

ARTÍCULO DE INVESTIGACIÓN

EL INVENTARIO DE CONCIENCIA METACOGNITIVA PARA DOCENTES (MAIT): ADAPTACIÓN CULTURAL Y VALIDACIÓN EN UNA MUESTRA DE DOCENTES COLOMBIANOS

THE METACOGNITIVE AWARENESS INVENTORY FOR TEACHERS (MAIT): CULTURAL ADAPTATION AND VALIDATION IN A SAMPLE OF COLOMBIAN TEACHERS

ANTONIO P. GUTIÉRREZ DE BLUME¹, DIANA MARCELA MONTOYA LONDOÑO²
 GEORGIA SOUTHERN UNIVERSITY. UNITED STATES OF AMERICA.
 UNIVERSIDAD DE CALDAS - UNIVERSIDAD DE MANIZALES. COLOMBIA

FECHA RECEPCIÓN: 24/10/2019 • FECHA ACEPTACIÓN: 1/3/2020

Para citar este artículo: Gutiérrez de Blume, A. P., & Montoya, D. (2020). El Inventario de Conciencia Metacognitiva para Docentes (MAIT): Adaptación cultural y validación en una muestra de docentes colombianos. *Psychología*, 14(2), 115-130 doi: 10.21500/19002386.4584

Resumen

Comprender cómo operan los constructos inobservables en función de la cultura y el lenguaje es un objetivo importante para los investigadores. Debido a que el idioma y la cultura ejercen una influencia tan poderosa en la forma en que se experimentan los fenómenos psicológicos, los constructos latentes pueden no funcionar universalmente entre culturas e idiomas. Por lo tanto, el propósito del presente estudio fue doble. Primero, se realizó un proceso de adaptación cultural del Inventario de Conciencia Metacognitiva para Docentes (MAIT) de 24 ítems, desarrollado originalmente por Balcikanli (2011). En segundo lugar, se probó y validó el instrumento traducido en una muestra de docentes colombianos ($N = 755$). Se presume una estructura de factor modificada a partir de la propuesta en el estudio de validación original. Los resultados revelaron que, después de la eliminación de tres ítems que no encajaban, la estructura de factores hipotética propuesta por el equipo para el MAIT-R de 21 ítems demostró un excelente ajuste a los datos.

- 1 Associate Professor, Georgia Southern University. Department of Curriculum, Foundations, and Reading. Estados Unidos. Address: PO. Box 8144, Statesboro, Georgia, USA. Telephone: 912-478-7831. Email: agutierrez@georgiasouthern.edu. orcid:0000-0001-6809-1728.
- 2 Magíster en Neuropsicología. Universidad San Buenaventura, Colombia. Docente Departamento de estudios educativos, Facultad de Artes y Humanidades, Universidad de Caldas. Dirección: Sede Principal Calle 65 # 26-10. Teléfono: 3146789695. Email: diana.montoya@ucaldas.edu.co. orcid:0000-0001-8007-0102.

Palabras clave: Metacognición, conocimiento metacognitivo, regulación metacognitiva, conciencia metacognitiva, docentes

Abstract

Understanding how unobservable constructs operate based on culture and language is an important objective for researchers. Because language and culture exert such a powerful influence on the way we experience psychological phenomena, latent constructs may not function universally between cultures and languages. Therefore, the purpose of the present study was twofold. First, we translated the Metacognitive Awareness Inventory for Teachers (MAIT) with 24 items, originally developed by Balcikanli (2011). Second, we test and validate the instrument translated to Spanish using a sample of Colombian teachers ($N = 755$). We hypothesized a modified factor structure based on the proposal in the original validation study. The results revealed that, after the elimination of three items that did not fit, our hypothesized factor structure for the MAIT-R of 21 items demonstrated an excellent fit to the data.

Keywords: Metacognition, metacognitive knowledge, metacognitive regulation, metacognitive awareness, teachers

Introducción

Desde diferentes disciplinas propias de las ciencias cognitivas, entre las que se destacan áreas como la psicología, la filosofía o las neurociencias, existe acuerdo en reconocer como una de las tendencias emergentes de investigación, en la relación entre cognición y educación, los estudios en metacognición (Proust, 2010; Zohar & Dori, 2012; Peña-Ayala & Cárdenas, 2015; Sawyer, 2014; Schraw & Gutiérrez, 2015), por cuanto esta se considera como ‘pensamiento sobre el pensamiento’ o, dicho de otra manera, como el proceso de conocimiento y regulación que realiza la persona de sus propios recursos cognitivos en función de favorecer su propio proceso de aprendizaje (Azevedo, 2009; Flavell, 2000).

Tradicionalmente, se han privilegiado los estudios sobre la metacognición del estudiante, con una limitada referencia a la metacognición del profesor en relación con el proceso de enseñanza (Kallio et al., 2017), aun cuando está suficientemente documentada la importancia de la metacognición para la formación de docentes que se conozcan y reflexionen sobre su enseñanza, lo que se espera afecte en sus propias aulas la formación de estudiantes capaces de autorregular su propio proceso de aprendizaje (Jiang, Ma & Gao, 2016; Kallio et al., 2017).

En relación con el abordaje de la metacognición del docente, el instrumento más empleado lo constituye El Inventario de Conciencia Metacognitiva para Docentes (MAIT) (Balcikanli, 2011), elaborado con base en el patrón de oro que más se ha utilizado en el mundo para evaluar la conciencia metacognitiva, prueba conocida

como Inventario Clásico para la Evaluación de la Conciencia Metacognitiva en Adultos (MAI) (Schraw & Sperling-Dennison, 1994). Actualmente, las investigaciones realizadas con el MAIT reportan un alfa de Cronbach que oscila entre 0.79 y 0.85 (Balcikanli, 2011; Mai, 2015; Palantis et al., 2017).

En su versión original, el MAIT fue desarrollado por Balcikanli (2011) en una investigación realizada con 674 docentes de inglés en Turquía; tuvo como objetivo realizar una versión adaptada del Inventario de Conciencia Metacognitiva (MAI) (Schraw & Sperling-Dennison, 1994) para ayudar a los docentes en ejercicio a darse cuenta de sus niveles metacognitivos de enseñanza (MAIT). En su versión, el investigador realizó ajustes al inventario para que pudiera adaptarse al proceso de conocimiento y regulación del docente frente a las situaciones de enseñanza, asimismo, modificó el formato de respuesta tipo Likert de 5 puntos, en el que la puntuación se estima de la siguiente manera: totalmente en desacuerdo (1), en desacuerdo (2), neutral (3), de acuerdo (4), totalmente de acuerdo (5).

Respecto a las propiedades psicométricas del instrumento en su versión original, se empleó una medida de adecuación de muestreo Kaiser-Meyer-Olkin (KMO) para identificar la validez del inventario (0.794), y el valor para Barlett TKest se identificó como significativo (2513.474). En relación con el análisis factorial, se confirmó la estructura de factores originales del MAI en relación con los factores de conocimiento y regulación metacognitiva; así se identificaron seis factores. El factor I incluyó los elementos 1, 7, 13, 19; el factor II, los

ítems 2, 8, 14, 20; el factor III, los reactivos 3, 9, 15, 21; el factor IV, los elementos 4, 10, 16, 22; el factor V, los ítems 5, 11, 17, 23, y, finalmente, el factor VI incluyó los reactivos 6, 12, 18, 24. Los valores de confiabilidad variaron de 0.79 a 0.85 (Balcikanli, 2011).

Entre los antecedentes de importancia en estudios de estandarización realizados con el MAIT, se consideran las investigaciones realizadas por Kallio et al. (2017), quienes desarrollaron un estudio en Finlandia con 208 docentes en servicio, vinculados con instituciones educativas de orientación vocacional, que tuvo como objetivo determinar la utilidad de una versión abreviada de 18 reactivos (forma MAIT-18) para evaluar la conciencia metacognitiva de docentes en servicio. En dicha versión ajustada de la prueba se eliminaron 6 reactivos porque se consideraron afirmaciones difíciles de interpretar o confusas para los participantes al momento de responder la prueba (ítems 3 y 20); asimismo, otros reactivos fueron eliminados dado que se consideró que ya estaban incluidos en otros elementos de la prueba (ítem 5) o eran difíciles de operacionalizar frente a algunos términos incluidos en el reactivo como “efectividad” (ítems 12 y 23). También fueron eliminados ítems con valores demasiado positivos para la asimetría (ítem 15). El alfa de Cronbach que se estableció para la versión abreviada del instrumento fue para el caso del conocimiento de la cognición de 0.80 (IC 95 % = 0.76-0.84) y de la regulación de la cognición fue de 0.86 (IC 95 % = 0.83-0.89). En general, puede indicarse que la consistencia interna de todos los factores fue buena, con un alfa del cuestionario completo de 0.90 (IC 95 % = .88-.92) (Kallio et al., 2017).

El estudio de Kallio et al. (2017) contribuye a confirmar el modelo teórico de metacognición señalado desde los primeros trabajos propuestos a partir del Inventario de Conciencia Metacognitiva General (MAI) (Schraw & Sperling-Dennison, 1994), desde el que se reconoce la existencia de dos grandes grupos de factores: el conocimiento de la cognición (declarativo, procedimental y condicional) y la regulación de la cognición (planificación, monitoreo y evaluación); del mismo modo, contribuye a confirmar el ajuste de la estructura factorial del instrumento con independencia del género o el nivel de formación, así como la consistencia interna del instrumento (Kallio et al., 2017).

Mai (2015) realizó un estudio en Malasia con 52 profesores de ciencias de la escuela primaria; tuvo como objetivo medir la percepción de estos sobre la conciencia metacognitiva. En dicho estudio, además, se estableció una adaptación del Inventario de Conciencia Metacognitiva para Docentes (MAIT) al idioma y se estimó el nivel de validez y confiabilidad para la versión adaptada de la prueba.

El inventario original del MAIT (Balcikanli, 2011) para esta investigación se modificó a una escala Likert de 5 puntos, que iban desde totalmente en desacuerdo (1) hasta totalmente de acuerdo (5). En el estudio de Mai (2015) se conservó la estructura de los dos grandes factores: conocimiento y regulación con sus subescalas originales. La fiabilidad de esta versión adaptada de la prueba se estableció mediante el coeficiente de consistencia interna alfa de Cronbach y el coeficiente de la mitad dividida. Los resultados del alfa de Cronbach permitieron establecer que la escala general tenía un alfa de 0.880 y un coeficiente de mitad dividida de 0.852.

Finalmente, como antecedente de la presente investigación, se señala un estudio realizado por Vanags, Pestovs y Peršņvica (2018) con 301 docentes de la ciudad de Letonia, que tuvo como objetivo establecer las propiedades psicométricas del inventario MAIT en su versión adaptada de 18 reactivos (MAIT-18) (Kallio et al., 2017), en el que se determinó la fiabilidad interna, la validez factorial y convergente de la prueba, con valores similares a la versión original del instrumento (Vanags et al., 2018).

La presente investigación tuvo como objetivo realizar el proceso de adaptación cultural y validación del MAIT en una muestra de docentes colombiana; así mismo, se constituye en el primer ejercicio de esta naturaleza para un instrumento de evaluación del constructo de metacognición en habla hispana.

Pregunta(s) de investigación e hipótesis

Con base en la tradición investigativa que se revisó, se plantearon las siguientes preguntas de investigación: ¿en qué medida el proceso de adaptación cultural del MAIT en una muestra de docentes colombiana puede contribuir a confirmar las propiedades psicométricas del instrumento en investigaciones reportadas en otras culturas?, y ¿cuál es la estructura de factores latentes del

Inventario de Conciencia Metacognitiva para Docentes (MAIT) previamente validado entre una muestra de docentes de Colombia?

Hipótesis: La estructura de factores del MAIT previamente adaptada y validada, que se muestra en la Figura 1a, se comporta diferente en una muestra de sujetos colombianos, dada la influencia de la cultura y del lenguaje en el aprendizaje humano (esto porque el estudio de validación original empleó una muestra de una cultura de habla inglesa). Además, dado lo que Schraw y Sperling-Dennison (1994) encontraron con respecto a la

estructura factorial del Inventario de Conciencia Metacognitiva (MAI) y lo que Gutiérrez et al. (2016) descubrieron con respecto a un factor de orden superior de monitoreo general en la metacognición, se predijo por parte del equipo investigador una estructura de factor de orden superior que muestra un mejor ajuste en esta muestra (ver Figura 1b). Así mismo, se concluye que el proceso de adaptación cultural del MAIT en una muestra de docentes colombiana sí puede contribuir a confirmar las propiedades psicométricas del instrumento.

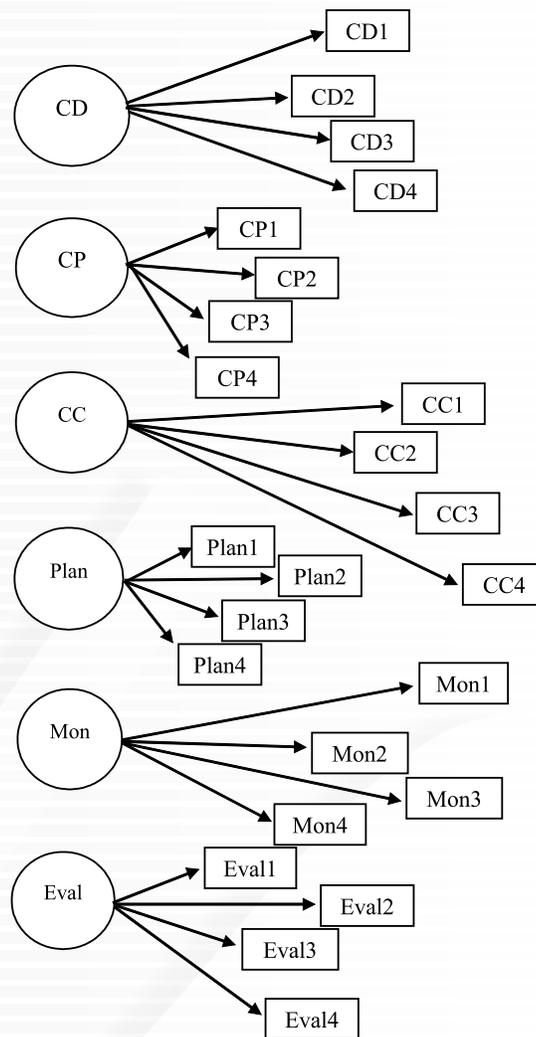


Figura 1a. La estructura factorial original del MAIT, validada por Balcikanli (2011). Nota: CD = conocimiento declarativo; CP = conocimiento procedimental; CC = conocimiento condicional; Plan = planificación; Mon = monitoreo; Eval = evaluación. Los cuadros representan variables manifiestas de los factores latentes. Cada variable está asociada con su parámetro de error/varianza residual para especificar el modelo, y se permitió que cada factor latente se correlacionara libremente entre sí. Sin embargo, los términos de error/residual y las correlaciones no se muestran en la figura para mantener la simplicidad.

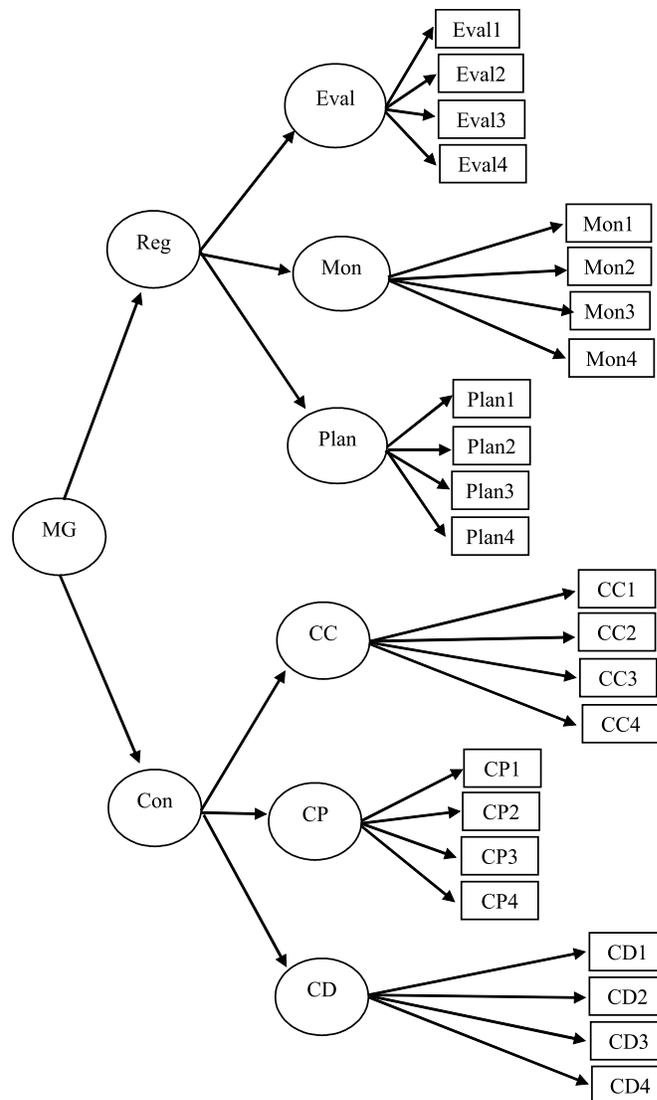


Figura 1b. Estructura factorial hipotética para el MAIT-R, teniendo en cuenta las conceptualizaciones de metacognición de Schraw y Sperling-Dennison (1994) y Gutiérrez et al. (2016). Nota: MG = metacognición general; Reg = regulación; Con = conocimiento; CD = conocimiento declarativo; CP = conocimiento procedimental; CC = conocimiento condicional; Plan = planificación; Mon = monitoreo; Eval = evaluación. Los cuadros representan variables manifiestas de los factores latentes. Cada variable está asociada con su parámetro de error/varianza residual para especificar el modelo, y el factor de metacognición general de tercer orden (MG) está asociado con un parámetro de varianza de perturbación/error. Sin embargo, los términos de error/residual no se muestran en la figura para mantener la simplicidad.

Método

Participantes y muestra

La selección de la muestra se realizó de manera intencional y de acuerdo con contactos establecidos con la unidad de calidad de la Alcaldía de Manizales y con una de las universidades públicas de la misma ciudad. Así, la muestra estuvo conformada por 816 individuos, una vez se realizó la revisión para establecer que todos

los instrumentos estuvieran por completo diligenciados, la muestra quedó conformada por 771 para el análisis final. Todas las personas participaron de la investigación de forma voluntaria y se contó con la respectiva firma del consentimiento informado para su vinculación.

De estos 771 participantes, 388 fueron mujeres (382 hombres; 1 no declarado); asimismo, puede indicarse que 142 eran estudiantes universitarios de último nivel de formación de diferentes programas de pregra-

do en educación, 303 eran docentes universitarios con formación de maestría y doctorado quienes se desempeñaban con grupos de estudiantes de pregrado en una universidad pública del centro del país, y 326 eran docentes adscritos a la alcaldía con titulación de maestría o doctorado, quienes desempeñaban su labor profesional con estudiantes de básica primaria, básica secundaria y media académica.

Materiales e instrumentos

MAIT. Inventario de Conciencia Metacognitiva para Docentes (Balcikanli, 2011). Este instrumento es una versión adaptada al contexto de la enseñanza, con base en el inventario clásico para la evaluación de la conciencia metacognitiva en adultos (MAI) (Schraw & Spersling-Dennison, 1994).

El cuestionario se estructura con un formato de respuesta tipo Likert de 5 puntos, en el que el grado de acuerdo varía desde totalmente en desacuerdo (1) hasta totalmente de acuerdo (5). La puntuación se proporciona de la siguiente manera: totalmente en desacuerdo (1), en desacuerdo (2), neutral (3), de acuerdo (4), totalmente de acuerdo (5).

El inventario consta de 24 declaraciones que permiten evaluar, en relación con el conocimiento metacognitivo, los componentes declarativo, procedimental y condicional y, en relación con la regulación metacognitiva, las subhabilidades de planificación, monitoreo y evaluación (Balcikanli, 2011). El alfa de Cronbach reportado para la prueba varía desde 0.79 para el caso de la habilidad de evaluación, hasta 0.85 para el caso del conocimiento declarativo (Balcikanli, 2011); asimismo, otros estudios en contextos culturales diversos también han reportado un alfa de Cronbach general que oscila entre 0.880 y 0.852 (Mai, 2015) hasta 0.894 (Palantis et al., 2017). Para la presente muestra, los coeficientes de confiabilidad de consistencia interna, los alfas de Cronbach variaron de 0.78 a 0.86 para las seis escalas del MAIT.

Procedimiento

Para la realización del proceso de adaptación y validación del instrumento MAIT en la muestra de docentes colombiana se tuvieron en cuenta los criterios dados por la Comisión Internacional de Test (ITC) y publicados en diferentes artículos sobre lineamientos internaciona-

les (Muñiz, Elosua & Hambleton, 2013; Muñiz, Hernández & Ponsoda, 2015; Muñiz, Elosua, Padilla & Hambleton, 2016).

En la etapa previa de traducción y revisión

Primero, es importante señalar que el instrumento MAIT corresponde a una versión adaptada del patrón de oro para evaluar la conciencia metacognitiva en adultos MAI, elaborado por Schraw y Dennison (1994). Tanto el MAI en su versión original como la versión adaptada del instrumento MAIT (Balcikanli, 2011) se han publicado como instrumentos completos dentro de revistas científicas de libre acceso, y, en particular, el MAI se ha traducido a más de 10 idiomas en el mundo.

En esta etapa del proceso de comprobación del registro de propiedad intelectual, es importante señalar que el primer autor del presente artículo se formó directamente con el profesor Schraw, quien fue su director de tesis, además formó parte directa durante su vida laboral de su equipo de trabajo. Asimismo, para contar con los debidos permisos, desde la Universidad participante en el estudio se escribió vía email al autor del MAIT (Balcikanli, 2011), quien autorizó que fuera usada la versión adaptada de su instrumento con fines investigativos.

Uno de los aspectos más destacados para tener en cuenta en la revisión de antecedentes es la documentación de la relevancia y universalidad del constructo, en cuanto la conciencia metacognitiva se ha evaluado en todo el mundo mediante el instrumento MAI.

En relación con la etapa sobre las directrices en el desarrollo del test

Para los procesos de traducción y revisión conceptual de las equivalencias entre la versión original y la versión retro-traducida del test se generó un comité de expertos, integrado por una estadística, tres psicólogos, una revisora de estilo y un experto en el constructo con competencia bilingüe en las dos lenguas. Durante el proceso se tuvo en cuenta, más que la traducción literal, el proceso de adaptación de significado al idioma destino, que fue el español, criterio que validó el experto en el constructo, dado que su primera lengua es el español. El proceso de corrección lingüística estuvo a cargo de la revisora de estilo.

El comité de expertos en el presente estudio tuvo en cuenta en la adaptación y en la traducción el respeto por la forma original del ítem, así como posibles dife-

rencias lingüísticas y culturales presentes entre las dos lenguas; asimismo, se conservó el formato del reactivo, la escala de respuesta, las reglas de corrección, las convenciones utilizadas y la forma de aplicación.

Se realizó una primera prueba piloto con una pequeña muestra de docentes, en la que se pudo establecer que todos los reactivos traducidos e incluidos en la prueba eran comprensibles para los docentes, con excepción del reactivo 4, que se modificó en función del sentido semántico del ítem en el contexto social. En la versión original del test, el reactivo 4 se escribe como “I pace myself while I am teaching in order to have enough time”, que parece traducirse “Me paso a mí mismo mientras estoy enseñando con el fin de tener tiempo suficiente”. Este ítem fue reemplazado por una equivalencia semántica, que se consideró en el comité de expertos como más pertinente, en la forma: “Regulo mi propio ritmo de trabajo para tener el tiempo suficiente”, antes de la fase operacional.

En general, en la aplicación previa del instrumento se estableció una buena reacción de las personas que realizaron la prueba; se aseguró que todos los ítems fueran comprendidos; se registró el tiempo aproximado que duro la aplicación (de 20 a 30 minutos).

Directrices de confirmación

Se revisaron las propiedades psicométricas de la versión adaptada del test y su equivalencia con respecto al test original. Este proceso se llevó a cabo dentro del estudio a través de un análisis factorial exploratorio con extracción de factor común y rotación oblicua, ya que se esperaba que los factores estuvieran correlacionados. Los resultados preliminares de esta etapa del proceso piloto, con la pequeña muestra piloto, revelaron que los ítems mostraron una validez de constructo adecuada, y las escalas, una adecuada consistencia interna (rango: .76-.85). Además, los factores se correlacionaron moderadamente y en la dirección teórica esperada (es decir, positiva).

Directrices sobre aplicación

De acuerdo con los lineamientos relacionados con la aplicación de pruebas, dados por Hambleton (1996), puede señalarse que la aplicación de la prueba se llevó a cabo por profesionales en psicología y educación, que laboran en la propia cultura, y tenían adecuada formación

y experiencia en la aplicación de tareas de evaluación en el ámbito de la educación.

Directrices sobre puntuación e interpretación

Para analizar los resultados finales se tuvo en cuenta que los estudios comparativos deben usarse para profundizar en el estudio de las semejanzas y diferencias entre los grupos analizados, pero no para establecer comparaciones sin tener en cuenta las diferencias culturales (Hambleton & Bollwark, 1991), y se ofreció, finalmente, información sobre las equivalencias del constructo, método e ítems, así como evidencias sobre la validez de la versión adaptada en la muestra con la que se trabajó.

Una vez se llevaron a cabo las etapas previas ya enunciadas, correspondientes al desarrollo del proceso de adaptación y validación del instrumento MAIT, se procedió a realizar el trabajo de campo. Así, en un primer momento, se realizó, con las tres muestras de docentes por separado (estudiantes de último nivel de formación de pregrado en educación, docentes adscritos a la alcaldía y docentes de la universidad), la presentación de los objetivos de la investigación por parte del equipo de trabajo, y después se llevó a cabo la firma del consentimiento informado. En un segundo momento, se procedió a realizar la aplicación del instrumento Inventario de Conciencia Metacognitiva para Docentes (MAIT) (Balcikanli, 2011) de manera grupal, con cada grupo de participantes, en diferentes momentos de evaluación de media hora aproximadamente cada uno.

Análisis de los datos

Todos los datos se examinaron a través del programa Paquete Estadístico para las Ciencias Sociales (SPSS), versión 23, para valores atípicos univariados y multivariados (Tabachnick & Fidell, 2013). Además, los datos se probaron para suposiciones univariadas y multivariadas, incluida la normalidad multivariada (asimetría y curtosis), multicolinealidad y singularidad, para proceder con la estimación de máxima verosimilitud (ML). Con respecto a la normalidad multivariada, los datos demostraron curtosis moderada; por lo tanto, las estadísticas de ML robusto (MLR) se solicitaron e interpretaron en lugar de las estadísticas de distribución normal de ML. Los procedimientos de MLR proporcionaron índices de ajuste precisos (p. ej., S-B χ^2 , * CFI, * NNFI, * IFI y * RMSEA y su * intervalo de confianza [*CI_{90%}]) que corri-

gen las violaciones de moderadas a severas de la normalidad multivariada. Además, los procedimientos de MLR ajustan/corriguen los errores estándar y la significación estadística de los coeficientes de regresión no estandarizados, considerando la falta de normalidad multivariada (Kline, 2005). Se cumplieron todos los demás supuestos de ML y no se detectaron valores atípicos extremos que socavarían la confiabilidad de los datos. El análisis de valores faltantes demostró que 16 casos (2.07 %) tenían datos faltantes, lo que arrojó 755 casos disponibles para el análisis final de la información.

Se calcularon estadísticas descriptivas para todas las medidas utilizando el software SPSS-23. El modelo de ecuaciones estructurales (SEM) se utilizó para explorar las relaciones hipotéticas entre los constructos en el presente estudio. El modelo hipotético (ver Figura 1b) se evaluó mediante el paquete de software estadístico EQS 6.3 (Bentler, 2005), especificando los efectos directos en los datos actuales.

Primero, estos datos se sometieron a un análisis factorial confirmatorio estándar (CFA) para examinar si las variables observadas especificadas (indicadores de los constructos latentes) eran en realidad productos de los constructos latentes. Para este propósito, se analizaron el ajuste general del modelo, las cargas de los factores y la varianza explicada que cada factor contribuyó a sus indicadores. Luego, se evaluó el modelo hipotético de medición de orden superior de la Figura 1b para el ajuste general del modelo, así como las cargas factoriales. Asimismo, se examinaron los índices de bondad de ajuste ajustados y el residuo cuadrático medio estandarizado (SRMR) entre el modelo hipotético y un modelo de solo varianza para determinar si el modelo propuesto se ajustaba a los datos. El modelo de medición en cada nivel (es decir, de primer orden, segundo orden y tercer orden por separado) se evaluó por separado para garantizar la precisión de los índices de ajuste del modelo (Kline, 2005; Tabachnick & Fidell, 2013). El rho de Dillon-Goldstein (ρ) (Werts, Linn & Jöreskog, 1974) también se utilizó para evaluar la confiabilidad global o compuesta del modelo. El rho mide qué tan bien las variables de manifiesto/indicador, como un bloque, representan la variable latente en la que se supone que se cargan. Al igual que la interpretación del alfa de Cronbach, los valores más altos para rho indican una mayor confiabilidad

del modelo, con .70 como límite inferior para la confiabilidad adecuada del modelo (Werts et al., 1974).

Finalmente, los procedimientos de respecificación *post hoc* (exploratorio), más específicamente, la prueba del multiplicador de Lagrange (LM) para la adición de parámetros y la prueba de Wald para la eliminación de parámetros, se realizaron en el modelo para mejorar su ajuste a los datos y representar más adecuadamente las variables manifiestas significativas de los constructos. Todas las respecificaciones del modelo se guiaron por consideraciones teóricas, así como por el conocimiento de los investigadores de los constructos.

Resultados

Modelo de primer orden

El modelo de medición inicial de primer orden especificó seis factores latentes con cuatro indicadores cada uno (24 en total): conocimiento declarativo, conocimiento de procedimiento, conocimiento condicional, planificación, monitoreo y evaluación. Además, cada indicador podría cargarse solo en su factor especificado, los términos residuales (error/varianza única) para los indicadores permanecieron sin correlación, los factores permitieron correlacionarse entre sí, y no se impusieron restricciones de igualdad en las cargas de factores. El CFA demostró que el modelo de medición hipotético era inadecuado, $\chi^2(237, N = 755) = 861.68, p < .001$, Bentler-Bonett * NNFI = .80, Bollen's * IFI = .83, * CFI = .83, SRMR = .06, * RMSEA = .06, y su asociado * CI_{90%} = .055, .063.

La prueba de LM sugirió que la adición de varias correlaciones de error dentro del conocimiento declarativo, de procedimiento y condicional, así como los factores de evaluación, respectivamente, se recomendaban para su inclusión en el modelo. Esto sugiere que estos indicadores comparten una varianza residual común. Debido a que estas correlaciones de errores tenían un sentido teórico sustantivo, estos parámetros se incluyeron en el modelo de medición de primer orden respecificado (Kline, 2005). Además, la prueba de Wald recomendó la eliminación de tres indicadores, CP1, CC1 y Plan1, correspondientes a los factores de conocimiento de procedimiento, conocimiento condicional y de planificación, respectivamente. La inspección de la solución

estandarizada reveló que las cargas factoriales para estos tres indicadores eran bastante bajas .133 (CP1), .178 (CC1) y .165 (Plan1); lo que indica que no estaban midiendo adecuadamente sus respectivos factores. Por lo tanto, para maximizar la interpretabilidad sustantiva de nuestros hallazgos, se decidió eliminar estos tres indicadores y dejar solo tres indicadores para los factores de conocimiento de procedimiento, conocimiento condicional y de planificación.

El modelo de medición final de primer orden con los tres indicadores eliminados y los parámetros de correlación de error agregados fue de mejor ajuste, $SB \chi^2 (173, N = 755) = 463.05, p < .001$, Bentler-Bonett *

NNFI = .90, Bollen's * IFI = .92, * CFI = .92, SRMR = .05, * RMSEA = .05, y su asociado * $CI_{90\%} = .042, .052$. Todos los factores de carga fueron de moderados a grandes y estadísticamente significativos en el nivel de significancia $p < .01$. Además, las varianzas explicadas (R^2) por los factores en sus respectivos indicadores fueron todas de pequeñas a moderadas, con un rango de .16 a .42 (Mediana = .30). Finalmente, el rho de Dillon-Goldstein, $\rho = .88$, mostró una alta confiabilidad del modelo. La Tabla 1 incluye estadísticas descriptivas (medias y desviaciones estándar) para el conjunto final de 21 ítems.

Tabla 1. Estadísticas descriptivas de los 21 ítems finales del MAIT-R

Indicador	M	DE
CD1	4.46	0.62
CD2	4.27	0.70
CD3	3.87	0.79
CD4	4.37	0.79
CP2	4.28	0.73
CP3	4.26	0.70
CP4	3.69	0.83
CC2	4.39	0.78
CC3	4.44	0.62
CC4	3.89	0.83
Plan2	4.36	0.77
Plan3	4.43	0.67
Plan4	4.24	0.73
Mon1	4.49	0.63
Mon2	4.13	0.78
Mon3	4.41	0.66
Mon4	4.19	0.78
Eval1	4.41	0.70
Eval2	4.17	0.78
Eval3	4.43	0.70
Eval4	3.94	0.85

N = 771

Nota. Los factores del conocimiento procedimental, el conocimiento condicional y de planificación tenían cada uno un indicador eliminado, según los procedimientos de respecificación del modelo. Clave: M = media; DS = desviación estándar; CD = conocimiento declarativo; CP = conocimiento procedimental; CC = conocimiento condicional; Plan = planificación; Mon = monitoreo; Eval = evaluación.

Las correlaciones de factores, que van desde $r = .60$ a $.79$, indicaron que había fuertes relaciones positivas entre los factores. Además, estas correlaciones indicaron que los constructos exhibieron una validez discriminante adecuada ($r < .90$). Las correlaciones entre el error y las varianzas únicas impuestas en el modelo final de medición de primer orden fueron de moderadas a fuertes, positivas y estadísticamente significativas.

Modelo de segundo orden

El modelo de segundo orden especificó seis factores latentes de primer orden con 21 indicadores (conocimiento declarativo, de procedimiento y condicional, así como planificación, monitoreo y evaluación) y dos factores de segundo orden: conocimiento y regulación (Schraw & Sperling-Dennison, 1994). En consecuencia, las correlaciones de factores en el CFA de primer orden fueron reemplazadas por rutas directas de factores desde el factor de segundo orden del conocimiento, hasta los factores de primer orden del conocimiento declarativo, de procedimiento y condicional y el factor de segundo orden de regulación para los factores de la planificación, el monitoreo y evaluación de primer orden. Al igual que con el CFA de primer orden, se permitió la correlación de los dos factores de segundo orden. El CFA de segundo orden demostró un ajuste adecuado a los datos, $SB \chi^2 (178, N = 755) = 424.58, p < .001$, Bentler-Bonett * NNFI = .91, Bollen's * IFI = .93, * CFI = .93, SRMR = .05, * RMSEA = .04, y su asociado * $CI_{90\%} = .038, .048$. Las cargas de factores para los factores de primer orden fueron fuertes y variaron de $.69$ a $.92$, y la correlación entre los factores de segundo orden de conocimiento y regulación fue $r = .75$. La prueba LM y la prueba Wald indicaron que no se justificaba ningún procedimiento de respecificación del modelo para el modelo de segundo orden y, por consiguiente, este se consideró el modelo final. La fiabilidad del modelo compuesto para este modelo también fue alta, $\rho = .87$.

Modelo de tercer orden

El modelo final de tercer orden fue idéntico al modelo de segundo orden, excepto que la correlación de factores de los factores de conocimiento y regulación de segundo orden fue reemplazada por una ruta de factor directa desde el factor de metacognición general de tercer orden hasta el conocimiento y la regulación. Se agregaron factores de segundo orden y una perturbación (término de error) al factor de metacognición general de tercer orden (Gutiérrez et al., 2016). El CFA de tercer orden demostró un excelente ajuste a los datos, $SB \chi^2 (177, N = 755) = 237.77, p = .001$, Bentler-Bonett * NNFI = .98, Bollen's * IFI = .98, * CFI = .98, SRMR = .03, * RMSEA = .02, y su asociado * $CI_{90\%} = .014, .028$. Las cargas de factores para los factores de segundo orden fueron fuertes, $.84$ para conocimiento y $.78$ para regulación. La prueba LM y la prueba Wald indicaron que no se justificaban procedimientos de respecificación del modelo para el modelo de tercer orden, y, por lo tanto, este se consideró el modelo final. La fiabilidad del modelo compuesto para este modelo también fue alta, $\rho = .89$ (en la Figura 2 se presenta el modelo final con todas las estimaciones de parámetros).

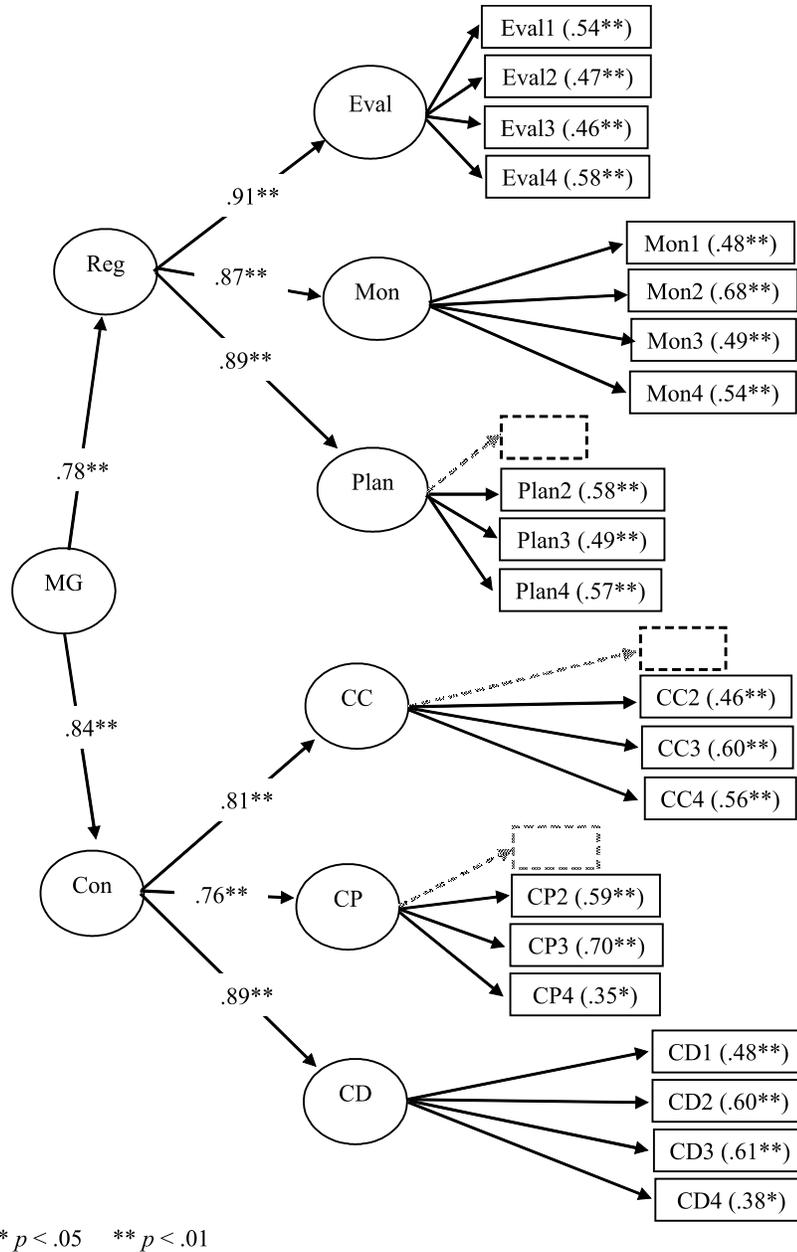


Figura 2. Modelo final para el MAIT-R. Las líneas discontinuas transparentes representan las tres variables de manifiesto eliminadas del modelo final. Los parámetros entre paréntesis representan las cargas factoriales estandarizadas de cada indicador a su factor respectivo. Nota. MG = metacognición general; Reg = regulación; Con = conocimiento; CD = conocimiento declarativo; CP = conocimiento procedimental; CC = conocimiento condicional; Plan = planificación; Mon = monitoreo; Eval = evaluación. Los cuadros representan variables manifiestas de los factores latentes. Cada variable está asociada con su parámetro de error/varianza residual para especificar el modelo, y el factor de metacognición general de tercer orden (MG) está asociado con un parámetro de varianza de perturbación/error. Sin embargo, los términos de error/residual no se muestran en la figura para mantener la simplicidad.

similar al factor de monitoreo general descubierto por Gutiérrez et al. (2016). Por lo tanto, el presente estudio respalda una estructura factorial más compleja para el MAIT que la conceptualizada originalmente por Balcikanli (2011). Además, los hallazgos revelaron, respecto al estudio original del MAIT, que tres de los reactivos, uno para el constructo de conocimiento procedural (factor 2), otro para el conocimiento condicional (factor 3) y, finalmente, uno para el constructo de planificación (factor 4), eran inadecuados y, en consecuencia, fueron eliminados del modelo final presentado en la propuesta para docentes en Colombia (MAIT-21). Si estas diferencias son realmente un producto de diferencias culturales y lingüísticas o un artefacto metodológico en la medida en que el estudio de validación inicial de Balcikanli (2011) no evaluó una estructura factorial más compleja y, por lo tanto, varios componentes de la varianza fueron relegados al error en ese modelo anterior, sigue siendo una vía fructífera para futuras investigaciones.

Es importante señalar que, de acuerdo con los hallazgos de esta investigación, fueron eliminados los reactivos correspondientes a la pregunta 2 del factor de conocimiento procedural (Trato de utilizar técnicas de enseñanza que me funcionaron en el pasado), la pregunta 3 del factor de conocimiento condicional (Utilizo mis fortalezas para compensar mis debilidades en el proceso de enseñanza) y la pregunta 4 de planeación (Regulo mi propio ritmo de trabajo para tener el tiempo suficiente). Estos hallazgos coinciden con la versión adaptada propuesta por Kallio et al., (2017) para docentes de Finlandia, dado que en su estudio también se eliminó el reactivo 3 por considerarlo como un ítem difícil de interpretar por los docentes y se distancia de la estructura original planteada por Balcikanli (2011) para docentes de Turquía respecto a la inclusión en la estructura factorial original de los reactivos 2 y 4 respectivamente.

Estos resultados parecen confirmar la universalidad de la estructura factorial de dos componentes de la metacognición: conocimiento y regulación validada de manera empírica desde la versión original del MAI (Schraw & Sperling-Denison, 1994) y parecen evidenciar diferencias culturales al interior de las diferentes subhabilidades metacognitivas, en especial a nivel de los conocimientos procedural y condicional, así como de la habilidad de planificación.

Implicaciones para la teoría, la investigación y la práctica

El principal aporte del presente trabajo lo constituye el poder contar con un instrumento pertinente para evaluar la metacognición del docente en docentes de habla hispana y, específicamente, en el contexto donde se realizó la estandarización de la prueba para docentes de una región al interior del país y con diferentes niveles de formación. Este aporte resulta significativo por la ausencia de instrumentos pertinentes en habla hispana para evaluar el constructo, y, especialmente, dada la necesidad de tener en cuenta la metacognición del docente en relación con las nuevas demandas de formación y actualización, asimismo, en relación con el desarrollo de investigaciones que se encaminen en esta perspectiva.

Al respecto, algunos investigadores han señalado que, aunque se invierte mucho tiempo promoviendo el concepto de formar docente como profesionales metacognitivos, se está aún lejos de realizar diferentes investigaciones encaminadas a determinar sistemáticamente si la metacognición del docente influye en el aprendizaje de los estudiantes de maneras deseables y, de ser así, de saber cómo las instituciones formadoras de docentes pueden contribuir con políticas y acciones concretas de formación a desarrollar la metacognición de los docentes (Duffy et al., 2009).

Contar con un instrumento válido y confiable, adaptado culturalmente para evaluar la metacognición del docente respecto a su proceso de enseñanza, sin duda, permitirá mejorar las políticas regionales y nacionales de formación continuada del docente desde los niveles más iniciales, en relación con la formación de docentes en escuelas normales y programas de licenciatura, con implicaciones hasta los niveles posgraduales de maestría y doctorado o, incluso, a nivel del desarrollo de políticas de formación estatales a nivel regional y nacional, con base en los resultados aportados desde el proceso de reflexión del docente, derivados de las puntuaciones reportadas desde el MAIT en su forma adaptada para la región (MAIT-21).

Avenidas para la investigación futura

Tomando en cuenta la riqueza cultural y nacional de Colombia, en una próxima investigación se debería realizar un proceso de estandarización del instrumento

en su forma (MAIT-21) a nivel nacional, con una muestra representativa de docentes de diferentes regiones del país, en cuanto a la realización de un estudio con una muestra más grande, esto permitiría el desarrollo de baremos para la prueba con representación nacional y, quizás, desde el establecimiento de diferencias por género, edad, años de experiencia docente, nivel de formación y área de dominio disciplinar.

Reflexiones metodológicas y limitaciones

Al igual que con cualquier investigación que involucre datos de encuestas con participantes humanos, la muestra de docentes colombianos puede no reflejar adecuadamente la verdadera magnitud de los efectos reales entre las variables latentes en el modelo por los investigadores para este estudio. De acuerdo con esto, otros modelos propuestos pueden mostrar con mayor precisión los efectos observados entre los constructos bajo investigación. Por lo anterior, es importante que los resultados de este estudio se interpreten solo en términos de las variables latentes examinadas y operacionalizadas, como las variables manifiestas utilizadas para medir los constructos. Finalmente, nuestros participantes se calificaron a sí mismos altamente entre las variables en estudio, lo que potencialmente puede crear efectos de techo.

A pesar de estas limitaciones, el presente estudio ofrece nuevos conocimientos sobre las investigaciones científicas con respecto a la validación de instrumentos de medición en varios idiomas y culturas. Nuestra sólida muestra de docentes colombianos nos ayudó a determinar que la medida inicial, validada con muestras de habla inglesa, funciona de manera diferente en las poblaciones de habla hispana, un hallazgo bastante práctico.

Conclusión

Los hallazgos de esta investigación demuestran la importancia de participar en más estudios interculturales que empleen validación cruzada de instrumentos desarrollados para una cultura e idioma. El MAIT original fue validado con una muestra de habla inglesa. Sin embargo, el instrumento no funcionó de manera similar para una muestra colombiana de docentes. Recurriendo a lo que han hecho otros que realizan investigaciones sobre metacognición (Gutiérrez et al., 2016; Schraw y Sperling-Dennison, 1994), se estudió la idoneidad de una estruc-

tura de factores de orden superior. Si bien esta estructura de factores de orden superior fue respaldada por los datos, tres de los ítems originales se consideraron no aptos para la muestra colombiana de docentes. Presumiblemente, la cultura y el idioma influyeron en las diferencias entre el MAIT original de 24 ítems y el desarrollado por el equipo de la presente investigación MAIT-R de 21 ítems. Se espera que este modelo adaptado no solo se emplee en otras muestras de Colombia, sino también en otras culturas de habla hispana en América Latina.

Referencias

- Azevedo, R. (2009). Theoretical, conceptual, methodological, and instructional issues in research on metacognition and self-regulated learning: A discussion. *Metacognition and Learning*, 4, 87-95. doi: 10.1007/s11409-009-9035-7
- Bentler, P. M. (2005). *EQS 6 structural equations program manual*. Encino, CA: Multivariate Software.
- Balcikanli, C. (2011). Inventario de Conciencia Metacognitiva para Docentes (MAIT). *Electronic Journal of Research in Educational Psychology*, 9(3), 1309-1332. doi:10.25115/ejrep.v9i25.1620
- Duffy, G., Miller, S., Parsons, S., & Meloth, M. (2009). Teachers as metacognitive professionals. In D. J. Hacker, J. Dunlosky & A. Graesser (Eds.), *Handbook of metacognition in education* (pp. 240-256). New York: Routledge.
- Flavell, J. (2000). *El desarrollo cognitivo*. Madrid: Visor Dis. S.A.
- Gutiérrez, A. P., Schraw, G., Kuch, F., Richmond, A. S. (2016). A two-process model of metacognitive monitoring: Evidence for general accuracy and error factors. *Learning and Instruction*, 44 1-10. doi: 10.1016/j.learninstruc.2016.02.006.
- Hambleton, R. K., y Bollwark, J. (1991). Adapting tests for use in different cultures: Technical issues and methods. *Bulletin of the International Test Commission*, 18, 3-32.
- Hambleton, R. K. (1996). Adaptación de tests para su uso en diferentes idiomas y culturas: fuentes de error, posibles soluciones y directrices prácticas. En J. Muñiz (Ed.), *Psicometría* (pp. 207-238). Madrid: Universitas.

- Hartman, H. (2002). Metacognition in learning and instruction. Theory, Research and Practice. In H. Hartman (Ed.), *Metacognition in learning and instruction: Theory, research and practice* (p. 125). London: Kluwer Academic.
- Jiang, Y., Ma, L., & Gao, L. (2016). Assessing teachers' metacognition in teaching: The Teacher Metacognition Inventory. *Teaching and Teacher Education*, 59, 403-413. doi:10.1016/j.tate.2016.07.014
- Kallio, H., Virta, K., Kallio, M., Virta, A., Hjärdenaal, F., & Sandven, J. (2017). The utility of the metacognitive awareness inventory for teachers among in-service teachers. *Journal of Education and Learning*, 6(4), 78-91. doi:10.5539/jel.v6n4p78
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: The Guilford Press.
- Lundeberg, M., & Mohan, L. (2009). Gender and cross-cultural differences in confidence. In D. J. Hacker, J. Dunlosky & A. Graesser (Eds.), *Handbook of metacognition in education* (pp. 221-239). New York: Routledge.
- Mai, M. (2015). Science Teachers self perception about metacognition. *Journal of Educational and Social Research*, 5(1), 77-86. doi:10.5901/jesr.2015.v5n1s1p77
- Martínez, M., Hernández, M. J., & Hernández, M. V. (2014). *Psicometría*. Madrid: Alianza Editorial.
- Muñiz, J., Elosua, P., & Hambleton, R. (2013). Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema*, 25(2): 151-157. doi:10.7334/psicothema2013.24
- Muñiz, J., Hernández, A., Ponsoda, V. (2015). Nuevas directrices sobre el uso de los tests: investigación, control de calidad y seguridad. *Papeles del Psicólogo*, 36(3): 161-173
- Muñiz, J., Elosua, P., Padilla, J.-L., & Hambleton, R. K. (2016). Test adaptation standards for cross-lingual assessment. In C. S. Wells & M. Faulkner-Bond (Eds.), *Educational measurement: From foundations to future* (p. 291-304). The Guilford Press.
- Palantis, N., Mohamed, J., Ibrahim, A., Ismail, S., Anuar, N., Ma'rof, A., & Buang, N. (2017). Patterns of metacognitive awareness among primary school teachers. *Jurnal Varidika. Kajian Penelitian Pendidikan*, 29(2), 141-146. Retrieved from <http://journals.ums.ac.id/index.php/varidika/article/view/5629/3675>
- Peña-Ayala, A., & Cárdenas, L. (2015). A conceptual model of the metacognitive activity. In A. Peña-Ayala (Ed.), *Metacognition: Fundamentals, applications and trends. A prolife of the current state -of-the-art* (pp. 39-64). New York: Springer.
- Proust, J. (2010). Metacognition. *Philosophy Compass*, 11, 989-998. doi:10.1111/j.1747-9991.2010.00340.x
- Sawyer, K. (2014). *The Cambridge handbook of the learning sciences*. New York: Cambridge University Press.
- Schraw, G., & Sperling-Dennison, R. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460-475. doi:10.1006/ceps.1994.1033
- Schraw, G., & Gutiérrez, A. (2015). Metacognitive strategy instruction that highlights the role of monitoring and control processes. In A. Peña-Ayala (Ed.), *Metacognition: Fundamentals, applications and trends. A prolife of the current state -of-the-art* (pp. 3-14). New York: Springer.
- Tabachnick, B. G., & Fidell L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.
- Vanags, E., Pestovs, P., & Peršēvica, A. (2018). Metacognitive awareness inventory for teachers (MAIT-18) reliability and validity. Academia de Tecnologías Rezekne (Presidencia). *International Scientific conference 'Society, Integration, Education—SIE2018'*. 'International Scientific Conference'. Rezekne Academy of Technologies, Rezekne, Latvia.
- Werts, C. E., Linn, R. L., & Jöreskog, K. G. (1974). Intraclass reliability estimates: Testing structural assumptions. *Educational and Psychological Measurement*, 34, 25-33. doi: 10.1177%2F001316447403400104
- Zohar, A., & Dori, Y. (2012). *Metacognition in science education. Trends in current research*. New York: Springer.

Apéndice 1. Versión original del MAIT para docentes de habla hispana usada con una muestra de docentes de la ciudad de Manizales. Forma (MAIT-24).

El instrumento MAIT es una lista de 24 declaraciones. No hay respuestas correctas o incorrectas en esta lista de afirmaciones acerca de usted mismo como docente. Sus respuestas expresan simplemente lo que es verdad para usted, o lo que usted considera importante en su ejercicio como docente. Lea cuidadosamente cada declaración y elija la opción que mejor lo describe.

Muchas gracias por su participación.

Opciones de respuesta:

1 = Totalmente en desacuerdo, 2 = En desacuerdo, 3 = Neutral, 4 = De acuerdo, 5 = Muy de acuerdo

1. Soy consciente de las fortalezas y debilidades a la hora de enseñar.	1	2	3	4	5
2. Trato de utilizar técnicas de enseñanza que me funcionaron en el pasado.	1	2	3	4	5
3. Utilizo mis fortalezas para compensar mis debilidades en el proceso de enseñanza	1	2	3	4	5
4. Regulo mi propio ritmo de trabajo para tener el tiempo suficiente.	1	2	3	4	5
5. Periódicamente me preguntó si estoy cumpliendo mis objetivos mientras estoy enseñando.	1	2	3	4	5
6. Me pregunto qué tan bien he logrado mis objetivos de enseñanza, una vez que he terminado.	1	2	3	4	5
7. Sé qué habilidades son las más importantes para poder ser un buen docente.	1	2	3	4	5
8. Tengo una razón específica para elegir cada unas de las técnicas de enseñanza que utilizo en clase.	1	2	3	4	5
9. Puedo motivarme a mí mismo cuando debo orientar una clase.	1	2	3	4	5
10. Establezco mis propias metas de enseñanza antes de comenzar una clase.	1	2	3	4	5
11. Generalmente evalué qué tan útiles son mis técnicas de enseñanza mientras oriento una clase.	1	2	3	4	5
12. Generalmente me pregunto si pude haber utilizado otras técnicas de enseñanza, después de cada clase.	1	2	3	4	5
13. Tengo control sobre qué tan bien enseño.	1	2	3	4	5
14. Soy consciente de qué técnicas de enseñanza utilizo mientras estoy enseñando.	1	2	3	4	5
15. Dependiendo de la situación, utilizo diferentes técnicas de enseñanza.	1	2	3	4	5
16. Me cuestiono acerca del material de enseñanza que voy a utilizar en clase.	1	2	3	4	5
17. Generalmente reviso qué tanto han comprendido mis estudiantes el tema, mientras estoy enseñando.	1	2	3	4	5
18. Después de orientar una clase, me pregunto si podría hacerlo mejor en una próxima ocasión.	1	2	3	4	5
19. Sé qué debo enseñar.	1	2	3	4	5
20. Empleo técnicas de enseñanza útiles automáticamente.	1	2	3	4	5
21. Sé cuándo cada técnica de enseñanza que utilizo será más eficaz.	1	2	3	4	5
22. Organizo mi tiempo para lograr mis metas de enseñanza.	1	2	3	4	5
23. Me hago preguntas acerca de qué tan bien lo estoy haciendo, mientras estoy enseñando.	1	2	3	4	5
24. Me hago preguntas acerca de si he considerado todas las técnicas posibles después de dar una clase.	1	2	3	4	5