

Modelización de una prueba de afecto hacia la matemática con la teoría de respuesta al ítem

Modelling Affect towards Mathematics scale using Item Response Theory

Facundo Juan Pablo Abal *
Sofía Esmeralda Auné **
Gabriela Susana Lozzia ***
Horacio Félix Attorresi ****

Resumen

Se aplicó el Modelo de Crédito Parcial (MCP) de la Teoría de Respuesta al Ítem (TRI) al análisis de ítems de una escala que mide Afecto hacia la Matemática. Esta variable describe el interés de los estudiantes de Psicología por involucrarse en actividades vinculadas a la matemática y los sentimientos asociados al uso de sus conceptos. La prueba consta de 8 ítems con

formato de respuesta Likert de 6 opciones. Participaron 1875 estudiantes de Psicología de la Universidad de Buenos Aires (Argentina) de los cuales un 82% fueron mujeres. El análisis de la consistencia interna brindó un índice altamente satisfactorio ($\text{Alfa} = .91$). Se verificó la condición de unidimensionalidad requerida por el modelo mediante un análisis factorial exploratorio. Todos los análisis basados sobre la TRI se realizaron con el programa Winsteps. La estimación de los parámetros del modelo se efectuó por Máxima

* Instituto de Investigaciones de la Facultad de Psicología de la Universidad de Buenos Aires, Argentina. La investigación que se presenta en este artículo fue realizada con subsidios de la Universidad de Buenos Aires (UBACyT 2011-14 N° 20020100100346BA y UBACyT 2014-17 N° 20020130100320BA), de la Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT PICT 2011 N° 0826) y en el marco de la Beca Postdoctoral otorgada al Dr. Facundo Juan Pablo Abal por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

Facundo Juan Pablo Abal. Dirección: Zuviría 5691, (1439), CABA.

Tel: (54 11) 4601-4331. Correo electrónico: fabal@psi.uba.ar.

** Sofía Esmeralda Auné. Dirección: Virrey Liniers 577 15° B (1220), CABA.

Tel: (54 11) 4931-3374. Correo electrónico: sofiaaune177@hotmail.com.

*** Gabriela Susana Lozzia. Dirección: Lope de Vega 1507 Dto 2. (1407), CABA.

Tel: (54 11) 4568-6172. Correo electrónico: glozzia@psi.uba.ar.

**** Horacio Félix Attorresi. Dirección: Rivera Indarte 160 1° A (1406), CABA.

Tel: (54 11) 4637-0923. Correo electrónico: hatorre@psi.uba.ar.

Fecha de recepción: 07 de octubre de 2014 - Fecha de aceptación: 15 de diciembre de 2014

Verosimilitud Conjunta. El ajuste del MCP fue satisfactorio para todos los ítems. La Función de Información del Test fue elevada en un rango amplio de niveles del rasgo latente. Un ítem presentó una inversión en dos parámetros de umbral. Como consecuencia, 1 de las 6 categorías del ítem no fue máximamente probable en ningún intervalo de la escala del rasgo latente. Se analizan las implicancias de este hallazgo en la evaluación de la calidad psicométrica del ítem. Los resultados de este estudio permitieron profundizar el análisis del constructo y aportaron evidencias de validez basadas en la estructura interna de la escala.

Palabras clave: Educación Matemática, Actitud hacia la Matemática, Dominio Afectivo, Modelo de Crédito Parcial.

Abstract

The Partial Credit Model (PCM) of Item Response Theory (IRT) was applied in the item analysis of a scale that evaluates Affect towards Mathematics. This variable describes the interest of psychology students to get involved in activities related to mathematics and the feelings associated with the use of its concepts. The scale comprises 8 items in polytomous response format (6-point Likert-type).

The sample was made up by 1875 students of the Psychology school of the University of Buenos Aires (Argentina), 82% of whom were women. The internal consistency analysis provides a highly satisfactory score ($\text{Alpha} = .91$). The unidimensionality assumption required by the model was confirmed through the exploratory factor analysis. All analyses based on IRT were performed by operating the Winsteps software. The parameters estimation was carried out through Joint Maximum Likelihood procedures. The PCM fitted to the data was satisfactorily for all items. Test Information Function was high across a wide range of latent trait. One item showed two reversed thresholds. Consequently, 1 of the 6 item

response was not most likely category for no range of latent trait scale. The implications of this finding in evaluating the psychometric properties of the item are analyzed. The results of this study allowed for deeper analysis of the construct and serve as validity evidence based on internal structure of test.

Key words: Mathematical Education, Attitude towards mathematics, Affective Domain, Partial Credit Model.

Parece una realidad universalmente aceptada que una parte considerable de los estudiantes de carreras universitarias humanísticas y sociales presentan una actitud desfavorable frente al aprendizaje de asignaturas de corte cuantitativo (Blanco, 2004; Narro, 1997; Rodríguez Feijóo, 2011). Tanto Matemática como Estadística son consideradas herramientas fundamentales para adquirir y construir conocimientos en carreras tales como Psicología, Sociología y Ciencias de la Educación pero para sus estudiantes éstas no suelen ser las materias más atractivas. Muchos ingresantes a estas carreras se sorprenden al descubrir que deben cursar materias que demandan capacidad para el cálculo precisamente porque suponen que se trata de carreras no-matemáticas (Coolican, 1994; Cortada, Rodríguez Feijóo & Kohan Cortada, 2008). Docentes e investigadores del área vienen poniendo de manifiesto en diversos contextos culturales la dificultad que tienen los alumnos para alcanzar una cabal comprensión de los conceptos matemáticos y estadísticos (Blanco, 2004, 2008). Las explicaciones y demostraciones matemáticas tienden a omitirse o adecuarse para evitar el rechazo de la materia, lo que va en detrimento de la formación (Ponsoda, 2000).

La enseñanza de asignaturas que incluyen nociones matemáticas en carreras no-matemáticas plantea un desafío que va más allá de la transmisión de los contenidos. El estudio de la Actitud hacia la Matemática en el nivel universita-

rio tiene una extensa tradición. Sin embargo, se ha llevado adelante principalmente en carreras de ciencias físico-químico-matemáticas o tecnológicas (Álvarez & Ruiz, 2010). Por ello, en el marco de una investigación más amplia que estudia diferentes componentes actitudinales en la enseñanza de Matemática y Estadística en estudiantes de Psicología se definió conceptualmente el constructo Afecto hacia la Matemática y se construyó un instrumento para su medición. Esta escala mide el interés del estudiante por involucrarse en actividades vinculadas a la Matemática y sentimientos asociados al uso de sus términos, símbolos y conceptos (Abal, 2013). Aunque con diferentes denominaciones, el componente afectivo de la Actitud hacia la Matemática se encuentra presente en gran parte de los modelos teóricos más destacados (Adelson & McCoach, 2011; Auzmendi, 1992; Bazán & Sotero, 1998; Palacios, Arias & Arias, 2013; Tapia & Marsh, 2004). Se trata de la dimensión con más resistencia al cambio frente a las intervenciones didácticas y la de mayor incidencia en el aprendizaje de los contenidos de la materia (Gómez-Chacón, 2005; McLeod, 1992). Así también, el afecto hacia la Matemática ocupa un lugar predominante en las aproximaciones teóricas que postulan una Actitud hacia la Matemática unidimensional (Brito, 1998; Muñoz & Mato, 2008; Vendramini, Silva & Dias, 2009). El factor común de estas teorizaciones es el énfasis en el carácter positivo de esta dimensión asociado al agrado, disfrute o gusto que experimenta el individuo al enfrentarse con la Matemática y que define su aceptación o rechazo.

La Escala de Afecto hacia la Matemática (Abal, 2013) cuenta con propiedades psicométricas óptimas obtenidas en el marco de la Teoría Clásica de Tests. No obstante, los estudios instrumentales actuales recomiendan incorporar evidencias de calidad psicométrica provenientes de fuentes que toman a los ítems como unidad de análisis (Elosua, 2003). Tal es el caso de los modelos de la Teoría de Respuesta al Ítem (TRI) cuyo aporte al estudio de los ítems enriquece la información suministrada por los indicadores globales de validez y confiabilidad de la teoría clásica

(Attorresi, Abal, Galibert, Lozzia & Aguerri, 2011). Los modelos de la TRI postulan la existencia de una relación directa entre el comportamiento de un individuo frente a un ítem (e.g. elegir una opción de respuesta) y el rasgo inobservable que genera esta conducta. Estos modelos establecen una relación funcional para describir la probabilidad que tiene una persona de elegir cada una de las opciones del ítem en función del nivel del rasgo latente que presente (Muñiz, 2010).

Recientemente Abal, Lozzia y Blum (2013) aplicaron los modelos logísticos dicotómicos de uno, dos y tres parámetros a los ítems politómicos de la escala de Afecto hacia la Matemática. Los resultados demostraron que la dicotomización de los datos perjudica notablemente a la capacidad discriminatoria de la escala y a la precisión de la medida. En consecuencia, se consideró más apropiado aplicar un modelo politómico para estudiar las respuestas de los individuos. Se eligió el Modelo de Crédito Parcial (MCP) de Masters y Wright (1997) por ser uno de los más utilizados para el análisis psicométrico de ítems que miden actitudes y rasgos de personalidad (e.g. Distefano, Morgan & Motl, 2012; Rojas & Pérez, 2001).

A partir de estas consideraciones teóricas y metodológicas se propone como objetivo general de este trabajo examinar la calidad de los ítems de la escala de Afecto hacia la Matemática a partir de la aplicación del Modelo de Crédito Parcial.

Método

Participantes

Se contó con la colaboración de 1875 estudiantes del segundo año de la Facultad de Psicología de Universidad de Buenos Aires. El 82% fueron mujeres y su edad osciló entre 18 y 62 años, con una media de 22.7 años ($DE = 6.33$).

Instrumento

Escala de Afecto hacia la Matemática (Abal, 2013). Consta de ocho enunciados con formato politómico de seis categorías de respuesta tipo Likert. Los anclajes lingüísticos para las opciones fueron: Totalmente en desacuerdo, En desacuerdo, Más bien en desacuerdo, Más bien de acuerdo, De acuerdo y Totalmente de acuerdo. El Alfa de Cronbach es de .91. Los ítems aparecen en la figura 1.

Procedimiento de recolección de datos

Se utilizó un diseño muestral no-probabilístico por accesibilidad. Los estudiantes respondieron el instrumento sin tiempo límite y en grupos reducidos. Previa administración se desarrolló una exposición motivadora que resumió a los participantes la finalidad de la actividad y la futura utilización de los datos.

Análisis de datos

Modelo de Crédito Parcial: La aplicación del MCP supone una segmentación adyacente del dato politómico para establecer una comparación de una puntuación con la inmediatamente anterior. Apelando a una denominación acorde con los test de rendimiento máximo, Master (2010) empleó la expresión paso para describir la transición que un individuo realiza de una categoría (h-1) a la siguiente h. Esto es, suponiendo que la resolución de un ítem de una prueba de habilidad involucra la realización de varios pasos sucesivos, el evaluado será puntuado con categorías de orden creciente considerando la cantidad de pasos completados correctamente. En ítems de tests de comportamiento típico no se obtiene un registro tangible de la transición, por lo tanto este paso se encuentra implícito y debe suponerse.

En el MCP la probabilidad que tiene un

individuo dar cada uno de los pasos de un ítem depende de su nivel de rasgo (θ) y del parámetro de umbral β_h correspondiente a ese paso. Cada parámetro β_h se define como la localización del paso en la escala del rasgo. Se trata de un valor de umbral en el que el evaluado tiene la misma probabilidad de responder la categoría h o en h - 1. Lógicamente, este parámetro está definido solamente para $h = 1, \dots, m$ porque no existe una categoría anterior a $h = 0$ y, por ende, tampoco existe el umbral β_0 . La transición de una categoría a la categoría adyacente supone que ya se han dado los pasos anteriores. El individuo con nivel de rasgo mayor a β_h tendrá una mayor probabilidad de realizar esa transición (e.g. tendrá más chance de elegir la categoría h-1 en lugar de elegir la h) y enfrentarse al siguiente paso. El evaluado que no cuente con una cantidad de rasgo suficiente como para superar el valor de umbral β_h tendrá menos probabilidad de hacer la transición de una categoría a otra.

Aunque la lógica que subyace al MCP es relativamente sencilla, la formulación general reviste cierta complejidad.

La probabilidad de un evaluado de elegir la opción h ($h = 0, \dots, m$) en el ítem i como:

$$P_i(h|\theta) = \frac{e^{\theta - \beta_h}}{\sum_{j=0}^m e^{\theta - \beta_j}}$$

$$\text{Donde se define: } \sum_{h=0}^0 (\theta - \beta_h) = 0$$

Básicamente, esta ecuación muestra cómo la probabilidad de responder a una categoría h se representa como una exponencial correspondiente a esta categoría que se divide por la sumatoria de las exponenciales correspondientes a cada una de las categorías (Embretson & Reise, 2000; Martínez Arias, Hernández Lloreda & Hernández Lloreda, 2006).

Estimación y Ajuste del MCP: Previa aplicación del MCP se verificó el supuesto de unidimensionalidad del constructo con el programa Factor 9.2

(Lorenzo-Seva & Ferrando, 2006). Se realizó un Análisis Factorial Exploratorio utilizando como procedimiento para la extracción de factores el método de mínimos cuadrados simple sobre la matriz de correlaciones de policóricas.

Los análisis basados en la TRI se realizaron con el programa Winsteps versión 3.63.0 (Linacre, 2006). La estimación de los parámetros se efectuó por el Método de Máxima Verosimilitud Conjunta. Considerando las características de la escala Likert usada en la prueba, el MCP requiere de cinco parámetros de umbral ($\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$) que separan las seis categorías de respuesta de los ítems; en consecuencia, fueron estimados un total de 40 parámetros. El ajuste del modelo a los datos se estudió ítem a ítem y también a nivel global considerando los valores de los estadísticos ajuste próximo (infit) como lejano (outfit). El infit detecta desajustes en las desviaciones cerca de la zona de medición del ítem. El outfit, en cambio, detecta desajustes en las desviaciones alejadas de la zona de medición del ítem. Se refiere a las anomalías detectadas en los extremos de la distribución muestral.

Los infit y outfit se interpretan como medias cuadráticas de los residuales no estandarizados (MNSQ) y adoptan un valor de 1 cuando se observa un ajuste perfecto entre los datos y el modelo. Se espera que la mayoría de los valores no excedan el intervalo 0.6 y 1.4 tanto para las personas como para los ítems (Wright & Linacre, 1994). Winsteps también ofrece las medias cuadráticas de los residuos estandarizados (ZSTD) pero este indicador no fue considerado porque resulta sensible para un tamaño de muestra elevado como el que tiene la presente investigación (Linacre, 2012).

Posteriormente, se graficaron con el programa Excel las curvas características de las categorías de respuestas para los ítems, la Función de Información del test y el error típico.

Resultados

Supuesto de unidimensionalidad

Mediante el estudio de Análisis Factorial (Prueba de esfericidad de Bartlett, $\chi^2=8488.3$; $gl= 28$; $p<.0001$, $KMO = .94$) se aisló un único factor con autovalor superior a 1. Este autovalor fue de 5.2 y describió el 65.4% de la varianza. El cociente entre el primer y segundo autovalor fue de 8.67. Las saturaciones factoriales de los ítems resultaron adecuadas dado que oscilaron entre .67 y .88. Estos resultados corroboran la unidimensionalidad de los datos considerando como criterios: a) la regla Kaiser (1960), b) un porcentaje de varianza explicada por el primer factor mayor a 60% (Hair, Anderson, Tatham & Black, 1999) y c) un cociente entre el autovalor del primer factor y del segundo mayor a 5 (Martínez Arias, 1995).

Calibración de los ítems

El proceso de estimación de los parámetros de los ítems alcanzó la convergencia en el ciclo 29. En la tabla 1 se exhiben los resúmenes estadísticos que permiten un estudio global del ajuste del modelo a los datos. Los niveles de Afecto hacia la Matemática de los individuos (θ) se distribuyeron con media 0.07 y desviación estándar de 1.37. Las medias de las MNSQ de infit y outfit en los sujetos se encontraron próximas a 1, reflejando que las respuestas se ajustan a los patrones pronosticados mediante el MCP. Existen, no obstante, valores mínimos y máximos que exceden los límites aceptables. Como afirmaron Revuelta et al (2006), es posible suponer que el número de individuos desajustados es pequeño en comparación con el total de la muestra y, por ende, no afectarán al ajuste del modelo. En relación al ajuste global de los ítems, la tabla muestra el promedio de los β_h de un mismo ítem (denominado β_{prom} en tabla 1), el cual revela la posición

relativa de cada ítem en la escala del rasgo latente al considerarlo como una totalidad. También para los ítems las medias de las MNSQ de infit y outfit fueron cercanas a 1 e incluso los valores extremos estuvieron incluidos en el intervalo de valores recomendado para señalar un ajuste satisfactorio. La tabla 2 amplía la información vinculada al ajuste

de los ítems detallando para cada uno de ellos los β_{Prom} , las MNSQ de Infit y Outfit y los parámetros de umbral β estimados. También aquí se aprecia que los valores de MNSQ se encuentran dentro de los límites aceptables, por lo que es posible considerar que el modelo resulta satisfactorio para representar los datos.

Tabla 1
Ajuste global del MCP

	Ajuste global de sujetos				Ajuste global de ítems			
	θ	<i>Se</i>	In	Out	β_{Prom}	<i>Se</i>	Infit	Outfit
			MNSQ	MNSQ			MNSQ	MNSQ
Media	0,07	0,41	1,01	1,01	0,00	0,03	1,01	1,01
DE	1,37	0,11	0,45	0,47	0,48	0,00	0,21	0,23
Máx	4,41	1,05	7,01	7,77	0,95	0,03	1,31	1,34
Min	-3,83	0,34	0,03	0,04	-0,75	0,02	0,70	0,69

Nota. θ = Nivel de Afecto hacia la Matemática; *Se* = Error de estimación; β_{Prom} = promedio de los β de un ítem; MNSQ = Media cuadrática de los residuales no estandarizados del ajuste interno (INFIT) y ajuste externo (OUTFIT).

Tabla 2
Estimación y ajuste de los ítems.

Ítem	β_{Prom}	<i>Se</i>	In	Out	Parámetros de umbral				
					MNSQ	MNSQ	β_1	β_2	β_3
1	-0,49	0,03	1,31	1,34	-2,01	-1,62	-0,72	-0,04	1,94
2	0,23	0,03	1,19	1,23	-1,57	-0,76	0,00	0,92	2,57
3	0,08	0,02	0,98	0,98	-1,42	-0,87	-0,09	0,58	2,21
4	0,95	0,03	1,22	1,22	-1,21	-0,36	1,06	1,96	3,27
5	0,14	0,03	1,08	1,09	-1,74	-0,78	-0,40	0,85	2,77
6	-0,01	0,02	0,70	0,69	-1,49	-0,73	-0,12	0,53	1,78
7	-0,16	0,03	0,75	0,76	-1,93	-1,19	-0,41	0,62	2,13
8	-0,75	0,03	0,85	0,80	-1,90	-2,05	-0,92	0,01	1,15

Nota. β_{Prom} = promedio de los β de un ítem; *Se* = Error de estimación; MNSQ = Media cuadrática de los residuales no estandarizados del ajuste interno (IN) y ajuste externo (OUT).

Los β se localizaron en un rango del rasgo latente desde -2.05 (β_2 del ítem 8) y hasta 3.27 (β_5 del ítem 4). En términos generales, los β presentaron valores dentro de un rango esperable entre -3 y 3. La escala adoptada tanto para θ como para los parámetros de umbral es arbitraria y no presenta unidades de medida intrínsecamente definidas. Sin embargo, se define por consenso una escala estandarizada de media de 0 y desviación estándar de 1. Así como la mayoría de los individuos se ubican a menos de 3 desviaciones por encima y por debajo de la media de θ , también es razonable esperar que los β se ubiquen en el mismo rango. Sólo el parámetro $\beta_5 = 3.27$ del ítem 4 excedió el rango esperable pero no se considera una desviación importante como para desestimar la calidad del ítem.

El β Prom de los ítems permite identificar una zona de actuación predominante para cada uno de ellos. Al mismo tiempo, permite profundizar el análisis de la relación existente entre el constructo Afectivo y la respuesta a la escala Likert de los ítems. Como los β prom se encuentran en la misma escala que el rasgo, admiten un ordenamiento de los indicadores de los reactivos en función del θ que demandan. La mayoría de los β prom se ubicaron en una zona de actuación en torno al valor central del rasgo con valores entre -0.16 y 0.23. Allí se localizan indicadores asociados a sentimientos de fastidio (ítem 6), desgano (ítem 3), desagrado (ítem 7), aburrimiento (ítem 2) o atracción (ítem 5) que pueden despertarse cuando el evaluado se vincula con la Matemática.

En la zona de actuación más baja (nivel medio-bajo de la variable) se observa que estar en desacuerdo con el ítem 8 (sentir desprecio por todo lo relacionado con los números y la matemática) demanda el menor nivel de Afecto dentro de los indicadores propuestos en la escala. Se trata de un ítem con redacción negativa, por lo que su β prom debe interpretarse como un punto central en la transición que los evaluados hacen de estar totalmente de acuerdo a totalmente en desacuerdo. También en la zona de actuación baja se localiza β prom = -0.49 correspondiente al ítem 1. En este caso, por tratarse de un ítem positivo, el parámetro

permite inferir que estar de acuerdo con la posibilidad de disfrutar de una clase de Matemática demanda un nivel bajo de Afecto. El ítem 4 fue el único con una zona de actuación más elevada, de lo que se deduce que experimentar alegría al utilizar nociones de matemática en otras materias es el indicador que requiere mayor nivel de Afecto.

Análisis de las Curvas Características

En la figura 1 se reproducen las Curvas Características de las categorías de respuestas de los ítems que describen la probabilidad de elección de cada opción de la escala Likert en función del nivel del rasgo latente.

Además de los β Prom, los parámetros β_3 también brindan información útil para describir el comportamiento general de los individuos con respecto a los ítems en virtud de su posición central. Los valores estimados de cada β_3 muestran el nivel de Afecto que un individuo necesita para que, con una probabilidad de .50, elija la opción Más bien de acuerdo en lugar de escoger la categoría Más bien en desacuerdo. Lógicamente, en el caso de los ítems inversos la transición se hará de Más bien en desacuerdo a Más bien de acuerdo. En consecuencia, el valor de β_3 representa un punto dentro de la escala del rasgo que separa la propensión del evaluado a responder en un sentido desfavorable o favorable respecto del contenido propuesto por el ítem.

A pesar de que los valores de β h logran cubrir una mayor extensión de la escala de la variable para los θ superiores a la media, se observa una tendencia de los umbrales a posicionarse en los niveles medio-bajos del rasgo latente. A excepción de los ítems 4 y 2, el β_3 del resto de los ítems fueron negativos dando cuenta de que los ítems del test de Afecto hacia la Matemática despertaron un nivel de adhesión latente. A excepción de los ítems 4 y 2, el β_3 del resto de los ítems fueron negativos dando cuenta de que los ítems del test de Afecto hacia la Matemática despertaron un nivel de adhesión considerable por parte de los estudiantes. Incluso

para el ítem 1 también el parámetro β_4 se ubicó ligeramente por debajo de la media del rasgo y sólo resultó positivo el parámetro β_5 . Esto señala que el contenido de este ítem requirió de menor nivel de Afecto para tender a estar De acuerdo.

Los parámetros β_h estimados para los ítems de la escala mostraron, en su gran mayoría, valores ordenados de forma creciente. Esto constituye una importante propiedad psicométrica de los ítems de la escala en tanto que garantiza que todas las categorías de la escala Likert son útiles para discriminar en algún rango específico de la variable. No obstante, en la tabla 1 se puede apreciar que la modelización del reactivo 8 registró una inversión entre las ubicaciones de los parámetros $\beta_1 = -1.9$ y $\beta_2 = -2.05$. Este resultado se ve reflejado en el gráfico correspondiente al ítem que aparece en la figura 1, donde la curva de la categoría De acuerdo no llega a ser máximamente probable para ningún rango del espectro del rasgo.

En principio, es posible concluir que, para este ítem, las categorías de respuesta de la escala Likert no funcionaron acorde a lo esperable. Un estudiante A cuyo nivel de Afecto θ es de -2.05 (o sea que $\theta_A = \beta_2$) tiene la misma probabilidad de responder De acuerdo o Más bien de acuerdo. En tanto que otro estudiante B con $\theta_B = -1.9$ tiene idéntica probabilidad de elegir la Totalmente de acuerdo o De acuerdo. Esto implica que A, aunque tenga un menor nivel de Afecto que B ($\theta_A < \theta_B$) podría optar por una categoría en la escala Likert que representa a un mayor nivel de Afecto. En consecuencia, la inversión demuestra que las diferentes categorías de la escala Likert no reflejan crecientes niveles del rasgo latente. Por ende, no se puede sostener el supuesto de monotonía creciente necesario para una medición válida entre la escala ordinal del ítem y la escala continua del rasgo.

Precisión de la medida

La figura 2 muestra la Función de Información del Test (FIT) y el error típico de medida obtenidos bajo el MCP. El test brinda una

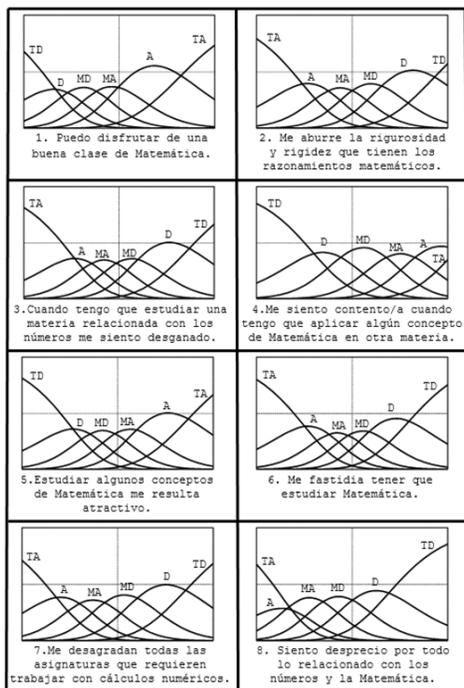


Figura 1. Curvas Características de las Categorías de Respuesta de los ítems. TD = Totalmente en desacuerdo; D = En desacuerdo; MD = Más bien en desacuerdo; MA = Más bien de acuerdo; A = De acuerdo; TA = Totalmente de acuerdo.

máxima información de 8.6 en $\theta = -0.5$, el cual podría considerarse como un nivel próximo al valor central de la distribución del rasgo latente. Dada su relación inversa, la función correspondiente al error típico mostró para el mismo nivel de la variable un valor mínimo de 0.34.

Otra característica de interés que mostró la FIT es su simetría respecto del valor θ próximo a 0. Esto demuestra que la prueba resultó precisa para medir en un nivel del rasgo latente en el que se encuentra la mayor cantidad de individuos de la población. En cambio, el error en la medición crece hacia los valores extremos del rasgo.

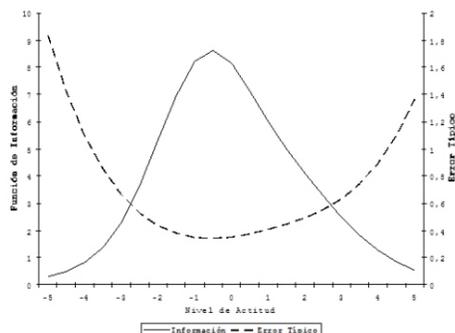


Figura 2. Función de Información del Test y Error Típico de medida.

Discusión

El estudio del afecto involucrado en el proceso de enseñanza-aprendizaje de la Matemática presenta una larga historia plagada de debates y reformulaciones (Gómez Chacón, 2005; Hernández, 2011; Mc Leod, 1992). Si bien es cierto que todas las intervenciones didácticas deben considerar la implicancia del afecto, en el contexto de las carreras no-matemáticas cobra mayor relevancia comparado con otras materias que cuentan con una actitud favorable por parte del estudiante (Cortada et al., 2008). En este sentido, la construcción de un instrumento de medición como el que se presenta en esta investigación contribuye a caracterizar un aspecto de esta dimensión afectiva en una población particular como es la de estudiantes universitarios de Psicología.

Por otro lado, la modelización de la escala de Afecto hacia la Matemática desde la perspectiva de la TRI mostró la utilidad que tiene la aplicación de esta teoría en el análisis de ítems de un test que mide aspectos no-cognitivos de la personalidad. En efecto, la TRI se utiliza con frecuencia en el campo educativo para la modelización de tests de habilidades, rendimiento o aptitudes pero las ventajas de su aplicación a tests de actitudes o rasgos de personalidad están poco difundidas (Abal, Lozzia, Aguerri, Galibert & Attorresi,

2010).

La aplicación del MCP a la escala de Afecto hacia la Matemática permitió acceder a una valiosa comprensión del rasgo medido a partir del análisis exhaustivo de los parámetros β_h de los ítems. Este estudio reveló información respecto de la relación que existe entre los indicadores propuestos por los reactivos y el constructo medido identificando la capacidad de las opciones de respuesta para discriminar rangos específicos del espectro de la variable. Así también, la FIT reveló que la escala brinda una información elevada a lo largo de una extensión considerable del recorrido del rasgo.

Aunque los ítems de la escala mostraron en general un óptimo comportamiento, el análisis del ítem 8 bajo el MCP dejó expuestas sus debilidades. Cabe destacar que la aparición de inversiones de los parámetros ha despertado una fuerte controversia aún no esclarecida. Para algunos autores no es necesariamente un problema que deba llevar a la pérdida del ítem sino que más bien refleja la distribución de los evaluados a través de las categorías de respuesta. Por esta razón, Adams, Wu y Wilson (2012) recomiendan colapsar las categorías cuyos parámetros resultaron invertidos. Sin embargo, otros autores han cuestionado este planteo mostrando la necesidad de profundizar las investigaciones (Andrich, 2013; Wetzel & Carstensen, 2014). Tanto los anclajes lingüísticos de las categorías de respuesta del ítem como la serie de números enteros sucesivos usados en la codificación presentan la propiedad de orden. Se espera, en consecuencia, que los rangos de la variable donde las curvas se tornan máximamente probables también adopten niveles crecientes (o decrecientes) de la variable. La aparición de β_h invertidos indica que el sistema utilizado para la codificación de los ítems no reúne los elementos que lógicamente deben cumplirse para una medición válida. Es decir, no es posible corroborar para este ítem una monotonía creciente entre la escala Likert y la escala del rasgo latente. Por lo tanto, bajo las condiciones de aplicación que impone el MCP el ítem 8 debería ser sustituido o eliminado.

Podría resultar de interés intentar comprender el poco atractivo que tuvo la opción de acuerdo entre los estudiantes de bajo nivel de Afecto; o bien, la tendencia de estos estudiantes a preferir la categoría Totalmente de acuerdo. Si se analiza el enunciado del reactivo es posible afirmar que el mismo describe un sentimiento de odio hacia la Matemática expresado en un nivel extremo (desprecio). A modo de hipótesis podría conjeturarse que es justamente este carácter extremo lo que originó que un grupo de estudiantes con nivel de Afecto medio-bajo, motivados por una reacción afectiva, hayan encontrado que la expresión Totalmente de acuerdo es la que mejor representaba su opinión. Por otra parte, otro conjunto de individuos reconoció, aun estando de acuerdo con lo que expresa el ítem, que la intensidad de odio reflejada por el enunciado era desmedida y se volcaron hacia una categoría más conservadora como es Más bien de acuerdo.

Los hallazgos recogidos en esta aplicación del MCP han servido para orientar las modificaciones que requiere la escala de Afecto hacia la Matemática. Pero además han proporcionado nuevos objetivos para futuras investigaciones que permitan perfeccionar el instrumento. Se propone estudiar el comportamiento de los reactivos de la escala (en particular el ítem 8) a partir de otros modelos politómicos. Así también, resultará de interés analizar si la reducción de algunas categorías de respuesta de la escala Likert o la implementación de una opción central podría afectar de manera considerable a indicadores de validez y confiabilidad.

Futuras investigaciones también se centrarán en realizar las adaptaciones correspondientes a los ítems para aplicar la escala en estudiantes de otras carreras humanístico-sociales. Un aumento en la heterogeneidad de la muestra permitiría obtener indicadores para estudiar la invariancia de los parámetros estimados en las distintas poblaciones dianas. Asimismo, se apuntará a la obtención de evidencias externas de validez para establecer las relaciones del rasgo evaluado con otros constructos o variables observables.

Referencias

- Abal, F. (2013). Comparación de modelos dicotómicos y politómicos de la Teoría de Respuesta al Ítem aplicados a un test de comportamiento típico. Tesis Doctoral no publicada. Universidad de Buenos Aires, Argentina.
- Abal, F., Lozzia, G. & Blum, G. (2013). Modelos Logísticos de la TRI aplicados a un test que mide el Afecto hacia la Matemática en estudiantes de Psicología. Trabajo presentado en el V Congreso Internacional de Investigación y Práctica Profesional en Psicología. XX Jornadas de Investigación y VIX Encuentro de Investigadores en Psicología del MERCOSUR. Buenos Aires, Argentina.
- Aval F., Lozzia, G., Aguerri, M., Galibert, M. & Attorresi, H. (2010). La escasa aplicación de la Teoría de Respuesta al Ítem en Tests de Ejecución Típica. *Revista Colombiana de Psicología*, 19 (1) 111-122.
- Adams, R.J., Wu, M.L. & Wilson, M. (2012). The Rasch rating model and the disordered threshold controversy. *Educational and Psychological Measurement*, 72 (4), 547-573.
- Adelson, J. L. & McCoach, D. B. (2011). Development and psychometric properties of the Math and Me Survey: Measuring third through sixth graders' attitudes towards mathematics. *Measurement and Evaluation in Counseling and Development*, 44, 225-247.
- Álvarez, Y. & Ruiz, M. (2010). Actitudes hacia las matemáticas en estudiantes de ingeniería en universidades autónomas venezolanas. *Revista de Pedagogía*, 31 (89), 225-249.
- Andrich, D. (2013). An Expanded Derivation of the Threshold Structure of the Polytomous Rasch Model That Dispels Any "Threshold Disorder Controversy". *Educational and Psychological Measurement*, 73 (1), 78-124.
- Attorresi, H., Abal, F., Galibert, M., Lozzia, G. & Aguerri, M. (2011). Aplicación del Modelo de Respuesta Graduada a los ítems de una escala

- de Voluntad de Trabajo. *Interdisciplinaria*, 28 (2), 231-244.
- Auzmendi, E. (1992). Las actitudes hacia la matemática-estadística en las enseñanzas medias y universitarias. Bilbao, España: Mensajero.
- Bazán J. & Sotero H. (1998). Una aplicación al estudio de actitudes hacia la Matemática en la UNALM. *Anales Científicos de la Universidad Nacional Agraria La Molina*, 36, 60-72.
- Blanco, A. (2004). Enseñar y aprender Estadística en las titulaciones universitarias de Ciencias Sociales. En J. C. Torre Puente & E. Gil Coria (Eds.), *Hacia una enseñanza universitaria centrada en el aprendizaje* (pp. 143-190). Madrid, España: Universidad Pontificia Comillas.
- Blanco, A. (2008). Una revisión crítica de la investigación sobre las actitudes de los estudiantes universitarios hacia la Estadística. *Revista Complutense de Educación*, 19 (2), 311-330.
- Brito, M. (1998). Adaptação e validação de uma escala de atitudes em relação à Matemática. *Zetetiké*, 6, 109-161.
- Coolican, H. (1994). *Métodos de investigación y estadística en psicología*. México: El Manual Moderno.
- Cortada, N., Rodríguez Feijóo, N. & Kohan Cortada, A. (2008). Análisis de las sugerencias para modificar las actitudes negativas hacia la estadística. Trabajo presentado en el I Encuentro de Docentes e Investigadores de Estadística en Psicología, Buenos Aires, Argentina.
- DiStefano, C., Morgan, G.B. & Motl, R.W. (2012). An examination of personality characteristics related to acquiescence. *Journal of applied measurement*, 13(1), 41-56.
- Elosua, P. (2003). Sobre la validez de los tests. *Psicothema*, 15(2), 315-321.
- Embretson, S. & Reise, S. (2000). *Item Response Theory for Psychologists*. Mahwah, NJ, EEUU: Erlbaum Publishers.
- Gómez-Chacón, I. (2005). *Matemática emocional. Los afectos en el aprendizaje matemático*. España: Narcea.
- Hair, J. F., Anderson, R. E., Tatham, R. L. & Black, W. C. (1999). *Análisis Multivariante*. Madrid, España: Prentice Hall.
- Hernández, G. (2011). Estado del arte de creencias y actitudes hacia las matemáticas. *Cuadernos de Educación y Desarrollo*, 3 (24). Extraído el 10 de abril de 2011 de <http://www.eumed.net/rev/ced/24/ghs.htm>
- Kaiser, H. F. (1960) The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20, 141-151.
- Linacre, J. M. (2006). *Winsteps® (Version 3.63.0) [Computer Software]*. Beaverton, Oregon, EEUU: Winsteps.com.
- Linacre, J. M. (2014). *Winsteps® Rasch measurement computer program User's Guide*. Beaverton, Oregon, EEUU: Winsteps.com.
- Lorenzo-Seva, U. & Ferrando, P. J. (2006). FACTOR: A computer program to fit the exploratory factor analysis model. *Behavioral Research Methods, Instruments and Computers*, 38(1), 88-91.
- Martínez Arias, M. R. (1995). *Psicometría: Teoría de los Tests Psicológicos y Educativos*. Madrid, España: Síntesis.
- Martínez Arias, M. R., Hernández Lloreda, M. V. & Hernández Lloreda, M. J. (2006). *Psicometría*. Madrid, España: Alianza Editorial.
- Masters, G. N. & Wright, B. D. (1997). The Partial Credit Model. En W. J. Van der Linden y R. K. Hambleton (Eds.), *Handbook of Modern Item Response Theory* (pp. 101-121). New York, EEUU: Springer.
- Masters, G. N. (2010). The Partial Credit Model. En M. Nering & R. Ostini (Eds.), *Handbook of Polytomous Item Response Theory Models* (pp. 109-122). New York, EEUU: Routledge Academic.
- McLeod, D. (1992). Research on affect in mathematics education: A reconceptualization. En D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 575-596). New York, EEUU: MacMillan.
- Muñiz, J. (2010). Las teorías de los tests: Teoría clásica y Teoría de respuesta a los ítems. *Papeles del Psicólogo*, 31 (1), 57-66.

- Muñoz, J. M. & Mato, M. D. (2008). Análisis de las actitudes respecto a las Matemáticas en alumnos de ESO. *Revista de Investigación Educativa*, 26(1), 209-226.
- Narro, A. E. (1997). Investigación sobre la concepción de la matemática en ciencias sociales en la UAM-XOCHIMILCO. *Política y Cultura*, 9, 249-280.
- Palacios, A., Arias, V. & Arias, B. (2014). Attitudes Towards Mathematics: Construction and Validation of a Measurement Instrument. *Revista de Psicodidáctica*, 19(1), 67-91.
- Ponsoda, V. (1990). Un punto de vista sobre la docencia de la Estadística en Psicología. *Estadística Española*, 31, 481-489.
- Revuelta, J., Abad, F. J. & Ponsoda, V. (2006). *Modelos Politémicos de respuesta al ítem*. Madrid, España: La Muralla.
- Rodríguez Feijóo, N. (2011). Actitudes de los estudiantes universitarios hacia la estadística. *Interdisciplinaria*, 28(2), 199-205.
- Rojas, A. J. & Pérez, C. (2001) *Nuevos Modelos para la Medición de Actitudes*. Valencia, España: Promolibro.
- Tapia, M. & Marsh, G. E. (2004). An instrument to measure mathematics attitudes. *Academic Exchange Quarterly*, 8, 16-21.
- Vendramini, C., Silva, M. & Dias, A. (2009). Avaliação de atitudes de estudantes de psicologia via modelo de crédito parcial da TRI. *Psico-USF*, 14(3), 287-298.
- Wetzel, E. & Carstensen, C. H. (2014, en prensa). Reversed Thresholds in Partial Credit Models. A Reason for Collapsing Categories? *Assessment*.
- Wright, B. D. & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transaction*, 8, 370.