

Revisión sistemática de instrumentos de medición acerca del manejo de la conducta en el aula

Systematic review of classroom behaviour management's measurement instruments

Andrea Pontex Benavides; Martín Alejandro, Márquez Marín; Lucero Miyomi Zevallos Ramos; Swan Sadith, Pacheco Díaz; Alessandra Belén Palacios Miguel; Vanessa Elizabeth Sopla Aliaga; Karen Noemí Arbañil García; Rocío Beatriz León Huerta; Caroline Millones Choquicondo; Estefani Aime Apaza Quispe

RESUMEN

Los instrumentos que miden el Manejo de la Conducta en el Aula (MCA) permiten conocer con qué estrategias se cuenta, cómo funcionan y cuáles se pueden incorporar al aula de acuerdo con la capacidad de los docentes, para crear entornos de aprendizajes positivos y funcionales. La presente revisión sistemática buscó analizar las investigaciones instrumentales acerca del MCA. Se abordaron y revisaron nueve estudios en función de criterios teóricos y de estándares de evaluación. Ninguno presenta como variable de medida al MCA, sino aproximaciones. Todos cumplen desde un 25 hasta un 87.5% con los criterios teóricos propuestos. Ocho reportan el proceso y resultado de la estimación de confiabilidad de los puntajes y 06 presentan evidencias de validez basadas en la estructura interna. Finalmente, se evidencia una necesidad de fomentar la investigación instrumental en el MCA, así como el acceso libre en el proceso de construcción de estos instrumentos.

Palabras clave: Manejo de conducta en el aula, Desempeño docente, Revisión sistemática.

ABSTRACT

The tests that measure Classroom Behaviour Management (CBM) allow us to know what strategies are available, how they work and which can be incorporated into the classroom according to the teachers' ability to create positive and functional learning environments. This systematic review sought to analyze instrumental research on CBM. Nine studies were addressed and reviewed based on theoretical criteria and evaluation standards. None had the CBM as a measurement variable, but approximations. All met from 25 to 87.5% with the proposed theoretical criteria. Eight report the process and results of the reliability estimate of the scores and six present evidence of validity based on internal structure. Finally, there is a need to promote instrumental research on CBM, as well as open-access in the process of building these instruments.

Keywords: Classroom Behaviour Management, Teacher performance, Systematic review.

Recibido: 09/07/2022

Aceptado: 15/12/2022

Publicado: 20/12/2023

Citar: Pontex Benavides, A. et al. (2023). Revisión sistemática de instrumentos de medición acerca del manejo de la conducta en el aula. *Investigaciones en Psicología*, 26(1), pp-7-15.

¹Universidad Nacional Mayor de San Marcos (UNMSM). (PsiE). Grupo de Investigación Psicología Educativa. República del Perú
E-mail psie.unmsm@gmail.com

Introducción

La construcción de entornos escolares saludables que favorezcan a un número de estudiantes con diferentes necesidades, conductas y alcances de atención es consecuencia de la labor que ejecuta el maestro al diseñar y poner en práctica distintas estrategias para el manejo del aula (Llywodraeth Cymru Welsh Government, 2012). Tal es así, que el manejo de la conducta en el aula, como competencia de los docentes, es aquel proceso de guía por parte de los profesores hacia los estudiantes, que requiere de una definición minuciosa de las conductas apropiadas que se esperan e inapropiadas que serán reducidas; utilizando para ello, estrategias antecedentes y consecuentes con el fin de (a) crear entornos de aprendizaje positivos y funcionales que disminuyan conductas disruptivas y recompensen el compromiso y el logro; (b) mantener y restaurar la clase como un clima efectivo para la enseñanza y el aprendizaje; y (c) establecer relaciones docente-estudiante de mayor calidez y sensibilidad, y menos enojo y dureza (Barrera y Valencia, 2008; Brophy, 1986; Li Grining, et al., 2010; Parsonson, 2012; Petty, 2001).

Brophy (1986) señala también que el manejo de aula efectivo es considerado como la habilidad para mantener y restaurar un entorno favorable para la enseñanza y el aprendizaje. Además, recoge acciones que llevan a cabo maestros que logran un efectivo manejo de aula; entre las que hacen referencia a la conducta, se resaltan:

- Conocer exactamente lo que se espera que los estudiantes hagan y cómo se quiere que lo hagan.
- Brindar instrucciones detalladamente o modelar los procedimientos y dar oportunidades para que los estudiantes practiquen y reciban retroalimentación.
- Anticipar las necesidades de guía e información de sus estudiantes.
- Monitorear el comportamiento de los estudiantes con una orientación instruccional y no punitiva.
- Expresar las reglas de conducta del aula y las expectativas de las conductas deseadas al inicio del año enfatizando en el *qué hacer* en vez del *qué no hacer*.
- Mantener seguimiento del progreso de los estudiantes a través de sistemas de conteo.
- Presentar información, instrucciones y objetivos claros, y ser capaz de descomponer tareas completas en procedimientos paso a paso.
- Introducir rápidamente las actividades, minimizando el tiempo en organizarse o en transiciones entre las tareas.

Walter & Frei (2007) presentan una serie de estrategias para el manejo de la conducta en el aula y la disciplina categorizadas en (1) proporcionar un entorno positivo para el aprendizaje, (2) preparar el aula para el éxito y (3) mantener a todos los estudiantes en la actividad. Implican una serie de técnicas conductuales adaptadas al contexto de aula y recomendaciones acerca de acciones que debe llevar a cabo el docente, antes o durante la clase.

Dada su implicancia en el ámbito educativo, se buscó estudiar dicho constructo, con el propósito de investigar la

realidad educativa a partir de la observación, la experiencia y los resultados, teniendo como medio la elaboración de revisiones sistemáticas, las cuales permiten la validación de prácticas, el diseño de políticas y la gestión de la educación (Páramo y Hederich, 2014). Algunas revisiones estuvieron enfocadas en la construcción de instrumentos, con la finalidad de ofrecer a la comunidad un conocimiento nuevo e innovador que permita entender mejor los problemas o fenómenos que se presentan en la práctica profesional. Asimismo, un instrumento confiable y validado nos ayuda a definir objetivos para encontrar un método que ayude a incorporar conductas deseadas dentro del aula (Stuhlman, Hamre, Downer & Pianta, 2010).

Por ello, resulta relevante conocer cómo ha variado la medición del manejo de la conducta y cómo estos instrumentos brindan una herramienta esencial para la práctica docente, por lo que importa conocer cuáles existen, cómo funcionan y aportan en el desarrollo de la calidad educativa, para así dar pie y conocer las bases para la construcción de un instrumento que tome en cuenta fortalezas y debilidades de los anteriores.

En el año 2019, Hepburn & Beamish realizaron una revisión de investigaciones referentes al manejo del aula en el contexto australiano y estadounidense, buscando las prácticas basadas en evidencia que se utilizaban en ese entonces, para ello se seleccionaron 27 artículos centrados específicamente en el tema. Se concluyó que hasta la fecha existe poca información sobre el conocimiento de los docentes acerca del manejo de la conducta en el aula, así como de la utilización de las prácticas basadas en evidencia, presentando dificultades para reducir las conductas problema en los estudiantes.

Con todo lo expuesto, se evidencia la necesidad de conocer el estado actual de estudios relacionados a la medición del manejo de la conducta en el aula. Por ello, el presente estudio tiene como finalidad revisar las investigaciones instrumentales referentes al manejo de conducta en el aula a nivel internacional, considerando su sustento teórico y estándares de evaluación. A partir de esto, se espera encontrar pocas investigaciones instrumentales referentes al MCA.

MÉTODO

El presente trabajo posee un diseño de investigación teórica, el cual se enfoca en seleccionar una serie de estudios acerca de una temática específica, a partir de los cuales se reporta los avances hasta ese momento obtenidos, siendo necesario realizar una exploración y análisis riguroso de las investigaciones recopiladas (Ato, López y Benavente, 2013).

Muestra

La selección de los artículos se realizó a partir de la búsqueda de bibliografía científica que cumpliera con los criterios de: publicación desde 1970, disponibilidad en inglés o español, acceso libre e indización en al menos una de 10 bases de datos revisadas (Dialnet, DOAJ, ERIC,

Jurn, Microsoft Academic, Redalyc, Refseek, Scielo, Science Research y Science Direct), artículos instrumentales o psicométricos y docentes de educación básica como población objetivo. Se utilizaron 22 términos de búsqueda que se muestran en la Tabla 1.

En la búsqueda de artículos se encontraron 111398 resultados para las palabras clave en español y 3711259 en inglés, obteniendo 3822657 artículos. Del total, 22 fueron investigaciones instrumentales de acceso libre y publicadas a partir de 1970 (10 en español y 12 en inglés). Posteriormente, se realizó el filtrado de investigaciones de acuerdo con los criterios de selección anteriormente expuestos, obteniendo al final 9 investigaciones instrumentales a revisar ubicadas en cuatro de las 10 bases de datos examinadas en la Tabla 2.

Análisis de la información

A través de una revisión bibliográfica acerca de las estrategias de Manejo de Conducta en el Aula (MCA) se seleccionaron las investigaciones y propuestas de diversos autores (Brophy, 1986; Petty, 2001; Walters & Frei, 2007 y Llywodraeth Cymru Welsh Government, 2012) para elaborar los criterios teóricos bajo los cuales se efectuaría la revisión instrumental. Posteriormente, la búsqueda de investigaciones instrumentales sobre el MCA permitió abordar nueve estudios, con base en los criterios de inclusión y exclusión propuestos. Luego, se elaboraron fichas de resumen individuales tomando en cuenta la muestra, las dimensiones y los análisis estadísticos. Además, se realizó una tabla de doble entrada comparando la estructura de cada instrumento en función a las dimensiones planteadas: (a) planifica el comportamiento adecuado y esperado de los estudiantes en el aula y (b) regula positivamente el comportamiento de los estudiantes hacia lo esperado y adecuado, las cuales se exponen en la Tabla 3.

Finalmente, se consideraron y compararon los estándares de evaluación de cada instrumento, entendiéndolos como pautas que deben ser consideradas al desarrollar y evaluar un test para garantizar una buena praxis por parte del profesional que lo construye y aplica. Para alcanzar los objetivos de esta investigación, se tomaron en cuenta los siguientes estándares de evaluación (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 2014):

- Descripción de la muestra: se especifican características sociodemográficas y de desarrollo relevante, en la medida que sea necesario, se incluyen antecedentes educativos, nivel de desarrollo, características de la comunidad o políticas de admisión a la escuela.
- Evidencias de validez:
 - a) Si la validez se logra mediante jueces, se especifica su experiencia, las instrucciones brindadas, si es que el trabajo de validación se realizó de manera conjunta o independiente y el nivel de acuerdo alcanzado.
 - b) Si se incluyen análisis estadísticos, se precisan las condiciones de recopilación de datos para juzgar la relevancia de los resultados.

Tabla 1
Términos de búsqueda

Español	Inglés
Manejo de la conducta en el aula	Classroom behavior management
Manejo de aula	Classroom management
Manejo de la conducta	Behavior management
Manejo instruccional	Instructional management
Disciplina	Discipline
Clima de aula	Classroom climate
Estrategias conductuales	Behavioral strategies
Estrategias didácticas	Teaching strategies
Organización del aula	Classroom organization
Estrategias de aprendizaje	Learning strategies

Fuente: Elaboración propia.

Tabla 2
Número de artículos encontrados según base de datos

Nombre	N° artículos seleccionados
Science direct	5
Redalyc	2
Eric	1
Microsoft Academic	1

Fuente: Elaboración propia.

Tabla 3
Dimensiones y capacidades de Manejo de Conducta del Aula (MCA)

Dimensiones	Capacidades
Planifica el comportamiento adecuado y esperado de los estudiantes en el aula	Estructura las sesiones de aprendizaje
	Diseña o selecciona estrategias para instaurar y mantener la conducta apropiada y/o elimina la conducta inapropiada
	Identifica la conducta apropiada e inapropiada
Regula positivamente el comportamiento de los estudiantes hacia lo esperado y adecuado	Domina la utilidad de las estrategias para instaurar y mantener la conducta apropiada y eliminar la conducta inapropiada
	Valora las conductas individuales, evolutivas y del grupo de estudiantes
Regula positivamente el comportamiento de los estudiantes hacia lo esperado y adecuado	Monitorea la conducta apropiada e inapropiada de los estudiantes en el aula
	Ejecuta estrategias para mantener la conducta apropiada
	Ejecuta estrategias para eliminar la conducta inapropiada

Fuente: Elaboración propia.

- Evidencias de validez de contenido: si la interpretación de los puntajes se basa en el contenido del instrumento, los ítemes responden a determinado dominio del constructo que se intenta medir. Además, se precisa el marco de referencia para la construcción del instrumento.
- Evidencias de validez basadas en procesos cognitivos: si la interpretación de los puntajes se basa en premisas acerca de la relación entre determinados procesos cognitivos que ejecuta el evaluado y el instrumento, la evidencia teórica o empírica que respalde ello se encuentra expuesta.
- Evidencias de validez basadas en estructura interna:
 - a) Si la interpretación de los puntajes depende de la relación entre los elementos o entre las partes del instrumento, se presenta evidencia de su estructura interna.
 - b) Si existen puntajes compuestos, se detalla la justificación y el proceso mediante el cual se obtuvieron dichos compuestos.
- Evidencias de validez basadas en la relación con otras variables
 - a) Selección de otras variables: si la validez se apoya en otras variables, se debe citar evidencia que justifique por qué se seleccionaron.
 - b) Validez convergente y discriminante: Se reportan evidencias experimentales o correlacionales entre las variables que se afirman que están relacionadas o no relacionadas.
 - c) Test-criterio: Cuando se afirma que un cierto nivel de rendimiento de la prueba predice un rendimiento de criterio adecuado o inadecuado, se informa sobre los niveles de rendimiento de criterio asociados con niveles dados de puntajes de prueba.
- Evidencias de validez basadas en las consecuencias del uso del test: si tras el uso del test se registran consecuencias, se reporta si estas surgieron debido a la sensibilidad del test a determinadas características que no se desean medir o un fallo para representar completamente la construcción prevista.
- Evaluación de la confiabilidad:
 - a) Se estima la confiabilidad tanto para el puntaje total como para los puntajes compuestos.
 - b) Los métodos para medir la confiabilidad son compatibles con la estructura interna del mismo.
- Reporte de la confiabilidad: cada método para cuantificar la confiabilidad de los puntajes se describe claramente mediante términos estadísticos apropiados.

RESULTADOS

a) Investigaciones instrumentales seleccionadas

1. *Problems in school Questionnaire (PIS)*

Alivernini, Lucidi & Manganeli (2012) realizaron el proceso de validación al contexto italiano de una versión reducida de este instrumento. Está compuesto por 6 preguntas, las cuales tienen una escala de 5 puntos con

2 ítemes, uno que representa *highly autonomy (HA)* y el otro, *highly controlling (HC)*. Los participantes fueron 443 docentes de primaria de 205 escuelas seleccionadas aleatoriamente y de donde se eligieron uno o dos salones de cuarto grado, también aleatorios. Para el análisis de datos se realizó un Análisis Factorial Confirmatorio Multi-grupo (AFCMG) obteniéndose dos factores en acuerdo con la estructura teórica. La invarianza de la medida se calculó a través de tres categorías de tiempo de servicio docente, para la invarianza métrica, se restringieron las cargas factoriales para que sean iguales en cada categoría. Los índices de ajuste demostraron que el modelo propuesto es apropiado para la explicación de los datos: $\chi^2 = 157,99$ ($df = 53$), $p < .001$, $\chi^2/df = 2.99$, $GFI = .94$, $AGFI = .92$, $RMSEA = .07$. La estimación de los índices de confiabilidad de las subescalas en términos de consistencia interna fue de .69 en HA y .61 en HC.

2. *Classroom Management Questionnaire (CMQ)*

Díaz, Gonzáles, Jara-Ramírez y Muñoz-Parra (2018) presentan una aproximación para la validación del CMQ, construido en una escala de tipo Likert y conformado por 60 ítemes que van desde "Raras veces" a "Usualmente". El instrumento está integrado por tres dimensiones de 20 ítemes cada una: (a) *Discipline*, (b) *Teaching and Learning* y (c) *Personal*. La muestra estuvo compuesta por 31 docentes entre las edades de 21 y 30 años. Las evidencias de validez del instrumento fueron obtenidas a través de las técnicas Delphi y Fleiss' Kappa, siendo necesarios 12 expertos para evaluar la claridad, coherencia y relevancia de cada ítem. En el análisis de la primera técnica, se obtuvo que en las dimensiones *Discipline* y *Personal* todos los ítemes fueron considerados apropiados, mientras que, en *Teaching and Learning* solo un ítem fue considerado inapropiado. Como resultado de la aplicación de la segunda técnica, se reportó que todas las dimensiones obtuvieron un regular, moderado, sustancial o casi perfecto acuerdo. La estimación de la confiabilidad se realizó a través del Coeficiente del Alfa de Cronbach, la dimensión *Discipline* obtuvo un coeficiente de .811; *Teaching and Learning*, un coeficiente de .860; y *Personal*, un coeficiente de .884.

3. *Tsang-Hester Observation Rubric (THOR)*

Contreras, Rittershausen, Correa, Solís, Núñez y Vásquez (2013) investigaron con respecto al instrumento original *Tsang-Hester Observation Rubric*, el cual fue elaborado por Henry Yeng-Chang Tsan en el 2003 y validado por Good en el 2006. Este permite medir el desempeño de los profesores en tres dimensiones: (a) evaluación de los aprendizajes, (b) conducción de la enseñanza y (c) manejo de la clase. Cada una de las dimensiones contempla un conjunto de criterios que presentan una rúbrica de cinco niveles, de los que se describen tres, los niveles 1, 3 y 5. El objetivo de su estudio fue validar el instrumento original en el contexto chileno teniendo en cuenta las normativas del Ministerio de Educación. Para ello realizaron dos juicios de expertos y una aplicación piloto. En once de los 25 ítemes del instrumento se encontró un 72 %

de coincidencia en la cualificación del desempeño atribuida al profesor observado para el proceso de validación.

4. *Teacher's Educational Practices Questionnaire (TEP-Q)*

Catalano, Perucchini & Vecchio (2014) realizaron un estudio mediante el cual analizaron la validez interna de este instrumento y, a su vez, verificaron si tenía relación con la enseñanza, creencias y autoeficacia. Dicho instrumento es un cuestionario de autoinforme de 36 ítems que evalúa las prácticas educativas. Se compone de tres dimensiones de la gestión del aula: (a) *relationship*, (b) *communication* y (c) *organization relationship*. La muestra estuvo compuesta de 307 docentes italianos, la estimación de la confiabilidad de las puntuaciones se realizó a través del Coeficiente del Alfa de Cronbach y se recogieron evidencias de validez basadas en la estructura interna que sustentaban un factor (*attention on relationships*, $\alpha = .68$) para la dimensión de *relationship* y dos (*effectiveness*, $\alpha = .66$, *self-regulation*, $\alpha = .65$) para *communication* y dos (*active methodologies*, $\alpha = .72$, *care on activities*, $\alpha = .60$) en *organization relationship*, estos factores se extrajeron por medio de una Factorización de Ejes Principales.

5. *Teacher Self-efficacy Scale (TSS)*

Karbasi y Samani (2016) consideraron importante contar con un instrumento que evalúe las capacidades relacionadas al buen desempeño docente. De esta forma, decidieron realizar un estudio para examinar las propiedades psicométricas del Teacher Self-efficacy Scale (TSS; Karbasi y Samani, 2016) para el contexto iraní; instrumento que, en un principio, estuvo conformado por 28 ítems distribuidos en 6 dimensiones: (a) *efficacy to influence decision making*, (b) *instructional self-efficacy*, (c) *disciplinary self-efficacy*, (d) *efficacy to enlist parental involvement*, (e) *efficacy to enlist community involvement* y (f) *efficacy to create a positive school climate*. La prueba fue administrada a 280 docentes iraníes para analizar la validez y confiabilidad del constructo del TSS. Los resultados mostraron que el TSS presenta un adecuado nivel de confiabilidad, obtenido a través del Coeficiente del Alfa de Cronbach ($= .77$ a $.85$) y suficientes evidencias de validez basadas en la estructura interna obtenidas a través de un Análisis de Componentes Principales con un método de rotación varimax, resultando cuatro factores que logran explicar el 61% de la varianza total de la escala: (a) *instructional self-efficacy*, (b) *efficacy to create positive school climate*, (c) *efficacy to enlist community involvement* y (d) *efficacy to influence decision making*.

6. *Classroom Management Questionnaire*

Akbari y Bolouri (2015) realizaron una investigación mediante el empleo de técnicas cuantitativas y cualitativas para desarrollar un cuestionario de gestión del comportamiento definiendo aspectos como (a) *language management*, (b) *structure management* y (c) *behavior management* a través de un cuestionario de 45 ítems. Dentro de sus propiedades psicométricas, la estimación de la confiabilidad se realizó a través del Coeficiente del Alfa de Cronbach. Para el instrumento se obtuvo un

coeficiente de $.89$; la dimensión *language management*, $.7$; *structure management*, $.83$ y *behavior management*, $.78$. En relación a las evidencias de validez basadas en la estructura interna, se llevó a cabo un Análisis Factorial Confirmatorio con el que se obtuvo un modelo de ajuste aceptable (CFI = $.988$, GFI = $.984$, RMSEA = $.091$).

7. *Behavior and Instructional Management Scale (BIMS)*

Martin y Sass (2010) presentaron la construcción de este instrumento cuyo objetivo es medir las creencias relacionadas al manejo del aula entre docentes y el rol importante de estas en las aulas. Consta de 12 ítems repartidos equitativamente en dos dimensiones: *Behavior Management (BM)* e *Instructional Management (IM)*. Se aplicó el instrumento a 550 docentes certificados de escuelas pública de Estados Unidos de América. En referencia al análisis de los ítems, se recogieron evidencias de validez basadas en el contenido, en la estructura interna y en la relación con otras variables. En cuanto al contenido, se consultó con 94 docentes en formación para conocer la claridad de los ítems y se solicitó su retroalimentación acerca de la poca claridad o incongruencia respecto al constructo, se indicó que estas evidencias fueron adecuadas. En relación a la estructura interna, como segunda parte del estudio, se llevó a cabo un Análisis Factorial Confirmatorio utilizando una matriz de correlaciones policóricas y por el método de estimación de mínimos cuadrados ponderados con media y varianza ajustada (WLSMV) con el que se obtuvo un modelo de ajuste de bueno a excelente ($\chi^2(28) = 106.637$, $p < .001$, CFI = $.013$, TLI = $.010$, RMSEA = $.010$, WRMR = $-.103$). En cuanto a la relación con otras variables, se recogió una adecuada evidencia de validez convergente y discriminante en torno a *Efficacy for Instructional Strategies (EIS)*, *Efficacy for Student Engagement (ESE)* y *Efficacy for Classroom Management (ECM)*, puesto que se halló que BM estaba ligeramente correlacionada de forma positiva con IM ($r = .24$), y de forma negativa con ECM ($r = -.19$) pero no con EIS ($r = .02$) o con ESE ($r = -.12$); por otro lado, se obtuvo que IM presentaba una alta y negativa correlación con ESE, ECM y EIS (r 's $> -.50$). Por último, la estimación de la confiabilidad de los puntajes se realizó a través de la consistencia interna con el Coeficiente del Alfa de Cronbach, obteniéndose un coeficiente de $.774$ para BM y de $.770$ para IM.

8. *Classroom Management Scale*

Temli-Durmus (2016) presentó la construcción del instrumento Classroom Management Scale. Se utilizó una muestra de 130 profesores de secundaria del curso de ciencias en Turquía para medir la variable denominada *classroom management*. Las dimensiones que la conformaron fueron: (a) *instructional management (IM)*, (b) *people management (PM)* y (c) *behavior management (BM)*. Para evaluar la validez basada en la estructura interna se realizó un análisis factorial confirmatorio ($\chi^2 = 231.9$; $gl = 87$; $p < .05$). Los valores resultantes de los índices de ajuste utilizados fueron: CFI = $.92$, TLI = $.89$ y RMSEA = $.081$ con un intervalo de confianza de $.05$ a

.07. Además, se utilizó el coeficiente del Alfa de Cronbach para evaluar la consistencia interna, siendo los valores obtenidos: (a) PM = .82, (b) BM = .75 y (c) IM = .80.

9. Protocol for classroom management styles assessment (PCMSA)

Djigic y Stojiljkovic (2012) utilizan el PCMSA con la finalidad de medir los estilos de manejo del aula. Para ello, el mencionado instrumento fue aplicado en 269 docentes distribuidos en ocho escuelas primarias ubicadas en zonas tanto urbanas como rurales de Serbia. El instrumento distingue los siguientes estilos: (a) *interventionist*, (b) *interactionist* y (c) *non-interventionist*; y está compuesto de 20 ítems incluidos en tres dimensiones: (a) *personality*, (b) *teaching* y (c) *discipline*. Con respecto a sus propiedades psicométricas, se evidencia que la confiabilidad de los puntajes, estimada con el Coeficiente del alfa de Cronbach y en evaluaciones de uno o dos observadores y en una o dos clases, obtuvo como mejor resultado utilizando dos observadores en dos clases ($\alpha = .902$).

b) Revisión teórica

No se encontraron instrumentos cuya variable de medida sea el MCA, sino aproximaciones desde la medición del manejo de aula, desempeño docente y manejo instruccional. Con respecto a los nueve artículos resultantes, en su mayoría publicados en el idioma inglés

(ocho), se observó que todos los instrumentos presentan por lo menos dos dimensiones que concuerdan teóricamente con las propuestas por los autores. Además, se evidencia que, a nivel de instrumento, llegan a cumplir desde un 25 % hasta un 87.5 % con los criterios propuestos, donde 7 de ellos sobrepasaron el 50 %; mientras que, a nivel de capacidad, los porcentajes fluctúan de 44.4 % a 88.9 %, siendo *diseña o selecciona estrategias para instaurar y mantener la conducta apropiada y/o eliminar la conducta inapropiada* la más considerada dentro de los criterios teóricos de construcción de los instrumentos analizados, ver Tabla 4 y Figura 1.

c) Revisión de estándares de evaluación

Con respecto a la revisión de los estándares de evaluación, se obtuvo que todos los estudios presentan una descripción específica de los participantes; ocho de ellos incluyen análisis estadísticos de los resultados de las pruebas y reportan el proceso y resultado de la estimación de confiabilidad de los puntajes, seis presentan evidencias de validez basadas en la estructura interna, cinco muestran evidencias de validez basadas en el contenido, uno presenta evidencias de validez basadas en la relación con otras variables y ninguno presenta evidencias de validez basadas en los procesos cognitivos ni en las consecuencias del uso del test (Tabla 5).

Tabla 4
Revisión de criterios teóricos

Dimensiones	Capacidades	Instrumentos									Porcentaje de artículos que consideran cada capacidad
		1	2	3	4	5	6	7	8	9	
Planifica el comportamiento adecuado y esperado de los estudiantes en el aula	Estructura las sesiones de aprendizaje.		Teaching and learning	Conducción de la enseñanza		Efficacy to create positive school climate	Structure management	Instructional Management	People Management Instructional Management	Teaching	77,8%
	Diseña o selecciona estrategias para instaurar y mantener la conducta apropiada y/o eliminar la conducta inapropiada		Discipline Teaching and learning	Conducción de la enseñanza	Relationship	Efficacy to create positive school climate	Structure management	Behavior Management	Behavior Management	Teaching	88,9%
	Identifica la conducta apropiada e inapropiada		Discipline	Manejo de la clase			Structure management	Behavior Management	Behavior Management	Discipline	66,7%
	Domina la utilidad de las estrategias para instaurar y mantener la conducta apropiada y eliminar la conducta inapropiada			Conducción de la enseñanza	Relationship Communication Organization		Structure management		Instructional Management		44,4%
Regula positivamente el comportamiento de los estudiantes hacia lo esperado y adecuado	Valora las conductas individuales, evolutivas y del grupo de estudiantes.	Highly Autonomy Supportive	Teaching and learning Personal						People Management	Personality	44,4%
	Monitorea la conducta apropiada e inapropiada de los estudiantes en el aula.			Manejo de la clase	Organization		Structure management	Instructional Management			44,4%
	Ejecuta estrategias para mantener la conducta apropiada	Highly Controlling	Discipline Teaching and learning Personal	Manejo de la clase Evaluación	Relationship Communication Organization	Instructional self-efficacy Efficacy to create positive school climate	Structure management			Discipline	77,8%
	Ejecuta estrategias para eliminar la conducta inapropiada		Discipline	Manejo de la clase Evaluación			Structure management		Behavior Management	Discipline	55,6%
Porcentaje de capacidades que considera cada artículo		25%	75%	87.5%	50%	37.5%	87.5%	50%	75%	75%	

Fuente: Elaboración propia.

Figura 1
 Porcentaje de coincidencia con los criterios teóricos planteados

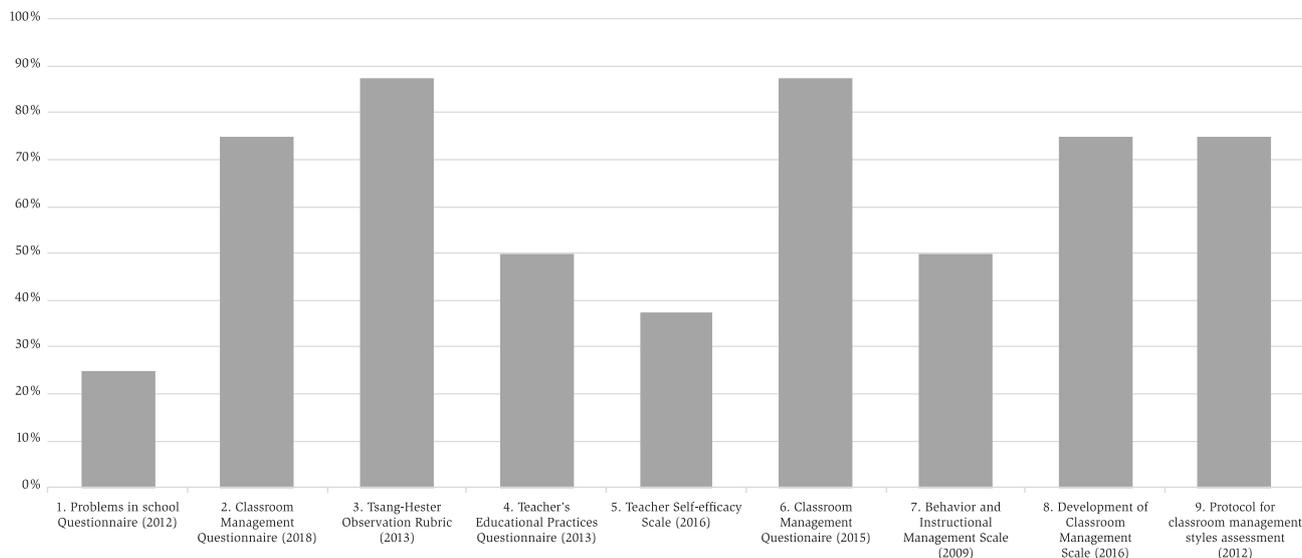


Tabla 5.
 Revisión de estándares de evaluación de los estudios instrumentales seleccionados

Estándares	1	2	3	4	5	6	7	8	9	
Muestra	Sí									
Evidencias de Validez										
	a	No aplica	Sí	Sí	No aplica	No aplica	Sí	Sí	No aplica	Sí
	b	Sí	Sí	No aplica	Sí	Sí	Sí	Sí	Sí	Sí
Evidencia de validez de contenido	No aplica	Sí	No	No aplica	No aplica	No aplica	Sí	Sí	No aplica	Sí
Evidencias de validez basadas en los procesos cognitivos	No aplica									
Evidencias de validez basadas en la estructura interna										
	a	Sí	No aplica	No aplica	Sí	Sí	Sí	Sí	Sí	No aplica
	b	Si	No aplica	No aplica	Sí	Sí	Sí	Sí	Sí	No aplica
Evidencias basadas en la relación con otras variables										
	a	No aplica	Sí	No aplica	No aplica					
	b	No aplica	Sí	No aplica	No aplica					
	c	No aplica								
Evidencias de validez basadas en las consecuencias del uso del test	No aplica									
Evaluación de confiabilidad										
	a	Si	Si	No aplica	Si	Si	Si	Si	Si	Si
	b	Si	Si	No aplica	Si	Si	Si	Si	Si	Si
Reporte de la confiabilidad	Si	Si	No aplica	Si						

Fuente: Elaboración propia con base en American Educational Research Association, American Psychological Association and National Council on Measurement in Education (2014).

DISCUSIÓN

Los resultados evidencian que no se encontraron instrumentos cuya variable de medida sea el MCA, sino aproximaciones desde la medición del manejo de aula, desempeño o eficacia docente y manejo instruccional. Esto indica que el MCA es mayoritariamente estudiado desde las variables previamente mencionadas, las cuales incluyen su concepto, sin embargo, las consideraciones que se tienen para medirlo no son suficientes, puesto que esta variable sigue siendo teóricamente amplia. Se reconoce

el esfuerzo por intentar incluir todas estas en un instrumento como los revisados en este estudio; sin embargo, se recomienda que, si se quiere medir el desempeño docente o manejo de aula, se realice mediante baterías y no a través de pruebas únicas.

Además, se observó que todos los instrumentos presentan como mínimo dos dimensiones que concuerdan teóricamente con las propuestas de los autores, en adición a ello, cumplen por lo menos dos estándares de evaluación planteados por American Educational Research Association, American Psychological Association & National Council on

Measurement in Education (2014). Esto puede explicarse a partir de la revisión conceptual del MCA, donde se consideraron diversos autores que incluyen aspectos teóricos similares en su definición. De acuerdo con la evidencia presentada, los criterios teóricos propuestos se consideran relevantes para la construcción de instrumentos de evaluación docente referidos al MCA puesto que hay dimensiones similares en la construcción de estos. En consecuencia, se podría indicar que, al momento de la construcción de cada uno de dichos instrumentos, los autores independientemente han utilizado criterios comunes.

A pesar de haber abordado al inicio una gran cantidad de artículos (pasado el millón), es importante resaltar la reducción a solo 22 debido a que éstas cumplían con dos de los criterios de inclusión: ser investigaciones instrumentales y de acceso libre. Esto denota la poca información gratuita que puede encontrarse acerca del MCA. Con respecto a la revisión de estándares de evaluación, se obtuvo que muchos de los estudios revisados presentan evidencias de validez de distintos tipos, sin embargo, no todos las incluyen; dentro de estos estándares se encuentran los análisis estadísticos, así como el proceso y resultado de la estimación de confiabilidad de los puntajes. Por tanto, para futuras revisiones sistemáticas que abarquen investigaciones instrumentales, se recomienda considerar dentro de los resultados un apartado donde se consideren los estadísticos utilizados, y además reportar si el artículo tiene como objetivo construir, adaptar o validar una prueba, ya que nos permitirá conocer de qué manera ha sido estudiada la variable con respecto a la cantidad de instrumentos que se encuentren disponibles en la red.

Aunado a ello, se observa la poca presencia de investigaciones instrumentales o psicométricas relacionadas al MCA. Por ello, es necesario continuar realizando investigaciones de este tipo, que permitan plantear objetivos con la finalidad de encontrar un método que ayude a incorporar conductas deseadas dentro del aula, así como también ofrecer a la comunidad un conocimiento innovador para entender mejor los problemas o fenómenos que se presentan en la labor docente, y a partir de ello validar las prácticas, diseñar políticas y gestionar la educación (Stuhlman, Hamre, Downer y Pianta, 2010; Páramo y Hederich, 2014).

Por otra parte, sobre los hallazgos referentes a los estándares de evaluación se evidencia que hay una preferencia con respecto a las evidencias de validez basadas en la estructura interna y la selección del coeficiente Alfa de Cronbach como estimador de la consistencia interna y, por consiguiente, el reporte de la confiabilidad de las puntuaciones de los tests. Esto puede deberse a que el proceso del análisis de dichas evidencias resulta ser más sencillo y accesible para los investigadores, debido a la cantidad de información disponible en bases de datos.

Asimismo, se encontró que no se reportan instrumentos que consideren todos los criterios planteados, tanto los teóricos como los de estándares de evaluación, de acuerdo a la variable de medida, por lo que se vuelve necesaria la construcción de un instrumento que los incluya, con el fin de que el docente tenga una herramienta que aporte a la mejora de la práctica educativa.

Finalmente, el presente estudio tuvo dentro de su proceso de elaboración un conjunto de limitaciones, a partir de las cuales se proponen recomendaciones para futuras investigaciones afines. La primera de ellas hace referencia al acceso libre de artículos de investigación, ya que muchos no estaban disponibles para la lectura completa sin un costo de por medio; por ello se recomienda que si los artículos cumplieron con los criterios propuestos por las revistas y la rigurosidad científica necesaria sea posible difundir la información de manera que los investigadores puedan acceder libremente, esto conllevaría a una mejora de las prácticas de investigación; además, se podrían proponer programas de financiamiento que faciliten el acceso a bases de datos pagadas, los cuales a su vez permitirían el respaldo del presupuesto de la investigación. Es así que la segunda limitación está relacionada con el proceso de selección de las bases de datos, la cual fue por conveniencia, por ello se sugiere definir criterios bibliométricos que guíen la búsqueda. En adición a ello, se observó que los diseños de búsqueda de las bases de datos revisadas eran distintos y muchos de ellos dificultaban el acceso específico a la información requerida, ante ello se propone que las bases de datos puedan incluir en sus procesos de búsqueda criterios más definidos que faciliten la revisión de sus artículos. Otra de las limitaciones encontradas, fue el manejo de solo dos idiomas dentro del equipo de investigación (inglés y español); por lo que se recomienda, en primer lugar, la promoción del aprendizaje de nuevos idiomas, así como la traducción formal de artículos en idiomas distintos al de origen y el establecimiento de alianzas entre equipos de investigación de diferentes países. Por último, la poca presencia de antecedentes sobre el tema estudiado limitó la comparación de los resultados del estudio realizado con otros afines, por lo que se sugiere continuar con la línea de investigación planteada para la construcción de un marco teórico sólido referido a la aplicación del MCA.

REFERENCIAS

- Akbari, R., & Bolouri, M. (2015). Development and Validation of an ELT Based Classroom Management Questionnaire. *Journal of Applied Linguistics and Language Research*, 2(4), 113-129.
- Alivernini, F., Lucidi, F. & Manganelli, S. (2012). The Validation of a Scale Measuring Teaching Styles in the Italian Context. *Procedia - Social and Behavioral Sciences*, 46, 1487-1490, 10.1016/j.sbspro.2012.05.326.
- American Educational Research Association, American Psychological Association and National Council on Measurement in Education (2014). *Standards for Educational and Psychological Testing*. Washington: American Educational Research Association.
- Ato, M. López, J y Benavente, A. (2013). *Un sistema de clasificación de los diseños de investigación en psicología*, 29(3), <https://doi.org/10.6018/analesps.29.3.178511>.

- Barrera, M. y Valencia, P. (2008) *Estrategias de manejo conductual en aula. Paz ciudadana*. Recuperado de <https://www.orientacionandujar.es/wp-content/uploads/2015/09/RECOPIACION-DE-ESTRATEGIAS-DE-MODIFICACION-DE-CONDUCTA-EN-EL-AULA.pdf>
- Brophy, J. (1986). Classroom Management Techniques [Técnicas del manejo del aula]. *Education and Urban Society*, 18(2), 182-194. Recuperado de: <http://journals.sagepub.com/doi/pdf/10.1177/0013124586018002005>
- Catalano, M., Perucchini, P. & Vecchio, G. (2014). The Quality Of Teacher's Educational Practices: Internal Validity and Applications of A New Self-Evaluation Questionnaire. *Procedia - Social and Behavioral Sciences*, 141, 459-464, doi: 10.1016/j.sbspro.2014.05.080
- Contreras, I., Rittershausen, S., Correa, E., Solís, C., Núñez, C. y Vásquez, N. (2013). IOC, un instrumento para cualificar desempeño docente en aula: Su generación y validación. *Estudios Pedagógicos*, 39(2), 85-96.
- Díaz, C., González, G., Jara-Ramírez, L. y Muñoz-Parra, J. (2018). Validación de un cuestionario de manejo del aula para profesores de inglés en formación y practicantes. *Revista Colombiana de Educación*, 75, 263-286.
- Djigić, G., & Stojiljković, S. (2012). Protocol for classroom management styles assessment designing. *Procedia-Social and Behavioral Sciences*, 45, 65-74. <https://doi.org/10.1016/j.sbspro.2012.06.543>
- Hepburn, L. & Beamish, W. (2019). Towards Implementation of Evidence-Based Practices for Classroom Management in Australia: A Review of Research. *Australian Journal of Teacher Education*, 44(2), 82-98. Recuperado de <https://ro.ecu.edu.au/ajte/vol44/iss2/6>
- Karbasi, S., & Samani, S. (2016). Psychometric properties of teacher self-efficacy scale. *Procedia-Social and Behavioral Sciences*, 217, 618-621. Recuperado de <https://cyberleninka.org/article/n/1365217.pdf>
- Li Grining, C., Cybele Raver, C., Champion, K., Sardin, L., Metzger, M. & Jones, S. (2010). Understanding and Improving Classroom Emotional Climate and Behavior Management in the "Real World": The Role of Head Start Teachers' Psychosocial Stressors. *Early Education and Development*, 21(1), 65-94. doi: 10.1080/10409280902783509
- Llywodraeth Cymru Welsh Government (2012). *Practical approaches to behaviour management in the classroom* [Enfoques prácticos para el manejo de la conducta en el aula]. Recuperado de: <http://learning.gov.wales/docs/learningwales/publications/140822-behaviour-management-handbook-for-primary-schools-en.pdf>
- Martin, N. K., & Sass, D. A. (2010). Construct validation of the behavior and instructional management scale. *Teaching and Teacher Education*, 26(5), 1124-1135. <https://doi.org/10.1016/j.tate.2009.12.001>.
- Páramo, P y Hederich, C. (2014). Educación basada en la evidencia. *Revista Colombiana de Educación*, (66), 13-16. <https://doi.org/10.17227/01203916.66rce13-16>.
- Parsonson, B. (2012). Evidence-based Classroom Behaviour Management Strategies [Estrategias de manejo de la conducta en el aula basadas en la evidencia]. *Kairaranga*, 13, 16-23. Recuperado de: <https://files.eric.ed.gov/fulltext/EJ976654.pdf>
- Petty, R. (2001). Manejo del aula, competencia básica del maestro líder. *Revista Electrónica Educare*, (1), 85-91. <https://doi.org/10.15359/ree.2001-1.7>
- Stuhlman, M., Hamre, B., Downer, J., & Pianta, R. (2010). Why should we use classroom observation. *Centre for Advanced of Teaching and Learning, University of Virginia*. Recuperado de https://curry.virginia.edu/uploads/resourceLibrary/CASTL_practitioner_Part1_single.pdf.
- Temli-Durmus, Y. (2016). Development of Classroom Management Scale for Science Teachers. *Universal Journal of Educational Research*, 4(9), 1950-1957. Recuperado de <https://eric.ed.gov/?id=EJ1113869>
- Walters, J., & Frei, S. (2007). Managing classroom behavior and discipline. *Shell Education*.