

Estudio de confiabilidad y validez de la Escala de Manejo de Compromisos Identitarios de Utrecht, adaptada para medir procesos de formación de identidad de estudiantes en su universidad

An Investigation of the Reliability and Validity of the Utrecht-Management of Identity Commitments Scale, Adapted to Measure Students' Identity Formation Processes at their University

Hanke Korpershoek

GION Education/Research, Universidad de Groningen, Países Bajos

Resumen

El objetivo de este estudio es demostrar la validez de constructo y predictiva del marco de medición de Crocetti, Rubini y Meeus (2008), originalmente desarrollado para medir los procesos de formación de identidad de individuos (como el de lograr compromisos) en distintos ámbitos. Se usó una versión adaptada de la U-MICS (Escala de Manejo de Compromisos Identitarios de Utrecht) para medir los procesos de formación de identidad de estudiantes universitarios en su universidad, que forman parte de su identidad personal. La U-MICS apunta a medir tres dimensiones de la formación de identidad: compromiso, exploración en profundidad y reconsideración de compromiso. Se aplicó el método de medición Rasch. Los resultados muestran que el modelo multidimensional representa de manera válida los procesos de formación de identidad en la universidad (validez de constructo), aunque algunos ítems requieren mejoras. Se plantean sugerencias de mejoramiento. Las escalas de compromiso y exploración en profundidad se relacionan positivamente con el rendimiento académico de los estudiantes, mientras que la escala de reconsideración de compromiso se relaciona negativamente con dichos resultados (validez predictiva).

Palabras clave: compromiso, procesos de formación de identidad, estudiantes universitarios, modelamiento multidimensional Rasch

Correspondencia a:

Hanke Korpershoek

GION Education/Research, University of Groningen, the Netherlands.

Grote Rozenstraat 3, 9712 TG, Groningen, the Netherlands.

Correo electrónico: h.korpershoek@rug.nl

La autora agradece a Mayra Mascareño por la traducción del resumen al castellano.

© 2015 PEL, <http://www.pensamientoeducativo.org> - <http://www.pel.cl>

ISSN: 0719-0409 DDI: 203.262, Santiago, Chile
doi: 10.7764/PEL.52.2.2015.5

Abstract

The objective of this paper was to demonstrate the construct validity and predictive validity of the measurement framework presented by Crocetti, Rubini, and Meeus (2008), which was originally developed to measure the identity formation processes (such as achieving commitments) of individuals in various domains. An adapted version of the Utrecht-Management of Identity Commitments Scale (U-MICS) was used to measure university students' identity formation processes, which are a part of students' personal identities, at their university. The U-MICS seeks to measure three dimensions of identity formation: commitment, in-depth exploration, and reconsideration of commitment. A Rasch measurement approach was applied. The results showed that the multidimensional model validly represented students' identity formation processes at university (construct validity), although some items need further improvement. Suggestions for improvement are discussed. The commitment and in-depth exploration scales were positively related to students' academic achievement, while the reconsideration of commitment scale was negatively related to the same (predictive validity).

Keywords: commitment, identity development processes, university students, multidimensional Rasch modelling

El objetivo del presente artículo es demostrar, en una muestra de estudiantes universitarios (Universidad de Groningen, Países Bajos) la validez de constructo y la validez predictiva del marco de medición presentado por Crocetti, Rubini y Meeus (2008), originalmente desarrollado para medir los procesos de formación de identidad de los individuos en distintos dominios. En el artículo se aplicó un enfoque de medición Rasch para demostrar las propiedades de medición del marco en el campo universitario.

Los procesos de formación de identidad importantes, como hacer compromisos, se han asociado con resultados educativos favorables en estudios anteriores, por ejemplo, con un funcionamiento exitoso en la universidad (Berzonsky & Kuk, 2005). El compromiso con la universidad y la sensación de estar conectado con la institución se asocian positivamente con el progreso en los estudios (Allen, Robbins, Casillas, & Oh, 2008). Además, mientras más se identifican los estudiantes con su universidad, más se involucran voluntariamente con otros estudiantes y profesores (Battista, Pivetti, & Berti, 2014). En términos más amplios, tener una sensación de pertenencia a la universidad, (por ejemplo, sentirse en casa dentro de ella, disfrutar su atmósfera) puede potenciar el progreso académico de los estudiantes (Meeuwisse, Severiens, & Born, 2010). Por el contrario, la falta de compromiso se ha asociado con resultados educativos desfavorables como el bajo rendimiento (por ejemplo, Hejazi, Shahraray, Farsinejad, & Asgary, 2009). De acuerdo con la relación positiva que se ha sugerido, entre los procesos de formación de identidad (como hacer compromisos) y buenos resultados educativos, investigar los procesos de formación de identidad de los estudiantes universitarios entrega la oportunidad de identificar estudiantes en riesgo de presentar bajo rendimiento o de abandonar sus estudios. Particularmente, en el caso de estudiantes de primer y segundo año, identificar a los individuos no comprometidos con sus universidades podría ayudar a evitar la deserción temprana.

El presente artículo emplea el marco teórico de Crocetti et al. (2008) para medir los procesos de formación de identidad de los estudiantes universitarios dentro de la universidad. Su propósito es demostrar la validez de constructo y la validez predictiva de dicho marco de medición en una muestra de estudiantes universitarios e investigar las propiedades de medición de las escalas (como su confiabilidad y el ajuste entre el modelo y los datos) empleando análisis multidimensional Rasch. El término validez se refiere al grado en el cual la evidencia y la teoría apoyan la interpretación de las puntuaciones arrojadas por un test (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014). En el presente artículo, recolectamos evidencia basada en procesos de respuesta y estructura interna (es decir, validez de constructo: el grado en el cual la operacionalización del constructo mide lo que sugiere la teoría) y en las relaciones con otras variables (es decir, validez predictiva: el grado en el cual la operacionalización se asocia con variables teóricamente relevantes). Se formularon dos preguntas de investigación:

1. ¿Produce el marco de medición de Crocetti et al. (2008) mediciones de los procesos de formación de identidad de estudiantes universitarios en su institución que sean confiables y que posean validez de constructo?
2. ¿Hasta qué punto se relacionan los procesos de formación de identidad de los estudiantes en la universidad con su logro académico (validez predictiva)?

Una ventaja de aplicar un modelo Rasch es que este tipo de análisis puede confirmar que los ítems usados evocan y definen la variable del modo esperado (Wright & Masters, 1982). Un modelo Rasch es un modelo estocástico. Aplica la transformación logarítmica para estimar log-odds (logit) para cada ítem y para cada persona evaluada. Por lo tanto, aplicar un modelo Rasch entrega información valiosa para establecer cuán apropiado es el uso de un cierto marco de medición en un contexto dado. Otra ventaja es que un modelo Rasch no presupone que los ítems son medidos a nivel intervalar, lo que sí sucede cuando se emplea un enfoque analítico factorial. Los modelos Rasch manejan los datos de escalas Likert de forma más apropiada, es decir, como datos ordenados en vez de como datos en un nivel intervalar (en otras palabras, distancias iguales entre las categorías de respuesta). Los datos brutos no intervalares se transforman en medidas de escala logit (más detalles en la sección Método). Además, el análisis Rasch entrega información adicional sobre los umbrales de los ítems (por ejemplo, si los umbrales de la escala Likert aumentan monótonamente para cada ítem) y evalúa distribuciones de ítems y de personas en mucho mayor detalle que el análisis factorial. Por lo tanto, los análisis Rasch pueden ser usados para generar sugerencias sobre cómo mejorar la medición de constructos psicológicos entre muestras específicas.

Durante la última década, el análisis multidimensional Rasch ha ingresado al campo internacional de la psicología educacional (véase Ackerman, Gierl, & Walker, 2003; Korpershoek, Xu, Mok, McInerney, & van der Werf, 2015; Lee, Zhang, & Yin, 2010; Liu, Minsky, Ling, & Kyllonen, 2009; Mok, McInerney, Cheng, & Lai, 2011, 2013; Mok & Xu, 2013; Yan & Mok, 2012). Los modelos multidimensionales de respuesta al ítem, que emplean correlaciones entre rasgos latentes, pueden usarse para aumentar la precisión de la medición de rasgos latentes individuales (Wang, Chen, & Cheng, 2004) mediante la mejora de los métodos de modelamiento y estimación (Wu & Adams, 2006). El análisis Rasch puede producir resultados distintos a los del análisis factorial cuando se examina la validez de constructo de un marco de medición. Por esta razón, en el presente artículo se aplicó el análisis Rasch para obtener nuevos antecedentes sobre la validez de constructo del constructo de los procesos de formación de identidad. De este modo, el presente artículo contribuye a la literatura al ejemplificar la aplicación del análisis multidimensional Rasch al constructo multidimensional de los procesos de formación de identidad. Esto se lleva a cabo en una muestra de estudiantes no comúnmente empleada en el campo de la teoría del desarrollo de la identidad: adultos jóvenes en el campo universitario.

Marco teórico

Desarrollo de la identidad: el proceso de hacer compromisos

El constructo de los procesos de formación de identidad se origina en la literatura sobre teoría del desarrollo de la identidad, la cual se basa fuertemente en el trabajo de Erikson (1950, 1968) y Marcia (1966, 1980, 1994). Como apuntan Côté y Levine (2014), Erik Erikson probablemente sea el teórico fundacional más influyente en el estudio de la formación de la identidad (p. 91). Erikson mostró, de forma convincente, que la formación de una identidad personal es una de las tareas de desarrollo fundamentales de los adolescentes. En su muy frecuentemente citado trabajo, describe una secuencia de ocho etapas de desarrollo psicosocial presentes en la formación de la identidad. Estas etapas representan una secuencia de crecimiento de una persona, durante la cual esta enfrenta y aborda los desafíos de los distintos períodos de la vida.

El trabajo de Erikson fue continuado por Marcia (1966, 1980, 1994), quien operacionalizó la identidad dentro del paradigma de los estadios de la identidad. Marcia apunta que, a nivel del comportamiento, el desarrollo de la identidad puede considerarse como el proceso de hacer compromisos (Marcia, 1980, p. 54). Marcia (1966) propuso cuatro estilos para abordar la tarea de formación de la identidad. Estos estadios son «modos de enfrentar los problemas de identidad característicos de los adolescentes tardíos»

(Marcia, 1980, p. 161). Los individuos son asignados a estos estadios de identidad sobre la base del grado en el cual han explorado o están explorando alternativas de identidad y considerando hasta qué punto están comprometidos con una de estas alternativas (Vleioras & Bosma, 2005). Los cuatro estadios de identidad son: Difusión de Identidad (bajo compromiso, baja exploración), Exclusión de Identidad (alto compromiso, baja exploración), Moratoria de Identidad (bajo compromiso, alta exploración) y Logro de Identidad (alto compromiso, alta exploración). Estos estadios son el resultado del proceso de desarrollo de la identidad. El estadio de difusión se considera el menos adaptado, mientras que el de logro se considera el más adaptado de los cuatro (Waterman, 1999).

Siguiendo los pasos de esta investigación original sobre formación y estadios de identidad, el número de estudios sobre desarrollo de la identidad en adolescentes y adultos jóvenes ha aumentado sostenidamente (para acceder a estudios de revisión y panoramas generales, véase Bosma, 1985; Klimstra, Hale, Raaijmakers, Branje, & Meeus, 2010; Kroger, 2000; Kroger, Martinussen, & Marcia, 2010; Meeus, 2011; Meeus, Iedema, Helsen, & Vollenbergh, 1999; Meeus, van de Schoot, Keijsers, Schwartz, & Branje, 2010; Meeus & de Wied, 2007; Waterman, 1999). En los últimos años, se ha prestado cada vez más atención al ambiente escolar como un dominio vital relevante para el desarrollo de la identidad. Es decir, el proceso de formación de identidad ocurre parcialmente en el ambiente escolar (Eccles & Roeser, 2011; Faircloth, 2009; Flum & Kaplan, 2012; Gee, 2001; Kaplan & Flum, 2012; Lannegrand-Willems & Bosma, 2006; Rich & Schachter, 2012). Debido a que los adolescentes pasan muchas horas al día en la escuela, se presume que el ambiente escolar tendrá una influencia sustancial sobre sus vidas. Como expresa Faircloth (2012), «La identidad puede ser vista como un tipo de negociación de la participación en el curso, modelada por el contexto, en la que ocurre» (p. 186). De modo similar, Flum y Kaplan (2012) indican que la escuela como comunidad social entrega «una red relacional, un sistema de relaciones que crea significado» (p. 241). El proceso de desarrollo de identidad continúa durante la adolescencia tardía (Klimstra, Hale, Raaijmakers et al., 2010; Meeus et al., 1999, 2010) e incluso más allá de la adultez joven (Kroger et al., 2010; véase también Frisen & Wangqvist, 2011; Luyckx, Klimstra, Duriez, Schwartz, & Vanhalst, 2012), hasta que los individuos desarrollan perfiles de dimensión de identidad más estables. Así como el ambiente escolar lo es para los adolescentes, el contexto universitario es considerado un campo vital importante para la formación identitaria en adultos jóvenes.

Procesos de formación de identidad de los estudiantes universitarios

El presente estudio busca expandir el foco al indagar en los procesos de formación de identidad observados en una muestra de estudiantes universitarios. El marco teórico de Crocetti et al. (2008) fue utilizado para medir los procesos de formación de identidad de los estudiantes universitarios durante sus estudios superiores. Meeus (1996) y Crocetti et al. (2008) continuaron el trabajo de Erikson y Marcia y desarrollaron un modelo tridimensional de formación de identidad (el llamado U-MICS), el cual puede usarse para evaluar la identidad de las personas en distintos ámbitos (por ejemplo, en la escuela, en las relaciones de amistad o en la religión). El modelo fue desarrollado para evaluar el contenido de los compromisos en múltiples áreas, la fuerza de dichos compromisos y el grado de exploración de cada área. En el presente artículo la universidad, como área de contenido, será el foco principal.

El modelo de formación de identidad de Crocetti et al. (2008) está conformado por tres dimensiones. La primera dimensión se denomina *compromiso* y se refiere a «una elección hecha en un área de identidad relevante y el alcance que posee dicha elección» (p. 218); en otras palabras, estar comprometido con la identidad que se ha elegido. La segunda dimensión se denomina *exploración en profundidad* y se refiere al grado en el cual los adolescentes exploran activamente sus compromisos actuales, es decir, la forma en que abordan sus compromisos existentes y hasta qué punto están activamente involucrados en investigar sus relaciones (véase también Klimstra, Hale III, Raaijmakers et al., 2010). Las dimensiones de compromiso y exploración en profundidad derivan del paradigma de los estadios de la identidad de Marcia (1966). Marcia describió cuatro estadios de identidad diferentes, sobre la base del grado de compromiso y exploración que experimentan los adolescentes. La tercera dimensión se denomina *reconsideración del compromiso* y se refiere a la comparación de compromisos actuales con otras alternativas posibles. Refleja falta de certeza con respecto a los compromisos. Por lo tanto, también incluye los esfuerzos de las personas de «cambiar compromisos presentes debido a que ya no son satisfactorios» (Crocetti et al., 2008, p. 209). La tercera dimensión no era parte del marco teórico de Marcia, pero fue introducida por Meeus (1996) para comprender cómo los adolescentes construyen y revisan su identidad en el tiempo. Los análisis

factoriales confirmatorios revelaron que el modelo de tres factores de Crocetti et al. (2008) entregaba un mejor ajuste que los modelos alternativos de uno y dos factores, y que el modelo de tres factores aplicaba a las submuestras de hombres y mujeres y a los grupos etarios de adolescentes tempranos y medios. Más aún, las tres dimensiones se asociaron con otros constructos psicológicos relevantes (por ejemplo, medidas de personalidad, problemas psicosociales y relaciones padres-adolescentes) en las direcciones esperadas.

En el presente estudio, los ítems incluidos en las tres escalas originales fueron adaptados al contexto universitario para así poder medir los procesos de formación de identidad de los estudiantes en el ámbito universitario. Como se expresó en la introducción, la indagación en los procesos de formación de identidad de los estudiantes universitarios puede ayudar a identificar a los estudiantes en riesgo de presentar bajo rendimiento o de desertar. Siguiendo el marco de medición de Crocetti et al., puntuaciones relativamente bajas en las primeras dos dimensiones (*compromiso* y *exploración en profundidad*) y una puntuación relativamente alta en la tercera dimensión (*reconsideración del compromiso*), se consideran generalmente como desfavorable. De observarse que la validez de constructo y la validez predictiva del marco de medición de Crocetti et al. fueran suficientes en los estudiantes universitarios incluidos en el presente estudio, las escalas podrían ser usadas por otros investigadores para identificar a los estudiantes en riesgo dentro del dominio universitario.

Método

Participantes

La muestra de participantes estuvo conformada por 80 estudiantes de la Facultad de Ciencias Sociales y del Comportamiento de la Universidad de Groningen, en los Países Bajos. Los estudiantes tuvieron la oportunidad de responder el cuestionario mediante la plataforma de estudio en línea disponible para todos los alumnos de la facultad (tasa de respuesta 2%). La muestra incluyó 68 mujeres (85%) y 12 hombres (15%). Cuarenta y cinco estudiantes eran de origen holandés (56%), 20 de origen alemán (25%), 11 de otros países europeos (14%) y cuatro de países no-europeos (5%). En promedio, los estudiantes tenían 22 años de edad. En total, 34 acababan de comenzar sus estudios universitarios, 16 habían estudiado en la universidad un máximo de dos años y 30 habían estudiado en la institución durante dos años y medio o más.

VARIABLES E INSTRUMENTOS

Procesos de formación de identidad de los estudiantes en la universidad. Se empleó la U-MICS (Crocetti et al., 2008) para medir los procesos de formación de identidad de los estudiantes. Los ítems fueron adaptados al contexto universitario (véase Apéndice A). El instrumento comprende tres subescalas: compromiso (cinco ítems), exploración en profundidad (cinco ítems) y reconsideración del compromiso (tres ítems). Los ítems de muestra incluyen: «Mi universidad me entrega certeza en la vida» (compromiso), «Pienso mucho sobre mi universidad» (exploración en profundidad) y «Frecuentemente pienso que sería mejor encontrar una universidad diferente» (reconsideración del compromiso). Se utilizó una escala Likert de cinco puntos para las respuestas. Las categorías se denominaron del modo siguiente: *completamente falso* (0), *falso* (1), *a veces verdadera* o *a veces no* (2), *verdadero* (3), *completamente verdadero* (4). El instrumento se administró en inglés para asegurar que los estudiantes no holandeses de la facultad también pudieran participar en el estudio. Puesto que en los cursos de la Universidad de Groningen se emplea frecuentemente literatura científica en inglés, consideramos razonable presumir que los estudiantes tendrían un dominio suficiente de esta lengua para comprender los ítems administrados.

Logro académico. Los estudiantes indicaron (auto-reporte) su promedio de notas (en una escala del 1 al 10, siendo un 6 el mínimo para pasar un examen y un 10 la nota más alta) en los cursos que habían tomado el año académico anterior. En los Países Bajos, esta escala de 10 puntos es la que se emplea oficialmente en las evaluaciones (pruebas escritas, exámenes finales) en todos los tipos de educación. Esta información estuvo disponible para 45 estudiantes, dado que los otros 35 aún no recibían los resultados de ninguna evaluación (ni de pruebas escritas ni de exámenes finales). La nota promedio de los 45 estudiantes fue de 7,3 (*DE* 0,7), con un rango de 5,0 a 9,5.

Análisis

Se presentan los resultados descriptivos (medias de puntuaciones logit y desviaciones estándar) para los tres factores (compromiso, exploración en profundidad, reconsideración del compromiso) sugeridos en el marco de medición de Crocetti et al. (2008). Estos valores de la escala logit (calculados sobre la base de datos brutos no intervalares transformados) tienen un significado intervalar constante. Las estimaciones de correlación no paramétricas (rho de Spearman) se reportan entre los tres factores y el logro académico de los estudiantes.

Como aseveran Adams, Wilson y Wang (1997), el Modelo Multidimensional de Coeficientes Aleatorizados Logit Multinomial (Multidimensional Random Coefficient Multinomial Logit, [MRCML]) integra varios modelos Rasch ya existentes, como el Modelo Logístico Simple (Simple Logistic Model) (Rasch, 1961), el Modelo Logístico Lineal de Rasgos Latentes (Linear Logistic Latent Trait Model) (Fischer, 1973), el Modelo de Escalas de Valoración (Rating Scale Model) (Andrich, 1978), el Modelo de Crédito Parcial (Partial Credit Model) (Masters, 1982) y el Modelo de Particiones Ordenadas (Ordered Partition Model) (Wilson, 1992). En el presente artículo, el modelo MRCML aplicado se basa en el Modelo de Escalas de Valoración de Andrich (1978).

El propósito general del modelo Rasch es extraer, a partir de datos adecuados, una definición útil de la variable estudiada (en este caso, los procesos de formación de identidad) y luego medir personas (en este caso, estudiantes universitarios) con respecto a esta variable (Wright & Masters, 1980). El Modelo de Escalas de Valoración original (Andrich, 1978) presupone unidimensionalidad, lo que implica que los ítems estarían trabajando juntos para definir una variable reconocible y significativa (es decir, con validez de constructo). Esto implica que las calibraciones de los ítems se extienden lo suficiente como para definir niveles distintos a lo largo de la variable medida (Wright & Masters, 1980). Adicionalmente, el modelo requiere que las respuestas de cada persona (sus valores theta) se acomoden en general al ordenamiento de los ítems (dificultad de los ítems) para la mayor parte de las personas; de no ser así, hay una falta de ajuste entre los valores theta estimados de las personas y la dificultad de los ítems. La idea general de aplicar un modelo Rasch a datos empíricos es, por lo tanto, realizar una evaluación del ajuste de los ítems (cuán bien se ajustan las respuestas a cada ítem a las expectativas del modelo de medición) y del ajuste de persona (cuán bien se ajustan las respuestas de cada persona a las expectativas del modelo) (Wright & Masters, 1980, p. 90).

El Modelo de Escalas de Valoración aplica una transformación logarítmica para estimar log-odds (logit) para cada ítem (parámetros de dificultad de los ítems) y cada persona (valores theta de las personas) incluidos en el análisis. El logit es la escala logarítmica natural del odds ratio, la cual convierte los valores de dos medidas (ítems y personas) a una escala común. Para datos de escalas de valoración, las escalas tienen tanto una estimación de dificultad como una serie de umbrales. La dificultad de cada ítem se basa en la manera en que las personas evaluadas efectivamente respondieron dicho ítem en la práctica (Bond & Fox, 2001). El umbral es el nivel en el cual la probabilidad de fracaso en una categoría de respuesta dada (bajo el umbral) se convierte en la probabilidad de éxito en dicha categoría (sobre el umbral; Bond & Fox, 2001, pp. 68-70). De este modo, el éxito en una escala de valoración puede interpretarse como apoyo a una categoría de respuesta dada (por ejemplo, *completamente verdadero*). El fracaso, entonces, se interpreta como el no apoyar una categoría de respuesta específica. Para los ítems, los enunciados con puntuaciones altas son aquellos que los estudiantes tendieron a apoyar, mientras que los enunciados con puntuaciones bajas fueron los que los estudiantes tendieron a no apoyar.

Un aspecto importante de los análisis es que algunos estudiantes de nuestra muestra alcanzaron la puntuación correspondiente a la categoría de respuesta más baja en todos los ítems de la tercera escala (reconsideración de compromiso). Debido al sesgo que podría causar este efecto suelo, estas puntuaciones fueron recodificadas como faltantes de acuerdo a la sugerencia de uno de los evaluadores del manuscrito. Los análisis de la tercera dimensión, por lo tanto, se basaron en una muestra de 56 estudiantes (70% de la muestra original). Los análisis de la primera y segunda dimensiones se basaron en la muestra completa de 80 estudiantes.

En el Modelo de Escalas de Valoración (Andrich, 1978), se presume que el espaciamiento de los umbrales es similar para todos los ítems incluidos en el análisis. En el artículo, se aplicó un Modelo de Escalas de Valoración (en lugar de un Modelo de Crédito Parcial), puesto que se conceptualizó que

las respuestas entre ítems de la misma escala tuvieron el mismo significado (sus escalas de valoración tienen igual estructura¹). En otras palabras, se esperaba que todos los ítems midieran el apoyo de los estudiantes a los enunciados sobre sus procesos de formación identitaria empleando las mismas rúbricas. Las propiedades de medición de los tres factores fueron analizadas empleando un modelo MRCML. Este es un modelo Rasch multidimensional para escalas politómicas (por ejemplo, de tipo Likert) (Adams et al., 1997). Un modelo de este tipo puede calibrar todos los factores simultáneamente y aumentar la precisión de la medición tomando en cuenta las correlaciones entre los factores (Wu & Adams, 2006). Se utilizó el programa ConQuest (versión 2.0; Wu, Adams, Wilson, & Haldane, 2007) para realizar los análisis.

Se emplearon medias cuadráticas de residuales (Mean Square [MNSQ]) para determinar cuán bien los datos empíricos cumplieron los requerimientos del modelo Rasch. Estos estadísticos de ajuste (prueba chi-cuadrado) son los valores medios de los residuales al cuadrado; en otras palabras, de las diferencias entre los valores observados y los valores predichos por el modelo Rasch. Se utilizaron dos pruebas chi-cuadrado para evaluar el ajuste de los ítems: estadísticos MNSQ infit (ponderados) y outfit (no ponderados). El estadístico infit se calcula sobre la base de cuánto se alinean los niveles de apoyo de los individuos evaluados con la dificultad de los ítems (es decir, hasta qué punto es fácil o difícil apoyar un cierto ítem) y le da más peso a los ítems con mejor alineamiento. El estadístico outfit no es ponderado. Los estadísticos infit y outfit tienen un valor esperado de 1 y pueden ir desde 0 hasta infinito (Linacre, 2002). Los valores entre 0,50 y 1,50 indican «ajuste útil» (useful fit) (Linacre, 2002) en los Modelos de Escalas de Valoración (Andrich, 1978); sin embargo, otros investigadores sugieren un rango de valores más estrecho, de 0,75 a 1,30 (por ejemplo, Bond & Fox, 2007). Se presentan los estadísticos de ajuste MNSQ basados en el modelo multidimensional y los basados en tres modelos unidimensionales separados. Se entregan las confiabilidades de separación de Person para cada escala. Se añaden mapas Wright para mostrar la estimación del parámetro de dificultad del ítem y el grado en el cual los ítems se alinean con los individuos evaluados (valores theta de los estudiantes).

Resultados

Con respecto a la primera pregunta de investigación (*¿Produce el marco de medición de Crocetti et al. (2008) mediciones de los procesos de formación de identidad de estudiantes universitarios en su institución que sean confiables y posean validez de constructo?*), la validez de constructo del cuestionario U-MICS adaptado fue considerada aceptable. Los ítems de las tres escalas capturaron la naturaleza multidimensional del constructo de los procesos de formación de identidad, lo que quedó demostrado mediante los índices de ajuste al modelo y los parámetros estimados. Los detalles se presentan a continuación.

Las calibraciones de pasos para todos los ítems aumentaron monótonamente de -3,390, -0,786 y 0,760 a 3,417 logits. Se observó una tendencia lineal en las escalas Likert, con distancias entre los pasos de 2,604, 1,546 y 2,657 respectivamente. Este resultado indica que cada paso define una posición distinta en la escala. Los ítems funcionaron adecuadamente en el modelo de tres factores, puesto que todas las distancias entre los pasos fueron superiores a uno (Linacre, 2002). La Tabla 1 muestra las confiabilidades de separación de personas de acuerdo a la Estimación de Máxima Probabilidad [Maximum Likelihood Estimation (MLE)] y la matriz de correlación de factores para el modelo de tres factores.

Tabla 1
Confiabilidades de separación de personas MLE y matriz de correlación de factores

Modelo de 3 factores ^a	Confiabilidad ^b	Compromiso	Exploración en profundidad	Rec. del compromiso
Compromiso	0,82	1		
Exploración en profundidad	0,85	0,18	1	
Reconsideración del compromiso	0,72	-0,67	-0,01	1

Nota: ^a La desviación final del modelo fue 2188,034 (22 parámetros estimados). ^b Confiabilidad de separación de personas MLE.

¹ El Modelo de Crédito Parcial (Masters, 1982) especifica que cada ítem tiene una escala de valoración cuya estructura es única.

Las confiabilidades de separación de personas mayores a ,80 se consideran aceptables. La escala de compromiso (.82) y la escala de exploración en profundidad cumplieron este criterio (.85), mientras que la confiabilidad de la escala de reconsideración del compromiso fue ligeramente más baja (.72). Con todo, las escalas lograron diferenciar a los estudiantes en las variables medidas, aunque sería deseable una mejora de la confiabilidad de la separación de personas en la tercera escala.

Los resultados (véase Tabla 2) mostraron además que el infit MNSQ de los ítems (ajuste ponderado) basado en el modelo multidimensional varió entre 0,53 y 1,47 y que el outfit MNSQ (ajuste no ponderado) varió entre 0,54 y 1,44. El infit MNSQ basado en los modelos unidimensionales (por separado para cada dimensión) varió entre 0,69 y 1,28, mientras que el outfit MNSQ varió entre 0,69 y 1,27. Todos los valores de ítems tuvieron estadísticos de ajuste dentro de los márgenes del «ajuste útil» (useful fit) (es decir, entre 0,50 y 1,50; Linacre, 2002), incluso cuando se inspeccionaron los intervalos de confianza de 95%. Sin embargo, cuatro de los 13 ítems no cumplieron con los criterios de corte más estrictos sugeridos por Bond y Fox (2007) cuando se evaluaron los estadísticos de ajuste de los ítems basados en el modelo multidimensional (el análisis de los modelos unidimensionales mostró que solo el ítem 11 no cumplió estos criterios más estrictos). En estos casos, los datos no se ajustaron al modelo. Más específicamente, tres ítems cruzaron el límite inferior de 0,75. Esta fue la situación del Ítem 1 (infit 0,65 [$t = -2,6$]; outfit 0,64 [$t = -2,5$]) y del Ítem 2 (infit 0,74 [$t = -1,8$]; outfit 0,74 [$t = -1,8$]), ambos parte de la escala de compromiso, y del Ítem 11 (infit 0,53 [$t = 2,6$]; outfit 0,54 [$t = 2,7$]), parte de la escala de reconsideración del compromiso. En estos ítems, el ajuste de los datos al modelo fue mejor de lo esperado (Smith, 1996). Además, el ítem 10, que forma parte de la escala de exploración en profundidad, cruzó el límite superior de 1,30 (infit 1,44 [$t = -2,9$]; outfit 1,47 [$t = -3,0$]). Estos valores indican la existencia de varianza no modelada. Los otros nueve ítems se ajustaron bien al modelo Rasch, ya que cumplieron los requisitos más exigentes de este. Los ítems 1, 2, 10 y 11 (véase Anexo A) deben ser mejorados para mejorar su ajuste al modelo Rasch de tres factores. Este es el caso específico de los ítems 1 y 10, debido a valores t inferiores a -2 (menos varianza de la modelada), y del Ítem 11, a causa de valores t mayores a +2 (más varianza de la modelada). Puede interpretarse que estos ítems tienen menos compatibilidad de lo esperado con el modelo ($p < ,05$; Wright & Masters, 1982). La sección de discusión incluye sugerencias para mejorar estos ítems.

Tabla 2

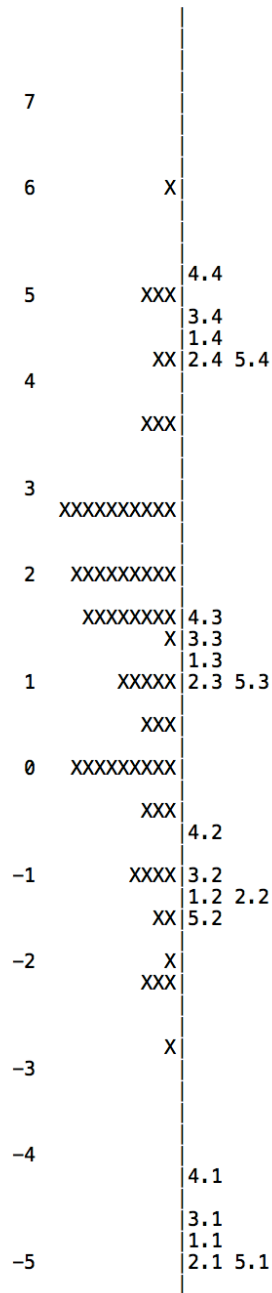
Infit y Outfit MNSQ de los ítems (Estadísticos basados en el modelo multidimensional y en los modelos unidimensionales)

	Infit MNSQ [intervalo de confianza 95%]	Outfit MNSQ [intervalo de confianza 95%]
Estadísticos basados en el modelo multidimensional:		
Compromiso:		
Ítem 1	0,65 [0,70 – 1,30]	0,64 [0,69 – 1,31]
Ítem 2	0,74 [0,69 – 1,31]	0,74 [0,69 – 1,31]
Ítem 3	0,75 [0,70 – 1,30]	0,75 [0,69 – 1,31]
Ítem 4	1,03 [0,70 – 1,30]	1,04 [0,69 – 1,31]
Ítem 5	0,89 [0,69 – 1,31]	0,90 [0,69 – 1,31]
Exploración en profundidad:		
Ítem 6	1,22 [0,71 – 1,29]	1,23 [0,69 – 1,31]
Ítem 7	1,18 [0,71 – 1,29]	1,20 [0,69 – 1,31]
Ítem 8	1,01 [0,70 – 1,30]	0,99 [0,69 – 1,31]
Ítem 9	1,30 [0,70 – 1,30]	1,29 [0,69 – 1,31]
Ítem 10	1,44 [0,71 – 1,29]	1,47 [0,69 – 1,31]
Reconsideración del compromiso:		
Ítem 11	0,53 [0,63 – 1,37]	0,54 [0,69 – 1,31]
Ítem 12	0,99 [0,64 – 1,36]	0,99 [0,69 – 1,31]
Ítem 13	1,07 [0,63 – 1,37]	1,03 [0,69 – 1,31]
Estadísticos basados en los modelos unidimensionales:		
Compromiso:		
Ítem 1	0,82 [0,69 – 1,31]	0,80 [0,69 – 1,31]
Ítem 2	0,87 [0,69 – 1,31]	0,87 [0,69 – 1,31]
Ítem 3	0,93 [0,69 – 1,31]	0,93 [0,69 – 1,31]
Ítem 4	1,27 [0,70 – 1,30]	1,27 [0,69 – 1,31]
Ítem 5	1,04 [0,69 – 1,31]	1,07 [0,69 – 1,31]
Exploración en profundidad:		
Ítem 6	1,08 [0,71 – 1,29]	1,08 [0,69 – 1,31]
Ítem 7	0,98 [0,71 – 1,29]	1,00 [0,69 – 1,31]
Ítem 8	0,96 [0,70 – 1,30]	0,94 [0,69 – 1,31]
Ítem 9	1,04 [0,70 – 1,30]	1,03 [0,69 – 1,31]
Ítem 10	1,05 [0,71 – 1,29]	1,06 [0,69 – 1,31]
Reconsideración del compromiso:		
Ítem 11	0,69 [0,60 – 1,40]	0,69 [0,63 – 1,37]
Ítem 12	1,07 [0,61 – 1,39]	1,07 [0,63 – 1,37]
Ítem 13	1,28 [0,65 – 1,35]	1,26 [0,63 – 1,37]

Los mapas Wright (mapas ítem-persona) para los tres factores se presentan en las Figuras 1 a 3. Estos mapas dan una idea visual de cuán bien los niveles de dificultad de los ítems (parámetros de dificultad de los ítems) alcanzan los niveles de apoyo de los estudiantes (valores theta de los estudiantes). Ambas distribuciones están equiparadas a los dos lados de la escala de medición Rasch. En los gráficos, los valores del lado derecho refieren a los umbrales específicos de los ítems, por ejemplo, 3.4 significa Ítem 3, umbral 4. Los ítems que fueron difíciles de apoyar se ubicaron en la parte superior, mientras que los más fáciles de apoyar se ubicaron más cerca del fondo. Cada x del lado izquierdo se refiere a un número específico de individuos evaluados.

=====
 ConQuest: Software de modelamiento de respuesta generalizada al ítem
 MAPA DE ESTIMACIONES MLE UMBRALES
 =====

Umbrales generalizados de ítems (Dimensión 1: Compromiso)



Cada 'X' representa 1.2 casos

Figura 1. Mapa Wright de la dimensión 1: Compromiso (N = 80).

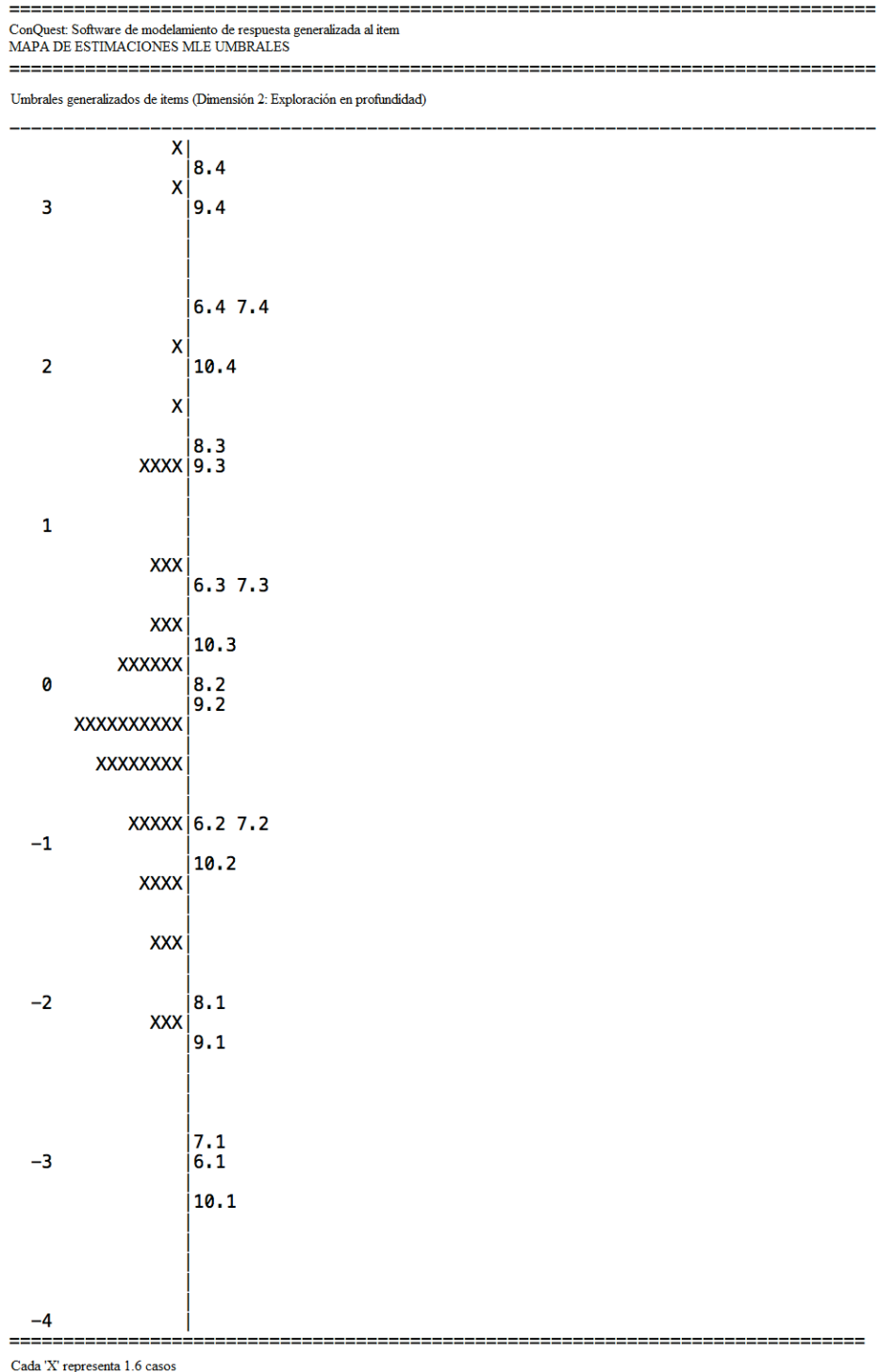


Figura 2. Mapa Wright de la dimensión 2: Exploración en profundidad (N = 80).

=====
 ConQuest: Software de modelamiento de respuesta generalizada al ítem
 MAPA DE ESTIMACIONES MLE UMBRALES
 =====

Umbral generalizado de ítems (Dimensión 3: Reconsideración del compromiso)

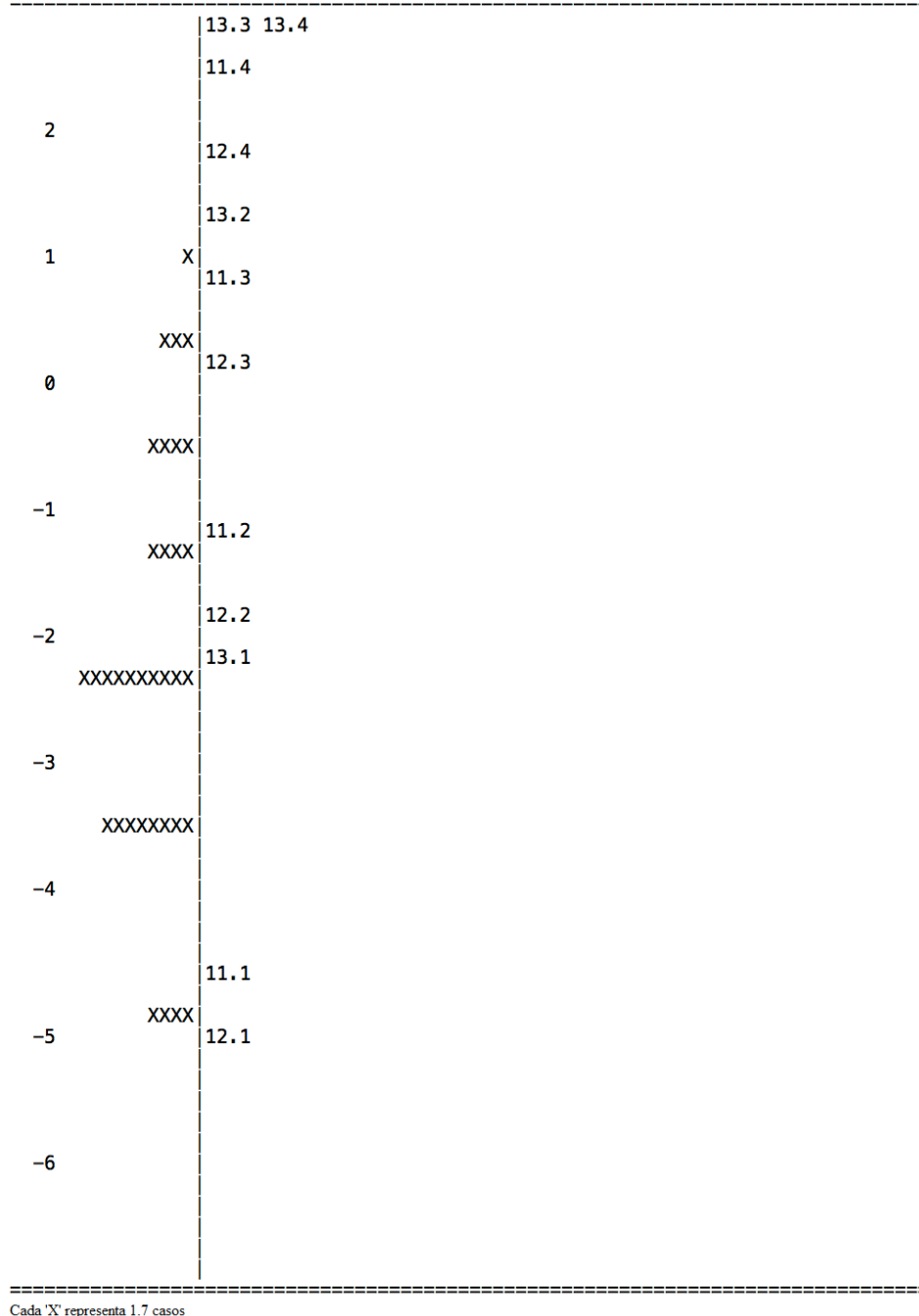


Figura 3. Mapa Wright de la dimensión 3: Reconsideración del compromiso (N = 56).

En las tres escalas, los mapas Wright muestran un buen nivel de coincidencia entre las distribuciones de ítems y de personas. En estas tres dimensiones, la distribución de los ítems y sus umbrales cubren bien los niveles de apoyo.

Con respecto a la segunda pregunta de investigación (*¿Hasta qué punto se relacionan los procesos de formación de identidad de los estudiantes en la universidad con su logro académico (validez predictiva)?*), se obtuvieron los siguientes resultados. En primer lugar, la Tabla 3 muestra las puntuaciones medias logit y las desviaciones estándar en las tres escalas de procesos de formación de identidad.

Tabla 3
Puntuaciones logit medias y desviaciones estándar de los tres factores

	Puntuaciones logit medias	Desviaciones estándar
Compromiso	1,21	0,20
Exploración en profundidad	0,95	0,26
Reconsideración del compromiso	0,67	0,26

Las correlaciones entre los tres factores fueron estimadas con el software ConQuest (véase Tabla 1). La rho de Spearman se utilizó para estimar las correlaciones entre las puntuaciones logit de tres factores y la variable de logro académico (promedio de notas obtenido).² Los coeficientes de correlación revelan que la asociación entre los tres factores y el logro académico fue, en general, baja y que la asociación más fuerte fue la de la escala de compromiso ($r = ,25$; $p < ,10$). Mientras más alto fuese el compromiso de los estudiantes (y hasta cierto punto su exploración en profundidad), más alta fue la nota promedio obtenida (o viceversa). Este hallazgo es coherente con un estudio anterior sobre el compromiso de los estudiantes con su universidad en relación con la dirección de la asociación, es decir, una asociación positiva entre compromiso y funcionamiento exitoso en la universidad (por ejemplo, Allen et al., 2008 hallaron una correlación de ,13 entre el compromiso de los estudiantes con la universidad y su promedio de notas en el primer año). La correlación entre la exploración en profundidad y la nota promedio obtenida también fue positiva ($r = ,22$). En concordancia con el contenido de la tercera escala (reconsideración del compromiso), la correlación entre esta y la nota promedio obtenida fue negativa ($r = -,19$). Para las últimas dos escalas, no hay información de estudios previos que permita comparar la dirección y la magnitud de estas correlaciones entre muestras de estudiantes universitarios, pero la dirección de las asociaciones parece ser coherente con la teoría de desarrollo de la identidad sobre la base de la cual se desarrollaron las escalas.

La estructura de tres factores del constructo sugerida por Crocetti et al. (2008) fue replicada en la muestra de estudiantes empleada. En coherencia con el estudio de validación de la U-MICS, en el cual se empleó una muestra de estudiantes universitarios suizos (Zimmermann, Mahaim, Mantzouranis, Genoud, & Crocetti, 2012), se observó una correlación negativa entre el compromiso y la reconsideración del compromiso, lo que indica que los estudiantes que establecen compromisos más fuertes son menos propensos a reconsiderarlos. La correlación negativa fue más fuerte en la presente muestra ($r = -,68$) que en el estudio suizo ($r = -,26$; Zimmerman et al., 2012), lo que puede explicarse por el hecho de que, en el estudio aquí descrito, muchos estudiantes acababan de iniciar sus estudios en la facultad y, por lo tanto, no habían reconsiderado su compromiso con la universidad al mismo grado en que lo harían otros estudiantes con más experiencia. La correlación positiva entre compromiso y exploración en profundidad fue menos fuerte en nuestro estudio ($r = ,17$), aunque tuvo la misma dirección que la observada por Zimmermann et al. (2012), quienes obtuvieron una correlación de ,43. La relación inexistente entre exploración en profundidad y reconsideración del compromiso coincide exactamente con el estudio de Zimmerman et al. (2012) ($r = ,04$). Es menos probable que quienes se encuentran reconsiderando sus compromisos exploren alternativas de compromiso simultáneamente.

² Los coeficientes de correlación de Spearman estimados pueden tener un cierto sesgo, dado que estos no eran parte de las estimaciones producidas por el programa ConQuest (y, por lo tanto, no se basaron en la covarianza estimada entre las distribuciones de población).

Discusión

En los análisis Rasch, los datos ordenados se manejan apropiadamente, entregando información adicional con respecto a los umbrales de los ítems. Los resultados de este estudio mostraron que todos los umbrales de los ítems en el modelo de tres factores aumentaron monótonamente, en concordancia con las categorías de respuesta de la escala Likert. Sin embargo, varios ítems necesitan mayor refinamiento para aumentar la validez de constructo de las escalas (Ítems 1, 10 y 11). Por ejemplo, el Ítem 10 (y también el 9) se refiere a la exploración en profundidad del compromiso de los estudiantes con su universidad mediante conversaciones con otras personas, mientras que los otros ítems de la escala de exploración en profundidad no necesariamente consideran este aspecto de la exploración. Esos otros ítems se centran en la exploración que los estudiantes realizan individualmente. Aunque los ítems se cargaron en el mismo factor, la escala parece cubrir dos aspectos distintos de la exploración. La primera sugerencia para mejorar las escalas sería, por lo tanto, investigar si la exploración en profundidad debiera dividirse en dos aspectos (sobre la base de la teoría de formación de la identidad) y, de acuerdo a esto, desarrollar ítems adicionales para medir estos dos aspectos por separado. Más aún, el Ítem 10 («*Frecuentemente hablo con otras personas sobre la Universidad de Groningen*») fue relativamente fácil de apoyar por parte de los estudiantes en comparación con los otros ítems de la escala de exploración en profundidad. Una redacción más concreta, como «*Frecuentemente hablo con personas que no pertenecen a la Universidad sobre asistir a la Universidad de Groningen*», podría resolver este problema, porque es imaginable que los estudiantes hablen frecuentemente con otros estudiantes sobre sus experiencias en la universidad. La segunda sugerencia para mejorar la escala se refiere al Ítem 1 («*La Universidad de Groningen me da seguridad en la vida*») de la escala de compromiso, el cual puede interpretarse de múltiples formas (;seguridad con respecto a qué aspecto de la vida?). Dado que 42 de 80 estudiantes respondieron «*verdadero*» y solo tres «*completamente verdadero*» (y ninguno «*completamente falso*»), tal parece que la redacción de la categoría no se ajusta a este ítem. Por lo tanto, se sugiere cambiar la redacción del ítem por un enunciado menos estricto («*me da*» es, en mi opinión, demasiado fuerte, porque los estudiantes nunca podrán saber eso a ciencia cierta), como por ejemplo «*La Universidad de Groningen me hace sentir seguro sobre mi vida futura*». La tercera sugerencia para mejorar las escalas se refiere al Ítem 11 («*Frecuentemente pienso que sería mejor intentar encontrar una universidad distinta*») de la escala de reconsideración del compromiso, aunque también al Ítem 13 («*De hecho, estoy buscando una universidad distinta*»), para los cuales se encontró un fuerte efecto suelo. Casi el 70% de los estudiantes respondió «*falso*» o «*completamente falso*» en el Ítem 11, y lo mismo ocurrió con casi el 90% de los estudiantes en el Ítem 13. Estos hallazgos indican que los ítems fueron demasiado difíciles de apoyar. Los estudiantes no mostraron intención de apoyar estos enunciados de reconsideración del compromiso. Presumiblemente, esto se debe al hecho de que un gran número de estudiantes de la muestra acababa de iniciar sus estudios en la universidad. Sería poco probable que ya estuvieran buscando una nueva universidad. Por lo tanto, se requerirían ítems más fáciles de apoyar para esta dimensión cuando se use el marco de medición de Crocetti et al. (2008) con muestras de estudiantes universitarios de primer año. Las sugerencias para mejorar esta tercera dimensión incluyen: (a) aumentar a cinco el número de ítems (en concordancia con las otras dos dimensiones) para así mejorar la confiabilidad de la escala, añadiendo, por ejemplo, «*A veces dudo si elegí la universidad correcta*» y «*A veces no sé si la universidad me entrega lo que necesito*», ambos más centrados en las razones detrás de la reconsideración de los compromisos; (b) reescribir los Ítems 11 y 12 reemplazando la palabra «*frecuentemente*» por «*a veces*» para que así los estudiantes usen las categorías más extremas con mayor frecuencia; y (c) reescribir el Ítem 13 para hacerlo menos estricto, por ejemplo «*De hecho, estoy buscando oportunidades para cambiarme a otra universidad*», de modo que sea más fácil apoyarlo.

Este estudio no carece de limitaciones. En primer lugar, el tamaño de la muestra (80 estudiantes) fue bastante pequeño. Linacre (2002) sugirió que se requieren al menos $25^*(m + 1)$ sujetos para obtener resultados estables, donde m es el número de categorías presentes en la escala de valoración. Por lo tanto, se necesitaría una muestra de 150 participantes para obtener resultados estables. La tasa de respuesta fue muy baja (2%), dado que los estudiantes fueron invitados a través de una plataforma en línea de la universidad que la mayoría de los estudiantes no revisa frecuentemente. Lamentablemente, esta era la única manera de recolectar datos de los estudiantes, porque no recibimos autorización para contactarlos directamente. Aunque el grupo que respondió es homogéneo con respecto a sus intereses académicos (ciencias sociales y del comportamiento), no está claro si estos estudiantes forman una muestra representativa en cuanto a características de origen. Adicionalmente, el hecho de que solo fuera posible emplear las respuestas de 56 estudiantes para la tercera dimensión debido a un fuerte efecto suelo, dificulta evaluar adecuadamente la validez de esta dimensión. El efecto suelo podría deberse tanto a la manera en que se formularon

los ítems y las categorías de respuesta correspondientes (como lo indican los índices de ajuste), como también por la muestra usada. Por lo tanto, se necesita más investigación que emplee tamaños muestrales mayores y diversas muestras de estudiantes universitarios para evaluar adecuadamente la validez de esta dimensión. En segundo lugar, como consecuencia del pequeño grupo de respuesta, los datos entregaron menos de 10 observaciones para una de las categorías de respuesta de algunos de los ítems (de la escala de reconsideración del compromiso), lo cual violó las presunciones del modelo Rasch. Linacre (2002) presentó lineamientos que indican que, al usar datos de escalas de valoración, se necesitan al menos 10 observaciones por categoría. Cuando hay menos observaciones disponibles en los datos, normalmente se recomienda combinar categorías adyacentes (lo que da como resultado una escala Likert de tres o cuatro puntos), para obtener una estructura robusta de categorías de alta frecuencia. Sin embargo, en esta muestra, las confiabilidades de separación de personas para las tres escalas se redujeron considerablemente cuando se siguió este procedimiento, razón por la cual se decidió mantener la escala Likert original de cinco puntos. Por lo tanto, es recomendable replicar los hallazgos presentados empleando una muestra de estudiantes más numerosa. Estas limitaciones subrayan el hecho de que los resultados descritos deben interpretarse con precaución. Los datos empíricos fueron útiles para ejemplificar la aplicación del análisis multidimensional Rasch, pero se recomienda decididamente replicar el estudio.

Una sugerencia adicional para mejorar aún más el marco de medición podría implicar experimentar con los nombres utilizados en las categorías de respuesta y con el número de categorías de respuesta. Aunque en la muestra encontré distancias entre las categorías mayores a uno, puede decirse que las personas no pueden distinguir fácilmente entre, por ejemplo, *verdadero* y *completamente verdadero*. Dado que los estudiantes universitarios en los Países Bajos usualmente están acostumbrados a las escalas Likert de cinco puntos, al parecer lograron de todos modos diferenciar adecuadamente las categorías; sin embargo, podría ser útil probar distintos nombres en múltiples muestras (véase también Arce-Ferrer, 2006). Otras categorías de respuesta (distintos nombres o menos categorías) podrían producir resultados diferentes. La presencia de estilos comunes de respuesta (véase, por ejemplo, Baumgartner & Steenkamp, 2001), como las respuestas extremas (usar los puntos extremos de la escala de valoración), el asentimiento (apoyar ítems sin importar el contenido) o las respuestas de punto medio (emplear solo categorías intermedias), necesitan investigarse en mayor profundidad para recoger más evidencia sobre la validez de constructo de los procesos de formación de identidad. Por ejemplo, una escala Likert de cuatro puntos podría producir un mejor ajuste de los ítems, porque eliminar la categoría central reduce en cierto grado las diferencias en la tendencia de los estudiantes de dar respuestas extremas o moderadas (de punto medio).

La correlación positiva encontrada entre el compromiso y el logro académico entrega indicios de la validez predictiva del constructo de los procesos de formación de identidad. Sin embargo, sería necesario contar con más información sobre los logros académicos de los estudiantes para indagar más profundamente en esta relación, dado que la información sobre las notas de algunos de ellos no estuvo disponible. Idealmente, en futuros estudios debería usarse un diseño de investigación longitudinal, en el cual los procesos de formación de identidad de los estudiantes en su institución y su logro académico se midieran en múltiples ocasiones. Esto es particularmente importante para medir los procesos de formación de identidad por separado, puesto que en la literatura se ha sugerido la existencia de fluctuaciones de corto plazo en los compromisos y en la reconsideración de los compromisos (Klimstra, Hale III, Frijns et al., 2010).

Más allá de estas limitaciones y sugerencias para investigaciones posteriores y de las formas mencionadas para mejorar las escalas, el presente artículo entregó nuevos antecedentes sobre la confiabilidad, la validez de constructo y la validez predictiva del constructo multidimensional de los procesos de formación de identidad en una muestra de estudiantes universitarios.

El artículo original fue recibido el 6 de noviembre de 2014

El artículo revisado fue recibido el 3 de junio de 2015

El artículo fue aceptado el 4 de junio de 2015

Referencias

- Ackerman, T. A., Gierl, M. J., & Walker, C. M. (2003). Using multidimensional item response theory to evaluate educational and psychological tests. *Educational Measurement: Issues and Practices*, 22, 37-51. doi: 10.1111/j.1745-3992.2003.tb00136.x
- Adams, R. J., Wilson, M. R., & Wang, W. C. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement*, 21, 1-24. doi: 10.1177/0146621697211001
- Allen, J., Robbins, S. B., Casillas, A., & Oh, I.-S. (2008). Third-year college retention and transfer: Effects of academic performance, motivation, and social connectedness. *Research in Higher Education*, 49, 647-664. doi: 10.1007/s11162-008-9098-3
- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME) (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43, 561-573. doi: 10.1007/bf02293814
- Arce-Ferrer, A. J. (2006). An investigation into the factors influencing extreme-response style. Improving meaning of translated and culturally adapted rating scales. *Educational and Psychological Measurement*, 66, 374-392. doi: 10.1177/0013164405278575
- Battista, S. D., Pivetti, M., & Berti, C. (2014). Engagement in the university context: Exploring the role of a sense of justice and social identification. *Social Psychology of Education*, 17, 471-490. doi: 10.1007/s11218-014-9255-9
- Baumgartner, J. E. M., & Steenkamp, H. (2001). Response style in marketing research: A cross-national investigation. *Journal of Marketing Research*, 38, 143-156. doi: 10.1509/jmkr.38.2.143.18840
- Berzonsky, M. D., & Kuk, L. S. (2005). Identity style, psychosocial maturity, and academic performance. *Personality and Individual Differences*, 39, 235-247. doi: 10.1016/j.paid.2005.01.010
- Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum.
- Bosma, H. A. (1985). *Identity development in adolescence. Coping with commitments*. Disertación doctoral no publicada. Groningen, Países Bajos: University of Groningen.
- Côté, J. E., & Levine, C. G. (2014). *Identity formation, agency, and culture. A social psychological synthesis*. Philadelphia, PA: Psychology Press.
- Crocetti, E., Rubini, M., & Meeus, W. (2008). Capturing the dynamics of identity formation in various ethnic groups. Development and validation of a three-dimensional model. *Journal of Adolescence*, 31, 207-222. doi: 10.1016/j.adolescence.2007.09.002
- Eccles, J. S., & Roeser, R. W. (2011). Schools as developmental contexts during adolescence. *Journal of Research on Adolescence*, 21, 225-241. doi: 10.1111/j.1532-7795.2010.00725.x
- Erikson, E. H. (1950). *Childhood and society*. Nueva York: Norton.
- Erikson, E. H. (1968). *Identity, youth and crisis*. Nueva York: Norton.
- Faircloth, B. S. (2009). Making the most of adolescence. Harnessing the search for identity to understand classroom belonging. *Journal of Adolescent Research*, 24, 321-348. doi: 10.1177/0743558409334248
- Faircloth, B. S. (2012). «Wearing a mask» vs. connecting identity with learning. *Contemporary Educational Psychology*, 37, 186-194. doi: 10.1016/j.cedpsych.2011.12.003
- Fischer, G. H. (1973). The linear logistic model as an instrument in educational research. *Acta Psychologica*, 37, 359-374. doi: 10.1016/0001-6918(73)90003-6
- Flum, H., & Kaplan, A. (2012). Identity formation in educational settings: A contextualized view of theory and research in practice. *Contemporary Educational Psychology*, 37, 240-245. doi: 10.1016/j.cedpsych.2012.01.003
- Frisen, A., & Wangqvist, M. (2011). Emerging adults in Sweden: Identity formation in the light of love, work, and family. *Journal of Adolescent Research*, 26, 200-221. doi: 10.1177/0743558410376829
- Gee, J. P. (2001). Identity as an analytic lens for research in education. *Review of Research in Education*, 25, 99-125. doi: 10.2307/1167322
- Hejazi, E., Shahraray, M., Farsinejad, M., & Asgary, A. (2009). Identity styles and academic achievement: mediating role of academic self-efficacy. *Social Psychology of Education*, 12, 123-135. doi: 10.1007/s11218-008-9067-x
- Kaplan, A., & Flum, H. (2012). Identity formation in educational settings: A critical focus for education in the 21st century. *Contemporary Educational Psychology*, 37, 171-175. doi: 10.1016/j.cedpsych.2012.01.005

- Klimstra, T. A., Hale III, W. W., Frijns, T., Luyckx, K., Lier, P. A. C. van, & Meeus, W. H. J. (2010). Short-term fluctuations in identity: Introducing a micro-level approach to identity formation. *Journal of Personality and Social Psychology, 99*, 191-202. doi: 10.1037/a0019584
- Klimstra, T. A., Hale III, W. W., Raaijmakers, Q. A. W., Branje, S. J. T., & Meeus, W. H. J. (2010). Identity formation in adolescence: Change or stability? *Journal of Youth and Adolescence, 39*, 150-162. doi: 10.1007/s10964-009-9401-4
- Korpershoek, H., Xu, J. K., Mok, M. M. A., McNerney, D. M., & van der Werf, G. (2015). Testing the multidimensionality of the Inventory of School Motivation in a Dutch student sample. *Journal of Applied Measurement, 16*, 41-59.
- Kroger, J. (2000). Ego identity status research in the new millennium. *International Journal of Behavioral Development, 24*, 145-148. doi: 10.1080/016502500383250
- Kroger, J., Martinussen, M., & Marcia, J. E. (2010). Identity status change during adolescence and young adulthood: A meta-analysis. *Journal of Adolescence, 33*, 683-698. doi: 10.1016/j.adolescence.2009.11.002
- Lannegrand-Willems, L., & Bosma, H. (2006). Identity development-in-context: The school as an important context for identity development. *Identity, 6*, 85-113. doi: 10.1207/s1532706x0601_6
- Lee, J. C.-K., Zhang, Z., & Yin, H. (2010). Using multidimensional Rasch analysis to validate the Chinese version of the Motivated Strategies for Learning Questionnaire (MSLQ-CV). *European Journal of Psychology of Education, 25*, 141-155. doi: 10.1007/s10212-009-0009-6
- Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3*(1), 85-106. doi: 10.1080/02701367.2008.10599493
- Liu, O. L., Minsky, J., Ling, G., & Kyllonen, P. (2009). Using the standardized letters of recommendation in selection: Results from a multidimensional Rasch model. *Educational and Psychological Measurement, 69*, 475-492. doi: 10.1177/0013164408322031
- Luyckx, K., Klimstra, T. A., Duriez, B., Schwartz, S. J., & Vanhalst, J. (2012). Identity processes and coping strategies in college students: Short-term longitudinal dynamics and the role of personality. *Journal of Youth and Adolescence, 41*, 1226-1239. doi: 10.1007/s10964-012-9753-z
- Marcia, J. E. (1966). Development and validation of ego-identity status. *Journal of Personality and Social Psychology, 3*, 551-558. doi: 10.1037/h0023281
- Marcia, J. E. (1980). Identity in adolescence. En J. Adelson (Ed.), *Handbook of adolescent psychology* (pp. 159-187). Nueva York: Wiley.
- Marcia, J. E. (1994). The empirical study of ego identity. En H. A. Bosma, T. L. G. Graafsma, H. D. Grotevant, & D. J. de Levita (Eds.), *Identity and development. An interdisciplinary approach* (pp. 67-80). Thousand Oaks, CA: Sage Publications, Inc.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*, 149-174. doi: 10.1007/bf02296272
- Meeus, W. (1996). Studies on identity development in adolescence: An overview of research and some new data. *Journal of Youth and Adolescence, 25*, 569-598. doi: 10.1007/bf01537355
- Meeus, W. (2011). The study of adolescent identity formation 2000-2010: A review of longitudinal research. *Journal of Research on Adolescence, 21*, 75-94. doi: 10.1111/j.1532-7795.2010.00716.x
- Meeus, W., Iedema, J., Helsen, M., & Vollebergh, W. (1999). Patterns of adolescent identity development: Review of literature and longitudinal analysis. *Developmental Review, 19*, 419-461. doi: 10.1006/drev.1999.0483
- Meeus, W., van de Schoot, R., Keijsers, L., Schwartz, S. J., & Branje, S. (2010). On the progression and stability of adolescent identity formation: A five-wave longitudinal study in early-to-middle and middle-to-late adolescence. *Child Development, 81*, 1565-1581. doi: 10.1111/j.1467-8624.2010.01492.x
- Meeus, W. H. J., & de Wied, M. A. (2007). Relationships with parents and identity in adolescence: A review of 25 years of research. En M. Watzlawick, & A. Born (Eds.), *Capturing identity: Quantitative and qualitative methods* (pp. 131-147). Lanham: University Press of America.
- Meeuwisse, M., Severiens, S. E., & Born, M. Ph. (2010). Learning environment, interaction, sense of belonging and study success in ethnically diverse student groups. *Research in Higher Education, 51*, 528-545. doi: 10.1007/s11162-010-9168-1
- Mok, M. M. C., McNerney, D. M., Cheng, R. W.-Y., & Lai, M. P. Y. (2011). *Qualitative and quantitative differences in primary and secondary students' learning motivation: Multivariate multilevel and multidimensional Rasch analysis*. Trabajo presentado en la American Educational Research Association 2011 Annual Meeting, Nueva Orleans, LA, EE.UU.

- Mok, M. M. C., McNerney, D. M., & Cheng, R. W.-Y. (2011). *Using multidimensional Rasch modelling to enhance measurement precision: The case of self-processes scales*. Trabajo presentado en la European Association for Research on Learning and Instruction 2011 Biannual meeting, Exeter, Reino Unido.
- Mok, M. M. C., & Xu, K. (2013). Using multidimensional Rasch to enhance measurement precision: Initial results from simulation and empirical studies. *Journal of Applied Measurement, 14*, 27-43.
- Rasch, G. (1961). On general laws and the meaning of measurement in psychology. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, 4*, 321-334.
- Rich, Y., & Schachter, E. P. (2012). High school identity climate and student identity development. *Contemporary Educational Psychology, 37*, 218-228. doi: 10.1016/j.cedpsych.2011.06.002
- Smith, R. M. (1996). Polytomous mean-square fit statistics. *Rasch Measurement Transactions, 10*, 516-517.
- Vleioras, G., & Bosma, H. A. (2005). Are identity styles important for psychological well-being? *Journal of Adolescence, 28*, 397-410. doi: 10.1016/j.adolescence.2004.09.001
- Wang, W.-C., Chen, P.-H., & Cheng, Y. Y. (2004). Improving measurement precision of test batteries using multidimensional item response models. *Psychological Methods, 9*, 116-136. doi: 10.1037/1082-989x.9.1.116
- Waterman, A. S. (1999). Identity, the identity statuses, and identity status development: A contemporary statement. *Developmental Review, 19*, 591-621. doi: 10.1006/drev.1999.0493
- Wilson, M. R. (1992). The partial order model: An extension of the partial credit model. *Applied Psychological Measurement, 16*, 309-325.
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago: MESA Press.
- Wu, M. L., Adams, R. J., Wilson, M. R., & Haldane, S. A. (2007). *ACER ConQuest, Version 2.0* [Computer software]. Camberwell, Victoria, Australia: ACER Press.
- Wu, M. L., & Adams, R. J. (2006). Modelling mathematics problem solving item responses using a multidimensional IRT model. *Mathematics Education Research Journal, 18*, 93-113. doi: 10.1007/bf03217438
- Yan, Z., & Mok, M. M. C. (2012). Validating the Coping Scale for Chinese Athletes using multidimensional Rasch analysis. *Psychology of Sport and Exercise, 13*, 271-279. doi: 10.1016/j.psychsport.2011.11.013
- Zimmermann, G., Mahaim, E. B., Mantzouranis, G., Genoud, P. A., & Crocetti, E. (2012). Brief report: The Identity Style Inventory (ISI-3) and the Utrecht-Management of Identity Commitments Scale (U-MICS): Factor structure, reliability, and convergent validity in French-speaking university students. *Journal of Adolescence, 35*, 461-465. doi: 10.1016/j.adolescence.2010.11.013

Apéndice A

Los ítems adaptados de la U-MICS (Crocetti et al., 2008):

Compromiso

- (1) La Universidad de Groningen me da seguridad en la vida.
- (2) La Universidad de Groningen me da confianza en mí mismo/a.
- (3) La Universidad de Groningen me hace sentirme seguro de mí mismo/a.
- (4) La Universidad de Groningen me entrega seguridad para el futuro.
- (5) La Universidad de Groningen me permite enfrentar el futuro con optimismo.

Exploración en profundidad

- (6) Trato de buscar mucha información sobre la Universidad de Groningen.
- (7) Frecuentemente reflexiono sobre la Universidad de Groningen.
- (8) Me esfuerzo mucho para seguir descubriendo cosas nuevas sobre la Universidad de Groningen.
- (9) Frecuentemente trato de descubrir qué piensan otras personas sobre la Universidad de Groningen.
- (10) Frecuentemente hablo con otras personas sobre la Universidad de Groningen.

Reconsideración del compromiso

- (11) Frecuentemente pienso que sería mejor tratar de buscar otra universidad.
- (12) Frecuentemente pienso que una universidad diferente haría que mi vida fuera más interesante.
- (13) De hecho, estoy buscando otra universidad.