

Invarianza de medición de una escala de ciudadanía en estudiantes latinoamericanos

Federico Viertel Arze

Trabajo de grado para postular a magister en medición y evaluación de programas  
educacionales

Juan Carlos Castillo

Supervisor

## Resumen

En las últimas décadas se ha renovado la investigación relativa al rol que puede asumir la educación cívica en el fortalecimiento de las democracias latinoamericanas. Las mediciones realizadas a gran escala, como ICCS 2009, aportan datos importantes sobre conocimiento cívico y ciudadanía de estudiantes y escuelas, incluyendo el modo cómo implementan la educación cívica, acorde con el currículo definido por cada sistema educativo. Las particularidades de los currículos, así como diferencias culturales entre los países, imponen que las comparaciones a partir de mediciones, requieran instrumentos probados en distintos niveles de equivalencia de medida para evitar sesgo. El presente estudio evalúa la invarianza de medición de una escala de ciudadanía entre países de latinoamérica participantes del ICCS 2009, y entre escuelas de la región que implementan la asignatura de educación cívica o no. Los resultados indican que las comparaciones entre países no pueden realizarse sin asumir un importante nivel de sesgo, mientras que la evaluación de invarianza entre escuelas que implementan educación cívica o no, se logra en su nivel más estricto. Se discuten aspectos metodológicos de las posibles fuentes de sesgo.

*Keywords:* ciudadanía; invarianza; educación; análisis factorial confirmatorio; medición

Invarianza de medición de una escala de ciudadanía en estudiantes latinoamericanos

La escuela es un motor importante para la adquisición de los principales aprendizajes que articulan el ejercicio ciudadano. En su práctica cotidiana dentro de los establecimientos, se busca preparar a los alumnos en la comprensión e interacción con el contexto social, y su participación política activa como ciudadanos adultos. Sin embargo los resultados de estudios internacionales indican una brecha en los países entre el currículo formal de educación cívica y su implementación en las escuelas, (Bolzendahl y Coffé, 2013; Schulz, Ainley, y Fraillon, 2013; Schulz, Ainley, Fraillon, Kerr, y Losito, 2010b) enfatizando el conocimiento cívico, por sobre otros dominios de la ciudadanía, como las actitudes de los estudiantes.

En las últimas décadas se ha renovado la investigación relativa al rol que puede asumir la educación cívica en el fortalecimiento de las democracias latinoamericanas (Hahn, 2010; Sherrod, Torney-Purta, y Flanagan, 2010). Este renovado impulso ha generado la implementación de la educación cívica en variadas formas por parte de los sistemas educativos de la región: como asignatura independiente, en conjunto con otras asignaturas, como un tema transversal, o incluso como un tema extracurricular (Cox, 2010; Cristian Cox; Martin Bascopé; Juan Carlos Castillo; Daniel Miranda; Macarena Bonhomme., 2015; Schulz, Ainley, Fraillon, Kerr, y Losito, 2010c).

Además se ha observado que la ciudadanía es un concepto dinámico, siendo susceptible a los cambios en los sistemas políticos, sociales y comunicacionales (Kymlicka, 2011; Kymlicka y Norman, 1997).

La Asociación Internacional para la Evaluación de los Logros Educativos (IEA) realizó el año 2009 el Estudio Internacional sobre Educación Cívica y Ciudadana (ICCS 2009), cuyo propósito fue investigar en un conjunto de países, el modo en que los jóvenes son preparados para asumir sus roles como ciudadanos en el siglo XXI. Para lograrlo, el estudio informa sobre el desempeño de los estudiantes y también sobre las actitudes, percepciones y actividades de éstos en relación con la educación cívica y la formación ciudadana. La población evaluada está conformada por estudiantes de octavo año de educación básica, con una edad promedio entre 13 y 14 años, participando Chile,

Colombia, Guatemala, Paraguay, República Dominicana y México en el módulo Latinoamericano.

Las dimensiones de ciudadanía medidas en el ICCS se conceptualizan como percepciones de los estudiantes de diferentes tipos de comportamiento de un buen ciudadano adulto (Schmitt y Kuljanin, 2008; Schulz, Ainley, y cols., 2010c; Torney-Purta, 2004), basados en gran medida en las conclusiones de Almond y Verba (1963) y su descripción de la cultura política. Esta definición concibe la ciudadanía como dos conjuntos interrelacionados de valores y creencias con respecto a la conducta ciudadana manifestada de forma convencional o relacionada con la participación en movimientos sociales, que se articulan en la interacción con otros y las instituciones formales (Dalton, 2006; Schulz, Ainley, y cols., 2010c). La ciudadanía convencional se refiere al comportamiento ciudadano asociado con el cumplimiento de normas sociales o derechos democráticos como votar en las elecciones nacionales, afiliarse a un partido político o seguimiento de problemas políticos en los medios de comunicación. Por otro lado, la participación en movimientos sociales se relaciona con formas más activas de participación social, tales como participación en protestas pacíficas, actividades en beneficio de la comunidad, promoción de derechos humanos y protección del medio ambiente. Estudios recientes (Alazzi, 2012; Bolzendahl y Coffé, 2013; Geboers, Geijsel, Admiraal, y ten Dam, 2013; Hooghe, Oser, y Marien, 2014; Pykett, Saward, y Schaefer, 2010) han explorado ambas dimensiones de la noción de ciudadanía, concluyendo que ésta ha cambiado dramáticamente en las sociedades contemporáneas, implicando que la ciudadanía convencional estaría siendo reemplazada por formas más participativas.

En el contexto latinoamericano del ICC2009, Cox (2015) compara la organización curricular de los países participantes, concluyendo que un rasgo común y fundamental entre estos sistemas educativos es que la educación ciudadana se ofrece a través de distintas áreas del currículo y no en una sola asignatura, cubriendo el conjunto de la secuencia formativa escolar. La diferencia estaría entre países que definen el currículo nacional bajo la forma de estándares o programas de estudios. En el caso de Colombia y Chile los currículos se organizan en base a estándares, no existiendo una asignatura de

educación cívica, siendo los profesores encargados en abordar los contenidos del área de historia y ciencias sociales. Los restantes cuatro países de la región presentarían programas de estudios con una asignatura de educación cívica o ciudadana con un docente especialista, además de estar presente en otras áreas curriculares.

Esto repercutiría en el tiempo dedicado a la educación cívica en la escuela, la preparación de docentes especialistas, y evaluación de competencias cívicas y ciudadanas en los alumnos (Bolzendahl y Coffé, 2013). Por lo tanto determinar los contenidos y objetivos en educación cívica requiere la definición de los componentes que conforman la ciudadanía y el significado de ser un buen ciudadano. En un examen crítico a las evaluaciones a larga escala, Zurstrassen (2011) centra su foco en el ICCS 2009, advirtiendo sobre posibles fuentes de sesgo dentro del estudio como el particularismo curricular de cada país en educación cívica, la ausencia de un modelo teórico-empírico adaptable internacionalmente que permita evaluar proeficiencia en educación cívica, falta de consenso interdisciplinario en investigación en educación ciudadana, así como las diferencias culturales entre y dentro de los países con un mismo idioma (Cox, 2010; Hahn, 1999, 2010).

Ante este escenario, por ejemplo ¿es posible comparar las actitudes de los estudiantes hacia la ciudadanía convencional o relacionada con movimientos sociales entre países latinoamericanos, considerando distintos contextos de enseñanza en educación cívica? o ¿comparar el efecto de docentes especialistas en el área, en escuelas con o sin asignatura de educación cívica?

Estudios a gran escala como el ICCS requieren asegurar la comparabilidad de sus resultados, por lo cual se hace necesario evaluar las propiedades psicométricas de las escalas que lo conforman, con la finalidad de aportar evidencia para su validez.

Cuando se realiza una medición a través de un instrumento a individuos que pertenecen a diferentes grupos, se asume que éste mide el mismo constructo bajo diferentes condiciones. Las poblaciones de interés habitualmente se definen por variables demográficas como el sexo, la edad, la raza, el país de origen o el idioma. Estas condiciones incluyen la estabilidad de la medida en diferentes culturas, evaluadores, o en

distintos procedimientos de administración del instrumento. El establecimiento de la equivalencia de medida es relevante para prácticamente la totalidad del empleo de las puntuaciones de pruebas, en distintas poblaciones, sea el objetivo simplemente evaluar diferencias individuales o utilizar las puntuaciones de la medición como predictores. Un instrumento posee equivalencia o invarianza de medida entre grupos cuando individuos pertenecientes a distintos grupos tienen idéntico nivel de rasgo en el constructo medido, y consecuentemente la misma probabilidad de obtener igual puntuación en la medición.

El Análisis Factorial Confirmatorio (AFC) constituye un caso particular de Modelo de Ecuaciones Estructurales que se ocupa específicamente de los modelos de medición y de las relaciones entre las variables observadas y las variables latentes. De los métodos factoriales disponibles, el AFC multigrupo es actualmente una herramienta ampliamente utilizada en el estudio de invarianza de medida en múltiples grupos (Brown, 2008; Millsap, 2012).

El objetivo de este estudio es evaluar la invarianza de medida en dos subescalas, CITCON (6 ítems) que indaga en la percepción de los estudiantes sobre la importancia de ciudadanía convencional, y CITSOC (4 ítems) sobre la percepción de la importancia de ciudadanía relacionada con movimientos sociales, las que conforman la escala de "Buen ciudadano adulto" dentro del ICCS 2009. En primer lugar se comparará entre los países latinoamericanos participantes para evaluar la consistencia de las escalas, considerando las particularidades curriculares de cada uno. A continuación se evaluará la invarianza de medida entre escuelas de la región que presenten o no la asignatura de educación cívica.

## Método

### Instrumento y muestra

El Estudio Internacional sobre Cívica y Educación Ciudadana (ICCS 2009) es un proyecto de la Asociación Internacional para la Evaluación de los Logros Educativos (IEA), una organización que por más de cincuenta años ha realizado estudios internacionales comparativos en el campo del desempeño educativo. El último estudio

de 2009 incluyó 38 países, 140000 estudiantes, 62000 docentes, y 5.300 escuelas (Schulz, Ainley, Fraillon, Kerr, y Losito, 2010a). En el ICCS de 2009 participaron seis países latinoamericanos (N=29298): Chile (n=5159), Colombia (n=6130), Guatemala (n=3959), México (n=6488), Paraguay (n=3277), y República Dominicana (n=4285). En el Cuadro 1 se detalla el número de escuelas participantes en la región por grupos de comparación para la evaluación de invarianza.

Dentro del marco de evaluación del ICCS 2009 se mide conocimientos cívicos y actitudes hacia la ciudadanía, teniendo como hipótesis principal que los jóvenes aprenden sobre la educación cívica y la ciudadanía a través de sus interacciones con múltiples comunidades (Schulz, Fraillon, Ainley, Losito, y Kerr, 2010; Torney-Purta, 2004). Asimismo el estudio reunió información contextual relacionada a sus características individuales, sus experiencias de aprendizaje en familia, la escuela y la comunidad en general.

El estudio mide las actitudes de los estudiantes en relación a la ciudadanía en una escala que contiene 10 ítems donde se les pidió evaluar una serie de posibles conductas ciudadanas en orden de importancia (likert de 4 alternativas). Se recoge información sobre dos componentes de ciudadanía, CITCON (6 ítems) que indagan en la percepción de los estudiantes sobre la importancia de la ciudadanía convencional, y CITSOC (4 ítems) sobre la percepción de la importancia de la ciudadanía relacionada con movimientos sociales en un ciudadano adulto (Cuadro 2).

### **Plan de análisis**

El objetivo de establecer la invarianza de medición es comprobar que bajo diferentes condiciones de observación y estudio del fenómeno, el instrumento de medida realmente mide el mismo constructo (Millsap, 2012).

La definición más utilizada del término equivalencia es la proporcionada por (Drasgow y Kanfer, 1985), según la cual un test o una subescala posee invarianza, o equivalencia de medida en varios grupos o poblaciones si personas con puntuaciones idénticas en el rasgo latente subyacente tienen la misma puntuación esperada a nivel de

ítem, a nivel de puntuación total en la escala o ambos. Una definición formal de equivalencia se enuncia de la siguiente manera:

Supongamos un conjunto de indicadores obtenidos de una muestra de sujetos y que además estas mediciones son una función estadística de otro conjunto de variables aleatorias  $\theta$ . Dada una variable  $x$  que indica el grupo (o población) al que pertenece el sujeto, podremos afirmar que nuestro conjunto de mediciones  $Y$  es invariante o equivalente con respecto a  $x$  si:

$$\Pr(Y|\theta = t, X = x) = \Pr(Y|\theta = t)$$

para todos los valores de  $x$  y  $t$ . Esto es, si la probabilidad de observar un conjunto de mediciones  $Y$  (un conjunto de variables dependientes) para un nivel fijo de predictores  $\theta = t$ , es independiente del grupo al que pertenezca el sujeto. En otras palabras, un conjunto de mediciones  $Y$  es invariante con respecto a  $x$  si la relación entre  $Y$  y  $\theta$ , dada por  $\Pr(Y|\theta = t)$  es la misma con independencia del grupo al que pertenezca el sujeto (Meredith, 1993; Millsap y Everson, 1993).

Esta definición general, indica que las mediciones  $Y|\theta$  pueden ser unidimensionales o multidimensionales, así como continuas o categóricas, y su relación dada por  $\Pr(Y|\theta = t)$  puede ser lineal o no lineal. A partir de ambas definiciones se puede concluir que, en caso de falta de equivalencia de medición en los grupos, es equívoco compararlas, ya que las diferencias encontradas pueden reflejar tanto diferencias verdaderas entre los grupos, como una diferencia en la relación entre la variable latente y la puntuación observada que no es igual en ambos grupos.

En un análisis factorial multigrupo, el modelo teórico se compara con la estructura observada en dos o más muestras. Habitualmente se sigue la estrategia de Jöreskog de comparación de estructuras de covarianzas (Joreskog, 1973) para comprobar la invarianza de medición. En esta estrategia se organizan modelos anidados en un orden jerárquico, con la disminución sucesiva del número de parámetros o el aumento de los grados de libertad, lo que implica que se van añadiendo restricciones al modelo, forzando la igualdad de parámetros entre los grupos de manera sucesiva. Estos modelos

cada vez más restrictivos se evalúan en términos del ajuste de sus datos al modelo (Milfont y Fischer, 2010; Vandenberg y Lance, 2000b).

**Modelo 1. Invarianza configural.** En este modelo se pone a prueba la hipótesis nula de que existe el mismo patrón de factores de carga fijos y libres en cada grupo, analizando los grupos de manera simultánea, dejando invariante el patrón de cargas factoriales. En otras palabras, cada ítem tiene que pertenecer al mismo factor en todos los grupos, pero se permite que todos los parámetros estimados varíen entre los grupos. Por tanto, los índices de ajuste de este modelo base de igualdad de patrones factoriales, reflejan el ajuste de los parámetros en ítems estimados separadamente para cada grupo. Si se da esta equivalencia, el número de factores y el patrón de matrices de cargas factoriales es similar entre los grupos, por lo que podemos decir que los ítems definen los mismos factores, aunque los pesos de los ítems sobre las escalas pueden variar entre grupos, lo que significa que aceptar esta hipótesis de equivalencia implica que los grupos asocian los mismos ítems con los mismos constructos. Este modelo se considera el modelo base con el que se evalúan los modelos de invarianza posteriores.

**Modelo 2. Invarianza métrica.** En este modelo, además del cumplimiento de la invarianza configural, se requiere que las saturaciones factoriales sean iguales entre grupos. Por tanto, no solo la composición de los factores debe ser constante, sino también el peso de cada variable en la constitución de cada factor. En la invarianza métrica se pone a prueba la hipótesis nula de que las cargas factoriales para cada ítem son invariantes en los grupos, poniendo así a prueba un modelo que añade la restricción de las cargas factoriales que ahora obligatoriamente tienen que ser iguales en los dos grupos. Aceptar esta hipótesis de equivalencia implica que la fuerza de la relación entre cada ítem y su constructo o factor subyacente es idéntica entre los grupos. En este caso, se puede decir que los factores significan lo mismo en los grupos y es legítimo comparar a los grupos en las varianzas (covarianzas) de los factores latentes incluidos en el modelo.

**Modelo 3. Invarianza escalar.** Este modelo pone a prueba la hipótesis nula de que los interceptos de las ecuaciones de regresión de los ítems sobre las variables latentes no varían en los grupos, aunque puede haber diferencias en las medias de los factores.

La invarianza escalar incluye la igualdad de las cargas factoriales y de los interceptos conjuntamente. Este modelo contrasta si las diferencias de medias entre los grupos en las variables se explican por las diferencias de medias en los factores latentes, por lo que al aceptar esta hipótesis de equivalencia se considera probada la invarianza de medición.

El AFC se basa en los supuestos de normalidad y linealidad. La utilización del AFC en ítems de un test de respuesta dicotómica o politómica supone la violación de este segundo supuesto, ya que los modelos subyacentes son modelos lineales y las relaciones entre los ítems dicotómicos o politómicos no lo son (Byrne y Watkins, 2003; Millsap y Yun-Tein, 2004).

En el AFC, la literatura sugiere que cuando los datos ordinales se analizan por el método de estimación de máxima verosimilitud, analizando la matriz de correlaciones de Pearson, la estimación de parámetros puede resultar sesgada (Millsap, 2012) concluyendo que el ajuste del modelo está deteriorado. Estos errores son menores en algunas condiciones: cuando las variables categóricas se aproximan a una distribución normal, los ítems tienen un alto número de categorías (5 o más) y las variables son simétricas (Kankaras, Vermunt, y Moors, 2011; Kaplan y Elliott, 1997; B. O. Muthén y Asparouhov, 2009). Una solución a este problema consiste en utilizar para el análisis la matriz de correlaciones policóricas donde se estima la relación lineal entre variables latentes continuas que subyacen a variables observadas ordinales que son indicadores manifiestos de aquellas (Kim y Yoon, 2011).

Un método de extracción apropiado cuando se analiza la matriz de correlaciones policóricas en variables ordinales es el método de mínimos cuadrados ponderados (WLS) y su versión robusta con varianza ajustada (WLSMV). Este procedimiento proporciona errores típicos correctos en muestras grandes y se recomienda, además, por sus mayores tasas de convergencia (Kankaras y cols., 2011; Kim y Yoon, 2011; Millsap y Yun-Tein, 2004)

Dado que no existe una única medida aceptada para determinar la bondad de ajuste, se recomienda un uso conjunto de varios índices globales en la evaluación de dicho ajuste. Los índices de ajuste pueden dividirse en dos clases: absolutos e

incrementales (Hu y Bentler, 1999). Los índices de ajuste absoluto expresan el grado de exactitud en que el modelo global predice satisfactoriamente la matriz de covarianzas observada. Por su parte, las medidas de ajuste incremental comparan el modelo analizado con un modelo de base habitualmente denominado modelo nulo.

El estadístico  $\chi^2$  de bondad de ajuste se utiliza para comparar el ajuste de las diversas estructuras factoriales puestas a prueba, calculando las diferencias en los valores de  $\chi^2$  entre los modelos para determinar si hay diferencias significativas entre ellos.

Para la interpretación de ajuste del modelo se utilizarán como índices absolutos CFI y TLI considerándose indicadores de buen ajuste los valores superiores a .90, y RMSEA con valores  $\leq$  .08 indicativos de un ajuste razonable (Hu y Bentler, 1999). Como índices de ajuste incremental se utilizará  $\Delta$ CFI,  $\Delta$ TLI y  $\Delta$ RMSEA considerándose apropiado variaciones menores a .01 (Cheung y Rensvold, 2002; Millsap, 2012; Widaman y Grimm, 2009)

Se utiliza Mplus 7.3 (B. Muthén, 1994) para el análisis, usando los pesos muestrales estimados para la muestra estratificada del ICCS 2009 (Schulz y Fraillon, 2009), con DIFFTEST para la comparación de  $S - B\Delta\chi^2$ .

## Resultados

En el Cuadro 3 se muestran los resultados del análisis de invarianza para los modelos a comparar entre países y los resultados para los modelos entre escuelas con o sin la asignatura de educación cívica en el currículum escolar.

En la estimación libre de los parámetros, el modelo comparativo entre países no logra la invarianza configural, con ajuste deteriorado en todos los índices. Esto es requisito para las pruebas subsecuentes de invarianza, por lo que no es posible asegurar la comparación libre de sesgo entre países latinoamericanos en la escala evaluada.

En el modelo que compara la implementación de educación cívica, en primera instancia se evalúa la invarianza configural, con base a índices de ajustes adecuados, según lo sugerido por Hu y Bentler (1999). Se observa sobrestimación de  $\chi^2$ , por lo que se prosigue según lo sugerido por Cheung y Rensvold (2002) en la evaluación de los

índices CFI, TLI y RMSEA y su criterio de variación incremental entre modelos anidados ( $\Delta$ CFI,  $\Delta$ TLI,  $\Delta$ RMSEA). El modelo muestra ajuste aceptable en todos los índices señalados, logrando invarianza configural, lo que indica que las variables latentes son especificadas por los mismos indicadores en los dos grupos, permitiendo evaluar los siguientes modelos de invarianza (Vandenberg y Lance, 2000a).

Al agregar al modelo configural restricciones sobre las cargas factoriales de los indicadores, se obtiene invarianza métrica. El valor de  $S - B\Delta\chi^2$  es significativo lo que puede estar influido por la sobrestimación de  $\chi^2$ , de forma que no se considerará en la evaluación invarianza como lo sugiere Cheung (2002). Como se observa en el Cuadro 3, los índices incrementales ( $\Delta$ CFI,  $\Delta$ TLI,  $\Delta$ RMSEA) no presentan aumento sustantivo, reforzando la aceptación de la equivalencia entre los dos grupos de escuelas en este nivel de invarianza. Esto permite que las diferencias de puntaje obtenidas en los indicadores entre los grupos, reflejen diferencias de puntaje en los constructos latentes comparables, dado que la relación entre las cargas factoriales con los constructos latentes son similares en ambos grupos.

Se restringe la estimación de los interceptos en los indicadores, así como sus cargas factoriales, igualándolas entre los grupos. El modelo de invarianza escalar ajusta de manera aceptable como se observa en el Cuadro 3, no resultando en un deterioro significativo en el ajuste del modelo en relación al modelo de invarianza métrica. Este nivel es requerido para comparar las medias en las variables latentes entre escuelas que presentan la asignatura de educación cívica y las que no la imparten como asignatura dentro del currículum. Esto implica que los valores observados en los indicadores conservan su relación con las medias en las variables latentes en los distintos grupos.

Los resultados obtenidos sustentan la invarianza de medición del modelo de dos factores de ciudadanía (CITSOC / CITCON) entre las escuelas que presentan o no la asignatura de educación cívica. No obstante, el nivel base de invarianza configural entre los países latinoamericanos que participaron del ICCS 2009 no se obtiene, por lo que la secuencia de comparación de modelos anidados para evaluar invarianza no permite sustentar la hipótesis de comparabilidad en ningún nivel.

## Discusión

El objetivo de este estudio es evaluar la invarianza de medida en dos subescalas que conforman el concepto de ciudadanía, en relación a las conductas esperables en un buen ciudadano adulto dentro del ICCS 2009. La evaluación con Análisis Factorial Confirmatorio Multigrupo entre los países latinoamericanos participantes no logra un nivel de invarianza configural, la forma más elemental de invarianza, y el modelo básico de análisis en todo estudio de equivalencia. Este nivel exige que los factores latentes queden especificados por los mismos indicadores observables en las dos subpoblaciones. El rechazo de la hipótesis de invarianza configural implica la falta de equivalencia sustantiva de constructos entre subpoblaciones o culturas. Para la evaluación de invarianza entre las escuelas que implementan o no la asignatura de educación cívica, se logra el nivel más exigente de restricción en la estimación de parámetros, que permite la comparación de las medias de las variables latentes entre los grupos.

Una presunción común en las pruebas para la equivalencia de un instrumento de medición a través de grupos culturales es que las fuentes de posibles sesgos descansan exclusivamente en los indicadores que componen las escalas. El debate existente sobre la estrategia de invarianza parcial ha decantado últimamente en que su utilización solo es recomendable con respaldo sustantivo que avale las comparaciones con modelos parciales de medición (Milfont y Fischer, 2010). Además, para escalas con ítems ordinales - categóricos, la inspección de índices de modificación se torna engorrosa y compleja al considerar la estructura policórica de cada indicador, complejidad que aumenta con el número de grupos de comparación. Abordajes como el Funcionamiento Diferencial del Ítem, propuesta desde los modelos de respuesta al ítem, resultan en estrategias que han mostrado efectividad en la detección de fuentes de sesgo (Elousa, 2005). Asparouhov y Muthén (2014) han realizado estudios para implementar una aproximación Bayesiana a los Modelos de Ecuaciones Estructurales, aunque su implementación para indicadores ordinales-categóricos se encuentra en fase de desarrollo.

En ese sentido los resultados obtenidos en ambas evaluaciones son consonantes con lo planteado por Byrne y van de Vijver (2010) desde la perspectiva de diseño general del

estudio, proponiendo que la elección de los grupos de comparación se base en una hipótesis sustantiva, argumentando que en lugar de seguir la tendencia común de comparación de grupos culturales geográficamente, tendría más sentido pensar en la cultura como la representación de las diferentes subpoblaciones dentro de las unidades geográficas (país, región), considerando entonces las particularidades que distinguirían el constructo en esa población.

Por otra parte Cox (2015) advierte que la interpretación de los resultados del ICCS 2009 debe contrastarse con el análisis curricular de los países participantes, ya que existe una brecha importante entre el énfasis del marco de evaluación de la prueba, los contenidos y objetivos declarados por los sistemas educativos y las diferencias en el modo de implementación de la educación cívica o ciudadana tanto dentro, como entre los países. Se puede sostener entonces, que cuando los grupos en estudio representan diferentes países, el sesgo puede ser determinado por el grado en que a los estudiantes de un país en particular le han enfatizado el concepto de ciudadanía, dependiendo por ejemplo, de la presencia de docentes especialistas en educación cívica.

En contextos en que resurge la discusión sobre la reimplementación de la asignatura de educación cívica, como ocurre en el caso particular de Chile, la investigación en torno al efecto de docentes especialistas en educación ciudadana adquiere relevancia. La evaluación de escalas asociadas a los componentes que conforman la ciudadanía resulta de utilidad dentro de estos contextos, ya que permite la definición de contenidos y objetivos curriculares consonantes con los modelos de ciudadanía emergentes, especialmente los relacionados a movimientos sociales. La investigación derivada también permitiría abordar de manera efectiva las dificultades y limitaciones que conlleva implementar la enseñanza de este tipo de modelos.

## Referencias

- Alazzi, K. (2012). Students perceptions of good citizenship: A study of middle and high school students in Jordan. *European Journal of Social Sciences*, 31(2), 223–230.
- Almond, G. A., y Verba, S. (1963). *The Civic Culture*. doi:  
<http://dx.doi.org/10.1016/B0-08-043076-7/01106-2>
- Asparouhov, T., y Muthén, B. (2014, julio). Multiple-Group Factor Analysis Alignment. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(4), 495–508. doi: 10.1080/10705511.2014.919210
- Bolzendahl, C., y Coffé, H. (2013, abril). Are ‘Good’ Citizens ‘Good’ Participants? Testing Citizenship Norms and Political Participation across 25 Nations. *Political Studies*, 61, 45–65. doi: 10.1111/1467-9248.12010
- Brown, T. (2008). *Confirmatory Factor Analysis for Applied Research*. doi:  
10.1198/tas.2008.s98
- Byrne, B. M., y van de Vijver, F. J. R. (2010). *Testing for Measurement and Structural Equivalence in Large-Scale Cross-Cultural Studies: Addressing the Issue of Nonequivalence* (Vol. 10). doi: 10.1080/15305051003637306
- Byrne, B. M., y Watkins, D. (2003). *The Issue Of Measurement Invariance Revisited* (Vol. 34). doi: 10.1177/0022022102250225
- Cheung, G. W., y Rensvold, R. B. (2002). *Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance* (Vol. 9).
- Cox, C. (2010). Oportunidades de aprendizaje escolar de la ciudadanía en América Latina: currículos comparados. *Sistema Regional de Evaluación y Desarrollo de ...*, 1–81.
- Cristian Cox; Martin Bascopé; Juan Carlos Castillo; Daniel Miranda; Macarena Bonhomme. (2015). Educación Ciudadana en América Latina: prioridades de los currículos escolares. En C. Cox y J. C. Castillo (Eds.), *Aprendizaje de la ciudadanía: contextos, experiencias y resultados* (1ª Edición ed., pp. 321–370). Ediciones Universidad Católica de Chile.
- Dalton, R. (2006). *Citizenship norms and political participation in America: The good*

*news is ... the bad news is wrong.*

- Dragow, F., y Kanfer, R. (1985). Equivalence of psychological measurement in heterogeneous populations. *The Journal of applied psychology*, *70*, 662–680. doi: 10.1037/0021-9010.70.4.662
- Elousa, P. (2005, diciembre). Evaluación progresiva de la invarianza factorial entre las versiones original y adaptada de una escala de autoconcepto. *Psicothema*, *17*(Número 2), 356–362.
- Geboers, E., Geijsel, F., Admiraal, W., y ten Dam, G. (2013, junio). Review of the effects of citizenship education. *Educational Research Review*, *9*, 158–173. doi: 10.1016/j.edurev.2012.02.001
- Hahn, C. L. (1999). *Citizenship Education: An empirical study of policy, practices and outcomes* (Vol. 25). doi: 10.1080/030549899104233
- Hahn, C. L. (2010). *Comparative civic education research: What we know and what we need to know* (Vol. 6).
- Hooghe, M., Oser, J., y Marien, S. (2014). A Comparative Analysis of “Good Citizenship”: A Latent Class Analysis of Adolescents’ Citizenship Norms in 38 Countries. *International Political Science* . . . .
- Hu, L. L., y Bentler, P. M. P. (1999, enero). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. Descargado de <http://www.tandfonline.com/doi/abs/10.1080/10705519909540118> doi: 10.1080/10705519909540118
- Joreskog, K. G. (1973). A general method for estimating linear structural equation systems. En *Structural equation models in the social sciences*.
- Kankaras, M., Vermunt, J. K., y Moors, G. (2011). *Measurement Equivalence of Ordinal Items: A Comparison of Factor Analytic, Item Response Theory, and Latent Class Approaches* (Vol. 40). doi: 10.1177/0049124111405301
- Kaplan, D., y Elliott, P. R. (1997). *A Model-Based Approach to Validating Education Indicators Using Multilevel Structural Equation Modeling* (Vol. 22). doi:

10.3102/10769986022003323

Kim, E. S., y Yoon, M. (2011). *Testing Measurement Invariance: A Comparison of Multiple-Group Categorical CFA and IRT* (Vol. 18). doi:

10.1080/10705511.2011.557337

Kymlicka, W. (2011). *Multicultural citizenship within multination states* (Vol. 11). doi:

10.1177/1468796811407813

Kymlicka, W., y Norman, W. (1997). El retorno del ciudadano. Una revisión de la producción reciente en teoría de la ciudadanía. *Ágora*(7), 5 – 42.

Meredith, W. (1993, diciembre). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525–543. doi: 10.1007/BF02294825

Milfont, T. L. T., y Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. . . . *Journal of psychological research*, 3(1), 111–121. doi: 10.1007/s11135-007-9143-x

Millsap, R. E. (2012). Investigating measurement invariance using confirmatory factor analysis. *Handbook of structural equation modeling.*, 380–392. New York.

Millsap, R. E., y Everson, H. T. (1993). *Methodology Review: Statistical Approaches for Assessing Measurement Bias* (Vol. 17). doi: 10.1177/014662169301700401

Millsap, R. E., y Yun-Tein, J. (2004). *Assessing Factorial Invariance in Ordered-Categorical Measures* (Vol. 39).

Muthén, B. (1994). *Multilevel Covariance Structure Analysis* (Vol. 22). doi: 10.1177/0049124194022003006

Muthén, B. O., y Asparouhov, T. (2009). Beyond Multilevel Regression Modeling: Multilevel Analysis in a General Latent Variable Framework. En *The handbook of advanced multilevel analysis* (pp. 1–42).

Pykett, J., Saward, M., y Schaefer, A. (2010, noviembre). Framing the Good Citizen. *The British Journal of Politics & International Relations*, 12(4), 523–538. doi: 10.1111/j.1467-856X.2010.00424.x

Schmitt, N., y Kuljanin, G. (2008, diciembre). Measurement invariance: Review of practice and implications. *Human Resource Management Review*, 18(4), 210–222.

doi: 10.1016/j.hrmr.2008.03.003

- Schulz, W., Ainley, J., y Fraillon, J. (2013). Student participation at school and future civic engagement: Results from ICCS 2009. *iea.nl*(June), 26–28.
- Schulz, W., Ainley, J., Fraillon, J., Kerr, D., y Losito, B. (2010a). *ICCS 2009 International Report: Civic Knowledge, Attitudes, and Engagement among Lower-Secondary School Students in 38 Countries*.
- Schulz, W., Ainley, J., Fraillon, J., Kerr, D., y Losito, B. (2010b). *ICCS 2009 International Report: Civic knowledge, attitudes and engagement among lower secondary school students in thirty-eight countries*. (Inf. Téc.).
- Schulz, W., Ainley, J., Fraillon, J., Kerr, D., y Losito, B. (2010c). *Initial findings from the IEA International civic and citizenship education study* (Inf. Téc.).
- Schulz, W., y Fraillon, J. (2009). The Iea International Civic and Citizenship Education Study (ICCS): Concept and design. *Cadmo*, 17, 21–34.
- Schulz, W., Fraillon, J., Ainley, J., Losito, B., y Kerr, D. (2010). *Estudio Internacional sobre Educación Cívica y Ciudadana. Marco de la evaluación* (Vol. 623).
- Sherrod, L. R., Torney-Purta, J., y Flanagan, C. (2010). *Handbook of Research on Civic Engagement in Youth* (L. R. Sherrod, J. Torney-Purta, y C. A. Flanagan, Eds.). Hoboken, NJ, USA: John Wiley & Sons, Inc. doi: 10.1002/9780470767603.ch0
- Torney-Purta, J. (2004). An overview of secondary analysis of the IEA Civic Education Study, its impact and directions for the future. *First IEA International Research Conference, Nikosia, . . . .*
- Vandenberg, R. J., y Lance, C. E. (2000a). A review and synthesis of the measurement invariance literature. *Organizational Research Methods*, 3, 4–70. doi: 10.1177/109442810031002
- Vandenberg, R. J., y Lance, C. E. (2000b, enero). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4. doi: 10.1177/109442810031002
- Widaman, K. F., y Grimm, K. J. (2009, enero). Invariance, or Noninvariance: That Is

the Question. *Measurement : interdisciplinary research and perspectives*, 7(1), 8–12. doi: 10.1080/15366360802715338

Zurstrassen, B. (2011). A Sceptical Look at the Quantitative Education Research in Civic and Citizenship Education. *Journal of Social Science*, 10(3), 6–15.

Cuadro 1

*Muestra*

País	Escuelas		N total
	con cívica	sin cívica	escuelas*
Chile	23	152	177 (2)
Colombia	56	139	196 (1)
Guatemala	44	98	144 (2)
México	139	74	214 (1)
Paraguay	113	35	149 (1)
Rep. Dominicana	66	78	144 (0)
Total	441	576	

Fuente: Confeccion propia basado en Schulz (2010) ; \* escuelas omitidas en ()

Cuadro 2

*Escala e indicadores*

CITCON	Que tan importantes son estas conductas para ser un buen ciudadano adulto?	Escala
IS2P21A	Votar en todas las elecciones nacionales	
IS2P21B	Unirse a un partido político	1. Muy importante
IS2P21C	Aprender la historia de su país	2. Importante
IS2P21D	Seguir los temas políticos en el diario, la radio, la televisión o internet.	3. No muy importante
IS2P21E	Mostrar respeto por los representantes del gobierno.	4. No importante
IS2P21F	Involucrarse en discusiones políticas.	
CITSOC		
IS2P21G	Participar en protestas pacíficas contra de leyes que considera injustas.	
IS2P21H	Participar en actividades para beneficiar a la gente de su comunidad (barrio, comuna o ciudad)	
IS2P21I	Participar en actividades que promuevan los derechos humanos.	
IS2P21J	Participar en actividades para la protección del medio ambiente.	

Cuadro 3

*Comparación de niveles de invarianza entre modelos Países / Ed. Cívica*

Modelo Países	$\chi^2$	gl	$S - B\Delta\chi^2$	gl	CFI	$\Delta CFI$	TLI	$\Delta TLI$	RMSEA	$\Delta RMSEA$
Configural	2876.746*	204			0.893		0.858		0.052	
Métrica	2774.943*	244	244.96*	40	0.899	0.006	0.888	0.03	0.046	- 0.006
Esalar	3999.145*	334	1764.613*	90	0.853	- 0.046	0.881	- 0.007	0.047	0.001
Modelo Ed. Cívica										
Configural	517.706*	68			0.985		0.981		0.021	
Métrica	523.675*	76	64.181*	8	0.986	0.001	0.983	0.002	0.02	- 0.001
Esalar	548.216*	94	160.757*	26	0.985	- 0.001	0.986	0.003	0.018	- 0.002

\*  $p < .01$ .