DESIGN AND VALIDATION OF TWO MEASURES: COMPETENCE AND SELF-EFFICACY IN MATHEMATICAL MODELING

DISEÑO Y VALIDACIÓN DE INSTRUMENTOS DE MEDICIÓN DE COMPETENCIAS Y AUTOEFICACIA DE MODELIZACIÓN MATEMÁTICA

Jennifer A. Czocher	Sindura Subanemy Kandasamy	Elizabeth Roan
Texas State University	Texas State University	Texas State University
Czocher.1@txstate.edu	s_k252@txstate.edu	Eaw109@txstate.edu

In this article, we share the design and validation processes of two instruments measuring aspects of the mathematical modeling process – one that measures competency and one that measures students' self-efficacy to do modeling. The study evaluates both instruments to establish their validity and reliability, using classical test theory.

Keywords: Measurement and Evaluation, Post-Secondary Education, Advanced Mathematical Thinking

Despite calls over recent decades to increase the number of graduates in STEM fields, these numbers have not grown sufficiently. Learning to apply knowledge in these majors implies integration of experiences leveraging design theories, scientific inquiry, technological literacy, and mathematical thinking (Kelley & Knowles, 2016). These goals can be realized through mathematical modeling. Modeling is of utmost importance for students pursuing STEM majors because modeling skills are of critical import to solving society's problems – whose solutions have global consequences. Today's students also take great interest in solving them (Eccles & Wang, 2016; Su, Rounds, & Armstrong, 2009). Further, research suggests that learning mathematics through modeling, as a pedagogical approach, has potential to increase student interest, proficiency in mathematics, robustness of mathematical knowledge, and self-efficacy for doing mathematics (Czocher, 2017; Czocher, Melhuish, & Kandasamy, 2019; Lesh, Hoover, Hole, Kelly, & Post, 2000; Rasmussen & Kwon, 2007; Sokolowski, 2015). Taken together, these factors are positively associated with persistence in mathematics and therefore in majors with high mathematics requirements. One aspect of incorporating more modeling in undergraduate mathematics classrooms is being able to demonstrate the efficacy of instructional interventions by measuring gains in students' modeling skills. This information would help refine programmatic innovations that focus on augmenting students' modeling experiences. Despite the need, there are presently no validated, reliable instruments to measure students' modeling skills available for undergraduates. In this article, we share two such instruments and their psychometric properties: one for modeling competencies and one for self-efficacy to carry out those competencies.

Conceptual Framework

For this project, we adopt a view of mathematical modeling as a cognitive process of rendering a non-mathematical problem about a real-world phenomenon of interest, such as those common to STEM fields, as a well-posed mathematical problem to be solved. It is a cyclic process realized as a suite of mathematical activities and cognitive processes (e.g., Kaiser, 2017). The mathematical problem can be expressed as an equation, a graph, a table, etc. The modeler solves the mathematical problem and interprets its solution in terms of the real-world context. The modeler validates and verifies each step of the process, evaluating whether the model correctly represents the situation and whether the solution makes sense (Czocher, 2018). Table 1 summarizes the conceptual framework, called a mathematical modeling cycle (MMC) (Blum & Leiss, 2007; Czocher, 2016; Maaß, 2006), and also defines the competencies that constitute the modeling process.

Table 1 Modeling competencies

Competency	Description
Understanding	Forming an idea of the real World problema or identifying a real world phenomenon worth investigating
Structuring	Identifying (ir)relevant quantities and variables; making assumptions to simplify the problem
Mathematizing	Expressing relations among the variables using a mathematical representation
Working mathematically	Solving the mathematical problema, using techniques learned in mathematics classes
Interpreting	Interpreting the mathematical results with reference to the context of the real world problem
Validating	Evaluating whether the model represents the situation; verifying the analysis; establishing limitatinos

We operationalize self-efficacy *about* a task as an individual's self-assessed capacity to successfully carry it out (Bandura, 2006; Betz & Hackett, 1983; Hackett & Betz, 1989). In this study, self-efficacy is always evaluated with reference to a specified task. We operationalize the construct *self-efficacy for mathematical modeling* as an individual's self-assessed capacity to successfully carry out the interrelated competencies of the mathematical modeling process. In this way, we can, for example, consider a student's self-efficacy to identify the most important variables involved in estimating the spread of smart homes in the 21st century. The conceptual frameworks are compatible and we used them together to guide the design of the modeling self-efficacy and modeling competency scale items.

Methods

This study has a quantitative nature and is situated within the development of the two instruments, with the purpose of establishing evidence in support of their validity and reliability. The population under study was university STEM majors in the United States. The modeling self efficacy (MSE) instrument went through four rounds of design and testing. In each field test, we used a sample of STEM majors who participated in an international modeling competition called SCUDEM[1], which focuses on modeling with differential equations. In the first round of field testing, there were 6 related items for students to report their self-efficacy for the modeling competencies. In the second round, we created an additional item asking about *establishing limits* (a competency of *validating*, see Table 2, item 6) and we clarified previous items. We used pre- and post- forms of the MSE to measure change in students' self-efficacy from before to after competing. We found gains of moderate effect size $d = 0.545$ $(t(92) = -6.663, p < 0.001)$. In the third round, we created a new item targeting *working mathematically* (Table 2, Item 4). Previously, this competency was excluded because it is traditionally the focus of mathematics instruction, and is complementary to modeling. In the third round, we measured statistically significant positive gains in self-efficacy for those participants who answered both the pre- and post-survey $(t = 4.202\,, df = 51, p < 0.001)$. The final round was carried out concurrently with field testing of the Modeling Competency Questionnaire (MCQ), detailed be low. In each round, we carried out a principal component analysis (Abdi & Williams, 2010) to estimate variance and calculated Cronbach's α as a measure of internal consistency. Summary statistics are in Table 3. Our analyses, together with the instrument's construction based in theories of mathematical modeling, suggest that the MSE is unidimensional with high internal consistency, fase, content, and construct validity.

[1] The annual SCUDEM challenge is hosted by SIMIODE, https://www.simiode.org/scudem. SIMIODE is a professional organization of educators who advocate teaching differential equations from a modeling perspective.

Table 2 Final MSE instrument.

Rate your level of confidence by recording a number from 0 to 100 using the scale given below 0 10 20 30 40 50 60 70 80 90 100 Cannot do at all Moderately can do Highly certain can do.	Competencies
Create a differential equation model for the spread of smart home appliances in the United States during the twenty-first century.	Mathematize
In (1) identify the important variables leading to a reasonably accurate prediction.	Identify variables
In (1) make simplifying assumptions to reduce the number of important variables.	Make assumptions
In (1) select an appropriate numerical, graphical, or analytic technique to solve the resulting differential equation	Work mathematically
In (1) consult appropriate resources to check whether your model was reasonable.	Validate
In (1) list the real-life and mathematical limitations of your model.	List limitations
In (1) create a short presentation to convince a smart appliance manufacturer that they could rely on your model to develop their business plan.	Communicate findings
Given a differential equation which describes the rate of formation of material A, $A'(t) = \alpha A(t)^{\beta}$ and a data set of observations for time, t, amount of material A at each time t, you could estimate the parameters α and β.	Estimate parameters

Tabla 3 Summary of analysis of MSE

Round	N	Varience (ACP)	α	Round	N	Variance (ACP)	α
1	38	62.5%	0.822	3	198	61.5%	0.908
2	276	67.1%	0.917	4	226	69.0%	0.935

Design and testing for the MCQ was carried out in three rounds (feasibility, difficulty, and discrimination) with distinct samples drawn from a large, southwestern university in the United States. We imposed four restrictions on the design: (1) items should be drawn from authentic and relevant contexts (e.g., radioactive decay or analysis of a recycling program), (2) items should draw on knowledge from STEM content or everyday knowledge, (3) items should target the aspects of the modeling competencies, and (4) distractor choices should capture decisions and justifications common to students' reasoning. We created 118 multiple choice items belonging to 9 real-world situations, selected from instructional and research materials from STEM education. Mathematics content included arithmetic, algebra, calculus, and differential equations. For each item, we created one correct answer and four distractors that would appear reasonable to the students but would not help to model the situation. To establish content and construct validity, we invited two mathematicians who teach differential equations to STEM students and three mathematics education researchers who specialize in teaching and learning of mathematical modeling to evaluate the items for appropriateness, correctness, and aptness to the MMC. In the first round, 14 students answered the MCQs and gave us reasoning to justify their choices. We eliminated items that did not make sense to the student. In cases where a student selected a distractor but had sensible reasoning, we modified the item. In the second round, 78 students answered 63 items, distributed among two forms that balanced contexts and competencies. For each item, we calculated the mean difficulty. The majority (76%) of the items had moderate difficulty ($0.20<p<0.70$). We eliminated items outside this range as either too difficult or too easy, restructuring some of the too-difficult items. To analyze distractor efficiency, we calculated the proportion of students that selected each option. At least 5% of the students selected each of the 253 distractors (one item had 5 distractors). For 17 items, a distractor was selected more frequently than the correct answer. These items were flagged as potentially strong discriminators among students with varying levels of modeling competencies. After restructuring problematic items, we chose 30 items (15 items for each of 2 forms). The two forms were administered to a sample of $n = 314$ volunteers who participated in the SCUDEM competition, $n = 135$ responded to Form 1 and $n = 139$ responded to Form 2. For each item, we calculated the mean difficulty. Form 1 had mean difficulty 0.359 ($SD = 0.126$), with $0.177 < p < .0595$. Form 2 had mean difficulty 0.369 ($SD = 0.129$), with $0.147 < p < 0.580$. Four items were too difficult. We conducted another analysis of distractors and concluded that they were functioning adequately. We used point-biserial correlations (rPBIS) to conduct discrimination analysis. Only one item from Form 1 had a negative rPBIS; the remaining items had rPBIs > 0.20. We report the statistics Revelle's Omega Total (ω_T) as an estimate of internal consistency. This selection is appropriate in cases, like the MCQ, where the instrument is multidimensional and when those multiple dimensions contribute to the construct under investigation (Revelle & Zinbarg, 2009). Using the software package '`userfriendlyscience`' in R, we obtained $\omega_T = 0.59$ y $\omega_T = 0.63$, respectively, for Forms 1 and 2. The scales are approaching traditional estimates of 0.7.

Discussion

In this article, we have shared two instruments measuring mathematical modeling competencies and modeling self-efficacy. We also documented their design processes and their psychometric properties. The instruments are aligned with theories of mathematical modeling and have gone through several rounds of field testing. Future research will move into Item Response Theory as a means for constructing and calibrating parallel versions of the modeling competence questionnaire for use as pre/post or group comparison measures. In this way, the instruments can help to evaluate innovative educational interventions aimed at augmenting students' modeling skills. With such information, instructors, researchers, and academic units can improve modeling experiences for students and provide evidence of their efficacy. We are cautious but optimistic that the instruments can meet this goal as the evidence presented here suggests that both are reliable and valid for that purpose.

Acknowledgements

This material was developed with support of Grant No. 170813 from the National Science Foundation. We also wish to thank our marvellous colleagues, Alejandra Sorto, Monica Uribe, Gloria Velasquez, and Alba Melgar for their Assistance in editing the Spanish version of this article.

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley interdiciplinaty reviews: computational statistics, 2*, 433-459.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 3, pp. 307-337). United States of America: Information Age Publishing.

Betz, N., & Hackett, G. (1983). The relationship of mathematics self-efficacy expectations to the selection of science-based college majors. *Journal of Vocational Behavior, 23*(3), 329-345.

Blum, W., & Leiss, D. (2007). How do students and teachers deal with modelling problems? In C. Haines, P. L. Galbraith, W. Blum, & S. Khan (Eds.), *Mathematical Modelling (ICTMA 12): Education, Engineering and Economics* (pp. 222-231). Chickester: Horwood Publishing Limited.

Czocher, J. (2016). Introducing Modeling Transition Diagrams as a Tool to Connect Mathematical Modeling to Mathematical Thinking. *Mathematical thinking and learning, 18*(2), 77-106.

Czocher, J. (2017). How can emphasizing mathematical modeling principles benefit students in a traditionally taught differential equations coures? *The Journal of Mathematics Behavior, 45*, 78-94.

Czocher, J. (2018). How does validating activity contribute to the modeling process? *Educational Studies in Mathematics, 99*, 137-159.

Czocher, J., Melhuish, K., & Kandasamy, S. S. (2019). Building mathematics self-efficacy of STEM undergraduates through mathematical modeling. *International Journal of Mathematical Education in Science and Technology*, 1-28. doi:10.1080/0020739X.2019.1634223

Eccles, J. S., & Wang, M.-T. (2016). What motivates females and males to pursue careers in mathematics and science? *International Journal of Behavioral Development, 40*(2), 100-106.

Hackett, G., & Betz, N. (1989). An Exploration of the Mathematics Self-Efficacy/Mathematics Performance Correspondence. *Journal for Research in Mathematics Education, 20*(3), 261-273.
Kaiser, G. (2017). The teaching and learning of mathematical modeling. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 267-291).
Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM Education, 3*(11). doi:10.1186/s40594-016-0046-z
Lesh, R., Hoover, M., Hole, B., Kelly, A., & Post, T. (2000). Principles for Developing Thought-Revealing Activities for Students and Teachers. In A. Kelly & R. Lesh (Eds.), *Research Design in Mathematics and Science Education* (pp. 591-646). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Maaß, K. (2006). What are modelling competencies? *ZDM: Mathematics Education, 38*(2), 113-142.
Rasmussen, C., & Kwon, O. N. (2007). An inquiry-oriented approach to undergraduate mathematics. *Journal of Mathematical Behavior, 26*, 189-194.
Revelle, W., & Zinbarg, R. (2009). Coefficients alpha, beta, omega and the glb: comments on Sijtsma. *Psychometrika, 74*(1), 145-154.
Sokolowski, A. (2015). The Effects of Mathematical Modelling on Students' Achievement-Meta-Analysis of Research. *The IAFOR Journal of Education, 3*(1), 93-114.
Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and thinks, women and people: A meta-analysis of sex differences in interest. *Psychological Bulletin, 1135*(6), 859-884.

DISEÑO Y VALIDACIÓN DE INSTRUMENTOS DE MEDICIÓN DE COMPETENCIAS Y AUTOEFICACIA DE MODELIZACIÓN MATEMÁTICA

DESIGN AND VALIDATION OF TWO MEASURES: COMPETENCE AND SELF-EFFICACY IN MATHEMATICAL MODELING

Jennifer A. Czocher	Sindura Subanemy Kandasamy	Elizabeth Roan
Texas State University	Texas State University	Texas State University
Czocher.1@txstate.edu	s_k252@txstate.edu	Eaw109@txstate.edu

En este artículo, describimos el diseño y la validación de dos instrumentos – uno que mide la autoeficacia y otro que mide las competencias del proceso de la modelización matemática. La investigación consiste en la evaluación de ambas para establecer la validez y la confiabilidad utilizando técnicas de teoría clásica de validación.

Palabras clave: Valoración y Evaluación, Educación Postsecundaria, Modelización

A pesar de que una de las metas de la educación postsecundaria ha sido el aumentar el número de los graduados universitarios en las carreras de ciencia, tecnología e ingeniería, estos números no se han incrementado suficientemente. Para adquirir conocimiento útil, es necesario que el aprendizaje se fundamente en la combinación de la práctica y teoría de diseño, indagación científica, y en el pensamiento matemático (Kelley & Knowles, 2016). A través de la modelización matemática, se pueden lograr estas metas. Las habilidades de modelización matemática son de suma importancia al cursar carreras universitarias que requieren técnicas aplicadas a las matemáticas. También son importantes para resolver problemas sociales cuyas soluciones conllevan consecuencias mundiales y tangibles. La posibilidad de resolver los problemas sociales llama la atención de los estudiantes (Eccles & Wang, 2016; Su, Rounds, & Armstrong, 2009). Además, las investigaciones empíricas sugieren que aprender matemáticas a través de la modelización es beneficiosa para obtener una autoeficacia y un conocimiento matemático más robusto (Czocher, 2017; Lesh, Hoover, Hole, Kelly, & Post, 2000; Rasmussen & Kwon, 2007; Sokolowski, 2015). La modelización matemática, guiada por las innovaciones educativas, aumenta el interés, la competencia, y la autoeficacia de los estudiantes hacia las matemáticas (Czocher, Melhuish, & Kandasamy, 2019). Conjuntamente, esos factores también están asociados positivamente con la perseverancia en los campos disciplinarios que requieren las matemáticas. Para evaluar atentamente las intervenciones educativas y mostrar su eficacia, es necesario medir el aprendizaje. Esto ayuda a refinar programas que se enfoquen en las habilidades de la modelización matemática. A pesar de su necesidad tangible no existen instrumentos válidos ni confiables para evaluar las habilidades de modelización de los estudiantes universitarios. Aquí, compartimos dos instrumentos de medición y sus propiedades psicométricas: uno de competencias de modelización y otro de autoeficacia en realizarla.

Marco de Referencia

En este trabajo de investigación, se plantea el supuesto de que la modelización matemática es un proceso iterativo y cíclico que puede ser conceptualizado como un conjunto de actividades matemáticas y procesos cognitivos (e.g., Kaiser, 2017). El proceso comienza con un problema de la vida real – como los que son comunes en los estudios de ciencia, ingeniería o en la vida cotidiana – y desemboca en un problema matemático. El problema matemático se puede expresar como una ecuación, un gráfico, o una tabla de valores. El modelador resuelve el problema matemático y desde la solución matemática él interpreta el significado de los resultados al problema original planteado. El modelador valida y verifica cada etapa del proceso para evaluar si el modelo representa correctamente la situación real y si la solución tiene sentido (Czocher, 2018). La Tabla 1 presenta el marco de referencia que se denomina "ciclo de modelización matemática" (CMM) (Blum & Leiss, 2007; Czocher, 2016; Maaß, 2006) y define las competencias que constituyen el proceso de modelización matemática. Definimos la autoeficacia de realizar una tarea como la confianza de una persona en sí misma y en su capacidad para lograr resolver la tarea exitosamente (Bandura, 2006; Betz & Hackett, 1983; Hackett & Betz, 1989). En esta investigación, la autoeficacia siempre es evaluada con referencia al objetivo de la tarea. Definimos el constructo *autoeficacia de modelización matemática* como la confianza de una persona en sí misma y en su capacidad de realizar las actividades interrelacionadas que constituyen el proceso de modelización. De esta manera podríamos medir la autoeficacia de un estudiante para identificar las variables más importantes involucrados en estimar la propagación de hogares inteligentes en el siglo 21. El CMM y la autoeficacia son compatibles, y los utilizamos en conjunto para guiar el diseño de los ítems.

Tabla 1 Competencias de modelización.

Competencia	Descripción
Comprender	Formación de una idea de lo que debe ser el problema o identificación de un fenómeno de la vida real que merece investigación
Establecer estructura	Identificar los factores y cantidades reales relevantes y la información que se puede ignorar; imponer restricciones o supuestos para simplificar el problema
Matematizar	Expresar las relaciones entre las cantidades en una representación matemática
Analizar	Resolver el problema matemático, usando técnicas aprendidas en la clase de matemáticas
Interpretar	Observar y entender los resultados matemáticos desde el contexto del problema real
Validar	Examinar si el modelo representa la situación; verificar el análisis; establecer limitaciones

Metodología

La investigación es de naturaleza cuantitativa y se enmarca dentro de un estudio de desarrollo para establecer evidencia en apoyo de la validez y la confiabilidad de los instrumentos. La población bajo estudio consistió en estudiantes universitarios que estudian carreras en ciencias, tecnología, ingeniería, y matemáticas. A continuación, se documenta el diseño de los ítems. La evaluación del instrumento de autoeficacia se realizó en cuatro rondas de pruebas. En cada prueba empírica, usamos una muestra de estudiantes universitarios inscritos en un concurso internacional de modelización basado en lo que se llama SCUDEM (por sus siglas en inglés). El concurso se lleva a cabo cada año y es parte de una organización de capacitación que apoya a los profesores de matemáticas a quienes les gustaría enseñar los conceptos de ecuaciones diferenciales desde una perspectiva de aplicaciones y modelización matemática. En la primera ronda, eran 6 ítems relacionados con la autoeficacia de modelización. En la segunda, creamos un ítem (Tabla 2, ítem 6) y modificamos los ítems anteriores para mejorar su claridad. En la segunda ronda, también medimos el cambio de autoeficacia antes y después de participar en el concurso y constatamos una ganancia de efecto moderato, $d = 0.545$ ($t(92) = -6.663, p < 0.001$). En la tercera, creamos un item nuevo de análisis matemático (Tabla 2, Ítem 4). Previamente fue excluido porque el enfoque eran las actividades complementarias de modelización. En la tercera ronda el instrumento midió el cambio positivo de autoeficacia ($t = 4.202$, $df = 51, p < 0.001$) de los participantes que contestaron las preguntas antes y después de participar en el concurso. En cada ronda de validación, realizamos un análisis de los componentes principales (Abdi & Williams, 2010), calculamos el Chronbach's α para estimar la consistencia interna, y medimos el cambio de autoeficacia antes y después de participar en el concurso. La Tabla 3 resume los resultados. Este análisis, en conjunto con su construcción basado en la teoría de modelización matemática, indica que el instrumento de autoeficacia es unidimensional con coherencia interna alta y tiene validez de diseño y de constructo.

Tabla 2 El instrumento final de autoeficacia.

Indica tu nivel de confiabilidad en cada uno de los escenarios siguientes, elegiendo un numero de 0 a 100 usando la siguiente escala: **0 10 20 30 40 50 60 70 80 90 100** **No puedo hacer. Tengo dudas… Con certeza elevada.**	**Competencias**
Crear un modelo de ecuaciones diferenciales para estimar la propagación de hogares inteligentes en el siglo 21.	Matematizar
En (1), identificar las cantidades importantes que aseguran una predicción razonablemente precisa.	Establecer estructura
En (1), establecer los supuestos que reducen la cantidad de factores importantes.	Establecer estructura.
En (1), elegir un método apropiado de tipo numérico, gráfico ó analítico para resolver la ecuación diferencial que resulta de (1).	Analizar
En (1) consultar a los recursos apropiados para verificar si el modelo matemático es razonable.	Validar
En (1) enumerar las limitaciones del modelo matemático, incluyendo restricciones de la vida real y restricciones matemáticas.	Validar
En (1), crear una presentación breve para persuadir un fabricante de aparatos inteligentes que podrían depender en tu modelo matemático para fomentar un plan de negocios.	Comunicar
Proporcionando una ecuación diferencial que modela la tasa de formación del material A, $A'(t) = \alpha A(t)^{\beta}$ y los datos de observaciones en tiempo t, la cantidad de material A por cada punto de tiempo t, podría estimar los parámetros α y β.	Establecer estructura

Tabla 3 El resumen del análisis del instrumento de autoeficacia de modelización

Ronda	N	Varianza (ACP)	α	Ronda	N	Varianza (ACP)	α
1	38	62.5%	0.822	3	198	61.5%	0.908
2	276	67.1%	0.917	4	226	69.0%	0.935

La evaluación del instrumento de competencias de modelización se realizó en tres rondas de pruebas con muestras distintas de una universidad de más de 40,000 estudiantes en los EEUU: viabilidad, dificultad, y discriminación. Para diseñar el instrumento tomamos en cuenta cuatro restricciones: (1) los ítems parten de contextos auténticos y relevantes (por ejemplo, la desintegración radioactiva o un programa de reciclaje). (2) Los ítems evocan conocimientos de matemática, ciencia, ingeniería, y sentido común. (3) Los ítems abordan aspectos de las competencias. Por ejemplo, un ítem aborda la competencia de establecer la estructura que se requiere al utilizar la habilidad de identificar cantidades importantes. (4) Los distractores son basados en las decisiones y justificaciones comunes al pensamiento de estudiantes actuales. Elaboramos 118 ítems de tipo selección múltiple (ISM) que pertenecen a 9 situaciones de la vida real elegido de materiales de cursos de matemáticas, física, biología, química e ingeniería. El contenido matemático incluye aritmética, álgebra, cálculo diferencial e integral y ecuaciones diferenciales. Por cada ítem, elaboramos una respuesta correcta y cuatro distractores que parecieran razonables a los estudiantes pero que no ayudaran a modelizar la situación. Para establecer la validez de contenido y la validez de los constructos, invitamos a revisar los ítems a dos investigadores matemáticos que se enfocan en la investigación de ecuaciones diferenciales y tres profesores universitarios de matemáticas que se especializan en realizar investigaciones sobre el aprender y enseñar la modelización. Aplicamos los cambios que sugirieron los expertos y eliminamos los ítems que resultaron no válidos. En la primera ronda, 14 estudiantes nos dieron su razonamiento para justificar sus elecciones. En el caso de que un estudiante eligiera un distractor y su razonamiento tuviera sentido, el ISM fue ajustado. Eliminamos los que no tenían sentido para los estudiantes. En la segunda ronda, 78 estudiantes contestaron 63 ISM en 2 versiones, equilibrando ítems de acuerdo a las distintas competencias de modelización. Por cada ISM, calculamos la dificultad media. La mayoría (76%) de los ISM tenían dificultad moderada (0.20<p<0.70). Eliminamos los ítems que eran demasiado fáciles (p>0.7) y restructuramos los ítems que fueron demasiado difíciles (p<0.20). Para analizar la eficacia de los distractores, calculamos la proporción de los estudiantes que eligieron cada opción. De los 253 distractores (62 ítems contaban con 4 distractores y 1 contaba con 5), El 5% de los participantes eligieron la mayoría de estos distractores. En 17 de los ítems, los distractores fueron elegidos más frecuentemente que las respuestas correctas. Estos fueron identificados de acuerdo con su potencial de discriminar entre estudiantes de distintas habilidades o como ítems que necesitaban ser reestructurados. Después de reestructurar los ISM según el análisis de distractores, elegimos 30 ítems (2 versiones de 15 ítems). Las dos versiones se administraron a una muestra de $n = 314$ voluntarios que participaron en el concurso SCUDEM, incluyendo $n = 135$ que contestaron a la versión 1 y $n = 139$ que contestaron a la versión 2. Por cada ISM, calculamos la dificultad media. La versión 1 obtuvo dificultad media de 0.359 ($SD = 0.126$), con $0.177 < p < .0595$. La versión 2 obtuvo dificultad media de 0.369 ($SD = 0.129$), con $0.147 < p < 0.580$. Cuatro ítems eran demasiado difíciles. Se realizó un análisis de detractores y concluimos que los distractores funcionaban adecuadamente. Para realizar el análisis de discriminación, usamos la correlación point-biserial (rPBIS por sus siglas en ingles). Un solo ítem de la versión 1 tenía rPBIS negativo. El resto tenían rPBIS > 0.20. Reportamos la estadística Revelle's Omega Total (ω_T) para estimar la consistencia interna. La selección fué apropiada en casos donde el instrumento era multidimensional y cuando múltiples dimensiones contribuían a predecir el constructo bajo investigación (Revelle & Zinbarg, 2009). Usando el paquete de software 'userfriendlyscience' del programa R, obtenemos $\omega_T = 0.59$ y $\omega_T = 0.63$ para la versión 1 y la versión 2, respectivamente. Las escalas se acercan al estimado tradicional 0.7.

Discusión

En este artículo, presentamos dos instrumentos de medición, uno de autoeficacia de modelización y uno de competencia de modelización. Así mismo, documentamos los procesos de construcción y diseño y las propiedades de ambos. Los instrumentos se alinean con las teorías de modelización y han pasado múltiples rondas de pruebas. Se planea emplear la Teoría de Repuesta al Ítems para componer versiones que sean paralelas para medir los cambios positivos de las competencias de modelización de los estudiantes con el propósito de evaluar programas educativos que se enfoquen en ensenar la modelización. Con esta información investigadores y docentes pueden mejorar las experiencias de modelización o proporcionar evidencia de su éxito. Estamos cautas pero optimistas que los instrumentos alcancen este objetivo ya que la evidencia expuesta aquí sugiere que los instrumentos son confiables y válidos para su propósito.

Agradecimientos

Este material se ha desarrollado gracias al apoyo de la donación No. 1750813 de la National Science Foundation. Además, estamos agradecidas a nuestras maravillosas colegas Alejandra Sorto, Monica Uribe, Gloria Velasquez y Alba Melgar por el apoyo brindado en la revisión del presente artículo.

Referencias

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley interdiciplinaty reviews: computational statistics, 2*, 433-459.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 3, pp. 307-337). United States of America: Information Age Publishing.

Betz, N., & Hackett, G. (1983). The relationship of mathematics self-efficacy expectations to the selection of science-based college majors. *Journal of Vocational Behavior, 23*(3), 329-345.

Blum, W., & Leiss, D. (2007). How do students and teachers deal with modelling problems? In C. Haines, P. L. Galbraith, W. Blum, & S. Khan (Eds.), *Mathematical Modelling (ICTMA 12): Education, Engineering and Economics* (pp. 222-231). Chickester: Horwood Publishing Limited.

Czocher, J. (2016). Introducing Modeling Transition Diagrams as a Tool to Connect Mathematical Modeling to Mathematical Thinking. *Mathematical thinking and learning, 18*(2), 77-106.

Czocher, J. (2017). How can emphasizing mathematical modeling principles benefit students in a traditionally taught differential equations coures? *The Journal of Mathematics Behavior, 45*, 78-94.

Czocher, J. (2018). How does validating activity contribute to the modeling process? *Educational Studies in Mathematics, 99*, 137-159.

Czocher, J., Melhuish, K., & Kandasamy, S. S. (2019). Building mathematics self-efficacy of STEM undergraduates through mathematical modeling. *International Journal of Mathematical Education in Science and Technology*, 1-28. doi:10.1080/0020739X.2019.1634223

Eccles, J. S., & Wang, M.-T. (2016). What motivates females and males to pursue careers in mathematics and science? *International Journal of Behavioral Development, 40*(2), 100-106.

Hackett, G., & Betz, N. (1989). An Exploration of the Mathematics Self-Efficacy/Mathematics Performance Correspondence. *Journal for Research in Mathematics Education, 20*(3), 261-273.

Kaiser, G. (2017). The teaching and learning of mathematical modeling. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 267-291).

Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM Education, 3*(11). doi:10.1186/s40594-016-0046-z

Lesh, R., Hoover, M., Hole, B., Kelly, A., & Post, T. (2000). Principles for Developing Thought-Revealing Activities for Students and Teachers. In A. Kelly & R. Lesh (Eds.), *Research Design in Mathematics and Science Education* (pp. 591-646). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Maaß, K. (2006). What are modelling competencies? *ZDM: Mathematics Education, 38*(2), 113-142.

Rasmussen, C., & Kwon, O. N. (2007). An inquiry-oriented approach to undergraduate mathematics. *Journal of Mathematical Behavior, 26*, 189-194.

Revelle, W., & Zinbarg, R. (2009). Coefficients alpha, beta, omega and the glb: comments on Sijtsma. *Psychometrika, 74*(1), 145-154.

Sokolowski, A. (2015). The Effects of Mathematical Modelling on Students' Achievement-Meta-Analysis of Research. *The IAFOR Journal of Education, 3*(1), 93-114.

Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and thinks, women and people: A meta-analysis of sex differences in interest. *Psychological Bulletin, 1135*(6), 859-884.