

Validación de la Versión en Español del *Mathematics Value Inventory* (MVI) entre Estudiantes Universitarios Uruguayos

Validation of the Spanish Version of the Mathematics Value Inventory (MVI) among Uruguayan University Students

Maria Noel Rodríguez-Ayan* & María Eugenia Sotelo Rico
Universidad de la República, Montevideo, Uruguay

Resumen

Con base en la teoría de expectativa-valor, el *Mathematics Value Inventory* (MVI) fue diseñado para medir la valoración de la matemática en cuatro áreas: interés, utilidad, logro y costo. En este trabajo se presenta la validación de la versión en español entre estudiantes universitarios uruguayos: validez factorial mediante análisis factorial confirmatorio (AFC), diferencias según género y validez predictiva del rendimiento académico. Se realizó un muestreo aleatorio estratificado simple según el rendimiento en matemática para un nivel de confianza de 95% y un error de 5%, participando voluntariamente 218 estudiantes (77% mujeres) de entre 18 y 29 años ($M = 20.2$, $SD = 1.8$). Los resultados confirman la estructura de cuatro factores relacionados y sugieren un factor global de segundo orden, que puede resultar más útil para una evaluación continua; el modelo de cuatro factores sería una mejor estrategia para detectar áreas problemáticas y comparar instituciones. Mayor interés y mayor necesidad de logro están asociados con un mejor rendimiento futuro; la utilidad y el costo emocional no tendrían impacto. No se encontró evidencia de diferencia según género. En suma, el MVI resulta adecuado para medir actitudes estudiantiles hacia la matemática, lo que puede orientar el diseño de estrategias de intervención.

Palabras clave: Matemática, interés, utilidad, logro, costo personal.

Abstract

Based on the expectancy-value theory, the Mathematics Value Inventory (MVI) was intended to measure math value in four areas: interest, utility, achievement and personal cost. This paper presents the validation of the Spanish version among Uruguayan university students: factorial validity through confirmatory factor analysis (CFA), gender related differences and predictive validity of academic achievement. A stratified random sampling according to Mathematics achievement was performed, with a confidence level of 95% and an error margin of 5%, and 218 students (77% women) voluntarily participated. Ages ranged from 18 to 29 years ($M = 20.2$, $SD = 1.8$). Results confirm the four-factor structure as well as suggest a second order factor. The latter is recommended for continuous assessment of general attitude towards math, while the four-factor model could be a better strategy to identify problematic areas and career comparison. Higher interest and need for achievement are related to greater academic performance, while utility and personal cost seem to have no impact on it. No gender differences on math value were found. In sum, the MVI proved to be a suitable instrument to measure students' attitudes towards mathematics, which may help the design of intervention strategies.

Keywords: Mathematics, interest, utility, achievement, personal cost.

Una preocupación compartida por la comunidad educativa es lograr que los alumnos valoren la matemática, tanto aquellos matriculados en programas de estudio en los que matemática no es una asignatura troncal como quienes provienen de carreras de orientación científico-tecnológica. En Iberoamérica en el segundo ciclo de

educación secundaria la matemática resulta ser un factor altamente discriminante en el momento de decidir cuál orientación realizar en el bachillerato (Polino, 2012). En Uruguay, el bajo rendimiento en matemática en la universidad es un problema que existe desde hace varios años, en particular en carreras de orientación científico-tecnológica, en las cuales la carga de matemática es alta. Las pruebas diagnósticas que se realizan al ingreso a las carreras universitarias de química han mostrado que el 70% de la prueba es resuelto por menos del 5% de los estudiantes (Rodríguez Ayán, Sotelo, Núñez, & González, 2011).

* Dirección postal: Universidad de la República, Unidad Académica de Educación Química, Facultad de Química, Isidoro de María 1620, 2° piso, CP 11800, Montevideo, Uruguay. E-mail: mayan@fq.edu.uy

Asimismo, los niveles más altos de reprobación de cursos y/o exámenes en estas carreras se ubican principalmente en las asignaturas de primer año del área Físico-Matemática (Rodríguez Ayán & Sotelo, 2012).

Debido a estos antecedentes se decidió explorar cuál es la valoración que hacen de la matemática los estudiantes de las cinco carreras profesionales de química de la Universidad de la República, todas con un contenido importante de matemática. El marco teórico para este trabajo está tomado de la teoría de expectativa-valor (Eccles et al., 1983; Eccles, Adler, & Meece, 1984; Wigfield & Eccles, 1992, 2000), según la cual las expectativas de los estudiantes sobre su éxito en las actividades y el valor que atribuyen a estas ejercen influencia sobre sus elecciones, sobre la persistencia en los estudios y sobre su rendimiento académico. El concepto global de valor atribuido se refiere a la percepción de los estudiantes sobre el interés que para él representa la actividad, su valor utilitario, su importancia para la concreción de logros y el costo personal asociado.

El modelo de Eccles et al. (1983; Eccles et al., 1984) se ha aplicado ampliamente entre estudiantes de enseñanza media y tanto las expectativas de éxito como la valoración parecen estar vinculados al rendimiento académico (e.g., Eccles & Wigfield, 2002; Pintrich & Schunk, 2002). Pero su aplicación en contextos universitarios es más bien escasa y entre los antecedentes disponibles no hemos encontrado resultados concluyentes para los efectos de las dimensiones de valoración sobre el rendimiento. En algunos trabajos se evaluó el efecto de la valoración global, por lo que no es posible discriminar el aporte de cada dimensión (e.g., Hood, Creed, & Neumann, 2012; Neuville et al., 2007). Bong (2001) cuantificó el efecto de la valoración global, mencionando (sin cuantificar) a las dimensiones utilidad y logro como mejores predictores que el interés. Cole, Bergin y Whittaker (2008) identificaron efectos de las componentes utilidad y logro mediados por el esfuerzo, quedando recogido el efecto del interés en las correlaciones con utilidad y logro. Hulleman, Durik, Schweigert y Harackiewicz (2008) encontraron efectos únicamente para la dimensión utilidad. No debe dejar de mencionarse que en las investigaciones se han empleado distintas técnicas de modelado, con la inclusión de distintas variables además de las valoraciones, con la consiguiente dificultad para comparar los resultados.

Tomando como base la teoría de expectativa-valor, con énfasis en que ya desde el primer año escolar los estudiantes valoran las distintas disciplinas de distinta manera (Eccles, Wigfield, Harold, & Blumenfeld, 1993), Luttrell et al. (2010) desarrollaron el *Mathematics Value Inventory* (MVI), para medir el valor atribuido a la matemática por parte de estudiantes universitarios de carreras en las que la matemática no es una asignatura troncal. Examinaron la relación de las puntuaciones con el género y con el número de cursos de matemática en que los estudiantes elegían matricularse. El MVI fue conceptualizado como compuesto por cuatro dimensiones: interés, utilidad, logro y costo personal. Las puntuaciones de la versión final de 28 ítems fueron sometidas a análisis factorial exploratorio

para hallar la estructura latente y obteniendo las cuatro dimensiones pretendidas, cada una compuesta por 7 ítems, que explican el 71% de la varianza total. Los autores no encontraron diferencias significativas de valoración según género, en ninguna de las dimensiones. Asimismo, encontraron que aquellos estudiantes matriculados en tres o más cursos de matemática obtuvieron puntuaciones más altas en las escalas de interés y utilidad, a la vez que percibieron un menor costo personal asociado. El tamaño del efecto de estas asociaciones, medido a través del estadístico *d* de Cohen, varió de .04 a .59.

La adaptación al español y validación del MVI nos parece de interés para los investigadores. Primero, a diferencia de otros cuestionarios que miden las valoraciones de los estudiantes universitarios en un sentido más general (Battle & Wigfield, 2003; Luttrell, 2000), el MVI fue diseñado específicamente para obtener valoraciones en el área de matemática. También existe el *Mathematics Anxiety Rating Scale* de Richardson y Suinn (1972), ampliamente utilizado, pero no captura las restantes dimensiones de valoración. Segundo, no hemos encontrado resultados de Análisis Factorial Confirmatorio (AFC) de las puntuaciones del MVI que confirmen su estructura. Tercero, tampoco hemos encontrado investigaciones que empleen el MVI entre estudiantes de universidades de habla hispana, por lo que no disponemos de evidencia de la estructura subyacente en esos colectivos. Cuarto, el MVI fue empleado entre estudiantes de carreras donde la matemática no es esencial y sería interesante comparar si los resultados pueden extenderse a estudiantes que han elegido carreras científicas.

En este trabajo nos proponemos adaptar el MVI al contexto universitario uruguayo, aplicarlo a estudiantes de carreras de Química, confirmar su estructura teórica y comparar nuestros resultados con los reportados por sus autores. El primer objetivo es confirmar la estructura teórica del MVI descrita por Luttrell et al. (2010) en una muestra de estudiantes universitarios uruguayos mediante AFC. Se analizó la posibilidad de calcular una puntuación total para el MVI como constructo global, consistente con la existencia de un factor de segundo orden subyacente a las cuatro dimensiones teóricas, o bien con un modelo de un factor simple de primer orden, más parsimonioso. En segundo lugar, se evaluó si existen diferencias en las medidas de las puntuaciones del MVI y de sus dimensiones según el género. Si bien existen estereotipos masculino y femenino, de mayor y menor inclinación por la matemática y las ciencias (Plante, de la Sablonnière, Aronson, & Téorêt, 2013; Polino, 2012), los antecedentes muestran que la valoración de la matemática tiende a equipararse a partir de la secundaria (Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002), no habiéndose encontrado diferencias en contextos universitarios (Luttrell et al., 2010). Con base en estos antecedentes se plantea que no hay diferencias de valoración de la matemática entre hombres y mujeres, ni a nivel de la escala total MVI ni de sus dimensiones. En tercer lugar, se examinó la relación de las puntuaciones con el rendimiento del alumno en los cursos universitarios de

matemática realizados previamente y con el rendimiento general futuro en la universidad, medido tres años después. Dado que la bibliografía sugiere la existencia de efectos positivos de las metas de aprendizaje sobre el rendimiento de estudiantes universitarios de carreras científicas (Grant & Dweck, 2003; Rhee, Zusho, & Pintrich, 2005; Rodríguez Ayán, 2007; Rodríguez Ayán & Ruiz, 2012; Zusho, Pintrich, & Coppola, 2003) extendemos tales hallazgos a la valoración de la matemática en su componente de interés, por considerar a esta dimensión equiparable a las metas de aprendizaje. Así, postulamos que la dimensión interés estará positivamente asociada al rendimiento. Con base en trabajos previos (Bong, 2001; Cole et al., 2008; Hulleman et al., 2008) se plantea que la utilidad y el logro también estarán positivamente asociados al rendimiento. Asimismo, nos planteamos que los estudiantes que perciben un mayor costo asociado a la matemática tendrán un rendimiento más bajo.

Método

Participantes

La población objetivo fueron las cohortes 2009 y 2010 de carreras universitarias de Química del Uruguay (806 estudiantes). La selección se realizó dividiendo a la población en tres estratos, según la situación curricular de los estudiantes a marzo de 2011 respecto a los dos primeros cursos obligatorios de Análisis y Álgebra. En cada estrato se seleccionó una muestra aleatoria simple para estimación de prevalencia, para un nivel de confianza del 95% y un error del 5%. El tamaño de la muestra fue de 401 casos divididos en: estudiantes con un 100% de aprobación ($n = 232$), estudiantes con un 50% de aprobación ($n = 59$) y estudiantes con 100% de reprobación ($n = 110$). Participaron voluntariamente 218 estudiantes (77% mujeres y 23% varones), el 50% perteneciente a cada cohorte. El 80% tenía 20 años o menos al momento de la encuesta. Mediante correo electrónico se los invitó a contestar el cuestionario en línea. Como la tasa de respuesta fue de 54%, en lugar de utilizar los expansores originales para las estimaciones poblacionales se utilizaron ponderadores calibrados (Deville, Särndal, & Sautory, 1993). Estos se obtienen modificando los ponderadores originales provenientes del diseño muestral mediante el uso de información auxiliar y permiten estimar sin error los totales poblacionales para las variables auxiliares. En este caso la información auxiliar empleada fue la situación en los cursos de Análisis y Álgebra. Ello compensa el no haber considerado pérdidas en el cálculo de la muestra.

Aspectos Éticos

Los estudiantes fueron informados acerca del alcance del estudio y de su libertad de participar o no de la investigación. Se aseguró confidencialidad respecto a la identidad de los participantes. Una vez procesados los datos se envió informe de devolución de la evaluación realizada.

Medidas

MVI. Instrumento desarrollado por Luttrell et al. (2010) para medir cuatro dimensiones relacionadas, de siete ítems cada una. El cuestionario original fue traducido al español por las autoras del presente trabajo y revisado por traductora pública inglés español. El interés fue definido como la importancia atribuida a la matemática debido a un genuino interés, equiparable a la motivación intrínseca de Deci y Ryan (1985) y a la motivación por el aprendizaje de Grant y Dweck (2003), por ejemplo, “Estudiar matemática me resulta placentero”. La utilidad se refiere a la importancia de aprender matemática porque considera que lo ayudará en su desarrollo profesional o personal, aun cuando pueda no resultarle interesante (Husman & Lens, 1999; Kauffman & Husman, 2004), por ejemplo (ítem por la negativa): “Tener una sólida formación en matemática no sirve para nada”. La importancia para la concreción de logros, o simplemente la dimensión logro, es la importancia de obtener buenos resultados en matemática, “Para mí es importante que me vaya bien en matemática”. La dimensión costo fue definida como la percepción estudiantil de la pérdida sufrida como resultado de intentar comprender matemática o intentar obtener buenos resultados (tiempo, dificultad, ansiedad, miedo al fracaso), “Tratar de hacer ejercicios de matemática me pone muy ansioso/a”. Cuando el costo supera los beneficios, ello resulta en una devaluación de la tarea. Al igual que Feather (1988) y que Luttrell et al. (2010), consideramos el interés, la utilidad y el logro como percepciones que incrementan el valor de estudiar matemática, en tanto que el costo contribuye a su desvalorización. Las respuestas fueron medidas en una escala Likert de 5 puntos (1 = *totalmente en desacuerdo*, 5 = *totalmente de acuerdo*). Los puntajes de la escala de utilidad excepto uno y de todos los ítems de costo personal fueron revertidos; así puntuaciones altas en todas las dimensiones reflejan un valor positivo de la matemática, en tanto que puntuaciones bajas reflejan un valor negativo. Las cuatro dimensiones explican el 71% de la varianza total contenida en los datos. Están todas correlacionadas entre sí, con coeficientes de correlación que varían entre .40 y .58, y presentan fiabilidades elevadas: .95 (interés), .92 (utilidad), .92 (logro) y .91 (costo). Los autores también informan la fiabilidad de la escala total, $\alpha = .95$ (Luttrell et al., 2010), si bien no proporcionan evidencia de que las puntuaciones puedan nuclearse en torno a un constructo unitario.

El rendimiento previo en matemática se estimó mediante la relación entre el número de cursos de matemática aprobados y el total de cursos de matemática realizados hasta el momento del relevamiento de opiniones. Como dicha proporción no pudo aproximarse a una variable cuantitativa fue categorizada, definiendo así cuatro grupos de rendimiento según los siguientes porcentajes de aprobados en matemática: ninguno, menos del 50%, menos del 75% y 75% o más. El rendimiento futuro general se estimó mediante dos indicadores: la relación entre el total de cursos aprobados y el total de cursos rendidos y la re-

lación entre el total de exámenes aprobados y el total de exámenes rendidos, contabilizados desde el momento del relevamiento hasta 3 años después.

Análisis

Para cumplir con el primer objetivo se empleó el AFC con estimación de máxima verosimilitud. Para evaluar el ajuste se emplearon varios criterios: la significación de las estimaciones ($p < .05$), distintos índices de ajuste global (χ^2/df , RMR, RMSEA, GFI y CFI) y los residuos estandarizados. El índice χ^2/df debería ≤ 3 para un buen ajuste, pero siendo un indicador muy sensible al tamaño de la muestra y a la complejidad de los modelos se sugiere emplearlo conjuntamente con otras medidas (Hu & Bentler, 1995). RMR y RMSEA son índices especialmente sensibles a errores de especificación de modelos simples y complejos respectivamente, siendo los puntos de corte recomendados para un buen ajuste de .08 y .06 respectivamente (Hu & Bentler, 1998). GFI y CFI se recomiendan por su sensibilidad a errores de especificación en modelos complejos, siendo .95 el punto de corte recomendado para ambos (Hu & Bentler, 1999). Puesto que existe controversia respecto al uso de puntos de corte estrictos para aceptar o desestimar un modelo (Marsh, Hau, & Grayson, 2005; Marsh, Hau, & Wen, 2004; Taylor & Pastor, 2007), otro insumo fueron los residuos estandarizados. Estos representan para cada par de ítems la diferencia entre la covarianza muestral y la covarianza estimada y son útiles para identificar un ajuste pobre local cuando > 3 (Byrne, 1988). Para comparar modelos anidados se empleó la significación del cambio en χ^2 y para comparar modelos no anidados el estadístico AIC (Akaike, 1973, 1987; Kline, 2005).

Se ajustó primero el modelo confirmatorio de cuatro factores relacionados de acuerdo a los resultados de Luttrell et al. (2010; modelo 1). Los resultados sugirieron el traslado de un ítem previsto para la dimensión costo a la escala de logro (modelo 2). Los resultados de los modelos 1 y 2 sugirieron la existencia de saturaciones cruzadas y errores correlacionados, por lo que se recurrió a la estrategia metodológica de agrupamiento de ítems o *item parceling* (modelo 3). Con ello se logra minimizar las varianzas específicas de los ítems, que en este caso no son de interés, y mejorar el ajuste. Ello se sustenta en que nuestro objetivo es comprender el constructo valoración de la matemática y su relación con el rendimiento académico y no analizar la naturaleza de los ítems y su relación con las dimensiones que se proponen medir (Little, 2013; Little, Rhemtulla, Gibston, & Schoemann, 2013). Siguiendo las recomendaciones de Little et al. (2013) se construyeron tres grupos por dimensión con el siguiente criterio: incluir en un mismo grupo ítems con saturación cruzada de signo opuesto en el mismo segundo factor. La puntuación de cada grupo se estimó como el promedio de las puntuaciones de los ítems. Como el ajuste del modelo 3 fue muy bueno, también se ajustó un modelo de un factor de segundo orden (modelo 4, donde saturaron cuatro factores, cada uno determinado por tres grupos) y un modelo de un único factor simple (modelo 5, donde saturaron nueve grupos).

Para el segundo objetivo se analizaron las diferencias de medias de las puntuaciones de las cuatro escalas y de la puntuación total según el género mediante el test *t* de Student. Entre los varones las puntuaciones en cada dimensión así como la puntuación total pueden aproximarse a una distribución normal de acuerdo al test de Kolmogorov-Smirnov (KS): $p = .200$ (interés, utilidad y logro), $p = .050$ (costo) y $p = .132$ (MVI). Entre las mujeres el test de KS fue consistente con el supuesto de normalidad para las dimensiones costo y logro y para el MVI ($p = .200$), en tanto que se rechazó la hipótesis nula para interés ($p = .011$) y utilidad ($p = .009$). Por lo tanto se complementaron los resultados del test de Student con el test no paramétrico de Mann-Whitney. El tamaño mínimo de la muestra para estas comparaciones específicas para una potencia estadística de 80%, un error de 5% y una relación de grupos de 3.3 sería de 208 casos (156 mujeres y 52 varones) para un tamaño del efecto moderado $d = .4$, en tanto que sería de 826 (619 mujeres y 207 varones) para un tamaño del efecto pequeño $d = .20$ (Faul, Erdfelder, Lang, & Buchner, 2007). El tamaño de la muestra con la que se trabajó (168 mujeres y 50 varones) permite identificar con una probabilidad de .80 diferencias moderadas entre los sexos, no así diferencias pequeñas.

Para el tercer objetivo se analizó la relación entre las puntuaciones y el rendimiento previo en todos los cursos de matemática que los estudiantes hubieran tomado (ANOVA de un factor, previa comprobación del supuesto de normalidad) y la relación entre las puntuaciones y el rendimiento futuro (modelos de regresión lineal de los indicadores empleando las dimensiones del MVI como variables explicativas). La variable rendimiento previo en los cursos de matemática presenta cuatro categorías: ningún curso aprobado, porcentaje de aprobados inferior a 50%, porcentaje de aprobados inferior a 75% y porcentaje de aprobados de 75% o más. El test de KS permite sustentar el supuesto de normalidad de las dimensiones y de la escala total en todos los grupos de rendimiento: $p = .075$ (costo, grupo sin cursos de matemática aprobados), $p = .158$ (interés, grupo con menos del 50% de cursos de matemática aprobados) y $p = .200$ para los restantes grupos y dimensiones, por lo que no se recurrió a tests no paramétricos. El año de ingreso se introdujo en los modelos como covariable para equiparar a los participantes en el tiempo de permanencia en la universidad. El tamaño del efecto para las diferencias de medias fue estimado mediante el estadístico *d* de Cohen. Cabe señalar que para las comparaciones entre los cuatro grupos de rendimiento el tamaño mínimo de muestra para una potencia estadística de 80% y un error de 5% sería de 73 para un tamaño del efecto moderado $d = .40$ y de 180 para un efecto pequeño $d = .25$ (Faul et al., 2007). Por lo tanto con nuestros datos es posible identificar diferencias pequeñas entre los cuatro grupos con una potencia del 80%.

Se emplearon los paquetes estadísticos *G*Power3*, *R i386*, *PASW 18* y *AMOS 7.0*.

Resultados

AFC

Los resultados de ajuste de los cinco modelos confirmatorios se muestran en la Tabla 1. Los modelos 1 y

2 (cuatro factores relacionados) y el modelo 5 (un factor simple) arrojaron índices de ajuste pobre. El modelo 3 (cuatro factores basados en agrupamientos) y el 4 (factor de segundo orden) alcanzaron un buen ajuste.

Tabla 1
Modelos Confirmatorios

Modelo	χ^2 / df	RMR	RMSEA	IC 90 %	GFI	CFI	AIC
1	8.582	.067	.105	[.101,.108]	.774	.692	3223.622
2	8.184	.064	.102	[.098,.105]	.778	.709	2939.206
3	5.630	.024	.082	[.072,.091]	.940	.940	330.239
4	7.249	.041	.095	[.086,.104]	.923	.915	418.448
5	31.585	.062	.210	[.202,.219]	.673	.537	1753.593

Nota. *df* = grados de libertad; 1 = modelo teórico de cuatro factores relacionados; 2 = modelo de cuatro factores relacionados con migración de un ítem de la escala de costo a la escala de logro; 3 = modelo de cuatro factores con *item parceling*; 4 = modelo de un factor de segundo orden; 5 = modelo de un factor simple.

Los índices RMSEA, GFI y CFI, particularmente sensibles a errores de especificación de modelos complejos, son claros indicadores de la falta de ajuste de los modelos 1 y 2. El índice estandarizado de Mardia (1970) fue de 10, lo que muestra vulneración del supuesto de normalidad multivariante en el que se sustenta el método de estimación de máxima verosimilitud. Sin embargo, algunas investigaciones han demostrado que se trata de un método de estimación robusto y se han obtenido resultados adecuados cuando el modelo está correctamente especificado, aun con valores del índice de Mardia superiores a 50 (Rodríguez Ayán & Ruiz, 2008). Ello sugiere que el ajuste pobre se debe más bien a problemas en la especificación del modelo más que a la vulneración del supuesto. La saturación del ítem “Me preocupa tener notas bajas en matemática”, teóricamente previsto en la escala de costo fue significativa ($p < .001$) pero muy baja, $\lambda = .120$. Analizando su contenido es posible interpretarlo bien como una medida de costo -la preocupación- pero también como una medida de necesidad de logro -el evitar calificaciones bajas. Los índices de modificación del modelo 1 también fueron sugestivos de incorporar dicho ítem a la dimensión logro. En base a estas consideraciones estadísticas y sustantivas se ajustó un segundo modelo de cuatro factores relacionados, con dicho ítem como indicador de logro y no de costo (modelo 2). Todas las estimaciones del modelo 2 resultaron significativas ($p < .001$), siendo la saturación de dicho ítem en la escala de logro de .506, lo que sustenta la idea de que puede concebirse como un indicador de necesidad de logro. Al pasar del modelo 1 al 2 hay una disminución en el valor de AIC. Todas las correlaciones en el modelo 2 fueron significativas ($p < .001$).

De los 378 residuos posibles del modelo 2, 76 resultaron $> |3|$ y 146 resultaron $> |2|$. Al analizar en detalle los ítems involucrados en estos residuos excesivamente altos se encontró que cinco ítems de la escala de logro

presentaban la mayor cantidad de residuos $> |3|$. Para cada uno de estos ítems, de las 27 covarianzas estimadas por el modelo, entre 7 y 12 resultaron muy alejadas de las covarianzas observadas en la muestra. Los índices de modificación sugieren errores correlacionados entre los ítems de la escala de logro y también entre ítems de escalas distintas, si bien la máxima correlación sugerida fue de .153. Los índices también sugieren saturaciones cruzadas, algunas $> .40$.

No podemos saber con certeza si el resultado del ajuste pobre se debe a errores de especificación del modelo - haber desestimado parámetros que no son triviales en la población (e.g., saturaciones cruzadas) - o a errores de parsimonia - haber desestimado parámetros que son distintos de cero en la población pero cuyo valor es muy pequeño (Cheung & Rensvold, 2002). De hecho “todos los modelos están mal en alguna medida, incluso en la población, y lo mejor que podemos esperar es identificar un modelo parsimonioso, sustantivamente significativo, que se ajuste adecuadamente a los datos observados” (MacCallum & Austin, 2000, p. 218). En este caso, las saturaciones cruzadas altas sugieren problemas de especificación más que de parsimonia, por lo que se ajustó el modelo 3 mediante *item parceling*. En la Tabla 2 se muestran las saturaciones de los ítems en el modelo 2 así como las saturaciones cruzadas y las segundas dimensiones sugeridas por los índices de modificación. A partir de las saturaciones cruzadas se construyeron los grupos de ítems para el ajuste del modelo 3.

El modelo 3 arrojó un ajuste adecuado a los datos (Tabla 1) y todas las estimaciones resultaron significativas ($p < .001$). Las cuatro dimensiones explican el 62% de la varianza. Las fiabilidades fueron .87 (interés), .84 (utilidad), .80 (logro) y .61 (costo personal), lo que se considera muy adecuado teniendo en cuenta que cada dimensión ahora está compuesta solo por tres indicadores.

Tabla 2

Saturaciones Principales y Saturaciones Cruzadas Sugeridas por los Índices de Modificación. Modelo de Cuatro Factores Relacionados con Migración de un Ítem de la Escala de Costo a la Escala de Logro (modelo 2)

Dimensión	Grupo ^a	Ítem	λ	Segundo factor ^b	$\lambda_{\text{cruzada}}^b$
Interés	1	1	.786	-	-
Interés	1	5	.709	Utilidad	-.161
Interés	1	25	.480	Utilidad	.332
Interés	2	9	.743	Logro	.155
Interés	2	17	.806	Logro	-.153
Interés	3	13	.736	-	-
Interés	3	21	.750	-	-
Utilidad	1	2	.716	Costo personal	-.073
Utilidad	1	14	.788	Costo personal	.059
Utilidad	2	6	.676	Logro	-.133
Utilidad	2	10	.728	Logro	-.162
Utilidad	2	26	.599	Logro	.433
Utilidad	3	18	.660	Interés	-.076
Utilidad	3	22	.482	Interés	.123
Logro	1	3	.502	Utilidad	.262
Logro	1	23	.509	Utilidad	-.244
Logro	2	7	.663	Utilidad	.272
Logro	2	11	.519	Utilidad	-.453
Logro	2	16	.506	Utilidad	-.196
Logro	3	15	.553	Costo personal	.139
Logro	3	19	.507	Costo personal	.070
Logro	3	27	.667	Costo personal	-.069
Costo personal	1	4	.639	-	-
Costo personal	1	20	.380	-	-
Costo personal	2	8	.191	Logro	-.931
Costo personal	2	12	.860	Logro	.338
Costo personal	3	24	.485	Logro	-.178
Costo personal	3	28	.648	Logro	-.264

^a Grupo de pertenencia del ítem para el ajuste del modelo 3; ^b Dimensión secundaria y saturación cruzada sugeridas en base a los índices de modificación.

Las cuatro dimensiones están correlacionadas, siendo la más alta entre interés y utilidad ($r = .603$), en línea con lo obtenido por Luttrell et al. (2010). En menor grado el interés también correlaciona de manera considerable con el logro ($r = .369$) y el costo personal ($r = .360$). La dimensión utilidad también correlaciona con logro ($r = .359$) y costo ($r = .304$). Debe recordarse que la escala de costo fue revertida, de modo tal que puntajes altos reflejan escaso costo y por lo tanto valoración positiva. Opuesto a la dirección informada por Luttrell et al., el logro correlaciona negativamente con el costo ($r = -.227$), lo que implica que en nuestro caso una mayor necesidad de logro trae aparejado un mayor sufrimiento.

El modelo 4, factor de segundo orden, arrojó un ajuste aceptable. Respecto al modelo 3, $\Delta\chi^2 = 92.209$, $df = 2$, $p < .001$. Las saturaciones de cada dimensión en el MVI fueron todas significativas ($p < .001$): .858 (interés),

.724 (utilidad), .409 (logro) y .368 (costo personal). El modelo 5 presenta un ajuste muy pobre, lo que pone de manifiesto que pese a que los factores están fuertemente correlacionados entre sí, la valoración de la matemática en la muestra de estudiantes uruguayos resulta un constructo multidimensional, tal como fue definido y operativizado por los autores del cuestionario original, y no unitario.

Valoración de la Matemática y Género

En línea con lo hipotetizado, no se aprecian diferencias de valoración de la matemática entre hombres y mujeres (Tabla 3) en ninguna de sus dimensiones. La diferencia de valoración global entre hombres y mujeres, medida a través de la puntuación total MVI, no alcanza significación ($p = .335$). El test de Mann-Whitney fue consistente con estos resultados, arrojando valores de p de .656, .540, .708, .118 y .271 para interés, utilidad, logro, costo y MVI respectivamente.

Tabla 3
Dimensiones de la Valoración de la Matemática según Género

Escala	Género				<i>t</i>	<i>p</i>	<i>df</i>	<i>d</i> Cohen
	Femenino		Masculino					
	Media	<i>SD</i>	Media	<i>SD</i>				
Interés	7.61	1.98	7.44	1.89	.479	.632	186	.086
Utilidad	5.56	1.46	5.42	1.85	.432	.667	51.053 ^a	.078
Logro	8.09	1.42	8.03	1.76	.214	.831	186	.039
Costo ^b	9.53	1.95	9.10	1.61	1.258	.210	186	.226
MVI	30.78	4.46	29.99	4.95	.966	.335	186	.174

Nota. *SD* = desviación estándar; *df* = grados de libertad; MVI = escala total.

^a Grados de libertad estimados ante la vulneración del supuesto de homogeneidad de varianzas; ^b La escala de Costo fue revertida, por lo que puntuaciones más altas reflejan menor costo (mejor valoración).

Valoración de la Matemática y Rendimiento Académico

En la Tabla 4 se muestra la estadística descriptiva del MVI y de sus dimensiones según los cuatro grupos de rendimiento en los cursos de matemática: sin aprobados, con menos del 50% aprobado, con menos del 75% aprobado y con 75% o más aprobado. Los resultados de *F* sugieren que las puntuaciones medias en la escala total y en las dimensiones interés, logro y costo difieren entre los grupos. Se informa el tamaño del efecto *d* así como la significación *p* del contraste *post-hoc* de ANOVA para la diferencia de medias entre cada grupo y el grupo de mejor rendimiento (75% o más aprobados), que se tomó como referencia. Tales efectos en general fueron moderados o bajos, variando en valor absoluto entre .023 y .425, en línea con los resultados de Luttrell et al. (2010).

Los estudiantes del extremo superior de rendimiento (mayor proporción de cursos de matemática aprobados)

son los que tienen menos interés por la disciplina ($M = 6.74$), en tanto que aquellos ubicados en el extremo inferior, sin ninguna aprobación, son los que se muestran más interesados ($M = 7.93$). El tamaño del efecto para la diferencia entre ambos grupos es $d = -.126$. En cuanto a la utilidad no se encontraron diferencias estadísticamente significativas entre las medias. Respecto a la necesidad de logro, para los dos grupos de rendimiento bajo esta necesidad es más baja que para los dos grupos de rendimiento superior, entre los cuales no hay diferencias ($p = .603$). Finalmente, en la escala de costo personal hay diferencias importantes entre el extremo superior y los demás grupos, con tamaños del efecto que varían entre .33 y .43. Entre los tres grupos que no alcanzan el 75% de aprobados no se encontraron diferencias de medias significativas ($p = .078$). Dada la puntuación inversa de esta escala, valores más altos implican menor costo, por lo que se puede

concluir que el grupo de mejor rendimiento es el que más ha sufrido debido a la complejidad de la matemática y al esfuerzo que supone dedicarse a ella.

En la Tabla 5 se muestran los resultados de los modelos predictivos del rendimiento futuro.

Tabla 4
Valoración de la Matemática y Rendimiento Previo en Matemática

Escala	Grupo de rendimiento en cursos de matemática											
	Sin aprobados			Menos del 50%			Menos del 75%			75% o más		
	<i>M</i> (<i>SD</i>)	<i>d</i>	<i>p</i> _{post-hoc}	<i>M</i> (<i>SD</i>)	<i>d</i>	<i>p</i> _{post-hoc}	<i>M</i> (<i>SD</i>)	<i>d</i>	<i>p</i> _{post-hoc}	<i>M</i> (<i>SD</i>)	<i>F</i> (3, 187)	η^2
I	7.93 (2.29)	-.126	.026	7.61 (1.74)	-.092	.079	7.69 (1.95)	-.100	.181	6.74 (1.35)	2.901 *	.045
U	5.83 (1.79)	-.063	.357	5.43 (1.22)	-.023	1.000	5.47 (1.38)	-.027	1.000	5.20 (1.65)	1.341	.021
L	7.78 (1.43)	-.114	.033	7.63 (1.46)	.135	.015	8.47 (1.44)	-.023	1.000	8.64 (1.43)	5.285**	.079
CP	10.15 (1.86)	-.425	<.001	9.59 (1.57)	-.323	.001	9.61 (1.68)	-.327	<.001	7.80 (1.69)	14.345***	.190
MVI	31.68 (5.29)	-.109	.004	30.26 (4.03)	-.062	.205	31.25 (3.96)	-.094	.028	28.39 (3.86)	4.518**	.069

Nota. *SD* = desviación estándar; *d* = estadístico de Cohen para estimar el tamaño del efecto para comparaciones con el grupo que tiene más del 75% de aprobados; *p*_{post-hoc} = significación estadística de la comparación *post-hoc* con el grupo que tiene más del 75% de aprobados empleando la corrección de Bonferroni; I = interés; U = utilidad; L = logro; CP = costo personal; MVI = escala global; $\eta^2 = SS_{\text{efecto}} / (SS_{\text{efecto}} + SS_{\text{error}})$.
* *p* < .05; ** *p* < .01; *** *p* < .001.

Tabla 5
Modelos Predictivos del Rendimiento Futuro

Factor	Cursos aprobados / cursos rendidos						Exámenes aprobados / exámenes rendidos					
	Paso 1			Paso 2			Paso 1			Paso 2		
	B	β	<i>t</i>	B	β	<i>t</i>	B	β	<i>t</i>	B	β	<i>t</i>
RPC	.669	.589	15.144**	.681	.600	15.501**	.699	.658	17.612**	.667	.628	17.575**
RPE	.145	.154	4.130**	.113	.120	3.307**	.105	.223	6.089**	.115	.243	6.712**
Año	.011	.027	.764	.019	.044	1.262	.017	.013	.316	.019	.151	4.104**
Interés				.016	.137	3.705**				-.032	-.197	-5.416**
Utilidad				-.014	-.093	-2.567*				.031	.183	5.332**
Logro				.027	.177	5.120**				.002	.017	.471
Costo personal				.005	.046	1.258				.038	.028	.715
<i>R</i> ²	.459			.507			.466			.527		

Nota. RPC = rendimiento previo en cursos; RPE = rendimiento previo en exámenes; Año = año de ingreso a la universidad.
* *p* < .05; ** *p* < .001.

El rendimiento previo en matemática es un buen predictor del rendimiento futuro, siendo este efecto particularmente relevante cuando el indicador de rendimiento es el mismo en los dos puntos temporales ($\beta = .667$ exámenes; $\beta = .681$ cursos). La inclusión de las dimensiones del MVI en el segundo bloque produce un pequeño aumento en la fracción de varianza explicada por los modelos (5%-6%), lo que muestra una discreta contribución de la valoración

de la matemática en la explicación del rendimiento académico futuro. No se observan cambios importantes en los coeficientes de regresión del rendimiento previo entre un bloque y otro, por lo que se desestiman posibles efectos indirectos del rendimiento previo, mediatizados por el valor que los alumnos le atribuyen a esta asignatura.

Los efectos de las dimensiones interés y necesidad de logro son positivos y del mismo orden de magnitud,

aproximadamente 4 y 3 veces menores que los efectos del rendimiento previo. A mayor interés por la matemática y mayor necesidad de logro, mayor es el rendimiento general futuro. El efecto de la dimensión costo personal no alcanzó significación en ninguno de los dos modelos: $p = .209$ (rendimiento en cursos) y $p = .638$ (rendimiento en exámenes). El efecto de la dimensión utilidad resulta paradójico por su signo negativo, lo que supone que una mayor percepción de utilidad de la matemática redundaría en un peor rendimiento general futuro. Pero debe tenerse en cuenta que la presencia del rendimiento previo en los modelos tiene no solo un valor explicativo del indicador de rendimiento, sino que además su papel es el de una covariable que equipara a los sujetos en el modelo, restando su efecto parcial de la variabilidad disponible por explicar. El efecto negativo de la utilidad se puede interpretar como una tendencia a obtener un rendimiento menor cuando aumenta la percepción de utilidad de la matemática (efecto neto), pero teniendo en cuenta que el alumno se encuentra ya en un determinado nivel de rendimiento: los alumnos con rendimiento previo alto disminuyen ligeramente su rendimiento y lo mismo sucede con los alumnos con un rendimiento previo bajo. Podría ser interesante estudiar estos efectos con más detalle mediante modelos multinivel de efectos aleatorios.

Discusión

En relación al primer objetivo propuesto, todos los modelos confirmatorios propuestos están identificados y las soluciones son técnica y conceptualmente correctas. Atendiendo a la bondad de ajuste, la estrategia de *item parceling* mejoró notablemente el ajuste del modelo de cuatro factores relacionados, eliminando las saturaciones cruzadas y la correlación entre los residuos. Podemos afirmar que la estructura que mejor explica la valoración de la matemática corresponde a un modelo de cuatro factores que, si bien se encuentran relacionados entre sí, reflejan contenidos diferentes.

El significado de las dimensiones se corresponde con los hallazgos de Luttrell et al. (2010). La valoración de la matemática puede considerarse como una combinación entre el interés, la percepción de su utilidad, la percepción de los logros que se pueden obtener a través de su estudio y el costo personal que ello supone. El hecho de que el modelo confirmatorio de un factor simple no haya tenido un buen ajuste sugiere que un solo factor no sería suficiente para explicar las correlaciones observadas. Pero es posible agrupar las cuatro dimensiones y resumir la información en un factor de segundo orden, de valoración general. Cabe señalar que un ítem fue concebido por los autores como indicador de costo y en nuestro caso parecería ser un indicador de logro. Ello ilustra la importancia de la validación de los cuestionarios al aplicarse en otros contextos. Debe tenerse presente, no obstante, que habiendo empleado la estrategia de *item parceling* se eliminaron las saturaciones cruzadas y los residuos correlacionados, por

lo que no sería correcto pretender explicar mediante estos resultados la relación de los indicadores observables con las dimensiones, sino solamente las relaciones entre estas y con otros constructos.

El interés, la utilidad y el logro están positivamente correlacionados entre sí, siendo la mayor asociación entre interés y utilidad; estos además están asociados a un menor sufrimiento emocional, en línea con Luttrell et al. (2010). Esto es relevante para los docentes, pues sustenta la idea de que es posible aumentar la motivación intrínseca de los estudiantes y por lo tanto el goce del estudio mediante el énfasis en la utilidad de la matemática (Csikszentmihalyi, 2014; Middleton, 1995; Shernoff, Csikszentmihalyi, Schneider, & Shernoff, 2003); ello a su vez redundaría en un menor sufrimiento asociado al estudio. El logro, sin embargo, está negativamente asociado al costo, es decir, se percibe un mayor sufrimiento por parte de aquellos con mayor deseo de obtener buenos resultados. No disponemos de antecedentes que hayan puesto mayor énfasis en el estudio de la dimensión costo personal, pero una posible explicación para este hallazgo puede ser que el perfil de nuestros estudiantes difiere del de las poblaciones en que se ha estudiado esta dimensión, por cuanto se trata de personas que han elegido carreras muy exigentes respecto a matemática.

Sobre el segundo objetivo, tal como se previó, no se observan diferencias entre hombres y mujeres respecto a su actitud hacia la matemática. Hay investigaciones que sugieren que los hombres tienen una mejor disposición que las mujeres hacia los estudios científicos (e.g., Plante et al., 2013), pero debe recordarse que todos los participantes de esta investigación eligieron estudiar carreras científicas. Por su parte, Luttrell et al. (2010) obtuvieron el mismo resultado con estudiantes de orientaciones no científicas. Por lo tanto, una hipótesis puede ser que una vez equiparados los estudiantes en términos de sus preferencias por las carreras, no se advierten diferencias entre hombres y mujeres respecto a su valoración de la matemática, sean estudiantes de ciencias o de otras opciones. No obstante, debe tenerse en cuenta que el tamaño de la muestra con que se trabajó permite detectar con un β de .80 solamente diferencias moderadas entre los sexos. Por lo tanto, no disponemos de evidencia suficiente como para descartar la existencia en la población de un efecto real, aunque débil. Dejamos para un futuro el análisis de esta posibilidad.

En cuanto al tercer objetivo, nuestros resultados sugieren que un mejor rendimiento en matemática está asociado a menos interés, mayor necesidad de logro y mayor sufrimiento. La utilidad de la asignatura, en cambio, no parece tener efectos relevantes. Estos hallazgos parecen diferir de los de Luttrell et al. (2010), que sugieren que a mayor cantidad de cursos de matemática completados, mayor interés por la asignatura, mayor percepción de su utilidad y menor costo personal, en tanto que la necesidad de logro no varía con el rendimiento. Debe señalarse el contexto diferente en el que se llevaron a cabo las dos investigaciones. En nuestro caso hemos trabajado con estudiantes

de carreras científicas, con una alta carga obligatoria de matemática, por oposición a los estudiantes de Luttrell et al. Las investigaciones de Eccles y sus colaboradores con estudiantes de enseñanza secundaria muestran que las valoraciones que hacen los jóvenes resultan buenos predictores de los cursos en los que optan por matricularse en los últimos años de la secundaria y en la universidad (e.g., Chow, Eccles, & Salmela-Aro, 2012; Durik, Vida, & Eccles, 2006; Eccles, Vida, & Barber, 2004; Simpkins, Davis-Kean, & Eccles, 2006). No es de extrañar, pues, que entre estudiantes de carreras universitarias donde la matemática no es esencial, aquellos que eligen un mayor número de cursos de matemática sean los que tienen mayor interés y los que creen que les será más útil, o aquellos a los que les demande menos esfuerzo y por lo tanto tengan mayor expectativa de éxito. En la presente investigación los grupos de estudiantes no se distinguen unos de otros por el número de cursos de matemática que eligieron realizar, sino por el éxito alcanzado en los cursos de matemática que obligatoriamente deben completar. Si se tiene en cuenta la dificultad y el miedo por la matemática que se ha instalado en Uruguay (Polino, 2012) estos resultados pueden apreciarse desde otro ángulo. Un mejor rendimiento de nuestros participantes está asociado ya no a una mayor percepción del valor utilitario de la disciplina, como sugieren las investigaciones de otros países (Bong, 2001; Cole et al., 2008; Durik et al., 2006; Hulleman et al., 2008; Simons, Dewitte, & Lens, 2004), sino a una mayor necesidad de logro (dado que deben aprobar matemática para avanzar en la carrera) así como a un mayor sufrimiento personal, quizás en detrimento de un interés inicial. Sería importante estudiar cómo varían las puntuaciones del MVI a lo largo del tiempo para los mismos estudiantes, para determinar si se observa un declive del interés y un aumento de la ansiedad respecto al estudio de la matemática. Dejamos para un futuro la contrastación de esta hipótesis.

Nuestros resultados prueban que el MVI es adecuado para medir actitudes estudiantiles hacia la matemática, lo que puede orientar el diseño de estrategias de intervención. La decisión de adoptar la perspectiva de un factor de segundo orden o la de cuatro factores depende de los objetivos pretendidos. El factor global puede ser más útil para evaluar en forma continuada la actitud general hacia la matemática por parte de estudiantes de una misma carrera o servicio. La perspectiva de cuatro dimensiones sería una mejor estrategia para detectar áreas problemáticas, de escasa valoración, y comparar las percepciones de estudiantes procedentes de diferentes carreras o instituciones.

Referencias

- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki (Eds.), *Proceedings of the 2nd International Symposium on Information Theory* (pp. 267-281). Budapest, Hungary: Akademiai Kiado.
- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, *52*, 317-332. doi:10.1007/BF02294359

- Battle, A., & Wigfield, A. (2003). College women's value orientations toward family, career, and graduate school. *Journal of Vocational Behavior*, *62*, 56-75. doi:10.1016/S0001-8791(02)00037-4
- Bong, M. (2001). Role of self-efficacy and task-value in predicting College students' course performance and future enrollment intentions. *Contemporary Educational Psychology*, *26*, 553-570. doi:10.1006/ceps.2000.1048
- Byrne, B. M. (1988). The Self Description Questionnaire III: Testing for equivalent factorial validity across ability. *Educational and Psychological Measurement*, *48*, 397-406. doi:10.1177/0013164488482012.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233-255. doi:10.1207/S15328007SEM0902_5
- Chow, A., Eccles, J. S., & Salmela-Aro, K. (2012). Task value profiles across subjects and aspirations to physical and IT-related sciences in the United States and Finland. *Developmental Psychology*, *48*(6), 1612-1628. doi:10.1037/a0030194
- Cole, J., Bergin, D., & Whittaker, T. (2008). Predicting student achievement for low stakes tests with effort and task value. *Contemporary Educational Psychology*, *33*, 609-624. doi:10.1016/j.cedpsych.2007.10.002
- Csikszentmihalyi, M. (2014). Intrinsic motivation and effective teaching: A flow analysis. In J. J. Bass (Ed.), *Application of flow in human development and education* (pp. 173-187). Dordrecht, Netherlands: Springer.
- Deci, E., & Ryan, R. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Deville, J. C., Särndal, C. E., & Sautory, O. (1993). Generalized raking procedures in survey sampling. *Journal of the American Statistical Association*, *88*(423), 1013-1020. doi:10.1080/01621459.1993.10476369
- Durik, A., Vida, M., & Eccles, J. (2006). Task value and ability beliefs as predictors of high school literacy choices: A developmental analysis. *Journal of Educational Psychology*, *98*(2), 382-393. doi:10.1037/0022-0663.98.2.382
- Eccles, J., Adler, T., Futterman, R., Goff, S., Kaczala, C., Meece, J., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75-146). San Francisco, CA: W. H. Freeman.
- Eccles, J., Adler, T., & Meece, J. (1984). Sex differences in achievement: A test of alternate theories. *Journal of Personality and Social Psychology*, *46*, 26-43. doi:10.1037/0022-3514.46.1.26
- Eccles, J. S., Vida, M. N., & Barber, B. (2004). The relation of early adolescents' college plans and both academic ability and task-value beliefs to subsequent college enrollment. *Journal of Early Adolescence*, *24*, 63-77. doi:10.1177/0272431603260919
- Eccles, J., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*, 109-132. doi:10.1146/annurev.psych.53.100901.135153
- Eccles, J., Wigfield, A., Harold, R., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task-perceptions during elementary school. *Child Development*, *64*, 830-847. doi:10.1111/j.1467-8624.1993.tb02946.x
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191. Retrieved from <http://www.gpower.>

- hhu.de/fileadmin/redaktion/Fakultaeten/Mathematisch-Naturwissenschaftliche_Fakultaet/Psychologie/AAP/gpower/GPower3-BRM-Paper.pdf
- Feather, N. T. (1988). Values, valences, and course enrollment: Testing the role of personal values within an expectancy-value framework. *Journal of Educational Psychology, 80*, 381-391. doi:10.1037/0022-0663.80.3.381
- Grant, H., & Dweck, C. (2003). Clarifying achievement goals. *Journal of Personality and Social Psychology, 85*, 541-553. doi:10.1037/0022-3514.85.3.541
- Hood, M., Creed, P., & Newmann, D. (2012). Using the expectancy value model of motivation to understand the relationship between student attitudes and achievement in statistics. *Statistics Education Research Journal, 11*(2), 72-85. Retrieved from https://www.stat.auckland.ac.nz/~iase/serj/SERJ11%282%29_Hood.pdf
- Hu, L., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Issues, concepts and applications* (pp. 76-99). Newbury Park, CA: Sage.
- Hu, L., & Bentler, P. M. (1998). Fit Indices in Covariance Structure Modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424-453. doi:10.1037/1082-989X.3.4.424
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structural analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55. doi:10.1080/10705519909540118
- Hulleman, C., Durik, A., Schweigert, S., & Harackiewicz, J. (2008). Task values, achievement goals, and interest: An integrative analysis. *Journal of Educational Psychology, 100*(2), 398-416. doi:10.1037/0022-0663.100.2.398
- Husman, J., & Lens, W. (1999). The role of the future in student motivation. *Educational Psychologist, 34*, 113-125. doi:10.1207/s15326985Sep3402_4
- Jacobs, J., Lanza, S., Osgood, D., Eccles, J., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development, 73*, 509-527. doi:10.1111/1467-8624.00421
- Kauffman, D. F., & Husman, J. (2004). Effects of time perspective on student motivation: Introduction to a special issue. *Educational Psychology Review, 16*, 1-7. doi:10.1023/B:EDPR.0000012342.37854.58
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.
- Little, T. D. (2013). *Longitudinal structural equation modeling*. New York: Guilford Press.
- Little, T. D., Rhemtulla, M., Gibston, K., & Schoemann, A. M. (2013). Why the items versus parcels controversy needn't be one. *Psychological Methods, 18*(3), 285-300. doi:10.1037/a0033266
- Luttrell, V. (2000). Development of the Higher Education Values Inventory (HEVI): Score reliability and factor structure. *Dissertation Abstracts International, Section B. Sciences and Engineering, 61*(2), 1103.
- Luttrell, V., Callen, B., Allen, C., Wood, M., Deeds, D., & Richard, D. (2010). The Mathematics Value Inventory for General Education Students: Development and initial validation. *Educational and Psychological Measurement, 70*(1), 142-160. doi:10.1177/0013164409344526
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika, 57*, 519-530. doi:10.1093/biomet/57.3.519
- Marsh, H. W., Hau, K.-T., & Grayson, D. A. (2005). Goodness of fit in structural equation modeling. In A. Maydeu-Olivares & J. McArdl (Eds.), *Contemporary psychometrics: A festschrift for Rod McDonald* (pp. 275-340). Mahwah, NJ: Lawrence Erlbaum.
- Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*, 320-341. doi:10.1207/s15328007sem1103_2
- MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology, 51*, 201-226. doi:10.1146/annurev.psych.51.1.201
- Middleton, J. (1995). A study of intrinsic motivation in the mathematics classroom: A personal constructs approach. *Journal for Research in Mathematics Education, 26*, 254-279. doi:10.2307/749130
- Neuville, S., Frenay, M., Schmitz, J., Boudrenghien, G., Noël, B., & Wertz, V. (2007). Tinto's theoretical perspective and expectancy-value paradigm: A confrontation to explain freshmen's academic achievement. *Psychologica Belgica, 47*(1-2), 31-50. doi:10.5334/pb-47-1-31
- Pintrich, P., & Schunk, D. (2002). *Motivation in education: Theory, research and applications*. Englewood Cliffs, NJ: Merrill Prentice Hall.
- Plante, I., de la Sablonnière, R., Aronson, J., & Téorêt, M. (2013). Gender stereotype endorsement and achievement-related outcomes: The role of competence beliefs and task values. *Contemporary Educational Psychology, 38*, 225-235. doi:10.1016/j.cedpsych.2013.03.004
- Polino, C. (2012). Las ciencias en el aula y el interés por las carreras científico-tecnológicas: un análisis de las expectativas de los alumnos de nivel secundario en Iberoamérica. *Revista Iberoamericana de Educación, 58*, 167-191. Recuperado en <http://www.rieoei.org/rie58a09.pdf>
- Rhee, C., Zusho, A., & Pintrich, P. (2005, April). *Multiple-goals, multiple hypotheses: Reexamining the 2 x 2 achievement goal framework in Introductory Chemistry and Psychology classrooms*. Ponencia presentada en la American Educational Research Association Annual Meeting, Montréal, QC, Canada.
- Richardson, F., & Suinn, R. (1972). The Mathematics Anxiety Rating Scale: Psychometric data. *Journal of Counseling Psychology, 19*, 551-554. doi:10.1037/h0033456
- Rodríguez Ayán, M. N. (2007). *Análisis multivariado del desempeño académico de estudiantes de carreras universitarias de Química* (Tesis doctoral, Universidad Autónoma de Madrid, España). Recuperado en https://repositorio.uam.es/bitstream/handle/10486/1800/5491_rodriguez_ayan.pdf?sequence=1
- Rodríguez Ayán, M. N., & Ruiz, M. (2008). Atenuación de la asimetría y de la curtosis de las puntuaciones observadas mediante transformaciones de variables: incidencia sobre la estructura factorial. *Psicológica. Journal of Methodology and Experimental Psychology, 29*, 205-227. Recuperado en <http://www.uv.es/psicologica/articulos2.08/6RODRIGUEZ.pdf>
- Rodríguez Ayán, M. N., & Ruiz, M. (2012). Validation of Haya-mizu and Weiner's Achievement Goal Questionnaire among Uruguayan University students. *Estudios de Psicología, 33*(3), 371-378.
- Rodríguez Ayán, M. N., & Sotelo, M. E. (2012). *Evaluación multidimensional de un programa de formación en la Facultad de Química*. Montevideo, Uruguay: Departamento de Publicaciones, Unidad de Comunicación de la Universidad de la República.

- Rodríguez Ayán, M. N., Sotelo, M. E., Nuñez, I., & González, M. (2011, agosto). *Diagnósticos en situación de inicio y acciones correctivas en la Facultad de Química*. Ponencia presentada en el Primer Encuentro de la Red de Unidades de Apoyo a la Enseñanza de la Universidad de la República, Montevideo, Uruguay.
- Shernoff, D., Csikszentmihalyi, M., Schneider, B., & Shernoff, E. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly*, 18, 158-176. doi:10.1521/scpq.18.2.158.21860
- Simons, J., Dewitte, S., & Lens, W. (2004). The roles of different types of instrumentality in motivation, study strategies and performance: Know why you learn, so you'll know what you learn! *British Journal of Educational Psychology*, 74, 343-360. doi:10.1348/0007099041552314
- Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology*, 42, 70-83. doi:10.1037/0012-1649.42.1.70
- Taylor, M. A., & Pastor, D. A. (2007). A confirmatory factor analysis of the Student Adaptation to College Questionnaire. *Educational and Psychological Measurement*, 67(6), 1002-1018. doi:10.1177/0013164406299125
- Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review*, 12, 265-310. doi:10.1016/0273-2297(92)90011-P
- Wigfield, A., & Eccles, J. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25, 68-81. doi:10.1006/ceps.1999.1015
- Zusho, A., Pintrich, P., & Coppola, B. (2003). Skill and will: The role of motivation and cognition in the learning of college Chemistry. *International Journal of Science Education*, 25(9), 1081-1094. doi:10.1080/0950069032000052207