciones emocionales y en dificultades y/o discapacidades de aprendizaje.

Pero, ¿qué son las dificultades de aprendizaje? Se trata de un término muy genérico que se emplea para designar a un grupo heterogéneo de trastornos que son manifestados por dificultades significativas en la adquisición y el uso de la capacidad para entender, hablar, escribir, razonar, leer, o para las matemáticas, que son intrínsecas al individuo, y que están provocadas por una disfunción y/o alteración del sistema nervioso central. Asimismo, dentro de las dificultades de aprendizaje (DA) hay tres grandes tipos, las referidas a la lectura, escritura y/o matemáticas, que a su vez incluyen diferentes subtipos dentro de cada una de ellas.

Así, algo importante a destacar es que esta condición afecta a un gran número de personas, pero no se le da la relevancia que realmente se merece. Esto

es debido a que, a pesar de ser un tema muy estudiado, no se ha llegado a un consenso o acuerdo sobre cómo enfocarlo, por lo que cada disciplina implicada ha tomado su camino de manera independiente, no de forma interdisciplinaria, y generando así una gran controversia.

Por otro lado, esta dificultad que sufren los niños/as, normalmente se tiende a “normalizar” o se tiende a aminorar su importancia. Esto se convierte en algo perjudicial para la vida del niño/a, pues es muy complicado poder desenvolverse en la sociedad cuando se carece de ciertos conocimientos y/o habilidades que realmente son básicas e imprescindibles en una persona. Además, se tiene una idea errónea de que las DA solo son plasmadas a nivel académico, y que solo se tendría dificultades en ese ámbito, lo cual para nada es así, ya que también influyen a nivel profesional, emocional y social.

Del mismo modo, al tratarse de un problema de base biológica, podría ser detectado con bastante rapidez si fuera de notable importancia; es decir, cuando se tiene un grado bastante alto de DA, podría verse incluso desde los primeros meses de vida del bebé. El problema está cuando el grado de dicha dificultad no es muy severo, ya que cuesta más identificarlo, pudiendo retrasarse su detección hasta varios años después. A su vez, aunque el problema que presentara algún niño/a no fuera tan grave, no significa por tanto que no se tuviera que intervenir.

Como se ha mencionado anteriormente, uno de los tipos de dificultades de aprendizaje son las Dificultades Específicas de Aprendizaje en Matemáticas (DEAM), las cuales, tal y como afirma Jiménez (2019), están relacionadas con déficits en el sentido numérico o comprensión del número. El sentido numérico puede definirse como una destreza innata del ser humano, que funciona de manera prácticamente intuitiva, y que nos permite hacer estimaciones o modificaciones en un conjunto de cosas de manera rápida y eficaz. Además, el sentido numérico cuenta con dos etapas diferenciadas: la etapa previa a la incorporación al ámbito educativo (sentido numérico informal), y la etapa posterior a la misma (sentido numérico formal).

El sentido numérico informal se genera tempranamente, incluso antes del habla, y no tiene distinción en función de las características o el contexto donde se desarrolla la persona. Este conjunto de habilidades básicas de carácter no simbólico (no utiliza contenido abstracto) son adquiridas por los niños/as antes de iniciar su

escolarización, y son el punto base para la posterior adquisición de habilidades matemáticas más complejas.

Por su parte, el sentido numérico formal tiene una adquisición más paulatina, que se empieza a generar una vez que el/la niño/a se incorpora a un centro educativo, y que le permite, a partir de las habilidades innatas con las que ya contaba, aprender a manejar contenido simbólico, a saber, utilizar y entender los números, las cantidades, y por tanto a comprender y manejar mejor su realidad y el ambiente que le rodea. Varias son las habilidades que componen el sentido numérico formal: conteo, estimación, identificación numérica, valor de posición, comparación de magnitudes, patrones numéricos y cálculo, las cuales se verán con mayor detalle más adelante. El conteo, tal y como definen Koponen, Aunola, Ahonen y Nurmi (2007), es la capacidad de enumerar la secuencia numérica adecuadamente, para lo cual se necesita saber el nombre de los números y el valor de los mismos en la línea numérica. La estimación hace referencia a la capacidad de identificar el número de elementos aproximados que componen un conjunto. A su vez, la identificación numérica es la capacidad de saber determinar el nombre, el símbolo y la magnitud que representa un número, lo que posteriormente permite diferenciar ese número de cualquier otro. El valor de posición se relaciona con la comprensión del valor que adquiere un número en función de la posición que ocupe (unidades, decenas, centenas...). Por su parte, la comparación de magnitudes tiene que ver con la capacidad de comprender que un número simbólico representa una cantidad, y que esa cantidad es “mayor que” o “menor que” otra. La tarea de patrones numéricos se adquiere una vez que se domina la línea numérica, y permite contar a partir de un número determinado. Por último, la habilidad de cálculo es aquella que permite realizar operaciones aritméticas, que al principio serán más sencillas (adición o sustracción), y que se irán complicando con el desarrollo posterior.

Asimismo, las DEAM, al tener su causa en factores biológicos, es importante descartar que sean factores ambientales o de carácter instruccional (en concreto por una mala instrucción educativa).

Retomando lo comentado previamente, el sentido numérico constituye la base de la adquisición de la competencia matemática, que se ve deteriorada en las DEAM. De esta forma, se hace prácticamente imprescindible detectar de manera temprana a aquellos/as niños/as que muestran ciertas carencias o dificultades a la

hora de enfrentarse a tareas matemáticas, pues de lo contrario muy posiblemente lo seguirán haciendo durante el resto de su etapa educativa (Duncan et al., 2007).

Dicho esto, y destacando una vez más la gran importancia que tiene el detectar de forma temprana dichas dificultades, en el presente trabajo se ha llevado a cabo una revisión sistemática de las herramientas más empleadas para la detección del riesgo de presentar DEAM. Así, dicha herramienta es denominada Medida Basada en el Currículo (MBC), la cual sirve para que se pueda evaluar no solo el progreso del estudiantado, sino también la efectividad de los métodos de instrucción que se estén utilizando, así como hacer un análisis de aquellas áreas en las que no se alcance las competencias básicas establecidas en el currículo escolar, para la edad de la muestra de estudiantes con la que se esté trabajando. Por ello se la ha considerado como una medida de cribado que puede identificar a alumnos/as que puedan estar en riesgo de presentar futuras DA.

La MBC, originada a través del Instituto de Discapacidades de Aprendizaje de la Universidad de Minnesota, es definida como el conjunto de medidas que parte del currículo del alumno/a para llevar a cabo una valoración de su progreso con métodos como la observación directa o el registro de su desempeño. A su vez, la MBC se sitúa en el marco del Modelo de Respuesta a la Intervención (RtI en inglés), que pretende trabajar en la prevención de las DEA partiendo de las respuestas del propio alumno/a y modificando o generando a partir de ellas nuevas formas de instrucción.

Dentro de las características con las que deben contar las MBC se encuentran las siguientes: deben ser fiables y válidas, fáciles y rápidas de administrar, ser útiles para plantear intervenciones y para registrar el progreso de las mismas. Asimismo, tal y como defienden diversos autores, las MBC también deben ser: poco costosas, posibles de modificar o de reproducir, rápidas, fáciles (tanto a la hora de usarlas como de enseñarlas), capaces de monitorear el progreso del estudiante durante su aplicación y favorecer la toma de decisiones.

Asimismo, esta herramienta tiene bastantes beneficios ya que es de bajo costo y su aplicación es asequible como para que la puedan administrar tanto profesores como profesionales de la educación, pero, además, tiene una muy alta precisión diagnóstica en lo que respecta al estatus de riesgo. Así pues, también se puede añadir que este instrumento sirve para detectar los riesgos en las tres áreas que se conocen dentro de las DA (i.e., lectura, escritura y matemáticas).

En los últimos años, se ha incrementado el valor que se le da a la detección temprana de las posibles DA de los/as niños/as por sus beneficios posteriores. Tal y como comentan Polignano y Hojnoski (2012), las MBC constituyen una forma de detección temprana que permite no solo detectar a dichos alumnos/as sino también plantear una intervención adecuada a sus necesidades.

Varias MBC han sido creadas para evaluar el riesgo y progreso de niños/as en el ámbito matemático, donde muchas de ellas se basan en el sentido numérico. Por ejemplo, Gersten et al. (2012) realizaron una revisión donde a través de tareas como conteo, comparación numérica, identificación numérica, recuperación de hechos numéricos, etc., trataron de analizar las características a nivel psicométrico de la MBC para niños/as de Educación Infantil y 1º de Primaria.

Con todo lo expuesto hasta ahora, y siguiendo la línea del trabajo llevado a cabo por Gersten et al. (2012), el principal objetivo de la revisión sistemática que aquí se presenta, ha sido analizar si las MBC constituyen una forma de evaluación fiable y eficaz para detectar tempranamente a aquellos/as niños/as en edad preescolar (de 3 a 5 años) que cuentan con un perfil de riesgo de presentar una DEAM, y si efectivamente ese riesgo es real o se trata de un método de evaluación poco preciso.

## **Método**

Para la realización del presente trabajo se llevó a cabo una revisión sistemática de la bibliografía existente sobre las MBC en el área de matemática temprana basado en el método PRISMA.

De esta forma, los documentos que se incluyen en esta revisión sistemática fueron seleccionados a través de la búsqueda en cuatro bases de datos diferentes: WOS, EbscoHost, Scopus y ERIC.

Para la búsqueda de documentos se utilizaron tres grupos de descriptores, el primero basado en propiedades psicométricas, el segundo relacionado con las habilidades matemáticas y el tercero con el tipo de evaluación utilizada. Las palabras introducidas en las diversas bases de datos fueron utilizadas en inglés con fines de ampliar las opciones de búsqueda.

Las dos autoras de este trabajo, alumnas de 4º de Psicología de la Universidad de La Laguna, realizaron búsquedas independientes con los mismos descriptores en todas las bases de datos, siendo un total de nueve búsquedas, que quedan reflejadas en el Anexo. En las diversas búsquedas se fueron añadiendo descriptores o modificando aquellos que no otorgaban un gran número de resultados.

Del mismo modo, se utilizó el asterisco (\*) cuando fue necesario para aumentar los resultados de búsqueda, puesto que ese carácter colocado tras el lexema de una palabra permite a las distintas bases de datos ampliar la búsqueda (por ejemplo, se utilizó “*measur\**” en lugar de “*measurement*” o “*measures*”). Así, finalmente se seleccionó el perfil de búsqueda utilizado en la quinta búsqueda, pues fue el que reflejó mejores resultados. En el primer grupo de descriptores se utilizaron las palabras: *screen\**, *benchmark\**, *diagnostic accuracy*, *classification accuracy*, *psychometric characteristics*, *utilit\**, *technical adequacy*, *roc*, *validation*, *criterion test*, *predictive validity*, *construct validity*, *content validity*, y *concurrent validity*. En el segundo conjunto las palabras elegidas fueron *math\**, *number sense*, *value position*, *magnitude comparison*, *numerical comparison*, *counting*, *quantity estimate*, *numerical sequences*, *numerical identification* y *calculation*. Por último, en el tercer grupo de descriptores se utilizaron *cbm* y *curriculum based measur\**.

Fueron seleccionados finalmente todos aquellos documentos que cumplieran con el quinto perfil de búsqueda mencionado anteriormente, siendo el número total de archivos obtenidos 980.

Una vez realizado ese primer paso, se utilizó la herramienta de trabajo “Rayyan” (Intelligent Systematic Review) para volcar la información obtenida a través de las búsquedas y continuar con la investigación. Así, de los 980 documentos recogidos inicialmente, y tras llevar a cabo un proceso de detección y eliminación de archivos duplicados, se obtuvo un total de 658 documentos.

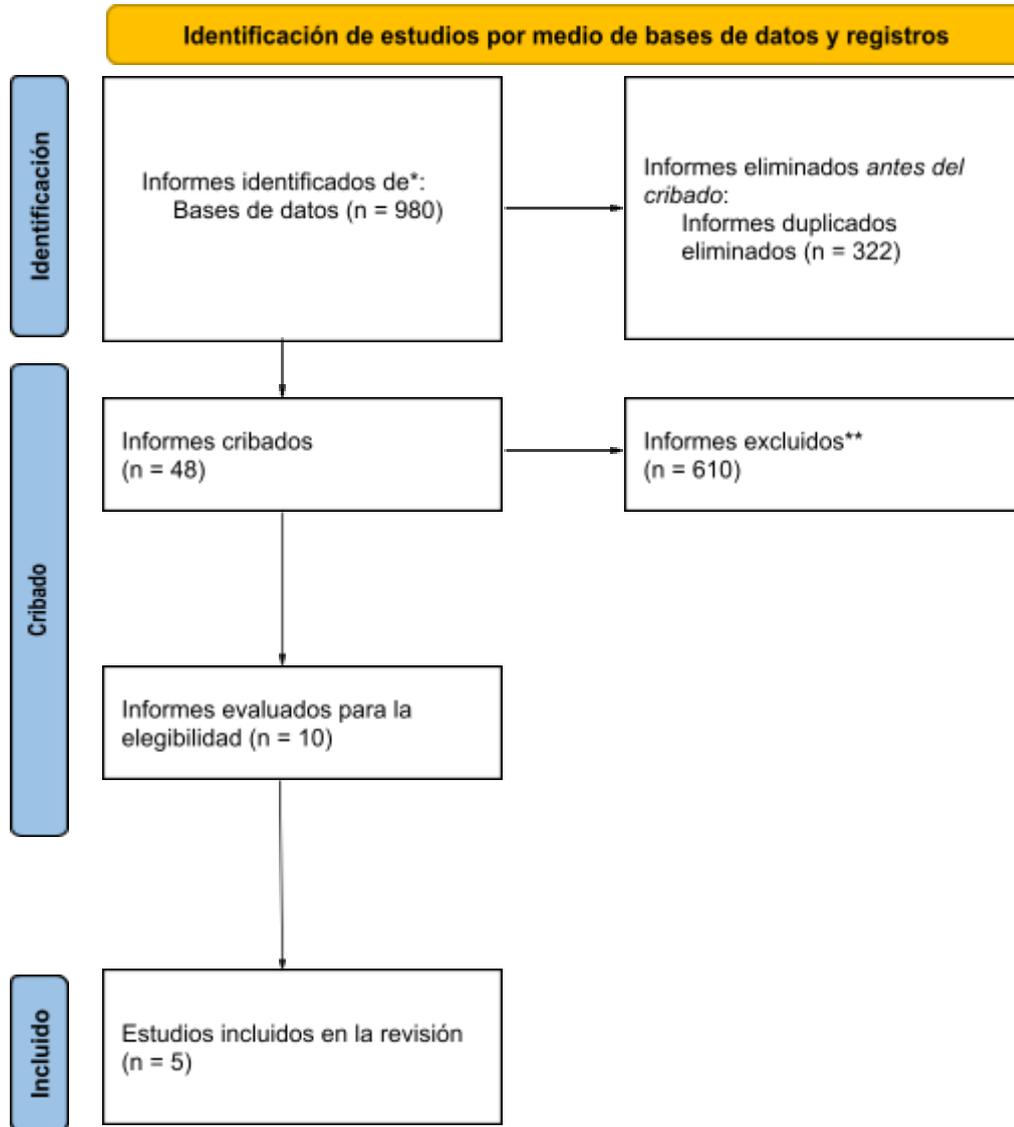
A continuación, comenzó un proceso de refinado de la búsqueda hecho de manera independiente, para luego calcular la fiabilidad interobservador, donde cada una de las autoras de este trabajo debió, a través de la lectura del título y el resumen de cada uno de los artículos previamente mencionados, decidir si era conveniente utilizarlo para el desarrollo de la revisión sistemática. Los criterios de inclusión utilizados fueron: 1) documentos escritos en inglés, 2) evaluaciones en el área de matemáticas, 3) evaluación de MBC en niños/as de preescolar, 1º y 2º de

Educación Primaria. Asimismo, los criterios de exclusión fueron los siguientes: 1) documentos no relacionados con las MBC en el área de matemáticas, 2) medidas recogidas en muestras de niños/as de 3º de Primaria o nivel educativo superior, 3) documentos que recogían información de MBC en otras áreas además de en matemáticas (lectura, escritura) y 4) documentos que no trataban sobre MBC.

Después de realizar cada búsqueda de manera individual, se procedió a la discusión conjunta de las decisiones tomadas sobre la inclusión y exclusión de los distintos documentos. Finalmente, y tras lograr un 100% de acuerdo, se decidió incluir un total de 48 documentos. Se adjunta a continuación, en la Figura 1, el diagrama de flujo que refleja el proceso de búsqueda utilizado.

**Figura 1**

Diagrama de flujo PRISMA 2020



Más adelante, para refinar aún más la búsqueda, se procedió a un nuevo proceso de exclusión conjunto, esta vez más exhaustivo. De este modo, los criterios de exclusión empleados en esta fase fueron: 1) evaluación de MBC en niños/as que no fueran exclusivamente de preescolar, 2) investigaciones realizadas a población no occidental, 3) documentos que incluyeran evaluaciones diferentes a las MBC y 4) investigaciones con una muestra inferior a 100 participantes, y así se seleccionaron y admitieron un total de 10 documentos.

Finalmente, una vez comenzada la recopilación de información para la redacción y análisis correspondientes a la revisión sistemática, y tras la lectura completa de los artículos evaluados para la elegibilidad, las autoras descartaron 5

de los 10 documentos seleccionados, pues detectaron que se habían incluido erróneamente archivos con una muestra inferior a la estipulada como criterio de exclusión, así como archivos donde la muestra no cumplía con el nivel educativo que se había establecido como límite.

En la Tabla 1 se puede observar más detalladamente la información perteneciente a los documentos que se utilizaron para la realización de este trabajo.

**Tabla 1**

*Características de los artículos incluidos en la revisión sistemática*

Título	Autor/es y año de publicación	Tipo de documento	Base de datos	Muestra	Estación del año	Lugar	Características psicométricas	Tareas	Prueba criterio
Assessing the classification accuracy of early numeracy curriculum-based measures using receiver operating characteristic curve analysis	Laracy, Seth D., Hojnoski, Robin L., Dever, Bridget V. (2016)	Artículo	Ebscohost	140 estudiantes	Otoño, invierno y primavera	Medio Oeste de los Estados Unidos	E: 0,49-0,91 (279) y 0,46-0,89 (140) S: 0,44-0,79 (279) y 0,36-0,86 (140) F(CC): 0,89-0,94 F(CO): 0,83-0,90 F(C11): 0,62-0,96 F(IN): 0,88-0,91 F(DC): 0,99 VC: - SC: -	CC CO C11 DN	-
Early Numeracy Assessment: The Development of the Preschool Early Numeracy Scales	Purpura, David J., Lonigan, Christopher J. (2015)	Artículo	Ebscohost	393 estudiantes	Primavera	Estados Unidos	-	CV C11 Ca CS CCo S CN ON ConjN PyC Cal	-

**Tabla 1** (continuación)

Título	Autor/es y año de publicación	Tipo de documento	Base de datos	Muestra	Estación del año	Lugar	Características psicométricas	Tareas	Prueba criterio
Further development of measures of early math performance for preschoolers	VanDerHeyden, Amanda M., Broussard, Carmen, Cooley, Amanda (2006)	Artículo	Ebscohost	127 estudiantes	Invierno y primavera	Sur de Luisiana	E: 0,88 (70%), 1 (62%), 1 (57%) S: 0,52 (70%), 0,22 (62%), 0,13 (57%) F: - VC: - SC: -	IN CVA COB DV	Brigance Screens
Investigating the incremental validity of cognitive variables in early mathematics screening	Clarke, Ben, Shanley, Lina, Kosty, Derek; Baker, Scott K., Cary, Mari Strand, Fien, Hank, Smolkowski, Keith. (2018)	Artículo	Ebscohost	458 estudiantes	Otoño e invierno	Noroeste del Pacífico	E: 70% (2007) y 66% (2010) S: 69% (2007) y 90% (2010) F: 0.78 - 0.99 VC: 0.46 - 0.72 SC: -	CO IN DC INuF TEMA-3 WASI Digit Span	TEMA-3

**Tabla 1** (continuación)

Título	Autor/es y año de publicación	Tipo de documento	Base de datos	Muestra	Estación del año	Lugar	Características psicométricas	Tareas	Prueba criterio
Using Curriculum-Based Measurement to Monitor Kindergarteners' Mathematics Development	Seethaler, Pamela M., Fuchs, Lynn S. (2011)	Artículo	Ebscohost	180 estudiantes	Otoño y primavera	Estados Unidos	E: 63,6% S: 90% F: 0,80 VC: - SC: -	C FA Sus TEMA-3	-

NOTA:

E = Especificidad, S = Sensibilidad, F = Fiabilidad, VC = Validez de Constructo, SC = Sensibilidad al Cambio.

CC = Comparación de Cantidades, CO = Conteo Oral, C11 = Conteo Correspondencia uno a uno, IN = Identificación Numérica, DC = Discriminación de Cantidad, DN = Denominación de Números, CV = Conteo Verbal, Ca = Cardinalidad, CS = Conteo de Subconjuntos, S = Subitización, CCo = Comparación de Conjuntos, CN = Comparación Numérica, ON = Orden de Números, ConjN = Conjunto a Números, PyC = Problemas Escritos y Combinaciones de números, Cal = Cálculo, CVA = Conteo en Voz Alta, COb = Contar Objetos, DV = Discriminación Visual, INuF = Identificación del Número que Falta, C = Conteo, FA = Fluidez en la adición, Sus = Sustracción.

## Resultados

Cinco documentos fueron finalmente seleccionados a través del proceso de búsqueda para esta revisión, siendo todos artículos publicados entre 2006 y 2018.

Todos ellos utilizaron una muestra de entre 127 y 458 niños/as en edad preescolar (de 3 a 5 años), que fueron evaluados en varias ocasiones a lo largo del año (otoño, invierno y primavera; invierno y primavera; otoño e invierno; otoño y primavera), excepto un único artículo que solo realizó una evaluación en primavera. Además, todas las investigaciones fueron realizadas en América del Norte, y más concretamente en Estados Unidos y el Noroeste del Pacífico (donde se incluyen regiones de Estados Unidos y de Canadá). Las muestras empleadas para los diferentes estudios incluyeron niños/as de diversos niveles socioeconómicos, así como de distintas nacionalidades (afroamericanos, caucásicos, latinos, asiáticos u otras etnias).

En cuanto a las tareas matemáticas pertenecientes a las MBC, estas se pueden observar en la Tabla 1.

Algo importante a destacar es que en el cuarto estudio se realizó otras evaluaciones como TEMA-3 (Test of Early Mathematics Ability), WASI (“Weschler Abbreviated Scales of Intelligence”) y “Digit Span Forward and Backward measures”. Así pues, las tareas evaluadas en TEMA-3 fueron el conteo, hechos numéricos, cálculo y conceptos matemáticos. En la prueba WASI se evaluó el procesamiento de la información y el razonamiento abstracto. En la prueba de “Digit Span” se analizó la memoria de trabajo. Además, en el quinto estudio también se llevaron a cabo otras evaluaciones como “Test of Early Mathematics Ability–3rd Ed.” (TEMA-3) que mide habilidades numéricas, facilidad en la comparación numérica, hechos numéricos, habilidades de cálculo, comprensión de conceptos.

A continuación, se detallan los datos estadísticos más relevantes de los diferentes estudios que han sido seleccionados, principalmente las variables de especificidad, sensibilidad, fiabilidad, validez del constructo y sensibilidad al cambio. Por eso, antes de presentar los resultados obtenidos en cada estudio, es importante definir en qué se basan estas variables.

La especificidad se refiere a la probabilidad de que los resultados de una prueba determinada salgan negativos si realmente lo son. Así, a medida que aumenta la especificidad en una prueba, disminuirá la cantidad de personas que sean negativas, pero que en los resultados obtenidos han salido como positivas. A

estos casos también se les puede denominar falsos positivos, que como se acaba de explicar, ocurren cuando una persona puntúa positiva en algo, en este caso en concreto en DA, cuando realmente es negativa.

Por otro lado, la sensibilidad hace referencia a la probabilidad que existe de que los resultados de una prueba determinada salgan positivos si realmente lo son. Por ello, a medida que aumenta la sensibilidad en una prueba, disminuirá la cantidad de personas que sean positivas, pero que en los resultados obtenidos han salido como negativas. A esto también se le puede denominar como falsos negativos, que serían aquellos casos en los que una persona aparece como negativa, en este caso en concreto en DA, cuando realmente es positiva.

Asimismo, cuando se habla de la fiabilidad de una prueba, se hace referencia a la precisión o el grado con el que se mide un rasgo psicológico determinado, independientemente de si es capaz o no. Mientras que la validez del constructo se trata de la adecuación de las inferencias que se hacen sobre la base de mediciones u observaciones.

Por último, en lo que respecta a la sensibilidad al cambio, se trata de una característica psicométrica que representa el grado con el que se obtienen distintos resultados en aplicaciones repetidas del mismo instrumento, en el momento en el que se ha producido un cambio real en el estado de salud. Por ello, también es conocida y definida como aquella capacidad que tiene un instrumento para poder detectar un cambio.

Por lo tanto, una vez explicado en qué consiste cada una de estas variables y las distinciones que hay entre ellas, se procederá a especificar los valores que se han obtenido. Los resultados, que también aparecen reflejados en la Tabla 1, fueron los siguientes:

- **Estudio 1:** Se ha encontrado que, para una muestra de 279 estudiantes, se ha obtenido una especificidad de entre 0,49 y 0,91, y una sensibilidad de entre 0,44 y 0,79. Mientras que, para una muestra de 140 estudiantes, la especificidad obtenida ha sido de entre 0,46 y 0,89, y una sensibilidad de entre 0,36 y 0,83. Además, la fiabilidad en este estudio fue determinada por tareas. Así, las medidas fueron: de 0.89 a 0.94 en comparación de cantidades, de 0.83 a 0.90 en conteo oral, de 0.62 a 0.96 en conteo de correspondencia uno a uno, de 0.88 a 0.91 en identificación

numérica y de 0.99 en discriminación de cantidad. No se encontraron datos sobre la validez de constructo ni sobre la sensibilidad al cambio.

- **Estudio 2:** No se encontró que se mencionara lo que son los términos de especificidad y sensibilidad, pero si es cierto, que se pudieron observar otros datos que también pueden ser relevantes. Tratándose de la fiabilidad en problemas y en cálculo. Por lo que se obtuvo una fiabilidad de 0,85 en problemas, y de 0,92 en cálculo.
- **Estudio 3:** En el 70% de los casos, se ha encontrado una especificidad de 0,88 y una sensibilidad de 0,52. Mientras que, en el 62% de los casos de este estudio, se obtuvo una especificidad de 1 y una sensibilidad de 0,22. Y, por otro lado, en el 57% de los casos, se mostró una especificidad de 1 y una sensibilidad de 0,13. No se encontraron valores sobre fiabilidad, validez de constructo o sensibilidad al cambio.
- **Estudio 4:** Fuchs y colegas (2007) lograron conseguir una especificidad del 70% y una sensibilidad del 69%. Sin embargo, Seethaler y Fuchs (2010), consiguieron una especificidad del 66% y una sensibilidad del 90%. A su vez, los valores de fiabilidad se encuentran en un intervalo de entre 0.78 a 0.99; y la validez de constructo tiene unos valores entre 0.46 y 0.72. No se obtuvieron resultados sobre la sensibilidad al cambio.
- **Estudio 5:** Se obtuvo una especificidad del 63,6%, lo cual equivale a la obtención de 24 falsos positivos, mientras que se obtuvo un 90% de sensibilidad, contando de este modo con 2 falsos negativos. Por otro lado, la fiabilidad encontrada fue de 0.80. No se obtuvieron resultados sobre validez de constructo o sensibilidad al cambio.

Una vez extraídos dichos datos, puede comprobarse como los resultados reflejan que efectivamente la MBC puede considerarse una buena herramienta de evaluación para los niños/as con DEAM.

Sin embargo, en general, se ha obtenido que en la mayoría de los estudios seleccionados el valor de la especificidad es mayor que el de la sensibilidad, por lo que se han registrado una mayor proporción de falsos negativos que de falsos positivos; es decir, que se han evaluado ciertos casos como “no riesgo” cuando en realidad sí que lo son. No obstante, esto no es algo alarmante pues es lo habitual en este tipo de estudios, más teniendo en cuenta que solo se han analizado 5 investigaciones, y los valores de ambas propiedades psicométricas siguen siendo

apropiados. Además, entre los pocos resultados que se han podido obtener en lo que respecta a la fiabilidad de estos estudios, se ha observado también como su valor por norma general es elevado. Mientras que, solo en un estudio se ha obtenido un valor de validez del constructo que no es favorable.

Al mismo tiempo, dentro de cada estudio hay matices diferentes que se procede a detallar.

En lo que respecta a lo comentado en el párrafo anterior, se encuentra por ejemplo que en el primer estudio (*“Assessing the classification accuracy of early numeracy curriculum-based measures using receiver operating characteristic curve analysis”*), hay dos intervalos tanto de especificidad como de sensibilidad dependiendo de si se trataba de una muestra u otra. Es por ello por lo que, si se trata de la especificidad, en la muestra mayor se observa que el intervalo es mucho más grande que en la muestra menor. Sin embargo, en cuanto a la sensibilidad pasa todo lo contrario, el intervalo de la muestra con menor número de participantes es mayor, pero empezando desde un nivel menor. Asimismo, en cuanto a los datos obtenidos en la fiabilidad de cada una de las tareas, se observa como en todas, exceptuando en C11, el intervalo empieza siempre en un valor superior a 0,80, lo que estaría indicando la calidad de dichas tareas. De igual modo, a pesar de en C11 empezar el intervalo con un valor inferior a 0,80, consigue alcanzarlo y superarlo.

En cuanto al segundo estudio (*“Early Numeracy Assessment: The Development of the Preschool Early Numeracy Scales”*), los autores no mencionaron en ningún momento la especificidad y la sensibilidad, pero sí el porcentaje de fiabilidad obtenido por los alumnos en las pruebas de cálculo y problemas, observándose, como era de esperar, que hay una mayor fiabilidad en cálculo que en problemas. Esto puede ser debido a que normalmente el alumnado sabe realizar las operaciones correctamente (sumar, restar, etc.), por lo que si le pides directamente que haga una operación de este tipo lo sabrá resolver. Pero, por el contrario, al pedirles que razonen y que intenten comprender qué es lo que se les está pidiendo en un problema, suelen tender a fallar, por la falta de comprensión y no porque no sepan realizar las operaciones, lo cual está vinculado a la mala comprensión lectora del alumnado.

Por otro lado, en el tercer estudio (*“Further development of measures of early math performance for preschoolers”*), se puede ver como varían los resultados dependiendo del porcentaje de casos que se tenga. Pero, si es cierto que,

independientemente del porcentaje de casos que haya, la especificidad siempre va a ser mayor que la sensibilidad. Además, se puede constatar cómo a medida que aumenta la especificidad, disminuye la sensibilidad.

Asimismo, en el cuarto estudio seleccionado (*“Investigating the incremental validity of cognitive variables in early mathematics screening”*), ocurre algo bastante curioso, y es que, al haber llevado a cabo la misma investigación en años diferentes (2007 y 2010), y siendo investigada por las mismas personas, los resultados obtenidos fueron completamente al revés. Es por ello por lo que en 2007 la especificidad era más alta que la sensibilidad, pero en 2010 la sensibilidad era más alta que la especificidad. Igualmente, es importante aclarar que realmente el porcentaje de especificidad no aminoró mucho en esos tres años, solamente un 4% por lo que, a pesar de haber sido más alta en 2007 que la sensibilidad, no lo era por mucho más. Lo que sí es cierto es que la sensibilidad aumentó en esos tres años un 21%, que es bastante. Esto hace que haya cambiado el enfoque de dicha investigación al haberla realizado por segunda vez. Por otro lado, en cuanto a la fiabilidad obtenida, se puede observar que el valor con el que comienza el intervalo está casi rozando el 0,80 por tan solo un 0,02 de diferencia, pero que de igual modo lograr superar el 0,80, así que se podría afirmar que es fiable. No obstante, en este estudio se ha podido observar la validez del constructo, se ha obtenido un intervalo con resultados bastante bajos, el valor más alto no lograr llegar al 0,80, por lo que en este caso no sería un resultado de buena calidad.

Por último, en cuanto al quinto estudio (*“Using Curriculum-Based Measurement to Monitor Kindergarteners’ Mathematics Development”*), se observa como es el único que desde el principio se encontró que los parámetros de sensibilidad son más altos que los de especificidad, existiendo una diferencia de un 26,4% entre ambos. Además, en cuanto a la fiabilidad, se ha encontrado justamente un valor de 0,80, por lo que se trata de un buen predictor de calidad y eficaz.

## **Discusión**

Las MBC, creadas a través del Instituto de Discapacidades de Aprendizaje de la Universidad de Minnesota en un estudio encabezado por Deno y Mirkin que tuvo lugar en los años 1970, y con el objetivo de generar una herramienta que permitiese

a los maestros y las maestras evaluar al alumnado basándose en su progreso académico, constituyen una de los instrumentos clave para la detección y evaluación de niños/as que necesitan algún tipo de intervención o ayuda más allá de la ya proporcionada en general por la escuela. Las MBC se enmarcan dentro del modelo Rtl, que ha tomado relevancia en los últimos años dados sus beneficios y su buena capacidad para la detección e intervención temprana para el alumnado con riesgo de presentar dificultades. Por esta razón, se planteó como objetivo de esta revisión sistemática analizar la efectividad de las MBC para la detección temprana de dificultades en el área de matemática temprana.

Teniendo en cuenta los resultados obtenidos tras revisar los diferentes documentos, puede asumirse que, en líneas generales, la CBM constituye una medida efectiva para la detección de los niños/as con DEAM.

En general, las MBC demuestran ser fiables y válidas, con buenas características psicométricas, y aptas para la detección de DEA en edad preescolar. Con esto se observa que es posible llevar a cabo un correcto cribado en las aulas de las escuelas, para detectar las necesidades del alumnado y así poder desarrollar intervenciones adaptadas a las mismas.

Además, las tareas más repetidas en los diferentes documentos analizados fueron conteo oral, conteo de correspondencia uno a uno e identificación numérica. Si bien es cierto que solo se han tenido en cuenta cinco artículos, por lo que sería necesario analizar un mayor número para poder hacer generalizaciones más certeras, sí que se puede entender que posiblemente estas sean las tareas más fiables para la evaluación del sentido numérico.

Por otro lado, cabe señalar algunas limitaciones que se han tenido a la hora de realizar esta revisión. En primer lugar, todos los estudios seleccionados fueron realizados con una muestra de niños/as residentes en Estados Unidos, por lo que no se pueden generalizar los resultados a otros países, pues hay que tener en cuenta la influencia del contexto sociocultural en el que se encuentra la población. Además, en algunos de los estudios las pruebas fueron pasadas por el propio profesorado, que puede no contar con la experiencia y/o el conocimiento preciso para la realización de este tipo de procedimiento, lo cual puede haber introducido errores en los resultados. A su vez, no pudieron detectarse datos sobre la sensibilidad al cambio en la mayor parte de los estudios analizados, y en los casos que sí se pudo, este valor era bajo. Finalmente, también cabe destacar que en

algunos de los estudios la muestra con la que se realizó la investigación contaba con un nivel socioeconómico muy bajo (en uno de los estudios incluso se menciona que debían incentivar a los niños/as comprándoles comida), por lo que parte de los resultados también podrían estar sesgados por este factor. Por ello, para próximas evaluaciones, se recomienda analizar un mayor número de investigaciones, contar con estudios con una población de diversos países, estudiar los resultados contando con la influencia del factor socioeconómico y cultural, así como revisar estudios que cuenten con datos sobre la sensibilidad al cambio u otros estadísticos relevantes.

### **Conclusiones**

Una vez finalizado el proceso de revisión sistemática llevado a cabo para la redacción de este manuscrito, se han obtenido diversas conclusiones.

Por un lado, se puede afirmar que las MBC constituyen una herramienta eficaz para la detección de niños/as con riesgo de padecer dificultades de aprendizaje en el área de matemáticas. Más específicamente, dado que esta revisión se ha hecho sobre MBC dirigidas a alumnado de entre 3 y 5 años, se puede afirmar que dichas medidas son adecuadas para evaluar la matemática temprana.

A su vez, tal y como se ha comentado, la detección temprana de las DEAM es de suma importancia, pues es lo que permitirá que el niño/a reciba una adecuada intervención adaptada a sus necesidades, y que así logre un desarrollo y progreso académico adecuado. Las MBC proporcionan esa posibilidad a las escuelas gracias a su formato de aplicación preventiva.

Sin embargo, también se han encontrado ciertas limitaciones por el reducido número de estudios evaluados, y por las características de los mismos, que hacen que se recomiende para futuras revisiones sistemáticas tomar en consideración los siguientes aspectos: evaluar un número más amplio de estudios, y que estos incluyan población de diversos países y con diferentes condiciones socioeconómicas y culturales, pues evitará posibles sesgos y dotará de mayor fiabilidad y validez a los resultados que se obtengan.

De cualquier manera, es más que notable la efectividad con la que cuentan las medidas basadas en el currículo como herramienta de cribado, como queda

reflejado en el análisis de sus propiedades psicométricas, donde la mayoría confirman intervalos con valores de especificidad y sensibilidad que oscilan entre 0,45 y 0,90, y que, si bien no son excesivamente elevados, sí que se encuentran en un rango aceptable.

Asimismo, se ha encontrado que, por norma general, los índices de especificidad son más altos, que los índices de sensibilidad. Esto indica que, generalmente, dentro del margen de error que se pueda cometer (como en cualquier estudio), se suelen encontrar más casos de falsos negativos que de falsos positivos.

### Referencias

- Clarke, B., Shanley, L., Kosty, D., Baker, S. K., Cary, M. S., Fien, H., y Smolkowski, K. (2018). Investigating the Incremental Validity of Cognitive Variables in Early Mathematics Screening. *School Psychology Quarterly*, 33(2), 264–271.  
<https://doi.org/10.1037/spq0000214>
- Deno, S. L. (2003). Developments in curriculum-based measurement. *The journal of special education*, 37(3), 184–192.  
<https://doi.org/10.1177%2F00224669030370030801>
- Duncan, G., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P. K., Pagani, L., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, H. y Japel, C. (2007). School Readiness and Later Achievement. *Developmental Psychology*, 43(6), 1428–1446.  
<http://dx.doi.org/10.1037/0012-1649.43.6.1428>

- Gersten, R., Clarke, B., Jordan, N. C., Newman-Gonchar, R., Haymond, K. y Wilkins, C. (2012). Universal Screening in Mathematics for the Primary Grades: Beginnings of a Research Base. *Exceptional Children*, 78(4), 423–445. <https://doi.org/10.1177%2F001440291207800403>
- Jiménez, J. E. (2019). Modelo de respuesta a la intervención y matemáticas: principales habilidades y detección temprana. *Modelo de respuesta a la intervención: Un enfoque preventivo para el abordaje de las dificultades específicas de aprendizaje (249–289)*. Ediciones Pirámide.
- Koponen, T. K., Aunola, K., Ahonen, T. y Nurmi, J. (2007). Cognitive predictors of single-digit and procedural calculation and their covariation with reading skill. *Journal of Experimental Child Psychology*, 97(3), 220–241. <http://dx.doi.org/10.1016/j.jecp.2007.03.001>
- Laracy, S. D., Hojnoski, R. L., y Dever, B. V. (2016). Assessing the Classification Accuracy of Early Numeracy Curriculum-Based Measures Using Receiver Operating Characteristic Curve Analysis. *Assessment for Effective Intervention*, 41(3), 172–183. DOI: 10.1177/1534508415621542
- Mateos, R. (2009). Dificultades de aprendizaje. *Psicología Educativa*, 15(1), 13–19.
- Polignano, J. C., y Hojnoski, R. L. (2012). Preliminary Evidence of the Technical Adequacy of Additional Curriculum-Based Measures for Preschool

Mathematics. *Assessment for Effective Intervention*, 37(2), 70–83.

<https://doi.org/10.1177%2F1534508411430323>

Purpura, D. J., y Lonigan, C. J. (2015). Early Numeracy Assessment: The Development of the Preschool Early Numeracy Scales. *Early Education and Development*, 26(2), 286–313.

<http://dx.doi.org/10.1080/10409289.2015.991084>

Seethaler, P. M., y Fuchs, L. S. (2011). Using Curriculum-Based Measurement to Monitor Kindergarteners' Mathematics Development. *Assessment for Effective Intervention*, 36(4), 219–229. DOI: 10.1177/1534508411413566

VanDerHeyden, A. M., Broussard, C., y Cooley, A. (2006). Further Development of Measures of Early Math Performance for Preschoolers. *Journal of School Psychology*, 44(6), 533–553. <http://dx.doi.org/10.1016/j.jsp.2006.07.003>

## Anexos

### Anexo

#### *Búsquedas realizadas*

búsquedas	1ª B		2ª B		3ª B		4ª B		5ªB		6ªB		7ªB		8ªB		9ªB	
	docs	desde																
WOS	50	1997	184	1984	218	1984	223	1984	268	1984	61	1997	225	1984	28	2006	38	1997
SCOPUS	54	1987	106	1987	124	1987	129	1987	147	1987	57	1987	121	1987	16	2006	28	1997
EBSCOH OST	94	2004	344	1974	368	1974	397	1974	477	1974	114	1997	410	1974	39	2006	105	1997
ERIC	18	2004	58	1994	61	1990	73	1990	88	1990	22	2004	69	1994	19	2006	19	2006

*NOTA:* Descriptores utilizados en cada una de las búsquedas.

1ªB: (screen\* OR benchmark\* OR "diagnostic accuracy" OR "classification accuracy" OR "psychometric characteristics" OR utilit\* OR "technical adequacy") AND ("number sense" OR "value position" OR "magnitude comparison" OR "numerical comparison" OR "counting" OR "quantity estimate" OR "numerical sequences" OR "numerical identification" OR calculation) AND (cbm OR "curriculum based measures").

2ªB: (screen\* OR benchmark\* OR "diagnostic accuracy" OR "classification accuracy" OR "psychometric characteristics" OR utilit\* OR "technical adequacy" OR roc) AND (math\* OR "number sense" OR "value position" OR "magnitude comparison" OR "numerical

comparison" OR "counting" OR "quantity estimate" OR "numerical sequences" OR "numerical identification" OR calculation) AND (cbm OR "curriculum based measures").

3<sup>a</sup>B: (screen\* OR benchmark\* OR "diagnostic accuracy" OR "classification accuracy" OR "psychometric characteristics" OR utilit\* OR "technical adequacy" OR roc OR validation OR "criterion test") AND (math\* OR "number sense" OR "value position" OR "magnitude comparison" OR "numerical comparison" OR "counting" OR "quantity estimate" OR "numerical sequences" OR "numerical identification" OR calculation) AND (cbm OR "curriculum based measures").

4<sup>a</sup>B: (screen\* OR benchmark\* OR "diagnostic accuracy" OR "classification accuracy" OR "psychometric characteristics" OR utilit\* OR "technical adequacy" OR roc OR validation OR "criterion test" OR "predictive validity" OR "construct validity" OR "content validity" OR "concurrent validity") AND (math\* OR "number sense" OR "value position" OR "magnitude comparison" OR "numerical comparison" OR "counting" OR "quantity estimate" OR "numerical sequences" OR "numerical identification" OR calculation) AND (cbm OR "curriculum based measures").

5<sup>a</sup>B: (screen\* OR benchmark\* OR "diagnostic accuracy" OR "classification accuracy" OR "psychometric characteristics" OR utilit\* OR "technical adequacy" OR roc OR validation OR "criterion test" OR "predictive validity" OR "construct validity" OR "content validity" OR "concurrent validity") AND (math\* OR "number sense" OR "value position" OR "magnitude comparison" OR "numerical comparison" OR "counting" OR "quantity estimate" OR "numerical sequences" OR "numerical identification" OR calculation) AND (cbm OR "curriculum based measur\*").

6<sup>a</sup>B: (screen\* OR benchmark\* OR "diagnostic accuracy" OR "classification accuracy" OR "psychometric characteristics" OR utilit\* OR "technical adequacy") AND ("number sense" OR "value position" OR "magnitude comparison" OR "numerical comparison" OR "counting" OR "quantity estimate" OR "numerical sequences" OR "numerical identification" OR calculation) AND (cbm OR "curriculum based measur\*").

7<sup>a</sup>B: (screen\* OR benchmark\* OR "diagnostic accuracy" OR "classification accuracy" OR "psychometric characteristics" OR utilit\* OR "technical adequacy" OR roc) AND (math\* OR "number sense" OR "value position" OR "magnitude comparison" OR "numerical comparison" OR "counting" OR "quantity estimate" OR "numerical

sequences" OR "numerical identification" OR calculation) AND (cbm OR "curriculum based measur\*").

8<sup>a</sup>B: (screen\* OR benchmark\* OR "diagnostic accuracy" OR "classification accuracy" OR "psychometric characteristics" OR utilit\* OR "technical adequacy" OR roc OR validation OR "criterion test" OR "predictive validity" OR "construct validity" OR "content validity" OR "concurrent validity") AND (math\* OR "number sense" OR "value position" OR "magnitude comparison" OR "numerical comparison" OR "counting" OR "quantity estimate" OR "numerical sequences" OR "numerical identification" OR calculation) AND (cbm OR "curriculum based measur\*") AND (kindergarten OR kg).

9<sup>a</sup>B: (screen\* OR benchmark\* OR "diagnostic accuracy" OR "classification accuracy" OR "psychometric characteristics" OR utilit\* OR "technical adequacy" OR roc OR validation OR "criterion test" OR "predictive validity" OR "construct validity" OR "content validity" OR "concurrent validity") AND (math\* OR "number sense" OR "value position" OR "magnitude comparison" OR "numerical comparison" OR "counting" OR "quantity estimate" OR "numerical sequences" OR "numerical identification" OR calculation) AND (cbm OR "curriculum based measur\*") AND (kindergarten OR kg OR "first-grade" OR "second-grade").