

Adaptación de un Instrumento para la Medición de la Convivencia Escolar en Escuelas de Educación Secundaria de México

Adaptation of an Instrument for Measuring School Coexistence in Middle School Students in Mexico

Cristina Vanessa Hernández De la Toba *
Joaquín Caso Niebla

¹ Instituto de Investigación y Desarrollo Educativo
de la Universidad Autónoma de Baja California

El presente estudio tuvo como propósito adaptar a la población mexicana el Cuestionario de Convivencia Escolar (CCE) elaborado por Díaz-Aguado, Martínez Arias y Martín (2010) dirigido a estudiantes del nivel secundaria, mediante el uso de procedimientos para la adaptación de instrumentos de medida (Díaz, 2015; Muñiz, Elosua y Hambleton, 2013); para la optimización de constructos complejos (Jornet, González-Such y Perales, 2012) y para la comprobación de invarianza de medición (Byrne, 2009; Cheung y Rensvold, 1999; Raju, Laffitte y Byrne, 2002). Se describen los resultados de cada una de las fases propuestas, producto de la aplicación de los CCE $k=74$ y CCE $k=39$ en muestras de Baja California $N=660$ y Querétaro $N=419$. El método se constituye de las fases: (a) Analítico-racional y (b) Empírica. A partir de la primera fase, se obtuvo el CCE adaptado $k=74$ con propiedades psicométricas aceptables (0.850 a 0,967); y como resultado de la segunda fase, se logró reducir el CCE a $k=39$ respetando los criterios de consistencia interna referidos por Jornet et al. (2012) y se comprobó la invarianza de la medición para las escalas de Fomento a la Convivencia Escolar, Acoso Escolar y Conductas Disruptivas del CCE adaptado y reducido.

Palabras clave: Adaptación, Optimización, Invarianza de medición, Validez, Constructo.

The present study was aimed to adapt the School Coexistence Questionnaire developed by Díaz-Aguado, Martínez Arias and Martín (2010) for middle school's students using procedures for adapting measuring instruments (Díaz, 2015; Muñiz, Elosua, & Hambleton, 2013); for optimizing complex constructs (Jornet, González-Such, & Perales, 2012) and to test measurement invariance (Byrne, 2009; Cheung & Rensvold, 1999; Raju, Laffitte, & Byrne, 2002). Describes the results obtained in each of the phases proposed in the method, resulting from the application of the CCE $k=74$ and CCE $k=39$ in samples of Baja California $N=660$ and Querétaro $N=419$. The method is made up of the following two phases: (1) Analytic-rational (2) Empiric. At the first phase, it was obtained the adapted CCE $k=74$ with acceptable psychometric properties (0.850 a 0.967); at the second phase, it was reduced the CCE to $K=39$ accordance with the criteria of internal consistency indices referred by Jornet et al. (2012), and it was tested measurement invariance of the School Coexistence's Promotion, Bullying and Disruptive Behavior Scales of adapted and reduced CCE.

Keywords: Adapting, Optimizing, Measurement invariance, Validity, Construct.

*Contacto: cristina.vanessa.hernandez.delatoba@uabc.edu.mx

delatoba@uabc.edu.mx

issn: 1989-0397

www.rinace.net/riee/

<https://revistas.uam.es/riee>

Recibido: 4 de julio de 2016

1ª Evaluación: 14 de septiembre de 2016

Aceptado: 1 de noviembre de 2016

1. Introducción

La convivencia escolar es un tema de interés internacional que está presente en diversos documentos normativos de políticas públicas (Jiménez, 2000). En el caso de México, desde 1982 a la fecha se han expedido acuerdos, leyes y programas ex profeso para el fomento de la convivencia en las escuelas. En los Acuerdos Secretariales 96, 97 y 98, se reglamentó la disciplina escolar para las escuelas primarias, secundarias técnicas y secundarias generales (Secretaría de Educación Pública, 1982). En 1993, se expidió la Ley General de Educación, en la que se cita en su artículo octavo, fracción tercera, que la educación contribuirá a la mejora de la convivencia humana (Secretaría de Educación Pública, 1993); en el 2007 se creó el Programa Nacional Escuela Segura con el propósito de consolidar a las escuelas públicas de educación básica como espacios seguros y confiables a través de la participación social y la formación ciudadana de los alumnos (Secretaría de Educación Pública, 2007). En el 2011 se llevó a cabo la reforma educativa a la educación básica en la que se incluyó el campo formativo de Desarrollo Personal y la Convivencia cuya finalidad es que los alumnos desarrollen el juicio crítico en favor de la democracia, la libertad, la paz, el respeto a las personas, a la legalidad y a los derechos (Secretaría de Educación Pública, 2011). En el 2011 y 2013 se realizaron importantes reformas al artículo tercero Constitucional asociadas con la convivencia en las escuelas (Cámara de Diputados del H. Congreso de la Unión, 2011; 2013). En el mismo 2013 se formuló el Programa Sectorial de Educación 2013-2018 (Secretaría de Educación Pública, 2013b), en el que se destacan líneas de acción asociadas al tema de la convivencia en las escuelas, tales como: la educación integral, la inclusión y la democracia; y por último, en el mismo 2013 la Secretaría de Educación Pública (SEP) puso énfasis en la construcción de ambientes donde los integrantes de la comunidad escolar convivan de manera pacífica, democrática e inclusiva (Secretaría de Educación Pública, 2013a).

Es innegable que el interés por mejorar la convivencia escolar en México es una preocupación sentida por las autoridades político-educativas. Sin embargo, un proceso de mejora implica el uso de la evaluación (Secretaría de Educación Pública, 2013b), puesto que la evaluación constituye una base de información para sensibilizar a la población sobre las condiciones en las que se presenta un fenómeno determinado o en las situaciones en las que se requiere poner atención (Rossi y Freeman, 1993), brinda orientación a los gobiernos para que puedan dirigir y ajustar estratégicamente sus políticas y programas hacia el logro de sus objetivos y metas (Cardozo, 2009; OCDE, 2011; Ponce, 2009).

No obstante, la evaluación de la convivencia es aún rudimentaria por ser un campo del conocimiento en construcción (Fierro et al, 2013) por ser un término utilizado por los actores del ámbito educativo no solamente para referirse a una gran diversidad de cuestiones; sino también, por ser estudiado desde perspectivas distintas (Fierro y Caso, 2013), por encontrarse en colindancia con otros referentes teóricos (Chaparro, Caso, Díaz y Urías, 2012) lo cual lo coloca en un constructo poliédrico, de múltiples caras (Ortega-Ruiz, Del Rey y Casas, 2013). Al respecto, en México se han documentado pocas experiencias de evaluación sobre la convivencia escolar a gran escala (Chaparro, Caso, y Fierro, 2012; Instituto Nacional para la Evaluación de la Educación, 2007, 2015; Ochoa y Salinas de la Vega, 2013; Valadez, 2008).

En este sentido, se reconoce la necesidad de contar con instrumentos de medición que ayuden a caracterizar este fenómeno de forma objetiva (Caso, Díaz y Chaparro, 2013), que permitan hacer distintas aproximaciones desde varias posturas teóricas y desde la perspectiva de diferentes actores (Fierro et al, 2013) y que generen conocimiento relacionado con los procesos implicados en el desarrollo de prácticas de convivencia inclusiva, democrática y no violenta que oriente la evaluación y el autodiagnóstico de escuelas, y que su vez motive el diseño de programas y propuestas de intervención (Fierro, 2011).

Por lo anterior, el objetivo central del presente estudio se basa en adaptar al contexto mexicano el Cuestionario de Convivencia Escolar para estudiantes de secundaria, el cual fue elaborado por el grupo de investigadores españoles integrado por Díaz Aguado, Martínez Arias y Martín (2010) con la mira de ser utilizado en estudios a gran escala, ya que estas evaluaciones tienen un peso cada vez mayor en temas de política nacional (Martínez, 2013). Sin embargo, al caracterizarse tales evaluaciones por recoger una mayor información, se pretende que sean utilizados instrumentos breves (Caso et al., 2013) a fin de simplificar el tiempo, el número de sesiones de la aplicación, el número de no respuestas (López-González, Tourón y Tejedor, 2012), y los efectos adversos como la fatiga y la falta de motivación derivados de un tiempo de administración excesivamente largo (Balluerka y Gorostiaga, 2012).

Bajo estas consideraciones, los objetivos específicos del presente estudio son: (a) Aplicar una metodología para la adaptación de instrumentos al Cuestionario de Convivencia Escolar; (b) Aplicar un conjunto de procedimientos para la optimización de la medida a la adaptación del Cuestionario de Convivencia Escolar; (c) Aportar evidencia de equivalencia métrica del Cuestionario de Convivencia Escolar que permita la comparación entre distintas muestras de México.

2. Método

De acuerdo con los objetivos del estudio, el método quedó conformado por las siguientes fases y etapas: 1) *fase analítico-razional*, misma que integra las etapas de adecuación del contenido, traducción y adecuación del formato, aplicación piloto y análisis psicométrico; 2) *fase empírica*, la cual está compuesta por las etapas de optimización, validación y de confirmación de la invarianza factorial (ver figura 1).

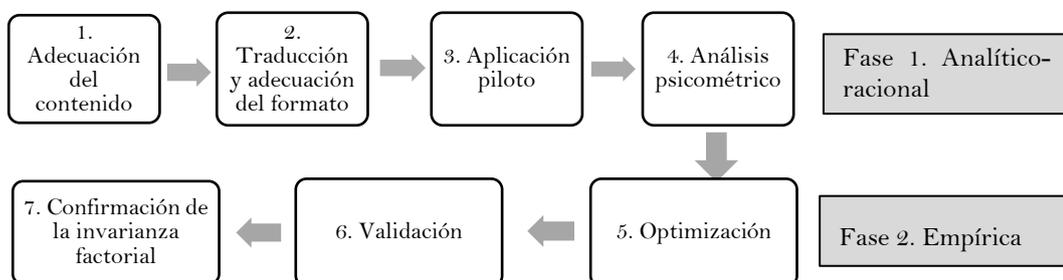


Figura 1. Fases y etapas del método del estudio
Fuente: Elaboración propia.

Las fases y etapas del método se describen a continuación.

2.1. Fase 1. Analítico-racional

La adaptación es un proceso mediante el cual es posible transformar en -otro- un instrumento existente para emplearse en una población particular y ser utilizado de forma equivalente en una población distinta (APA, 2013), bajo el supuesto que el constructo medido por el test orientará hacia las mismas o similares predicciones e interpretaciones en la nueva cultura (Geisinger, 1994; Tanzer y Sim, 1999) y que permitirá realizar mediciones válidas en cada contexto o cultura (Klerk, 2008).

La adaptación es un proceso complejo que se halla expuesto a una gran cantidad de fuentes de error (Hambleton, 1996, 2005; Muñiz y Hambleton, 1996; Sireci, Patsula y Hambleton, 2005; Van de Vijver y Hambleton, 1996; Van de Vijver y Poortinga, 2005) que pueden ser controladas mediante la observancia de un marco de referencia (Cachón, 2007). En este sentido, y con la finalidad de evitar dichas fuentes de error, en esta fase se conjugaron una serie de procedimientos y directrices asociados con procesos de adaptación (Díaz, 2015; Hambleton, 1994, 1996; Muñiz y Hambleton, 1996; Muñiz et al., 2013).

2.1.1. Etapa 1. Adecuación del contenido

La presente fase tuvo como objetivo eliminar a aquellos ítems del CCE original $k=244$ (Díaz-Aguado et al., 2010) que no cumplieran de manera satisfactoria con los criterios de pertinencia, congruencia y relevancia; así como determinar la suficiencia de los ítems para cada una de las escalas propuestas. El criterio de pertinencia hace referencia a la correspondencia conceptual entre los títulos de las dimensiones y los ítems que la integran. La congruencia es concebida como el grado en que el ítem es consecuente con la dimensión correspondiente, mientras que la relevancia se entiende como el grado en que la información contenida en el ítem es significativa para evaluar la dimensión a la que se adscribe; el criterio de suficiencia alude al grado en que se incluyen los ítems necesarios para evaluar cada dimensión.

- ✓ **Participantes.** Cuatro especialistas en temas y contenidos relacionados con la convivencia escolar, la psicometría y el diseño, desarrollo, validación y adaptación de instrumentos de medición psicológica y educativa.
- ✓ **Materiales.** La versión original del CCE para estudiantes (Díaz-Aguado et al., 2010), organizado en cinco dimensiones (ver tabla 1).

Tabla 1. Dimensiones y número de ítems del instrumento CCE original

DIMENSIÓN	K
Calidad de la convivencia escolar	42
Obstáculos a la convivencia en las relaciones entre estudiantes	35
Otros obstáculos a la convivencia	62
Condiciones del centro educativo para construir y mejorar la convivencia	37
Otros indicadores	68
<i>Total</i>	<i>244</i>

Fuente: Elaboración propia.

- ✓ **Procedimiento.** Bajo el juicio de expertos a través del método de consenso (Escobar-Pérez y Cuervo-Martínez, 2008), las decisiones respecto a los criterios de pertinencia, congruencia, relevancia y suficiencia de los ítems de cada escala, fueron tomadas con base en el criterio referido por Herrera Rojas (1993), que existiera acuerdo entre más del 60% de los participantes.

2.1.2. Etapa 2. Traducción y adecuación del formato

Esta etapa tuvo como propósito transformar las palabras y expresiones de los ítems, opciones de respuesta e instrucciones que conforman el instrumento al contexto mexicano; así como modificar el formato en línea a formato de lápiz y papel debido a la limitación de los recursos informáticos. Las tareas realizadas fueron: (a) identificar los ítems expresados en forma de pregunta para ser transformados en afirmaciones; (b) asegurar la congruencia entre la base del ítem y sus opciones de respuesta; (c) valorar la estructura gramatical de los ítems originales de acuerdo al contexto de México y corregirlos; y (d) verificar la semántica de los ítems modificados. Respecto a la estructura gramatical, se valoró que el ítem se encontrara libre de errores gramaticales y que la sintaxis del enunciado fuera simple y común para la población destino. En relación a la semántica, se evaluó que las ideas y los significados transferidos al ítem traducido fueran iguales a los del ítem en el idioma fuente.

- ✓ **Participantes.** Se continuó con la colaboración del grupo de trabajo conformado en la etapa anterior considerando el conocimiento que poseían del instrumento en cuestión y por las limitaciones temporales que imperaban en el desarrollo del proyecto para conformar un nuevo comité. Aunque algunos autores señalan la conveniencia de expertos en lingüista para llevar a cabo las tareas indicadas en esta etapa (Costa y Brito, 2002; Díaz, 2015; Gaité, Ramírez, Herrera y Vázquez-Barquero, 1997; Muñiz et al., 2013; Ramada, Serra y Delclós, 2013; Villalobos, Arévalo y Rojas, 2012), el propósito de la misma no se vio afectado.
- ✓ **Materiales.** Se utilizó la versión del instrumento generada en la etapa anterior.
- ✓ **Procedimiento:** Se continuó con el juicio del grupo de expertos a través del método de consenso (Escobar-Pérez y Cuervo-Martínez, 2008). En todos los casos, las decisiones se tomaron al existir acuerdo entre más del 60% de los participantes (Herrera Rojas, 1993).

2.1.3. Etapa 3. Aplicación

El objetivo de esta etapa fue valorar el funcionamiento inicial de las escalas, ítems, instrucciones y opciones de respuesta.

- ✓ **Participantes.** Una muestra de 660 estudiantes del tercer grado de secundaria de los Municipios de Ensenada, Mexicali y Tijuana del Estado de Baja California, seleccionados de manera no probabilística (ver tabla 2).
- ✓ **Materiales.** Debido a que el Cuestionario de Convivencia Escolar formó parte de una estrategia evaluativa más amplia a cargo de la Unidad de Evaluación Educativa del Instituto de Investigación y Desarrollo Educativo de la UABC, denominada “Estrategia Evaluativa 2015, Factores asociados al aprendizaje” (Rodríguez, en prensa) se utilizaron los siguientes materiales: (a) Cuadernillo que contiene la batería de instrumentos incluida la nueva versión del

Cuestionario de Convivencia Escolar k=74; (b) Hojas de respuesta; (c) Hoja de incidencias.

- ✓ **Procedimiento.** Se diseñó y utilizó un procedimiento para cuidar la estandarización de la aplicación.

Tabla 2. Participantes de Baja California

		N	%
SEXO	Masculino	314	47,6
	Femenino	346	52,4
MUNICIPIO	Ensenada	105	15,9
	Tijuana	320	48,5
	Mexicali	235	35,6
MODALIDAD	General	234	35,5
	Privada	321	48,6
	Técnica	27	4,1
	Telesecundaria	78	11,8

Fuente: Elaboración propia.

2.1.4. Etapa 4. Análisis psicométrico (de los coeficientes de confiabilidad)

El objetivo de esta etapa fue obtener evidencias de confiabilidad del instrumento en cuestión.

- ✓ **Participantes.** Misma muestra de estudiantes de la etapa anterior.
- ✓ **Materiales.** (a) Hojas de respuestas; (b) Hoja de cálculo del programa Microsoft Excel. (c) Programas de análisis estadístico *Statistical Package for the Social Sciences* (SPSS) versión 19 y Programa R: *Project for Statistical Computing* versión 3.1.1.
- ✓ **Procedimiento.** Mediante el uso de los programas SPSS y R se realizaron los análisis de confiabilidad (alpha de Cronbach) y de discriminación (correlación punto biserial).

2.2. Fase 2. Empírica

2.2.1. Etapa 5. Optimización de la medida

El objetivo de esta etapa fue reducir la extensión del instrumento conservando los ítems clave que aportan mayor información respecto al constructo medido, aplicando el procedimiento propuesto por Jornet et al. (2012). Este procedimiento consiste en conservar sólo a los ítems clave portadores de información suficiente que permiten medir un determinado constructo a través de una escala con un elevado grado de confiabilidad y validez, equivalente a la que se obtendría con la aplicación total del cuestionario o de la escala original. Aunque tal procedimiento consta de seis pasos, sólo fueron aplicados los primeros cuatro por considerar que de continuar con la reducción del instrumento podría afectarse el entramado conceptual. Los pasos aplicados fueron: 1) Exploración del comportamiento de los reactivos de la escala original; 2) Análisis de valores perdidos y decisiones sobre imputación; 3) Selección de reactivos a partir de su contribución a la confiabilidad de la escala; y 4) Estimación del puntaje total de la versión reducida.

2.2.2. Etapa 6. Validación

El objetivo de esta etapa fue obtener evidencias de validez de constructo del micro-instrumento (versión reducida del instrumento), ya que es mediante la validez que se prueban las hipótesis acerca de relaciones teóricas (Messick, 1980) y se obtienen las evidencias que soportan las inferencias basadas en los puntajes obtenidos en el test (Contreras, 2000).

- ✓ **Participantes.** Misma muestra de la etapa tres.
- ✓ **Procedimiento.** Haciendo uso de la base de datos del programa SPSS de la muestra de Baja California, se contrastaron los modelos de medida construidos por el equipo de investigadores españoles aplicando la técnica de análisis factorial confirmatorio (AFC) por ser una técnica utilizada por excelencia para la validación de constructos (Pérez-Gil, Chacón y Moreno, 2000).

2.2.3. Etapa 8. Determinación de la invarianza factorial

El objetivo de esta etapa fue asegurar la invarianza o equivalencia factorial del CCE adaptado y reducido, a través de dos muestras -Baja California y Querétaro-. Este procedimiento, invarianza o equivalencia factorial, fue elegido por ser el indicado cuando las muestras pertenecen a diversas poblaciones (Gunnell, Wilson, Zumbo, Mack y Crocker, 2012).

- ✓ **Materiales.** (a) CCE adaptado y reducido; (b) hoja de respuestas, (c) procedimiento para la aplicación del cuestionario adaptado del procedimiento utilizado en la Estrategia Evaluativa 2015 de la Unidad de Evaluación Educativa (UEE-UABC) a fin de garantizar la estandarización de la aplicación; (d) reporte de incidencias; y (e) base de datos de la muestra de Baja California.
- ✓ **Participantes.** Misma muestra utilizada en la etapa tres, y una muestra de estudiantes del tercer grado del Estado de Querétaro N=419, (ver tabla 2).

Tabla 3. Participantes de Querétaro

		N	%
SEXO	Masculino	203	48,4
	Femenino	216	51,6
MUNICIPIO	San Joaquín	144	34,4
	Jalpán de Serra	77	18,4
	Vizcarrón Cadereyta de Montes	96	22,9
	Ahuacatlán de Guadalupe	102	24,3

Fuente: Elaboración propia.

- ✓ **Procedimiento.** Se utilizó la técnica de análisis factorial confirmatorio multigrupo (AFCMG) ya que facilita la investigación del grado en que las medidas son invariantes entre los grupos (Chen, 2008). En concreto, esta técnica permite verificar si los niveles de las variables observadas de la variable latente entre los grupos tienen los mismos puntajes esperados en la medida (Drasgow y Kanfer, 1985).

La prueba de la invarianza de la medida consistió en la comparación de los índices de ajuste los modelos: base (*configural*), débil (*weak o metric*), fuerte (*strong o scalar*), y estricto (*strict*) aplicando restricciones de igualdad (Byrne, 2009; Cheung y Rensvold, 1999; Raju, Laffitte y Byrne, 2002) para cada una de las

escalas del CCE adaptado. Estos índices de ajuste se obtuvieron aplicando el AFCMG en cada uno de los modelos con apego a las siguientes restricciones sugeridas por (Hirschfeld y Von Brachel, 2014):

- El modelo base o *configural* se mantuvo libre de restricciones. En este modelo, se supone que el patrón de cargas es similar en todos los grupos, pero las magnitudes de los parámetros (cargas, interceptos, varianzas) pueden variar;
- En el modelo débil (*weak o metric*) se restringieron las cargas factoriales a ser iguales;
- En el modelo fuerte (*strong o scalar*) se restringieron las cargas factoriales y los interceptos;
- Dado que el modelo estricto (*strict*) exige que las cargas factoriales, los interceptos y las varianzas residuales deben ser restringidas y que la invarianza estricta requiere del uso de la theta-parametrización para identificar los parámetros del modelo (Hirschfeld y Von Brachel, 2014), éste no se configuró debido a que la identificación de los parámetros que ofrece la versión del programa R utilizado, era distinta a la sugerida por el autor.

Al realizar el AFCMG en cada uno de los modelos de las escalas se obtuvieron los índices de ajuste absolutos y comparativos RMSEA ($\leq 0,06$), CFI ($\geq 0,95$) y TLI ($\geq 0,95$) (Abad, Olea, Ponsoda, y García, 2011; Hu y Bentler, 1999).

Una vez confirmados los índices de ajuste de los modelos, se obtuvieron nuevamente los índices de ajuste de los modelos anidados (base, débil y fuerte), permitiendo la contrastación de los índices de ajuste (CFI y RMSEA) y la determinación de la invarianza mediante la diferencia entre los índices de ajuste $\leq 0,010$ (Hirschfeld y Von Brachel, 2014).

3. Resultados

A continuación se presentan los resultados obtenidos en cada una de las fases y etapas correspondientes del presente estudio.

3.1. Fase 1. Analítico-racional

3.1.1. Etapa 1. Adecuación del contenido

Al realizar el análisis de las dimensiones del CCE original mediante el criterio de pertinencia, se observó que al interior de las dimensiones se aglomeraban distintos constructos. Por lo que, bajo la definición de convivencia escolar como un conjunto de relaciones entre los actores de una comunidad educativa -alumnos, profesores, directivos y padres de familia- (Bazdresch, 2009; Díaz-Aguado, s.f.; Fernández, 1998; Fierro et al., 2013; Tuvilla, s.f.) que implican la comprensión de las diferencias, el aprecio a la interdependencia y la pluralidad, el aprendizaje para enfrentar los conflictos de una manera positiva y la promoción del entendimiento mutuo y la paz mediante la participación democrática (Carbajal, 2013), se optó por conservar de las dimensiones originales del CCE sólo cuatro rasgos para su operacionalización agrupados en dimensiones y subdimensiones (ver tabla 4). Con base en tal decisión, el comité juzgó

que la mayoría de los ítems no formaban parte del constructo - los cuatros rasgos - a medir y que además, otra buena parte aunque sí correspondían a la dimensión, había ítems suficientes que permitían operacionalizarlo. Por lo tanto, de los 244 ítems que conformaban el CCE original, se eliminaron 172, 121 por no cumplir con el criterio relevancia y 51 por no apegarse al criterio de congruencia, conservándose 72 de los ítems originales (ver tabla 5). En el caso de las dimensiones de Clima Escolar y Acoso Escolar se consideró la carencia de elementos del constructo indispensables. Así que, de acuerdo con el criterio de suficiencia, se determinó la inclusión de dos ítems (ver tabla 4).

Tabla 4. Escalas del instrumento original y del instrumento modificado

ORIGINAL		ADAPTADA		
K	Escala/dimensión	K	Escala/dimensión	Subdimensión
42	Calidad de la convivencia escolar	13 (+1)	Clima escolar	Relación entre profesores y estudiantes Relación entre la familia y la escuela
35	Obstáculos a la convivencia en las relaciones entre estudiantes	31 (+1)	Acoso escolar	Victimización de acoso escolar Participación en acoso escolar
62	Otros obstáculos a la convivencia	16	Conductas disruptivas	Conductas disruptivas en el aula Conductas agresivas entre profesores y estudiantes
37	Condiciones del centro educativo para construir y mejorar la convivencia	14	Fomento a la convivencia escolar	Respeto a la diversidad Enseñanza de preceptos de convivencia escolar
68	Otros indicadores	Se eliminó		

Nota: +1= inclusión de nuevo reactivo.

Fuente: Elaboración propia.

Si bien, la reducción del instrumento original fue considerable, no se descarta que se trata de un proceso de adaptación bajo la categoría de ensamblaje. Según Tanzer y Sim (1999) y Van de Vijver y Hambleton (1996), el ensamblaje es un enfoque de adaptación que equivale a compilar un nuevo instrumento por haber sido modificado a - profundidad- y en el que se resalta la pertinencia cultural del instrumento, al considerar aspectos del constructo que son importantes para la cultura objetivo pero no para el instrumento original.

Tabla 5. Cantidad de ítems eliminados por no reunir los criterios de relevancia y congruencia

DIMENSIÓN	ÍTEMS ELIMINADOS			Total
	k original	No relevancia	No congruencia	
Calidad de la convivencia escolar/Clima Escolar	42	28	2	30
Obstáculos a la convivencia en las relaciones entre estudiantes/Acoso Escolar	35	3	2	5
Otros obstáculos a la convivencia/Conductas Disruptivas	62	34	12	46
Condiciones del centro educativo para construir y mejorar la convivencia/Fomento a la convivencia escolar	37	16	7	23
Otros indicadores	68	40	28	68
<i>Total</i>	<i>244</i>	<i>121</i>	<i>51</i>	<i>172</i>

Fuente: Elaboración propia.

3.1.2. Etapa 2. Traducción y adecuación de formato

El comité observó que algunos de los ítems estaban expresados en forma de pregunta, lo que derivó en la modificación del ítem-pregunta a ítem-afirmación. También se identificó que todos los ítems eran objeto de modificación en cuestiones de gramática y semántica de acuerdo con el idioma castellano de México (ver tabla 6). En la tabla 7 se presentan algunos ejemplos de los cambios señalados.

Tabla 6. Número de ítems modificados

DIMENSIÓN	ÍTEMS MODIFICADOS			Total
	Por expresarse en pregunta	Por aspectos gramaticales	Por aspectos semánticos	
Clima escolar	0	12	2	12
Acoso escolar	8	22	3	30
Conductas disruptivas	0	20	0	20
Fomento a la convivencia	0	10	2	10
<i>Total</i>	<i>8</i>	<i>64</i>	<i>7</i>	<i>72</i>

Fuente: Elaboración propia.

Tabla 7. Ítems modificados por expresarse en pregunta, por aspectos gramaticales y semánticos

ÍTEM ORIGINAL	EXPRESADO EN PREGUNTA	ASPECTO GRAMATICAL	ASPECTO SEMÁNTICO	ÍTEM MODIFICADO
Los estudiantes se ayudan entre sí, aunque no sean amigos		1		Los estudiantes de esta escuela se ayudan entre sí, aunque no sean amigos(as)
Hay peleas entre estudiantes		1		En esta escuela es común que se peleen los estudiantes
Los estudiantes que pertenecen a distintos grupos o pandillas se llevan bien		1	1	Los estudiantes de esta escuela se llevan bien sin importar que pertenezcan a diferentes grupos
Hay bandas en el centro que perjudican la convivencia		1	1	En esta escuela hay estudiantes que afectan la convivencia.
Tiene manía a algunos estudiantes		1	1	A los profesores les caen mal algunos estudiantes
¿Has recibido mensajes a través de Internet o de teléfono móvil en los que te insultaran, amenazaran, ofendieran o asustaran?	Sí			Mis compañeros(as) me envían mensajes a través de internet o teléfono celular en los que me insultan u ofenden
¿Han difundido fotos o imágenes tuyas por Internet o teléfono móvil para Obligarte a hacer después algo que no querías con amenazas?	Sí			Mis compañeros(as) han difundido videos o imágenes mías por Internet o teléfono celular para utilizarlas en mi contra

Fuente: Elaboración propia.

En relación a la congruencia entre instrucciones-ítem-opciones de respuesta, debido a las modificaciones antes realizadas, fue necesario cambiar todas las opciones de respuesta e instrucciones del cuestionario.

3.1.3. Etapa 3 y 4. Aplicación y análisis de los coeficientes de confiabilidad del instrumento

La confiabilidad se ha utilizado para referirse a los coeficientes de confiabilidad de la Teoría Clásica de los Test (TCT) entendidos entre otros, como la consistencia interna del conjunto de ítems, si se analiza el grado en que los distintos ítems miden cierto constructo (American Education Research Association, American Psychological Association, National Council on Measurement in Education, 2014). Cuando un test es traducido o adaptado para su uso en un nuevo grupo lingüístico o cultural, los índices de confiabilidad deben calcularse y estudiarse (American Education Research Association et al., 2014; Geisinger, 1994).

A partir de la aplicación piloto, se calcularon los coeficientes de consistencia interna (alpha de Cronbach) para cada una de las escalas, obteniendo puntuaciones que van de 0.850 a 0.967, y se obtuvieron los índices de correlación $r_{bis} \geq 0.20$ (ver tabla 8). A partir de los criterios de Contreras (2000), $r_{bis} \geq 0.20$ y George y Mallery (2003), alfa de Cronbach $>.9$ excelente y $>.8$ bueno, los valores resultantes se consideraron como aceptables.

Tabla 8. Índices de consistencia interna y de discriminación de las escalas

ESCALAS/DIMENSIONES	K	ALPHA DE CRONBACH	$R_{bis} \leq 0.20$
Clima escolar	13	0,850	0
Acoso Escolar	31	0,967	0
Conductas disruptivas	20	0,932	0
Fomento a la convivencia escolar	10	0,884	0

Fuente: Elaboración propia.

3.2. Fase 2. Reductiva

3.2.1. Etapa 5. Optimización de la medida

Paso 1. Se identificaron los ítems que presentaban escasa variabilidad eliminándose aquellos elementos con una distribución atípica. Como criterio de exclusión se estableció una variabilidad mayor a 80%, por lo que se eliminaron siete ítems, tres de la Escala de Clima Escolar y cuatro de la Escala de Conductas Disruptivas (ver tabla 9).

Paso 2. En esta fase se analizaron los valores perdidos con la perspectiva de que su existencia fuera de carácter aleatorio y no a un patrón de no respuestas sistemático. En las variables o ítems donde se registraron valores perdidos menores al 30%, fueron sustituidos por la mediana, tal como lo sugirieron Jornet et al. (2012). Cabe mencionar, que en ninguno de los casos se observó que los valores perdidos excedieran el límite permitido.

Paso 3. En esta fase se buscó la reducción de los ítems a partir de la identificación de los ítems redundantes. Para poder identificarlos, se buscó a aquellos ítems con una correlación mayor a 0,50. Atendiendo estos criterios se eliminaron 28 ítems por presentar una correlación $\geq 0,50$. A pesar de haber identificado otros ítems que se apegaban al criterio se determinó conservarlos debido a que al hacerlo se afectaba la definición del constructo. El presente análisis dejó como saldo la eliminación de dos ítems de la Escala de Clima Escolar, 16 de la Escala de Acoso Escolar, seis de la escala

de Conductas Disruptivas y cuatro de la Escala de Fomento a la Convivencia Escolar (ver tabla 9).

Tabla 9. Optimización de la medida de las escalas

ESCALA CLIMA ESCOLAR					
Versión	K	alfa	rbis	ítems eliminados	Criterios de exclusión
1	13	0,8500	1	---	---
2	10	0,7800	0,970	3	variabilidad > 80%
3	8	0,8200	0,938	2	$r \geq .5$ y análisis conceptual
ESCALA ACOSO ESCOLAR					
Versión	K	alfa	rbis	ítems eliminados	Criterios de exclusión
1	31	0,967421	1	---	---
2	15	0,9502573	0,970	16	$r \geq .5$ y análisis conceptual
ESCALA CONDUCTAS DISRUPTIVAS					
Versión	K	alfa	rbis	ítems eliminados	Criterios de exclusión
1	20	0,9317074	1	---	---
2	16	0,9225304	0,983	4	variabilidad > 80%
3	10	0,8948949	0,967	6	$r \geq .5$ análisis conceptual
ESCALA FOMENTO A LA CONVIVENCIA ESCOLAR					
Versión	K	alfa	rbis	ítems eliminados	Criterios de exclusión
1	10	0,8837126	1	---	---
2	6	0,713786	0,896	4	$r \geq .5$ análisis conceptual

Nota: --- No aplica: r= correlación; rbis=correlación punto biserial.

Fuente: Elaboración propia.

Fase 4 y 5. Se estimó la puntuación total del cuestionario mediante la suma total de las puntuaciones de los ítems y se exploró la calidad de la versión dos a través del índice de consistencia interna en cada una de las escalas a fin de medir el tamaño de la pérdida en confiabilidad de la versión obtenida de la fase anterior, cuidando que ésta no excediera el 10% del alfa inicial - pérdida del alfa $\leq 10\%$ - (ver tabla 9).

Fase 6. En la metodología propuesta por Jornet et al. (2012) se sugiere esta segunda etapa de reducción. Sin embargo, se determinó obviarla debido a que se consideró que en caso de continuar con la reducción del cuestionario se podría afectar su entramado conceptual.

3.2.2. Etapa 6. Validación

Se aplicó el análisis factorial confirmatorio (AFC) utilizando el método de estimación de mínimos cuadrados no ponderados (ULS) sobre la matriz de correlaciones policórica, por motivo de la naturaleza ordinal de las variables observadas (Luo, 2011). A partir del AFC y en consideración a lo expuesto por Kim y Bentler (2006), se obtuvieron los índices de ajuste absoluto SRMR (*Standardized Root Mean Square Residual*) y el RMSEA (*Root Mean Square error approximation*), así como los índices de ajuste comparativo CFI (*Comparative Fit Index*) y TLI (*Índice de Tucker-Lewis*). El ajuste del modelo se determinó con base en los índices de bondad de ajuste absoluto SRMR ($\leq 0,08$) y RMSEA ($\leq 0,06$) y los índices de bondad de ajuste comparativo CFI ($\geq 0,95$) y TLI ($\geq 0,95$) (Abad, et al., 2011; Hu y Bentler, 1999).

Como se observa en la tabla 10, los índices para cada una de las escalas: Clima Escolar, Acoso Escolar, Conductas Disruptivas y Fomento a la Convivencia Escolar reflejan un buen ajuste en cada uno de los modelos [CFI ($\geq 0,95$) y TLI ($\geq 0,95$); SRMR ($\leq 0,08$); y RMSEA $\leq 0,07$]; aunque en las escalas de Acoso Escolar y Fomento a la Convivencia Escolar el RMSEA no haya resultado significativo. Al respecto, Kim y Bentler (2006)

sostienen que la valoración de ajuste de los datos puede realizarse mediante el uso de por lo menos dos tipos de índices diferentes (absoluto, incremental o de parsimonia), por ejemplo TLI y SRMR, RMSEA y SRMR o CFI y SRMR (Hu y Bentler, 1999).

Tabla 10. Índices de ajuste de las escalas del Cuestionario Convivencia Escolar adaptado y reducido

ÍNDICE DE AJUSTE	CLIMA ESCOLAR	ACOSO ESCOLAR	CONDUCTAS DISRUPTIVAS	FOMENTO A LA CONVIVENCIA ESCOLAR
Chi-cuadrado	172,487	180,500	127,292	2,951
Grados de libertad	19	89	34	4
CFI	0,950	0,997	0,987	1,000
TLI	0,926	0,996	0,983	1,001
SRMR	0,020	0,048	0,059	0,017
RMSEA	0,111	0,039	0,065	0,000
Intervalo de confianza	0,096-0,126	0,031-0,048	0,053-0,077	0,000-0,051
Valor p RMSEA	0,000	0,983	0,022	0,944

Fuente: Elaboración propia.

3.3. Fase 3. Equivalencia

3.3.1. Etapa 5. Determinación de la invarianza factorial

Con la finalidad de asegurar la invarianza o equivalencia factorial de las escalas entre las muestras de Baja California y Querétaro se generó una sola base de datos, misma que permitió el cálculo de los índices de ajuste absoluto, así como los índices de ajuste de los modelos anidados. Para determinar el ajuste de los modelos, se consideraron ajustes de tipo exacto o absoluto, (chi-cuadrada no significativa) (Hirschfeld y Von Brachel, 2014; Magnus, 2005) y ajustes de tipo aproximado o comparativo ($CFI \geq 0,95$, $TLI \geq 0,95$ y $RMSEA \leq 0,06$) (Abad et al., 2011; Hu y Bentler, 1999). No obstante, debido a que la chi cuadrada es sensible al tamaño de la muestra (Byrne y Van de Vijver, 2010), ésta fue utilizada con precaución.

Para el caso de la Escala de Clima Escolar, se observa que los índices de ajuste absolutos para los modelos *configural* y *weak* fueron aceptables. Sin embargo, aunque en el modelo *strong* el RMSEA estuvo muy por encima de los valores esperados, se obtuvo un $TLI \geq 0,95$ y un $CFI \geq 0,99$ el cual puede interpretarse como un criterio confiable (Hu y Bentler, 1999). Respecto a la Escala Fomento a la Convivencia Escolar, todos los índices de ajuste absoluto de los modelos (*configural*, *weak* y *strong*) resultaron aceptables, con excepción de la chi-cuadrada (ver tabla 11).

En la escala de Acoso Escolar, en los tres modelos el CFI fue aceptable y el TLI en los modelos *weak* y *strong* estuvieron dentro del rango ($\geq 0,95$), con excepción del modelo *configural*. En el caso de RMSEA, el modelo *configural* y el *strong* se mantuvieron en el índice esperado ($\leq 0,06$), con la excepción del modelo *weak*. Sin embargo, a pesar de los contrastes entre los índices resultantes, se considera que los modelos tienen un buen ajuste, ya que el CFI en todos los casos fue aceptable (Hu y Bentler, 1999); y por último, en la escala de Conductas Disruptivas, los índices de ajuste en los tres modelos del CFI y TLI fueron aceptables ($\geq 0,95$), y no para el caso del RMSEA, ya que en todos los modelos se obtuvieron índices superiores a lo esperado ($\leq 0,06$), ver tabla 11.

Tabla 11. Índices de ajuste absoluto de los modelos de las escalas del Cuestionario

ESCALA	MODELOS	CHISQ	DF	P	CFI	TLI	RMSEA
Clima Escolar	Configural	55,972	22	0,000	0,993	0,982	0,054
	Weak	67,707	35	0,001	0,993	0,989	0,042
	Strong	82,799	48	0,001	0,993	0,993	0,992
Fomento a la Convivencia Escolar	Configural	18,001	8	0,021	0,998	0,995	0,048
	Weak	16,18	11	0,135	0,999	0,998	0,030
	Strong	23,298	19	0,224	0,999	0,999	0,020
Acoso Escolar	Configural	955,522	178	0,000	0,955	0,946	0,090
	Weak	903,669	191	0,000	0,958	0,954	0,954
	Strong	949,942	219	0,000	0,957	0,959	0,079
Conductas Disruptivas	Configural	354,059	68	0,000	0,991	0,988	0,088
	Weak	409,304	76	0,000	0,989	0,987	0,090
	Strong	392,844	94	0,000	0,990	0,991	0,077

Nota: chisq=chi cuadrada; df=grados de libertad; p=nivel de significancia; CFI=Comparative Fit Index; TLI=Índice de Tucker-Lewis; RMSEA=Root Mean Square error approximation.

Fuente: Elaboración propia.

Una vez asegurados los índices de ajuste absoluto, se procedió al cálculo de los ajustes de los modelos anidados. En el caso de la Escala de Clima Escolar, para los modelos *weak* y *strong* se obtuvieron puntajes aceptables en todos los índices y una chi-cuadrada no significativa. Cabe mencionar que aunque en el modelo *configural*, el CFI resultó muy por debajo del puntaje esperado, el resto de los índices se mantuvieron en el rango esperado. Para el caso de la Escala de Fomento a la Convivencia, todos los índices resultantes se encuentran dentro de los criterios establecidos (ver tabla 12).

Respecto a la Escala de Acoso Escolar, se observa que en los tres modelos el CFI y TLI resultaron aceptables. En el caso del RMSEA sólo el modelo *strong* estuvo dentro del criterio (≤ 0.06); y en el caso de la Escala de Conductas Disruptivas, al igual que la escala anterior, los índices de ajuste de los modelos anidados resultaron aceptables para el CFI y TLI (≥ 0.95), mientras que el RMSEA se ubicó fuera de los parámetros establecidos para los tres casos restantes (ver tabla 12).

Tabla 12. Índices de ajuste de los modelos anidados de las escalas de Cuestionario

ESCALA	MODELOS	CHISQ	DF	P	CFI	TLI	RMSEA
Clima Escolar	Configural	24,620	22	0,316	0,000	0,999	0,015
	Weak	43,488	35	0,154	0,999	0,998	0,021
	Strong	52,057	48	0,319	0,999	0,999	0,013
Fomento a la Convivencia Escolar	Configural	6,439	8	0,598	1,000	1,001	0,000
	Weak	7,46	11	0,761	1,000	1,001	0,000
	Strong	11,844	19	0,892	1,000	1,001	0,000
Acoso Escolar	Configural	649,193	178	0,000	0,991	0,989	0,070
	Weak	772,487	191	0,000	0,989	0,988	0,075
	Strong	699,008	219	0,000	0,991	0,991	0,064
Conductas Disruptivas	Configural	354,059	68	0,000	0,991	0,988	0,088
	Weak	409,304	76	0,000	0,989	0,987	0,090
	Strong	392,844	94	0,000	0,990	0,991	0,077

Nota: chisq=chi cuadrada; df=grados de libertad; p=nivel de significancia; CFI=Comparative Fit Index; TLI=Índice de Tucker-Lewis; RMSEA=Root Mean Square error approximation.

Fuente: Elaboración propia.

Una vez analizados y aceptados los índices de ajuste anidados según los rangos esperados, se procedió a calcular las diferencias entre dichos modelos buscando asegurar la evidencia de invarianza entre el modelo configural y el débil o métrico, ya que dicha

evidencia es fundamental para establecer la invarianza de medición y la interpretación conceptual de un instrumento (Pickering y Plotnikoff, 2009).

Con base lo anterior, en la Escala de Clima Escolar se observó que la diferencia del CFI entre el modelo *configural-weak*, no presentó índices de ajuste aceptables, mientras que la diferencia entre los modelos *weak-strong* resultó aceptable (CFI.delta \leq 0,010). Para el caso de la Escala de Fomento a la Convivencia Escolar, la diferencia entre los modelos *configural-weak* y *weak-strong*, resultaron dentro de los parámetros establecidos (ver tabla 13).

Tabla 13. Comparación de modelos anidados de la Escala Clima Escolar

ESCALA	MODELOS	CHISQ. DELTA	DF. DELTA	P. DELTA	CFI. DELTA	TLI. DELTA	RMSEA. DELTA
Clima Escolar	Configural-Weak	18,868	13	-0,162	0,999	-0,001	0,006
	Weak-Strong	8,569	13	0,165	0,000	0,001	-0,008
Fomento a la Convivencia Escolar	Configural-Weak	1,021	3.000	0,163	0,000	0,000	0,000
	Weak-Strong	4,384	8.000	0,131	0,000	0,000	0,000
Acoso Escolar	Configural-Weak	123,294	13	0,000	-0,002	-0,001	0,005
	Weak-Strong	-73,479	28	0,000	0,002	0,003	-0,011
Conductas Disruptivas	Configural-Weak	55,245	8	0,000	-0,002	-0,001	0,002
	Weak-Strong	-16,46	18	0,000	0,001	0,004	-0,013

Nota: chisq.delta=diferencia entre la chi cuadrada; df.delta=diferencia entre los grados de libertad; p.delta=diferencia entre el nivel de significancia; CFI.delta=diferencia entre el *Comparative Fit Index*; TLI.delta= diferencia entre el *Índice de Tucker-Lewis*; RMSEA.delta= diferencia entre el *Root Mean Square error approximation*.

Fuente: Elaboración propia.

Acerca de la Escala de Acoso Escolar, en la tabla 13 se observa que las diferencias entre los modelos *configural-weak* y *weak-strong* se encuentran dentro del rango establecido CFI.delta (\leq 0,010); y por último, en la Escala de Conductas Disruptivas, una vez realizada la comparación de las diferencias del CFI.delta entre los modelos *configural-weak* y *weak-strong*, se observa que ambas comparaciones permiten asegurar la equivalencia de la medida (CFI.delta \leq 0,10), ver tabla 13.

4. Discusión

El objetivo central de este trabajo ha consistido en mostrar la aplicación de metodologías para la adaptación, optimización y validación del Cuestionario de Convivencia Escolar (Díaz-Aguado, et al., 2010). Con base en los principales resultados, es posible decir que mediante el procedimiento de adaptación aplicado, se logró obtener una versión del Cuestionario de Convivencia Escolar (adaptado) k=74, útil para caracterizar muestras de estudiantes en el contexto de Baja California, puesto que los índices de consistencia interna van desde un nivel bueno a excelente (de 0,850 a 0,967), según lo señalado George y Mallery (2003). De la misma manera, el procedimiento de optimización de la medida propuesto por Jornet, et al. (2012) resultó óptimo, ya que permitió reducir el Cuestionario a k=39 manteniendo los índices de consistencia interna dentro del rango esperado (\leq 10% al índice inicial).

Bajo estos resultados favorables, y con la finalidad de contar con un instrumento para ser aplicado no sólo en Baja California sino en nuestro país, el Cuestionario k=39 fue sometido a un procedimiento de invarianza de medición. Con base en los resultados

obtenidos en la fase de validación empírica, es posible asegurar que en el caso de las Escalas de Fomento a la Convivencia Escolar, Acoso Escolar y Conductas Disruptivas que integran el Cuestionario de Convivencia Escolar, pueden ser utilizadas en estudios a gran escala para describir de manera general la situación que se vive en las escuelas mexicanas respecto al tema de la convivencia. Tal aseveración, se sustenta en que las diferencias entre los modelos restringidos *configural-weak* se encuentran dentro del rango establecido CFI.delta ($\leq 0,010$), evidencia fundamental para establecer la invarianza de medición y la interpretación conceptual de un instrumento (Pickering y Plotnikoff, 2009). No obstante, se sugiere utilizar con mayor cautela los resultados que se obtengan a partir de la Escala de Clima Escolar, ya que aunque la diferencia entre los modelos *weak-strong* se mantuvo dentro del parámetro señalado, no fue el caso para los modelos *configural-weak*.

No obstante con todo lo alentador de los resultados obtenidos, queda aún pendiente por realizar la aplicación de ambos instrumentos $k=74$ y $k=39$ en muestras de estudiantes más amplias, puesto que resulta de gran relevancia el documentar sus propiedades psicométricas.

Referencias

- Abad, F., Olea, J., Ponsoda, V. y García, C. (2011). *Medición en Ciencias Sociales y de la Salud*. Madrid: Síntesis.
- American Education Research Association, American Psychological Association, National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, D.C.: American Education Research Association.
- American Psychological Association. (2013). *Going international: A practical guide for psychologists. Using translated/adapted measurement scales internationally*. Nueva York, NY: Autor.
- Baena, G. (1998). *Técnicas de investigación*. Ciudad de México: Editores Mexicanos Unidos.
- Balluerka, N. y Gorostiaga, A. (2012). Elaboración de versiones reducidas de instrumentos de medida: Una perspectiva práctica. *Intervención Psicosocial*, 21(1), 103-110. doi: 10.5093/in2012v21n1a7
- Bazdresch, M. (2009). La vida cotidiana escolar en la formación valoral, un caso. *REICE. Revista sobre Calidad, Eficacia y Cambio en Educación*, 7(2), 49-71.
- Byrne, B. (2009). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Nueva York, NY: Taylor and Francis Group.
- Byrne, B. y Van de Vijver, F. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of non equivalence. *International Journal of Testing*, 10(2), 107-132. doi: 10.1080/15305051003637306
- Cámara de Diputados del H. Congreso de la Unión. (2011). *Diario Oficial de la Federación*. Ciudad de México: Diario Oficial de la Federación.
- Cámara de Diputados del H. Congreso de la Unión. (2013). *Diario Oficial de la Federación*. Ciudad de México: Diario Oficial de la Federación.
- Cachón, C. (2007). *La función de la equivalencia en el proceso de la medición intercultural. Memoria electrónica COMIE*. Ciudad de México: COMIE.
- Carbajal, P. (2013). Convivencia democrática en las escuelas. Apuntes para una reconceptualización. *Revista Iberoamericana de Evaluación Educativa*, 6(2), 13-35.

- Cardozo, M. (2009). La institucionalización de una cultura de la evaluación en la administración pública mexicana: Avances y desafíos pendientes. *Convergencia. Revista de Ciencias Sociales*, 16(49), 175-198.
- Caso, J., Díaz, C. y Chaparro, A. (2013). Aplicación de un procedimiento para la optimización de la medida de la convivencia escolar. *Revista Iberoamericana de Evaluación Educativa*, 6(2), 137-145.
- Chaparro, A., Caso, J. y Fierro, C. (2012). *Validación psicométrica de indicadores de convivencia democrática, inclusiva y pacífica. Reporte de resultados*. Ciudad de México: CONCYTEG.
- Chaparro, A., Caso, J., Díaz, C. y Urías, E. (2012). *Instrumentos para el diagnóstico e intervención en escuelas basados en indicadores de convivencia democrática, inclusiva y no violenta*. Ensenada: Universidad Autónoma de Baja California.
- Chen, F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology*, 25(1), 1005-1018. doi: 10.1037/a0013193
- Cheung, G. y Rensvold, R. (1999). Testing factorial invariance across groups: A reconceptualization and proposed new method. *Journal of Management*, 1, 1-27. doi: 10.1177/014920639902500101
- CONEVAL. (2012). *Informe de pobreza y evaluación en el estado de Sonora*. Ciudad de México: Autor. Recuperado de goo.gl/bvhAFW
- CONEVAL. (2012). *Porcentaje de población en situación de pobreza extrema según entidad federativa*. Ciudad de México: Autor. Recuperado de goo.gl/vJBXct
- Contreras, L. A. (2000). *Desarrollo y pilotaje de un examen de español para la educación primaria* (Tesis de maestría). Universidad Autónoma de Baja California, México.
- Costa, N. y Brito, E. (2002). Adaptación cultural de instrumentos utilizados en salud ocupacional. *Revista Panamericana de Salud Pública*, 11(2), 109-111. doi:10.1590/S1020-49892002000200007
- Delval, J. (1992). *Aprender a aprender. I. El desarrollo de la capacidad de pensar*. Madrid: Alhambra Longman.
- Díaz Aguado, M. J., Martínez Arias, R. y Martín, J. (2010). *Estudio estatal de convivencia escolar de la educación secundaria 2007-2009*. Madrid: Ministerio de Educación.
- Díaz, C. D. (2015). *Adaptación de un instrumento para la medición de la convivencia escolar* (Tesis de maestría). Universidad Autónoma de Baja California, México.
- Drasgow, F. y Kanfer, R. (1985). Equivalence of psychological measurement in heterogeneous populations. *Journal of Applied Psychology*, 70(4), 662-680.
- Elosua, P., Mujika, J., Almeida, L. y Hermosilla, D. (2014). Procedimientos analítico-rationales en la adaptación de test. Adaptación al español de la batería de pruebas de razonamiento. *Revista Latinoamericana de Psicología*, 45(2), 117-126. doi: 10.1016/S0120-0534(14)70015-9
- Escobar-Pérez, J. y Cuervo-Martínez, A. (2008). Validez de contenido y juicio de expertos: Una aproximación a su utilización. *Avances en Medición*, 6, 27-36.
- Fernández, I. (1998). *Prevención de la violencia y resolución de conflictos: Clima escolar con factor de calidad*. Madrid: Narcea.

- Fierro, C. (2011). *Convivencia democrática e inclusiva. Una perspectiva para gestionar la seguridad escolar*. Recuperado de <http://basica.sep.gob.mx/escuelasegura/pdf/congresoBuenasPrac/convivencia.pdf>
- Fierro, C. y Caso, J. (2013). Presentación del monográfico. *Revista Iberoamericana de Evaluación Educativa*, 6(2), 7-12.
- Fierro, C., Tapia, G., Martínez-Parente, R., Macouzet, M. y Jiménez, M. (2013). Conversando sobre la convivencia en la escuela: Una guía para el autodiagnóstico de la convivencia escolar. *Revista Iberoamericana de Evaluación Educativa*, 2(6), 103-124.
- Fierro, C., Tapia, G., Fortoul, B., Martínez-Parente, R., Macouzet, M. y Jiménez, M. (2013). Conversando sobre la convivencia en la escuela: Una guía para el auto-diagnóstico de la convivencia escolar. *Revista Iberoamericana de Evaluación Educativa*, 2(6), 103-124.
- Gaite, L., Ramírez, N., Herrera, S. y Vázquez-Barquero, J. (1997). Traducción y adaptación transcultural de instrumentos en psiquiatría: Aspectos metodológicos. *Archivos de Neurobiología*, 10(2), 189-214.
- Geisinger, K. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment*, 6(4), 304-312. doi: 10.1037/1040-3590.6.4.304
- George, D. y Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference*. Boston, MA: Allen & Bacon.
- Gunnell, K., Wilson, P., Zumbo, B., Mack, D. y Crocker, P. (2012). Assessing psychological need satisfaction in exercise contexts: Issues of score invariance, item modification, and context. *Measurement in Physical Education and Exercise Science*, 16(3), 219-236. doi: 10.1080/1091367X.2012.693340
- Hambleton, R. K. (1996). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment*, 10, 229-240.
- Hambleton, R. K. (2005). Issues, designs and technical guidelines for adapting test into multiple languages and cultures. En R. Hambleton, P. Merenda y S. Spielberger (Eds.), *Adapting educational and psychological test for cross-cultural assessment* (pp. 3-38). Filadelfia, NJ: Lawrence Erlbaum Associates.
- Hambleton, R. K. y Patsula, L. (1999). Increasing the validity of adapted tests: Myths to be avoided and guidelines for improving test adaptation practices. *Journal of Applied Testing Technology*, 1(1), 1-30.
- Hernández, R., Fernández, C. y Baptista, P. (2010). *Metodología de la investigación*. Lima: McGraw-Hill.
- Herrera Rojas, A. (1993). *La medición en psicología*. Bogotá: Universidad de Bogotá.
- Herrero, J. (2010). El análisis factorial confirmatorio en el estudio de la estructura y estabilidad de los instrumentos de evaluación: Un ejemplo con el cuestionario de autoestima. *Intervención Psicosocial*, 19(3), 289-300. doi: 10.5093/in2010v19n3a9
- Hirschfeld, G. y Von Brachel, R. (2014). Multiple-Group confirmatory factor analysis in R - A tutorial in measurement invariance with continuous and ordinal indicators. *Practical Assessment, Research & Evaluation*, 19(7), 1-12.
- Hu, L. y Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Model*, 6, 1-55. doi: 10.1080/10705519909540118

- INEGI. (2013). *Estadísticas*. Recuperado de <http://www.inegi.org.mx/sistemas/olap/proyectos/bd/consulta.asp?p=17118&c=27769&s=est>
- Instituto Nacional para la Evaluación de la Educación. (2007). *Disciplina, violencia y consumo de sustancias nocivas a la salud en escuelas primarias y secundarias de México*. Ciudad de México: Autor.
- Instituto Nacional para la Evaluación de la Educación. (2015). *Comunicado de prensa n°1*. Ciudad de México: Autor.
- Instituto Nacional para la Evaluación de la Educación. (2015). *Plan Nacional para las Evaluaciones de los Aprendizajes*. Ciudad de México: Autor.
- International Test Commission. (2010). *International Test Commission guidelines for translating and adapting tests*. Nueva York, NY: International Test Commission.
- Jiménez, J. F. (2000). *Derechos de los niños. Colección nuestros derechos*. Ciudad de México: Instituto de Investigaciones Jurídicas.
- Jornet, J., González-Such, J. y Perales, M. (2012). Diseño de cuestionarios de contextos para la evaluación de sistemas educativos: Optimización de la medida de constructos complejos. *Bordón*, 64(2), 89-108.
- Kim, K. y Bentler, P. (2006). Data modeling: Structural equation modeling. En J. Green, G. Camili y P. Elmore (Eds.), *Handbook of complementary methods in education research* (pp. 161-175). Nueva York, NY: American Education Association.
- Klerk, G. (2008). *Cross-cultural testing. In online readings in testing and assessment*. Recuperado de <http://www.intestcom.org/Publications/ORTA.php>
- Luo, H. (2011). *Some aspects in confirmatory factor analysis of ordinal variables and generating non-normal data*. Recuperado de <http://www.divaportal.org/smash/record.jsf?pid=diva2:405108>
- López-González, E., Tourón, J. y Tejedor, F. (2012). Diseño de un micro-instrumento para medir el clima de aprendizaje en cuestionarios de contexto. *Bordón*, 64(2), 111-126.
- Magnus, L. (2005). Examining the validity of a Swedish version of the self-presentation in exercise questionnaire. *Measurement in Physical Education an Exercise Science*, 9(2), 113-134. doi: 10.1207/s15327841mpee0902_3
- Magnusson, D. (1982). *Teoría de los test*. Ciudad de México: Trillas.
- Martínez, F. (2013). El futuro de la evaluación educativa. *Sinéctica*, 40. Recuperado de http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-109X2013000100006
- Martínez-Arias, R. (1995). *Psicometría: Test de los test psicológicos y educativos*. Madrid: Síntesis.
- Messick, S. (1980). Test validity and ethics of assessment. *American psychologist*, 35, 1012-1027.
- México. (1982). *Acuerdo Secretarial 96 que establece la organización y funcionamiento de las escuelas primarias*. Diario Oficial de la Federación, de 7 de diciembre de 1982.
- México. (1982). *Acuerdo Secretarial 97 que establece la organización y funcionamiento de las Escuelas Secundarias Técnicas*. Diario Oficial de la Federación, de 3 de diciembre de 1982.
- México. (1982). *Acuerdo Secretarial 98 por el que se establece la organización y funcionamiento de las Escuelas de Educación Secundaria*. Diario Oficial de la Federación, de 7 de diciembre de 1982.

- México. (1993). *Ley General de Educación*. Diario Oficial (Separata), de 13 de julio de 1993.
- México. (2002). *Ley para la protección de los derechos de niñas, niños y adolescentes*. Diario Oficial de la Federación, núm. 19, de 29 de mayo de 2000, pp. 2-10.
- México. (2013). *Acuerdo 705 por el que se emiten las Reglas de Operación del Programa de Escuela Segura*. Diario Oficial de la Federación, de 28 de diciembre de 2013, pp. 52-71.
- Muñiz, J. y Hambleton, R. (1996). Directrices para la adaptación y traducción de instrumentos. *Papeles del Psicólogo*, 66, 63-70.
- Muñiz, J., Elosua, P. y Hambleton, R. (2013). Directrices para la traducción y adaptación de los test: Segunda edición. *Psicothema*, 25(2), 151-157. doi: 10.7334/psicothema2013.24
- OCDE. (2011). *Establecimiento de un marco para la evaluación de incentivos docentes: Consideraciones para México*. Ciudad de México: Autor.
- Ochoa, A. y Salinas, J. (2013). Diagnóstico de la convivencia escolar en escuelas de educación básica de la ciudad de Querétaro. En R. Hevia y J. S. Bravo (Orgs.), *Actas del V Congreso Iberoamericano de Violencia Escolar. Conversar la cultura escolar para construir convivencia* (pp. 56-75). Santiago de Chile: UDP - OEI - OVE.
- Ortega-Ruiz, R., Del Rey, R. y Casas, J. A. (2013). La convivencia escolar: Clave en la predicción del bullying. *Revista Iberoamericana de Evaluación Educativa*, 6(2), 91-102.
- Pickering, M. y Plotnikoff, R. (2009). Factor structure and measurement invariance of 10-item decisional balance scale: Longitudinal and subgroup examination within an adult diabetic sample. *Measurement in Physical Education and Exercise Science*, 13(4), 206-226. doi: 10.1080/10913670903260086
- Ponce, V. (2009). Investigación y políticas educativas. *Revista Electrónica Sinéctica*, 33, 1-33.
- Raju, N., Laffitte, L. y Byrne, B. (2002). Measurement equivalence a comparison methods based on confirmatory analysis and item response theory. *Journal Applied Psychology*, 87(3), 517-531. doi: 10.1037//0021-9010.87.3.517
- Ramada, J., Serra, C. y Delclós, G. (2013). Adaptación cultural y validación de cuestionarios de salud: Revisión y recomendaciones metodológicas. *Salud Pública de México*, 55(1), 57-66.
- Ramos, R. Y. (2001). *Educación integral. Una educación holística para el siglo XXI*. Madrid: Desclée de Brouwer.
- Reyes, E. P. (2014). *Validez del cuestionario de opinión de alumnos universitarios sobre la competencia docente* (Tesis doctoral). Universidad Autónoma de Baja California, México.
- Rodríguez, J. C. (en prensa). *Reporte técnico: Estrategia evaluativa 2015. Factores asociados al aprendizaje*. Ensenada: UEE-UABC.
- Rossi, P. y Freeman, H. (1993). *Evaluation. A systemic approach*. Londres: Sage.
- Secretaría de Educación Pública. (2007). *Programa Nacional Escuela Segura*. Ciudad de México: Autor.
- Secretaría de Educación Pública. (2011). *Programas de Estudio 2011. Guía para el maestro. Educación Básica. Secundaria. Formación Cívica y ética*. Ciudad de México: Dirección General de Desarrollo Curricular-Subsecretaría de Educación Básica.
- Secretaría de Educación Pública. (2013). *Programa Sectorial de Educación 2013-2018*. Ciudad de México: Autor.
- Silva, A. (1992). *Métodos cuantitativos en psicología*. Ciudad de México: Trillas.

- Sireci, S., Patsula, L. y Hambleton, R. (2005). Statistical methods for identifying flaws in the test adaptation process. En R. Hambleton, P. Merenda y S. Spielberger (Eds.), *Adapting educational and psychological test for cross-cultural assessment* (pp. 93-115). Filadelfia, NJ: Lawrence Erlbaum Associates.
- Tanzer, N. K. y Sim, C. Q. (1999). Adapting instruments for use in multiple languages and cultures: A review of the ITC Guidelines for Test Adaptations. *European Journal of Psychological Assessment*, 15(3), 258-269. doi: 10.1027//1015-5759.15.3.258
- Valadez, I. (2008). *Violencia escolar: Maltrato entre iguales en escuelas secundarias de la zona metropolitana de Guadalajara*. Ciudad de México: Mar-Eva.
- Van de Vijver, F. y Hambleton, R. (1996). Translating tests: Some practical guidelines. *European Psychologist*, 1(2), 89-99. doi: 10.1027/1016-9040.1.2.89
- Van de Vijver, F. y Poortinga, Y. (2005). Conceptual and methodological issues in adapting tests. En R. Hambleton, P. Merenda y S. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 39-63). Filadelfia, NJ: Lawrence Erlbaum Associates.
- Villalobos, F., Arévalo, C. y Rojas, F. (2012). Adaptación del Inventario de Resiliencia ante el Suicidio (SRI-25) en adolescentes y jóvenes de Colombia. *Revista Panamericana de Salud Pública*, 31(3), 233-239. doi: 10.1590/S1020-49892012000300008

Breve CV de los autores

Cristina Vanessa Hernández De la Toba

Doctora en Ciencias Educativas en el Instituto de Investigación y Desarrollo Educativo (IIDE) de la Universidad Autónoma de Baja California (UABC). Directivo-administrativo en el ámbito escolar en áreas de Planeación Institucional, Desarrollo Curricular, Servicios Escolares y de Subdirección. Profesora de educación secundaria, preparatoria y educación superior. Investigadora en Evaluación Educativa. Investigaciones recientes: *Desarrollo y validación de pruebas criterioles*. ORCID ID: 0000-0002-0264-1296. Email: cristina.vanessa.hernandez.delatoba@uabc.edu.mx

Joaquín Caso Niebla

Doctor en Psicología Educativa por la Universidad Nacional Autónoma de México (UNAM). Investigador del Instituto de Investigación y Desarrollo Educativo de la Universidad Autónoma de Baja California, en el área de evaluación educativa. Miembro del Sistema Nacional de Investigadores. Investigaciones recientes: *Desarrollo y validación de instrumentos de medición psicológica y educativa*, *Evaluación del aprendizaje de estudiantes de educación básica*, *Variables asociadas al rendimiento académico de adolescentes mexicanos*, y *Evaluación de la convivencia escolar*. Actualmente se desempeña como Director del Instituto de Investigación y Desarrollo Educativo de la UABC. ORCID ID: 0000-0002-3557-1722. Email: jcaso@uabc.edu.mx