

# Construcción de una lista de cotejo (*checklist*) de dificultades de aprendizaje del cálculo aritmético

por Antonio CORONADO-HIJÓN  
Universidad de Sevilla

## 1. Introducción

La competencia matemática, aun siendo una de las habilidades y conocimientos más necesarios en la sociedad actual, es a la vez una de las más difíciles de adquirir según nos indican evaluaciones internacionales al respecto (Mullis, Martin, Foy y Arora, 2012; OECD, 2013).

El cálculo es un componente esencial en la resolución de problemas aritméticos, y éste es uno de los contenidos más importantes de las Matemáticas, junto a la geometría, la medida o la probabilidad. Por este motivo, un gran porcentaje de las dificultades de aprendizaje de las matemáticas tiene que ver con la aritmética en la que el cálculo representa un papel esencial (Orrantía, 2000).

Los estudios sobre el tipo de trastornos que muestra el alumnado con dificultades en el aprendizaje del cálculo, reflejan dos tipos de déficits funcionales básicos: déficits procedimentales y déficits en la recuperación de hechos (Geary, 1990, 1993; Geary, Brown y Samaranayake, 1991;

Goldman, Pellegrino y Mertz, 1988; Kirby y Becker, 1988; Orrantía, 2000). Los déficits procedimentales tienen que ver con procedimientos aritméticos (estrategias de conteo) evolutivamente inmaduros, frecuentes errores y menor velocidad de conteo verbal. Los déficits en la recuperación de hechos conllevan una inadecuada representación de hechos aritméticos en la memoria, así como una importante proporción de fallos y tiempos de respuesta en la recuperación.

Actualmente, en las evaluaciones de la competencia matemática aritmética, del alumnado de los primeros niveles de escolarización, es más habitual la identificación de dificultades relativas a la recuperación de hechos matemáticos, mientras que la relacionada con déficits procedimentales, suele quedar relegada e incluso ausente.

Esta escasa evaluación de los déficits procedimentales en el cálculo aritmético, tampoco queda resuelta desde el ámbito del diagnóstico, donde el protocolo clásico

consiste en identificar una discrepancia, de alrededor de dos desviaciones típicas, entre los resultados obtenidos por el sujeto en una prueba normalizada de rendimiento matemático y un test estandarizado de inteligencia.

Esta desatención hacia los errores específicos de las dificultades de aprendizaje del cálculo aritmético, tanto en el ámbito de la evaluación como en el del diagnóstico, tiene importantes efectos desde el punto de vista de la intervención educativa adaptada, donde se muestra imprescindible identificar los errores y problemas concretos de cada sujeto, conocimientos que constituyen el prerrequisito básico para el diseño curricular de adaptaciones individuales o de aula (González-Pienda y González-Pumariega, 1998; Blanco y Bermejo, 2008). Por ello, la opción actual más aconsejada (Ortiz, 2004), implica tipos de evaluación variados y que incluyan test estandarizados, pruebas criterios de diagnóstico, pruebas basadas en contenidos curriculares, análisis de los errores mediante métodos observacionales y listas de cotejo o «checklist» (Horowitz, 2011), en la evaluación de las dificultades de aprendizaje (DA).

En los Estados Unidos de América, la Ley de Educación para personas con Dificultades (*Individuals with Disabilities Education Improvement Act*, IDEA, 2004, por su sigla en inglés), da un giro radical a los enfoques más clásicos de identificación de alumnado con DA. De tal manera que si, con anterioridad a esta Ley, se prescribía a los profesionales la utilización del modelo de discrepancia entre el cociente intelectual (CI) y el rendimiento escolar

en la identificación, a partir de la promulgación de esta norma, se insta a implementar lo que han denominado como, «Respuesta a la Intervención» (*Response to Intervention*, RtI), un método alternativo basado más que en procedimientos psicométricos estandarizados, en la evaluación educativa como procedimiento fundamental en la detección de los errores del alumnado en riesgo de fracaso escolar.

En esta dirección, posteriormente, la Oficina de Educación Especial y Servicios de Rehabilitación del Departamento de Educación en Estados Unidos (*Office of Special Education and Rehabilitative Services*, OSERS, 2006) en la regulación federal que desarrolla la normativa legislada para la identificación de alumnado con DA, afirma que: no es exigible la utilización del criterio diagnóstico de discrepancia en la determinación del alumnado con DA y que son adecuados otros procedimientos alternativos de identificación centrados en las respuestas o resultados de aprendizaje, contrastados y validados en estudios e investigaciones.

En relación a nuestro ámbito más cercano, también se ha venido promoviendo desde instituciones educativas, que la identificación de las dificultades de aprendizaje se realicen sobre las tareas contextualizadas de aprendizaje, desde una perspectiva fundamentalmente psicoeducativa centrada en los procesos y estrategias (Romero y Lavigne, 2006).

El objetivo pretendido con este trabajo es la creación de un instrumento o rejilla de observación de dificultades de aprendizaje en el cálculo aritmético, que diseñado

con los requerimientos científicos de la observación sistematizada, permita identificar los principales errores y dificultades con unos niveles estadísticamente aceptables de validez y fiabilidad.

### 2. Método

Se seguirá una metodología observacional sistematizada, en la que con antelación se determinan las categorías a observar, las producciones que interesa registrar se definen previamente y se cuantifican en una medida final.

Anguera (1990), define la metodología observacional como «un procedimiento encaminado a articular una percepción deliberada de la realidad manifiesta con su adecuada interpretación captando su significado, de forma que mediante un registro objetivo, sistemático y específico de la conducta generada de forma espontánea en un determinado contexto, y una vez que se ha sometido a una adecuada codificación y análisis, nos proporcione resultados válidos dentro de un marco específico de conocimiento».

La propuesta de esta investigación se enmarca dentro de lo que podemos considerar ya una tradición consolidada de combinar la perspectiva metodológica cualitativa y la cuantitativa, en el desarrollo de la metodología observacional (Anguera, 2010; Bakeman y Gottman, 1986), y que define la observación sistemática como una forma particular de cuantificar la conducta.

El procedimiento lógico de la metodología observacional posibilita y aconseja la

utilización secuenciada de dos enfoques, que comienza con un escaneado de la realidad desde una perspectiva procedimental de carácter cualitativo, para a continuación, plasmarse en un tipo de registro o rejilla de observación mediante algún sistema de puntuación, desde una perspectiva cuantitativa (Anguera e Izquierdo, 2006).

La utilización de un instrumento de registro es una técnica de fácil manejo, no intromisiva y que se puede realizar sobre distintos materiales o producciones matemáticas con un valor claramente contextual, posibilitando el análisis de una importante cantidad de unidades de información relevante.

En numerosos estudios se ha comprobado la eficacia de esta complementariedad (Jonsson *et al.*, 2006).

### 3. Participantes

En la prueba piloto, para llevar a cabo la recogida de información se entrenó a dos observadores, elegidos mediante muestreo aleatorio simple de entre estudiantes del último curso de Pedagogía de la Universidad de Sevilla, para que valoraran las producciones matemáticas de una muestra de conveniencia de 17 alumnos que presentaban bajos resultados en matemáticas, escolarizados en el nivel de 1º de ESO en un IES de Sevilla situado en una zona de nivel socioeconómico medio.

El estudio experimental, se llevó a cabo sobre las producciones matemáticas de 94 sujetos. Para la recogida de información se entrenó a dos grupo de ocho jueces observadores, elegidos mediante

muestreo aleatorio simple de estudiantes del último curso de Pedagogía y Psicopedagogía de la Universidad de Sevilla, que valoraron respectivamente, siguiendo los 20 indicadores *ad hoc* del instrumento, de una parte, la ocurrencia de dificultades de aprendizaje del cálculo (DAC), en las producciones aritméticas de una prueba objetiva de cálculo realizada por 48 alumnos de entre doce y trece años, escolarizados en 1º de Educación Secundaria Obligatoria de un instituto de Sevilla y de la que se obtuvo por tanto, 960 respuestas (48x20). El otro grupo de observadores siguiendo los mismos 20 indicadores del instrumento, analizó las producciones en una prueba de matemáticas realizada por cada uno de los 46 alumnos de entre nueve y diez años, escolarizados en 4º de Educación Primaria en un colegio de Sevilla, de los que se obtuvieron 920 respuestas (46x20), haciendo un total de 94 sujetos y 1.880 respuestas observadas (20 ítems en cada sujeto). Ambos centros escolares están situados en una zona de nivel socioeconómico medio de Sevilla.

Para el tamaño de la muestra se tuvo en cuenta el criterio clásico propuesto por Kline (1986) que concibe como suficiente una muestra a partir de tres veces el número de ítems de un instrumento.

#### 4. Instrumento

Se ha utilizado a modo de lista de chequeo o verificación, una rejilla de observación de dificultades de aprendizaje en el cálculo aritmético.

Su funcionamiento consiste básicamente en dedicar atención a la ocurrencia de

los errores de cálculo definidos con antelación, registrando su aparición y frecuencia. Este tipo de registro es el más ampliamente utilizado en educación: el *sistema de categorías*.

Este instrumento integra cuatro subrejillas que a modo de categorías y referidos a las cuatro operaciones aritméticas básicas (S, R, M y D), encuadran un total de 20 indicadores de errores o dificultades en el aprendizaje de éstas.

Los destinatarios del uso de este instrumento de observación son: profesorado de matemáticas, pedagogos, psicopedagogos y psicólogos de la educación.

Las situaciones que se han de observar son aquellas operaciones aritméticas de cálculo realizadas por alumnado en cualquier prueba objetiva o prueba de rendimiento matemático estandarizada. La única condición es que el sujeto a evaluar realice las operaciones en un papel, haciendo referencia al algoritmo relacionado.

### 5. Etapas de Construcción de la *checklist*

#### *1ª Fase: Búsqueda bibliográfica*

La fuente de información utilizada para delimitar y validar inicialmente el contenido, se ha basado en una revisión de la literatura científica publicada sobre las dificultades que aparecen con más frecuencia en el aprendizaje en las operaciones de cálculo aritmético básico (Brown y Burton, 1978; Maza, 1995; Miranda, 1987; Miranda, Fortes y Gil, 2000; Orrantía, 2000).

### 2ª Fase: Proceso de elaboración de los indicadores

Seleccionados los indicadores o ítems del instrumento de observación, y previo a la prueba piloto, éstos fueron revisados por seis expertos. Cuatro de ellos, profesorado de Educación Secundaria con un mínimo de diez años de experiencia docente, licenciados en matemáticas y especialistas en esa disciplina. Dos maestras especialistas en pedagogía terapéutica (educación especial), con un mínimo de cinco años de experiencia docente. Ninguno de los expertos de revisión participó en la previa elección de los ítems de observación.

El objetivo de esta revisión es analizar la validez de contenido del instrumento solicitando la opinión de un grupo de expertos sobre el grado en que las manifestaciones relevantes están recogidas en el mismo (Wilson, 2005).

La relevancia se estimó en función del grado en que los descriptores de dificultades de aprendizaje enunciados, presentaban una serie de condiciones convenientes (Padilla, 2002), a saber: objetivos, claros, completos, mutuamente excluyentes, exhaustivos y homogéneos.

Las primeras tres condiciones (objetivos, claros y completos) se estimaron para cada uno de los indicadores recogidos en las cuatro categorías (S, R, M y D) mientras que las cuestiones referidas a la exclusión conceptual, exhaustividad y la homogeneidad se referían al total de los indicadores del instrumento.

Siguiendo las recomendaciones de Hambleton (1980), se calculó la media

de los valores asignados por cada uno de los expertos a cada indicador.

Una vez revisados los ítems en función de esta valoración de los expertos y de sus aportaciones, se eliminaron los indicadores que estaban por debajo del punto medio de la escala Likert.

Como resultado de esta fase, el instrumento quedó compuesto por 20 indicadores distribuidos en 4 subrejillas:

1. Suma, 4 indicadores.
2. Resta, 6 indicadores.
3. Multiplicación, 6 indicadores.
4. División, 4 indicadores.

### 3ª Fase: Prueba piloto

Para estimar el funcionamiento general del instrumento, y realizar un estudio preliminar de la fiabilidad o concordancia entre observadores, se realizó una prueba piloto de tipo cualitativa y cuantitativa (Wilson, 2005), en la que tras pasar un examen de matemáticas correspondiente al nivel de final de 6 de Primaria a una muestra intencionada de 17 alumnos de 1º de ESO, se pidió a dos jueces (estudiantes del último curso de Pedagogía) que, utilizaran la *checklist*, para observar y detectar las dificultades en el cálculo de las operaciones básicas que presentaba el alumnado referido (Coronado-Hijón, 2010).

A los datos cuantitativos respecto a la frecuencia de dificultades observadas en los sujetos, por cada uno dos jueces, se

aplicó el Coeficiente Kappa de Cohen (Cohen, 1960).

Fruto de este estudio preliminar, también se estimó conveniente incluir una

ejemplificación «tipo» y una descripción más detallada de cada una de las distintas dificultades que a modo de indicadores se señalan en el instrumento, como objetivo de observación (TABLA 1).

TABLA 1: *Ejemplificación tipo del indicador D4.*

Indicador error división	Ejemplo	Descripción
(D4) Errores al bajar las cifras	658   6 058 19 4	Al bajar la siguiente cifra del dividendo se equivoca y baja más de una cifra o no baja ninguna y da por finalizada la operación

#### 4ª Fase: Estudio de campo

Una vez asegurada la validez de contenido en la prueba piloto, se siguió como guía las siguientes etapas de construcción propuestas por autores como Hill y Hill (2002), para llevar a cabo en el estudio de campo, los análisis de validez de constructo y de fiabilidad en sus vertientes de consistencia interna del instrumento, así como de la consistencia o concordancia entre observadores.

La validez de constructo, está relacionado con la medida en que un instrumento mide el constructo que pretende medir y es considerada como prueba mayor que comprende las de criterio y de contenido, considerándose estas últimas, evidencia de la validez del constructo (Padilla, 2002).

Aunque hay varios procedimientos de objetivar esta validez, una de las más habituales es el Análisis Factorial con el cual se confirma la estructura inicialmente asignada a un instrumento de observación y dentro de éste, el Análisis Factorial

Exploratorio, procedimiento para generar teorías más que para confirmarlas, debido a que en las ciencias sociales es difícil determinar con exactitud el valor de las correlaciones con cada factor (Kline, 1994:11).

Respecto a la valoración de la consistencia o fiabilidad de las medidas se ha utilizado el Alfa de Cronbach, por ser el indicador más ampliamente utilizado para este tipo de análisis. En cuanto a la concordancia entre observadores, ésta se ha analizado mediante el índice Kappa de Fleiss (Fleiss, Cohen y Everitt, 1969).

#### Resultados

Para establecer los distintos cálculos estadísticos se utilizó el programa informático SPSS (Versión 18.0).

Se comenzó con el análisis de la validez de constructo y el análisis factorial, tomándose como valores las medias aritméticas de las valoraciones de los jueces, sobre cada indicador para cada sujeto.

Para el tratamiento de los datos, especialmente el del análisis del índice de concordancia entre observadores que solo permite valores nominales, se ha tenido en cuenta, tan solo, si el observador ha constatado (valor 2), o no (valor 1), la ocurrencia de un tipo de error de cálculo.

En el análisis factorial exploratorio y cuando las variables (indicadores) son muchas (unas 20, o incluso más), Nunnally (1978:418-419) recomienda el Análisis de Componentes Principales, para explicar el 100% de la varianza observada y, por ello, todas las comunalidades iniciales deben ser

iguales a la unidad (valor de la varianza de una variable en puntuaciones típicas).

En la TABLA 2 se muestran las *comunalidades* asignadas inicialmente a las variables, así como las comunalidades reproducidas por la solución factorial (*extracción*) para el total de valoraciones y observaciones realizadas por los dos grupos de observadores.

Se considera que existe una estructura factorial clara cuando los ítems que definen un factor tienen pesos de ,50 o más en este factor (Nunnally y Bernstein, 1994:535).

TABLA 2: *Método de extracción: Comunalidades.*

	Inicial	Extracción
S1	1,000	,581
S2	1,000	,752
S3	1,000	,680
S4	1,000	,773
R1	1,000	,604
R2	1,000	,585
R3	1,000	,752
R4	1,000	,767
R5	1,000	,761
R6	1,000	,438

	Inicial	Extracción
M1	1,000	,648
M2	1,000	,565
M3	1,000	,611
M4	1,000	,596
M5	1,000	,720
M6	1,000	,722
D1	1,000	,664
D2	1,000	,480
D3	1,000	,761
D4	1,000	,580

En el estudio de campo, las comunalidades aportaron valores por encima de ,50, salvo una que con una puntuación de ,480 se situó muy cercana a este dato, por lo que podemos considerar que todas las variables están bien representadas y de manera bastante equilibrada y homogénea en el espacio de los factores.

Para determinar la estructura factorial necesaria, usamos el método de Kaiser que utiliza el programa estadístico SPSS y señala los factores con autovalores mayores que 1. En la TABLA 3 vemos el porcentaje de varianza explicada de cada uno de los componentes que se han extraído.

TABLA 3: *Varianza total explicada.*

Componentes	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	3,239	16,196	16,196	3,239	16,196	16,196
2	1,957	9,785	25,981	1,957	9,785	25,981
3	1,497	7,487	33,468	1,497	7,487	33,468
4	1,448	7,238	40,706	1,448	7,238	40,706
5	1,361	6,807	47,513	1,361	6,807	47,513
6	1,299	6,495	54,008	1,299	6,495	54,008
7	1,178	5,891	59,899	1,178	5,891	59,899
8	1,059	5,294	65,194	1,059	5,294	65,194
9	,976	4,880	70,074			
10	,851	4,256	74,330			
11	,810	4,049	78,379			
12	,758	3,792	82,171			
13	,746	3,729	85,899			
14	,630	3,152	89,052			
15	,555	2,777	91,828			
16	,508	2,539	94,368			
17	,380	1,901	96,269			
18	,321	1,604	97,873			
19	,270	1,349	99,221			
20	,156	,779	100,000			

Para el análisis de estos datos podemos tomar como referencia el estudio de revisión realizado por Henson (2006) sobre 60 estudios de análisis factoriales, que mostró que la proporción *media* de varianza explicada por los factores era del 52%.

En nuestro caso y del total del fenómeno que estamos estudiando, los datos de varianza encontrados se sitúan por encima de esta media ya que, concretamente, el 65,194% de la varianza total es explicada por 8 componentes principales o ítems, correspondientes a dificultades de apren-

dizaje en el cálculo de la suma y la resta, distribuidas sus varianzas de tal manera o peso:

— El componente 1 de la suma: «contar para hallar la suma», explica el constructo en un 16,196%.

— El componente 2 de la suma: «colocar erróneamente las cantidades en la suma», en un 9,785%.

— El componente 3 de la suma: «errores en las llevadas de la suma», en un 7,487%.

— El componente 4 de la suma: «empezar las operaciones de suma por la izquierda», un 7,238%.

— El componente 5, (1 de la resta): «restar la cifra mayor de la menor sin tener en cuenta su posición (arriba o abajo)», un 6,807%

— El componente 6, (2 de la resta): «errores en las “llevadas” de la resta», tiene un porcentaje de explicación del 6,495%.

— El componente 7, (3 de la resta): «colocar erróneamente las cantidades en la resta, explica un 5,891%.

— El componente 8, (4 de la resta): «empezar las operaciones de resta por la izquierda», explica el constructo en un 5,294%.

Respecto al análisis de la consistencia interna o fiabilidad de las medidas se ha utilizado el Alfa de Cronbach, que es el in-

dicador más ampliamente utilizado para este tipo de análisis.

El resultado obtenido es de 0,890. Hill y Hill (2002) consideran los valores entre 0,80 y 0,90 como un nivel de fiabilidad bueno.

Para estimar el grado de concordancia entre las valoraciones de los observadores, realizado en el programa estadístico SPSS (Versión 18.0), necesitó de la implementación de una nueva sintaxis para el cálculo del índice Kappa de Fleiss (1969), debido a que el índice Kappa de Cohen (1960) que era el ofrecido por el programa estadístico, sólo nos permite comparar la fiabilidad entre dos observadores.

El índice Kappa de Fleiss nos ha mostrado para las observaciones o valoraciones realizadas por los dos grupos de jueces sobre las producciones de cada uno de los grupos de alumnado, un grado de concordancia que se sitúa en un nivel moderado, por encima de aceptable, según Landis y Koch (1977).

— Respuestas observadas en el grupo de alumnado de 4º de Primaria, 920 respuestas (20 ítems x 46 sujetos). Índice Kappa: 0,5363

— Respuestas observadas en alumnado de 1º de la E.S.O., 960 respuestas (20 ítems por 48 sujetos). Índice Kappa: 0,4429

### Discusión

Este estudio muestra las etapas en la elaboración y validación de una lista

de cotejo o verificación de dificultades de aprendizaje del cálculo aritmético, que diseñada con los requerimientos científicos de la observación sistematizada, permite identificar los principales errores y dificultades con unos niveles estadísticamente aceptables de validez y fiabilidad.

Esta investigación se enmarca dentro de la ya tradicional combinación de los enfoques metodológicos cualitativo y cuantitativo, en el desarrollo de la metodología observacional (Anguera, 2010; Bakeman y Gottman, 1986), conceptualizando la observación sistemática como una forma particular de cuantificar la conducta.

Del análisis de la validez de contenido, procedente de la prueba piloto, ha resultado que los indicadores que componen la rejilla de observación y diagnóstico pueden considerarse como una muestra relevante y representativa de dificultades de aprendizaje en el cálculo aritmético, presentando los descriptores de DA enunciados, las ventajas de ser: (1) Objetivos, (2) Claros y (3) Completos.

También se ha demostrado que en relación a la globalidad del instrumento y a las distintas categorías, el instrumento se muestra bastante: (1) Homogéneo, (2) Exhaustivo y con una buena (3) Exclusión conceptual entre los indicadores presentados (Padilla, 2002).

El Análisis Factorial Exploratorio, nos revela una estructura factorial clara en la que todas las variables están bien representadas y de manera bastante equilibrada y homogénea en el espacio de los factores.

Una lista de cotejo tal es un instrumento que permite evaluar aspectos que no están suficientemente cubiertos en este momento por otros instrumentos, por lo que puede complementar a otras pruebas criteriosales y/o normativas en la identificación de discalculia, con la ventaja de su fácil utilización y aplicación a tareas en contextos naturales, ejercicios de clase, exámenes, etc., por lo que su uso abarca no solo al personal especializado en diagnóstico psicopedagógico sino también, al docente de matemáticas. Estas cualidades hacen de este instrumento de observación un material idóneo para un tipo de diagnóstico ligado a la evaluación continua y por tanto, al alumnado en riesgo y a la prevención desde la evaluación a la respuesta del alumnado a la intervención educativa (RtI).

Asimismo, los datos del análisis factorial que muestran a los indicadores de dificultades de aprendizaje de suma y resta, como los que mejor explican el constructo de dificultades en el aprendizaje del cálculo (DAC) tienen importantes implicaciones educativas que apuntan a la conveniencia de intervenir primariamente en la reeducación de estas dificultades por su clara incidencia en las cuatro operaciones básicas del cálculo aritmético, lo cual abre una promisorio dirección de investigación que ha de ser contrastada con nuevos estudios.

El instrumento que presentamos en esta investigación está aún en desarrollo, y los resultados de este estudio nos confirman en la conveniencia de seguir profundizando en su perfeccionamiento.

**Dirección para la correspondencia:**

Antonio Coronado-Hijón, Departamento de Métodos de Investigación y Diagnóstico en Educación, Universidad de Sevilla. C/ Pirotecnia, s/n. 41013 Sevilla. Email: acoronado1@us.es.

Fecha de recepción de la versión definitiva de este artículo: 10. VI. 2014.

**Bibliografía**

ANGUERA, M. T. (1990) Metodología observacional, en ARNAY, J., ANGUERA, M. T. y GÓMEZ, J. (eds.), *Metodología de la investigación en Ciencias del Comportamiento* (Murcia, Universidad de Murcia).

ANGUERA, M. T. (2010) Posibilidades y relevancia de la observación sistemática por el profesional de la Psicología, *Papeles del Psicólogo*, 31:1, pp. 122-130.

ANGUERA, M. T. e IZQUIERDO, C. (2006) Methodological approaches in human communication. From complexity of situation to data analysis, en RIVA, G., ANGUERA, M. T., WIEDERHOLD B. K, y MANTOVANI, F. (Coords.) *From Communication to Presence. Cognition, Emotions and Culture towards the Ultimate Communicative Experience* (Amsterdam, IOS Press) pp. 203-222.

ARIAS, E. y ANGUERA, M. T. (2004) Detección de patrones de conducta comunicativa en un grupo terapéutico de adolescentes, *Acción Psicológica*, 3:3, pp. 199-206.

BAKEMAN, R. y GOTTMAN, J. M. (1986) *Observing interaction. An introduction to sequential analysis* (Cambridge, Cambridge University Press).

BLANCO, M. y BERMEJO, V. (2008) ¿Nos permite la evaluación criterial por ciclos la detección precoz de las dificultades de aprendizaje en matemáticas?, en GONZÁLEZ-PIENDA, J. A. y NÚÑEZ, J. C. (Coords.), *Psicología y Educación: Un lugar de encuentro* (Oviedo, Ediuño).

BROWN, R. y BURTON, R. (1978) Diagnostic models for procedural in basic mathematical skills, *Cognitive Science*, 2, pp. 155-192.

COHEN, J. A. (1960) A coefficient of agreement for nominal scales, *Educational and Psychological Measurement*, 20, pp. 37-46.

CORONADO-HIJÓN, A. (2010) Evaluación criterial de las dificultades de aprendizaje en el cálculo. Un análisis de caso en evaluación inicial en la ESO, *UNO. Revista de Didáctica de las Matemáticas*, 54, pp. 88-102.

FLEISS, J. L., COHEN, J. y EVERITT, B. S. (1969) Large sample standard errors of kappa and weighted kappa, *Psychological bulletin*, 72, pp. 232-327.

GEARY, D. C. (1990) A componential analysis of an early learning deficit in mathematics, *Journal of Experimental Child Psychology*, 49, pp. 363-383.

GEARY, D. C. (1993) Mathematical disabilities: Cognitive, neuropsychological, and genetic components, *Psychological Bulletin*, 114, pp. 345-362.

GEARY, D. C., BROWN, S. C. y SAMARANYAKE, V. A. (1991) Cognitive addition: A short longitudinal study of strategy choice and speed-of-processing differences in normal and mathematically disabled children, *Developmental Psychology*, 27, pp. 787-797.

- GEARY, D. (2003) Learning Disabilities in Arithmetic: Problem-solving Differences and Cognitive Deficits, en GEARY, D., *Children's mathematical development: Research and practical applications* (Washington, DC, American Psychological Association).
- GOLDMAN S. R., PELLEGRINO, J. W. y MERTZ, D. L. (1988) Extended practice of basic addition facts: Strategy changes in learning disabled students, *Cognition and Instruction*, 5, pp. 223-265.
- GONZÁLEZ-PIENDA, J. A. y GONZÁLEZ-PUMARIEGA, S. (1998) Evaluación e intervención en las dificultades de aprendizaje de las matemáticas, en GONZÁLEZ PIENDA, J. A. y NÚÑEZ PÉREZ, J. C., *Dificultades de Aprendizaje Escolar* (Madrid, Pirámide).
- HAMBLETON, R. K. (1980) True score validity and standard setting methods, en BERK, R. A. (ed.), *Criterion referenced measurement: the state of the art* (Baltimore, Johns Hopkins University Press).
- HENSON, R. K. y KYLE, R. J. (2006) Use of Exploratory Factor Analysis in Published Research: Common Errors and Some Comment on Improved Practice, *Educational and Psychological Measurement*, 66, pp. 393-416.
- JONSSON, G. K., ANGUERA, M. T., BLANCO-VILLASEÑOR, A., LOSADA, J. L., HERNÁNDEZ-MENDO, A., ARDÁ, T., CAMERINO, O. y CASTELLANO, J. (2006) Hidden patterns of play interaction in soccer using SOF-CODER, *Behavior Research Methods*, 38:3, pp. 372-381.
- HILL M. y HILL A. (2000) *Investigação por Questionário* (Lisboa, Ediciones Sílabo).
- HOROWITZ, S. H (2011) *Lista de Verificación de las Discapacidades de Aprendizaje* (The National Center for Learning Disabilities, Inc. New York, NY). Ver: [http://www.nclld.org/images/content/files/checklist\\_spanish\\_rev\\_final.pdf](http://www.nclld.org/images/content/files/checklist_spanish_rev_final.pdf) (Consultado el 09. VI. 2014).
- IDEA (2004) *Individuals with Disabilities Education Improvement Act*, Public Law 108-466.
- JONSSON, G. K., ANGUERA, M. T., BLANCO-VILLASEÑOR, A., LOSADA, J. L., HERNÁNDEZ-MENDO, A., ARDÁ, T., CAMERINO, O. y CASTELLANO, J. (2006) Hidden patterns of play interaction in soccer using SOF-CODER, *Behavior Research Methods*, 38:3, pp. 372-381.
- JORDAN, N. C., HANICH, L. B. y KAPLAN, D. (2003). Arithmetic Fact. Mastery in young children: A longitudinal Investigation, *Journal of Experimental Child Psychology*, 85, pp. 103-119.
- JORDAN, N. C., LEVINE, S. C., y HUTTENLOCHER, J. (1994) Development of calculation abilities in middle and low income children after formal instruction in school, *Journal of applied developmental psychology*, 15, pp. 223-240.
- KIRBY, J. R. y BECKER, L. D. (1988) Cognitive components of learning problems in arithmetic, *Remedial and Special Education*, 9, pp. 7-16.
- KLINE, P. (1986) *A Handbook of Test Construction* (New York, Methuen).
- KLINE, P. (1994) *An Easy Guide to Factor Analysis* (Newbury Park, Sage).
- LANDIS, J. R. y KOCH, G. G. (1977) The measurement of observer agreement for categorical data, *Biometrics*, 33, pp. 159-174.

- MAZA, C. (1995) *Aritmética y representación. De la comprensión del texto al uso de materiales* (Barcelona, Paidós).
- MIRANDA, A. (1987) *Dificultades de aprendizaje en la lectura, escritura y cálculo* (Valencia, Promolibro).
- MIRANDA, A., FORTES, C. y GIL, M. D. (2000) *Dificultades del aprendizaje de las Matemáticas. Un enfoque evolutivo* (Málaga, Aljibe).
- MULLIS, I. V. S., MARTIN, M. O., FOY, P. y ARORA, A. (2012) *TIMSS 2011 International Results in Mathematics* (Chestnut Hill, Massachusetts, TIMSS & PIRLS International Study Center).
- NUNNALLY, C. (1978) *Psychometric Theory* (New York, McGraw-Hill).
- NUNNALLY, C. y BERNSTEIN, H. (1994) *Psychometric Theory* (New York, McGraw-Hill).
- OECD (2013) *PISA 2012 Results: What Students Know and Can Do - Student Performance in Mathematics, Reading and Science (Volume I)* (PISA, OECD Publishing).
- ORRANTIA, J. (2000) Las dificultades en el aprendizaje del cálculo desde el punto de vista cognitivo, *Premios nacionales de investigación e innovación educativa*, 1, pp. 75-102.
- ORTIZ, M. R., (2004) *Manual de Dificultades de Aprendizaje* (Madrid, Pirámide).
- OSERS (2006) Office of Special Education and Rehabilitative Services; Overview Information; National Institute on Disability and Rehabilitation Research (NIDRR). Ver <http://www2.ed.gov/about/offices/list/osers/news-2006.html> (Consultado el 21. I. 2014).
- PADILLA, M<sup>a</sup> T. (2002) *Técnicas e instrumentos para el diagnóstico y la evaluación educativa* (Madrid, Ed. CCS).
- ROMERO, J. y LAVIGNE, R. (2006) Unificación de Criterios Diagnósticos. II Procedimientos de Evaluación y Diagnósticos, *Materiales para la Práctica Orientadora*, Vol. 2 (Sevilla, Consejería de Educación de la Junta de Andalucía).
- SANTIUSTE, V. y GONZÁLEZ-PÉREZ, J. (2005) *Dificultades de aprendizaje e intervención psicopedagógica* (Madrid, Editorial CCS).
- VIANNA, M. H. (1983) *Los test en la educación* (Pamplona, EUNSA).
- WILSON, M. (2005) *Constructing measures: an item response modeling approach* (Mahwah, NJ, Lawrence Erlbaum Associates).

### Resumen: Construcción de una lista de cotejo (*checklist*) de dificultades de aprendizaje del cálculo aritmético

Los enfoques más clásicos en la identificación de alumnado con dificultades de aprendizaje, basados en el criterio psicométrico de discrepancia, están dando paso a nuevos paradigmas que, provenientes desde el ámbito de la investigación, recomiendan la utilización de procedimientos alternativos que basados en la evaluación criterial hayan sido contrastados y validados en resultados de investigación.

En esta línea de investigación se sitúa este trabajo, cuyo objetivo es describir el proceso de construcción, validación y confiabilidad de un instrumento de observación y cotejo de dificultades de aprendizaje del cálculo aritmético, diseñado desde los requerimientos científicos de la observación sistematizada, en la que con antelación se determinan las categorías a observar, las producciones que interesa registrar se definen previamente y se cuantifican en una medida final. Además, una lista de cotejo de estas características, es un instrumento de identificación criterial, con unos niveles estadísticamente aceptables de validez y fiabilidad, así como un instrumento de fácil utilización y aplicación a tareas en contextos naturales, ejercicios de clase, exámenes, etc., por lo que puede ser de uso generalizado, tanto por el personal especializado en diagnóstico psicopedagógico como por los docentes.

**Descriptores:** Evaluación pedagógica, lista de cotejo, dificultades de aprendizaje del cálculo.

**Summary:**  
**Drawing up a checklist of learning difficulties in arithmetic calculations**

The most traditional approaches to identifying students with learning diffi-

culties, based on the psychometric criterion of discrepancy, are giving way to new paradigms stemming from research recommending the use of alternative, criteria-based procedures contrasted and validated in research results.

Pooling from these new developments, this research aims to describe the construction, validation and reliability of a tool for observing and checking learning difficulties in arithmetic calculations. This tool is designed based on the scientific requirements of systematized observation where the categories to be observed and the production to be recorded have been previously defined and are subsequently quantified in a final measurement. Furthermore, a checklist of these characteristics serves both to identify the criteria that have statistically acceptable levels of validity and reliability and as an easy-to-use tool for tasks in natural contexts, class exercises, exams, and so forth. It therefore can be broadly used by both teachers and personnel specialized in psycho-pedagogical diagnosis alike.

**Key Words:** Educational evaluation, checklist, learning difficulties in arithmetic calculations.