

DISEÑO Y APLICACIÓN DE UNA PRUEBA PARA EVALUAR COMPETENCIAS EN EL ÁREA DE MEDICIÓN Y EVALUACIÓN EN ESTUDIANTES DE PSICOLOGÍA

Bertha-Lucía Avendaño, María del Pilar Fresneda,
Aida-Ingrith Martínez & Ingrid Santos
Universidad Católica de Colombia

Resumen

El objetivo general del presente estudio fue diseñar una prueba para evaluar competencias de los estudiantes de Psicología, en el área de Medición y Evaluación. Basados en los contenidos de las cuatro asignaturas de medición que ofrece el programa de psicología de una universidad privada de la ciudad de Bogotá, se elaboró la prueba, que fue sometida a juicio de expertos y aplicada a 40 estudiantes que voluntariamente la respondieron. El instrumento diseñado cumplió con los requisitos psicométricos básicos exigidos por la Teoría de Respuesta al Ítem (TRI). La confiabilidad más alta se encontró en los ítems de Medición y Evaluación I y II, con un valor de 0.85 y 0.80 respectivamente; la confiabilidad más baja se encontró en los ítems de Medición IV con un valor de 0.59. Bajo los lineamientos de la TRI, se puede concluir que se dio respuesta a la pregunta planteada, ya que el instrumento elaborado resultó adecuadamente ajustado y los ítems se encontraron bien distribuidos según los parámetros que plantea el modelo.

Palabras Clave: *prueba, medición, evaluación, calidad, competencias, educación, tri*

Abstract

The general objective of the present study was to design a test to evaluate the competences of Psychology students about the area of Measurement and Evaluation. Based on the contents of the four measurement subjects that the Psychology program of a private university of Bogotá city offers, a test was designed and it was given to experts on order to appreciate their opinion, after that it was applied to 40 voluntary students. The designed instrument fulfilled the basic psychometric requirements demanded by the Item Response Theory (IRT). The highest reliability was found in the analysis of the items of Measurement and Evaluation I and II, with a value of 0.85 and 0.80 respectively; the lowest reliability was found in the analysis of the Measurement IV items with a value of 0.59. Under the lineaments of the IRT, as a conclusion it could be said that an answer was given to the outlined question, since the instrument designed was appropriately adjusted and the items were correctly distributed according to the parameters that the pattern outlines.

Key Words: *test, measurement, evaluation, quality, competences, education, irt*

Introducción

El surgimiento de las instituciones educativas no está desligado del contexto, ya que nacen en momentos específicos buscando dar respuesta a las diferentes necesidades políticas, sociales y económicas vividas en esos momentos y en esos contextos; las universidades son

pues producto de la sociedad. Cada institución tiene su propia cultura, su misión y su visión, cada una busca aportar desde sus áreas, soluciones que contribuyan al beneficio del país; pero el tipo de problemas que cada universidad analiza y por ende cada programa dentro de la misma, varía de acuerdo con los antecedentes, historia y la propia dinámica que haya desarrollado.

La primera universidad que se fundó en Colombia fue la Santo Tomás de Aquino en 1580, y más adelante, a partir de 1622 se crearon las universidades Javeriana, Colegio Mayor de Nuestra Señora del Rosario, Colegio Mayor de San Buenaventura, Universidad de Antioquia, Nacional, Cauca, Cartagena, Externado de Colombia, Nariño y Libre en 1923. La mayoría de estas universidades impartían una educación basada en la formación religiosa y el conocimiento del latín; hasta ese momento la educación superior estaba prácticamente orientada a la formación de la clase dirigente. Más tarde, en 1936, se crea la Universidad Pontificia Bolivariana, en 1947 la Industrial de Santander y la Universidad de Los Andes en 1949. No es sino hasta 1952, cuando se marca una gran diferencia en la educación con la fundación de la Universidad Gran Colombia debido a que se constituye como la primera universidad nocturna, dirigida a la clase obrera y trabajadora. Más adelante, en 1970 se firma el acta de constitución de la Fundación Educacional Interamericana (FEI), conocida hoy como la Universidad Católica de Colombia, que actualmente junto con la Fundación Universitaria Konrad Lorenz es de las pocas universidades que ofrecen el programa de Psicología en jornada nocturna.

Posteriormente, se crea el instituto de crédito ICETEX, para facilitar la capacitación profesional de los obreros (Vargas, Scoppetta, Cristancho, Cubides, Flórez, Gamboa, Montenegro, Ortega, Pérez, & Romero, 1995).

En 1968, se crea el Instituto Colombiano para el Fomento de la Educación Superior ICFES, con el objetivo de coordinar y fomentar la educación superior y otorgar el reconocimiento oficial a los nuevos planteles oficiales y privados y los títulos que se expidan. En 1969, se crea COLCIENCIAS cuyo objetivo fundamental es fomentar la investigación universitaria. En 1970 se crea el Fondo Universitario Nacional FUN, que luego se transforma en ASCUN, como entidad encargada de la vigilancia de la educación superior.

Con la Ley 30 de 1992 se creó el sistema Nacional de Acreditación, con el objetivo fundamental de garantizar a la sociedad que las instituciones que hacen parte del sistema cumplan con los más altos requisitos de calidad y realicen sus propósitos y objetivos. Así, el Consejo Nacional de acreditación CNA, instalado en julio de 1995, orienta el proceso de acreditación, lo organiza, lo fiscaliza, da fe de su calidad y recomienda al Ministro de Educación Nacional acreditar los programas o instituciones que cumplen con los requisitos establecidos. Con el artículo 34 de la misma ley se crea el Consejo Nacional de Educación Superior (CESU), como organismo del gobierno nacional vinculado al Ministerio de Educación Nacional con funciones de coordinación, planificación, recomendación y asesoría.

En 1994 con la Ley 115, se expidió la Ley General de Educación que indica las normas generales para regular el servicio público de la educación; es decir, se señalaron las acciones y actividades necesarias para enriquecer el proceso de Acreditación. Se determinó la preparación del Plan Nacional de Desarrollo Educativo que contiene los requisitos de calidad y cobertura, el cual empezó a regir desde 1996 por un término de 10 años al cabo de los cuales el MEN, en coordinación con las entidades territoriales lo preparará nuevamente; así mismo se determinó que cada institución elaborara y pusiera en práctica un PEI (Proyecto Educativo Institucional). Se estableció la evaluación anual de las instituciones, el examen anual de idoneidad académica para los docentes en su área de especialidad docente y de actualización pedagógica y profesional cada 6 meses y la formación de educandos con los más altos estándares de calidad.

Con el fin de velar por la calidad, el cumplimiento de los fines de la Educación y por la formación moral, intelectual y física de los educandos, se determinó que el MEN estableciera un sistema Nacional de Evaluación de la Educación (SNEE), el cual operaría en coordinación con el Servicio Nacional de Pruebas ICFES y con las entidades territoriales.

En el Decreto 2790 de 1994 se dictaron las normas para la inspección y vigilancia de los programas académicos de pregrado en la educación superior. Mientras que con el Decreto 0620 de 2000 se establecen reglas y criterios para la evaluación de docentes en ejercicio, al servicio del Estado.

Con el Decreto 0917 de 2001, se establecen los estándares de calidad en programas académicos de pregrado en ciencias de la salud y con el 1527 de 2002 se establecen los estándares de calidad en los programas profesionales de pregrado en Psicología.

Es importante anotar que en el siglo XX Colombia experimentó grandes cambios en sus estructuras sociales y económicas que modificaron el estilo de vida, así como los valores y actitudes de la población. Desde 1950 la población se triplicó; aumentó la esperanza de vida al nacer y la juventud ingresó masivamente a los niveles secundario y terciario de la educación, lo que no garantizó su ingreso al mercado laboral (Aldana, E., Chaparro, L., García, G., Gutiérrez, R., Llinás, R., Palacios, M., Patarroyo, M., Posada E., Restrepo A. & Vasco E., 1996)

La violencia en el país, agudizada en los últimos 50 años, desgastó a las instituciones políticas y sociales, debilitó la presencia y control del Estado en todos los niveles, disminuyó la participación democrática de ciudadanos, incrementó y consolidó la corrupción y los grupos al margen de la ley. Como respuesta a esta problemática, el Estado y la sociedad han venido dando pasos en función de legitimar las instituciones gubernamentales y favorecer la equidad y las oportunidades de los ciudadanos. Entre las medidas más destacables está el nuevo contrato social a través de la Constitución de 1991, que ha permitido la modernización y la reforma de los organismos ejecutivo, legislativo y judicial (Aldana & otros, 1996).

En el plano económico, la apertura ha sido una respuesta a la tendencia mundial de globalización de los mercados y al mejoramiento de la productividad y competitividad de la economía. Esta realidad de cambio y modernización no ha sido ajena a la educación en donde las instituciones educativas tenían dificultades dadas por la ausencia de currículos actualizados, que promovieran la formación integral del estudiante, las deficiencias académicas y pedagógicas de los docentes, la lentitud en la actualización e implementación de los avances científicos y tecnológicos, la falta de procesos válidos de selección de estudiantes y la escasa producción de investigaciones.

Como solución a las problemáticas dadas en las instituciones y en la educación en general, la Misión de Ciencia, Educación y Desarrollo en su informe “Colombia: al filo de la oportunidad” (1995) propuso seis grandes temas para una reforma educativa que permitiera la solución a los problemas antes mencionados. Los temas tocaban indicadores relacionados con el contenido y alcance de la reforma en sus aspectos cualitativos intrínsecos (los valores, los métodos y estrategias pedagógicas, la promoción de los métodos de las ciencias modernas, su inserción en la cultura nacional), así como su concordancia con las metas constitucionales. Los recursos financieros mínimos requeridos, los cambios en el manejo financiero y administrativo de la educación, necesarios para que el sistema educativo pueda absorber eficientemente los recursos financieros. La efectividad de la educación, en sus aspectos de eficacia y eficiencia, los cambios institucionales que suponen la descentralización administrativa, la democracia local y la autonomía institucional. Los cambios en las instituciones educativas, que requieren un sistema con mayor integración entre sus niveles primario, secundario y postsecundario, de un lado; y del otro, entre el mundo escolar y el mundo laboral y productivo.

Las propuestas de la misión, conjuntamente con los principios enmarcados en la constitución del 91 han tenido eco en el Estado a través de las reformas a la educación, dadas a partir de la Ley 30 de 1992 para la Educación Superior, la Ley General de la Educación (Ley 115 de 1994) y de una serie de decretos y resoluciones que han marcado la entrada a una nueva visión de la educación en Colombia en función de la calidad, a partir de la evaluación permanente de sus educandos, las instituciones y los programas, entre otros. Dicha evaluación, ha sufrido diferentes cambios a lo largo de los años y sobretodo últimamente con las reformas educativas de cada gobierno. Particularmente se puede hablar de la evaluación psicológica en Colombia, la que fue importante a través de la Psicometría durante la última mitad del siglo XX y el primer año del siglo XXI, con la medición de atributos del ser humano. Colombia ha sido un país líder en el contexto latinoamericano en los avances implementados en este campo, al ser el primero en incorporar, desde inicios de la década del 90, nuevas tecnologías para el análisis de datos y discusiones conceptuales relacionadas con la validez, confiabilidad y evaluación. En el campo de la evaluación educativa inclusive ha asesorado a otros países de la región. En este aspecto vale reconocer el trabajo que ha realizado el ICFES (Pardo, 2001).

En pro de la Evaluación y mejoramiento de la educación superior, se crearon los Exámenes de Calidad de la Educación Superior ECAES, algunos de los cuales están basados en competencias y constituyen una modalidad de Examen de Estado para la evaluación externa de los estudiantes de últimos semestres, y forman parte con otros procesos y acciones de un conjunto de instrumentos que el gobierno nacional dispone para evaluar y mejorar la calidad de la educación. Los objetivos generales de estos exámenes son comprobar niveles mínimos de conocimientos de los estudiantes de último año de pregrado y homologar y convalidar títulos de estudios de Educación Superior realizados en el exterior. (Artículos 1 y 2 del capítulo I, Decretos 1716 y 2233 de 2001).

Llevar a la práctica los exámenes de Calidad de la Educación Superior (ECAES) se había intentado desde 1967, pero solo se comienzan a realizar a partir del año 2001 cuando se hace realidad este propósito tras la expedición de los decretos 1716 y 2233, que reglamentan los ECAES en las carreras de Medicina e Ingeniería Mecánica, respectivamente. De acuerdo con la norma hasta el momento, la presentación de los exámenes en el primer año será voluntaria pero a partir del segundo es obligatoria. Para la carrera de Derecho se programó la aplicación de los exámenes con carácter voluntario en el 2002. (Balance de Gestión 1998-2002, ICFES, 2002). Para los programas de Psicología fue voluntaria en el año 2003 y ASCOFAPSI (Asociación Colombiana de Facultades de Psicología) fue la organización encargada de elaborar el examen de Psicología.

Desde el ICFES se avanza actualmente en la fundamentación conceptual de exámenes de diferentes disciplinas, en las profesiones de Ingeniería eléctrica, Civil, Industrial, de Sistemas, Administración, Enfermería, Contaduría, Arquitectura, Odontología, entre otras; los cuales se irán implementando paulatinamente en los años sucesivos y tienen dentro de sus objetivos plantear los ítems de los exámenes ECAES en términos de competencias.

Evaluación de competencias

Respecto al concepto de competencia se puede establecer que fue trabajado desde la lingüística con Noam Chomsky a finales de los años 50, para explicar el carácter creativo o generativo del lenguaje y para dar cuenta de la facilidad con la que un niño se apropia del sistema lingüístico y llega al campo de la educación después de la relectura al interior de la Psicología cognitiva y cultural. La corriente cognitiva con la propuesta de competencia lingüística, postulada por Chomsky define la competencia como un conocimiento de las reglas o principios abstractos que regulan el sistema lingüístico que está representado en la mente de los hablantes y que es parcialmente innato, en el sentido de que no se deriva totalmente de la experiencia y sólo se evidencia a través de la actuación o desempeño lingüístico, es decir, del habla, la escritura y la lectura (Bogoya, Torrado, Vinent, Jurado, Pérez, García, Acevedo, Cárdenas, Granés & Sarmiento, 1999).

Torrado (1998) plantea que el concepto de competencia no se reduce a aspectos intelectuales o formales sino que incorpora la idea de “características personales (conocimientos, destrezas y actitudes) que producen resultados adaptativos en ambientes significativos” y cita a Raven para señalar que las competencias están conformadas por “habilidades motivadas” de tipo cognitivo, social o físico. Es decir, el individuo tiene la oportunidad de practicar nuevos estilos de conducta en diversos contextos. Bogoya et al (1999) la definen como un saber hacer en el ámbito de un contexto determinado, de modo que el concepto pueda ser determinado en múltiples actividades y en variadas situaciones (Bogoya et al, 2000).

Bogoya et al (2000) realizan una propuesta de la competencia graduada en tres niveles, en los cuales al pasar del nivel uno al tres se muestra un grado de dominio y profundidad cada vez mayor, una elaboración conceptual más rigurosa y exigente. El primer nivel como línea de base, da razón de la apropiación de un conjunto mínimo de conocimientos. El segundo nivel hace referencia al uso comprensivo de los objetos o elementos de un sistema de significación dirigido hacia la resolución de problemas. El tercer nivel comprende el control y la explicación del uso. En este nivel puede formularse un juicio, en el cual se interpreta, conjetura y generaliza.

La evaluación basada en competencias es un elemento esencial del enfoque basado en normas de competencia educativa. La evaluación por competencias se distingue de la evaluación tradicional porque en esta última la persona evaluada se observa de forma fragmentada, es decir, por un lado estaba su personalidad, por otro sus motivaciones, más allá de sus valores y por otro sus habilidades; de forma que se utilizaba una prueba por cada fragmento; luego el evaluador debía interpretar los resultados para concluir con ello en un perfil del candidato, independiente del contexto en que se encontrase. En la evaluación por competencias se observa a la persona como un ser integral, el cual despliega sus dimensiones de competencia en una área de desempeño específica.

La evaluación por competencias, de acuerdo con Torrado (1998), presenta algunas características que se resumen a continuación: a) no es posible evaluar directamente, siendo necesario entonces evaluarse a través de la actuación, la cual es variable, b) Debido a que el término contexto incluye los tipos de instrumentos simbólicos, procedimientos o herramientas mediadoras que hacen posible la ejecución, es importante multiplicar las posibilidades de uso de la competencia a través de la introducción de diferentes lenguajes y medios de representación del conocimiento y c) En la actuación se integran conocimiento generador de las estrategias para resolver un problema y habilidad.

Los modelos psicométricos utilizados para el análisis de evaluaciones por competencias son la Teoría Clásica de los Tests (TCT) y la Teoría de Respuesta al ítem (TRI). La TCT fue hasta finales del siglo XX, el modelo en el cual se basaron los investigadores para construir test psicométricos e interpretar las puntuaciones logradas con estos instrumentos.

Actualmente, la Teoría de Respuesta al Ítem (TRI), comienza a tomar auge en el campo de la medición aunque realmente la génesis de ésta teoría puede encontrarse a mediados de los años treinta y comienzos de los cuarenta (Hamblenton & Swaminathan, 1985).

En la teoría clásica, la medida de una variable está asociada con el instrumento utilizado para medirla, las propiedades de los ítems están en función de los sujetos a los que se aplican; es decir, la puntuación depende del tipo de prueba y las propiedades de la prueba están en función de las personas que la responden. La habilidad promedio o el rango de puntajes en un grupo de examinados, depende de la muestra a la cual se aplicó la prueba y de los ítems (Avendaño y Medellín, 2001).

La teoría Clásica de los Tests, (TCT) presenta varias debilidades, a saber: no proporciona información precisa acerca de la habilidad de cada una de las personas, ya que si dos examinados obtuvieron la misma puntuación, por ejemplo 50, no se puede establecer realmente si los dos tienen la misma habilidad; además, presenta problemas con la definición del término Fiabilidad el cual está relacionado con el concepto de formas paralelas de una prueba, bastante cuestionado. Por otro lado, el modelo presume que la varianza del error de medida es igual para todos los examinados y no determina que algunas personas son más consistentes en sus respuestas que otras y que ésta consistencia varía de acuerdo con la habilidad (Muñiz, 1990, citado por Avendaño, 2003).

Por todas las razones anteriores los psicómetras han buscado modelos alternativos para la medición de rasgos y constructos psicológicos que cumplan con las siguientes propiedades:

Que las características de los ítems no dependan del grupo en que se calculan

Que las puntuaciones que describan el comportamiento de los sujetos examinados con las pruebas, no dependan de los contenidos particulares de éste ni de su dificultad.

Que los modelos estén expresados a nivel del ítem y no de la prueba completa.

Que el modelo no requiera del supuesto estricto del paralelismo u otras formas de equivalencia para establecer la confiabilidad de la prueba.

Que el modelo proporcione medidas de precisión para cada nivel de habilidad, eliminando el supuesto de la homoscedasticidad de la varianza del error.

Que el modelo permita relacionar el rasgo o constructo con el rendimiento de la persona en el ítem (Martínez, 1996).

El objetivo central de la TRI será solucionar los anteriores problemas, lo que permitirá en términos generales: obtener mediciones que no varíen en función del instrumento utilizado, que sean invariantes respecto de los test empleados y disponer de un instrumento de medida

cuyas propiedades no dependan de los objetos medidos y no varíe respecto de los sujetos evaluados (Muñiz, 1990).

La TRI proporciona un sustento teórico basado en la estadística, la matemática, el álgebra y el cálculo, y dependiendo del modelo, maneja diferentes parámetros, que hace posible establecer un banco de ítems adaptado al nivel del examinado, lográndose con esto una evaluación más ajustada a las características de las personas que responden la prueba.

Pueden existir muchos modelos que se diferencian ya sea por la forma matemática de la curva característica del ítem o por el número de parámetros que considera. Todos los modelos tienen por lo menos un parámetro que describe al ítem y por lo menos uno que describe a la persona.

El modelo de Rasch, como su nombre lo indica fue formulado originalmente por Rasch y posteriormente fue trabajado en la Universidad de Chicago por Wright y sus colaboradores. Es tal vez el modelo más conocido de la TRI. El análisis de Rasch es un modelo que establece la probabilidad de respuesta de una persona ante un estímulo dado, en términos de la diferencia entre la medida del rasgo de la persona y la medida del estímulo utilizado (Tristán, 2001).

El modelo de respuesta estocástica de Rasch, describe la probabilidad del éxito de una persona en un ítem como una función de la habilidad de la persona y la dificultad de la pregunta, siendo una aproximación estadística al análisis de las respuestas a una prueba y de otros tipos de observación ordinal. Rasch derivó su modelo como una expresión logística simple y demostró que en esta forma, los parámetros de la persona y de la pregunta son estadísticamente independientes.

Adicionalmente al modelo de Rasch o de un parámetro, existen el de dos y el de tres parámetros, que se basan en principios similares a los planteados anteriormente. Hay algunas diferencias a nivel conceptual, en términos de los parámetros a partir de los cuales se hace la medición.

A partir de los aspectos expuestos anteriormente se plantea la siguiente pregunta de investigación ¿Cumple la prueba elaborada para evaluar competencias en el área de Medición y Evaluación, con los parámetros psicométricos exigidos por el modelo de la Teoría de respuesta al ítem?

Como objetivo general se propuso diseñar y aplicar una prueba para evaluar competencias en el área de Medición y Evaluación en estudiantes de Psicología. Para lograr tal objetivo, se elaboró una prueba cuyos contenidos fueron sometidos a Juicio de expertos quienes validaron la pertinencia, relevancia, gramática y representatividad de los ítems, de acuerdo con los contenidos de las asignaturas de Medición y Evaluación del programa de Psicología analizado. A partir de las recomendaciones y sugerencias hechas por ellos, se realizaron las

modificaciones necesarias. Posteriormente la prueba diseñada se aplicó a un grupo de estudiantes para conocer sus competencias en los diferentes temas evaluados. Se analizó también la calibración de los ítems bajo los parámetros psicométricos exigidos, con el fin de aportar con el instrumento a la preparación de los estudiantes al examen ECAES exigido por el gobierno.

Las variables tenidas en cuenta fueron las competencias y los temas de las asignaturas de los cuatros programas de Medición y Evaluación. Respecto a las competencias se tomó la definición dada por el ICFES (2000), que define la competencia como ‘saber hacer en contexto’. Las competencias son las acciones que un estudiante realiza en el contexto de una disciplina del conocimiento o de una problemática. Se evaluaron las acciones de tipo interpretativo, argumentativo y propositivo. Las acciones de tipo Interpretativo están orientadas a encontrar el sentido de un texto, de una proposición, de un problema, de una gráfica, de alguna representación o esquema; es decir, se pretende reconstruir local y globalmente un texto. Las acciones de tipo Argumentativo son aquellas acciones que tienen como fin dar razón de una afirmación y se expresan en el por qué de una proposición, en la articulación de conceptos y teorías, en la demostración matemática, en la organización de premisas para sustentar una conclusión, en el establecimiento de relaciones causales, entre otras, y las acciones de tipo propositivo, implican el planteamiento de hipótesis, la resolución de problemas, la elaboración de alternativas de explicación a un evento o a un conjunto de ellos.

Respecto a los cuatro programas, a continuación se enuncian los temas principales de cada uno de ellos. El programa de Medición y Evaluación I incluye los temas: visión histórica y actual de la medición y evaluación; conceptos de medición y evaluación; escalamiento; construcción de escalas sociales; evaluaciones alternativas - mapas conceptuales; marco de interpretación de resultados con referencia al criterio; teoría clásica de los tests, curva normal, confiabilidad y error de medición; Teoría General de la Validez de la medición; validez relacionada con el constructo, contenido y criterio; y validez integral.

Los temas fundamentales del programa de Medición y Evaluación II son: el diseño y validación de un instrumento de conocimientos, la Teoría de la Generalizabilidad, la Teoría de Respuesta al Ítem y los procesos de Orientación Vocacional. Medición y Evaluación III, revisa ampliamente los conceptos de inteligencia, competencia, aptitud y personalidad haciendo una evaluación de los problemas teóricos y evaluativos en la medición de las mismas; los estudiantes se preparan en el manejo de diferentes instrumentos que les serán muy útiles durante su vida profesional.

En lo referente a Medición y Evaluación IV, el programa comprende los siguientes temas: evaluación psicológica en diferentes contextos, procesos de evaluación en el contexto organizacional, competencias profesionales, proceso de selección de personal, proceso de evaluación de desempeño, evaluación psicométrica en clínica, presentación de instrumentos de evaluación en psicología clínica, organizacional y desarrollo.

Método

Diseño

Para el desarrollo del presente estudio se utilizó un diseño de tipo descriptivo con aplicaciones en el campo psicométrico. Se recurrió al juicio de expertos con el fin de elaborar un instrumento que permitiera realizar una prueba para evaluar las competencias en el área de Medición y Evaluación.

Participantes

Se trabajó con 40 estudiantes de un programa de Psicología quienes ya habían tomado las cuatro asignaturas de Medición y Evaluación que el programa ofrece. Los estudiantes participaron de manera voluntaria en la aplicación de las pruebas.

Instrumento

Se utilizó la prueba realizada por las investigadoras para evaluar competencias en el área de Medición y Evaluación en estudiantes de Psicología. La prueba abarcó los temas de Medición y Evaluación de una facultad de Psicología de la ciudad de Bogotá, el cual está conformado por cuatro asignaturas en donde se tratan los temas fundamentales de este campo del saber. El número de preguntas para cada uno de los programas se determinó a través del plan de prueba revisado por los jueces (Ver tabla 1)

Tabla 1. Plan de Prueba

	AC. TIPO INTERPRETATIVO	AC. TIPO ARGUMENTATIVO	AC. TIPO PROPOSITIVO	TOTAL
MEDICIÓN I	8	12	-	20
MEDICIÓN II	9	2	1	12
MEDICIÓN III	11	9	3	23
MEDICIÓN IV	9	1	2	12
TOTAL	37	24	6	67

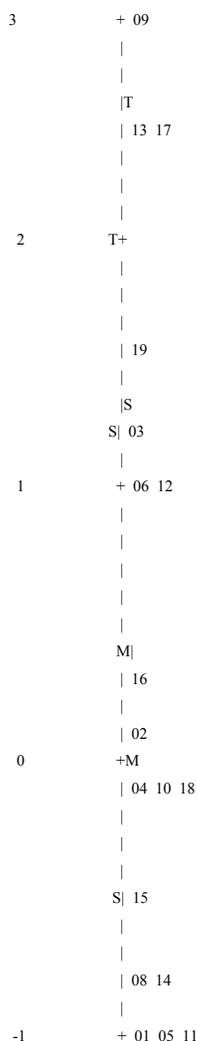
Resultados

El análisis de resultados se realizó con el paquete estadístico WINSTEPS, diseñado para examinar ítems bajo el modelo de Rasch. El WINSTEPS muestra gráficos y tablas que describen las medidas de tendencia central y variabilidad; presenta además, la dificultad, discriminación y posición de cada ítem dentro del continuo de la variable estudiada, con una estimación central para la medida de cada persona y calibración de ítems.

Dando respuesta a la pregunta planteada en esta investigación se analizó el comportamiento de los ítems, en cada una de las cuatro asignaturas del área de Medición y Evaluación. Sólo se describe el comportamiento de los ítems de la asignatura de Medición I, ya que fueron los que obtuvieron la confiabilidad más alta.

Análisis de los Ítems de Medición I

La figura 1 representa uno de los resultados más importantes que arroja el programa Winsteps; en la cual se pueden identificar los siguientes aspectos: la primera columna presenta los puntajes logit, que para el caso indica que los 20 ítems de la prueba se distribuyeron entre -2.18 y 3 desviaciones logit. La línea punteada tiene tanto a la izquierda como a la derecha las letras M, S y T. La M indica la media, la S una desviación y la T dos desviaciones, todas en escala logit. Los números de la derecha indican la distribución de los ítems. Se observa que el ítem número 9 es el más difícil, seguido por los ítems 13 y 17 que se encuentran al mismo nivel, le siguen en dificultad los ítems 19 y 3. Aparecen después los ítems 6 y 12 (con igual



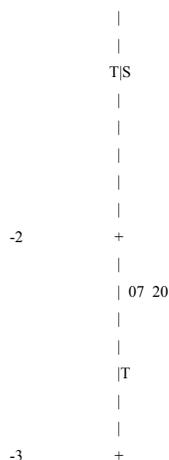


Figura 1. Calibración de los ítems a lo largo de la variable (Medición I)

nivel de dificultad), luego el 16 y el 2. A continuación aparecen los ítems 4, 10 y 18 (que tienen la misma dificultad) seguidos del ítem 15. Finalmente aparecen los ítems 8, 14, 1, 5 y 11 seguidos del 7 y 20 que son los más fáciles. Los espacios entre ítems en la figura indican que es necesario elaborar más ítems que permitan medir las porciones de la variable que no están siendo evaluadas. Así por ejemplo, deben diseñarse preguntas que evalúen el espacio entre el ítem 9 y los ítems 13 y 17. Por otro lado, se puede afirmar que las preguntas 2, 4, 10 y 18 tienen una dificultad promedio y que en términos generales el 50% de las preguntas se encuentran entre -1 y +1 desviación logit.

En los resultados que aparecen en la tabla 2 se presenta el resumen de los estadísticos que arrojó el modelo de los ítems respecto al instrumento; la confiabilidad de la totalidad de la prueba es de .85 reflejando un alto nivel de confiabilidad en la medición del atributo. El nivel de ajuste para los puntajes por medio de método de mínimos cuadrados estandarizados (MNSQ) alrededor de la media es 1.00 y este mismo valor se obtuvo para los extremos, esto indica que la prueba se ajusta perfectamente en los valores promedios y en los extremos; el modelo predice que los puntajes de los ítems deben encontrarse dentro de un rango de ajuste entre 0.7 y 1.3 puntuaciones logit. Por lo anterior, se puede asegurar que las preguntas se encuentran bien distribuidas según los parámetros que plantea el modelo. Los puntajes máximo y mínimo en desviaciones logit concuerdan con la figura 1 y muestran con precisión los valores logit superior e inferior que para el caso de los ítems de Medición I, indica que el ítem 9 se encuentra a 3.00 desviaciones logit por encima de la media y los ítems 7 y 20 a -2.18 desviaciones logit por debajo de la media. Cabe resaltar que el programa no tuvo en cuenta el ítem 9 para el análisis, ya que no contribuye al ajuste en el modelo por su nivel de dificultad.

Tabla 2. Resumen general de los puntajes obtenidos en los ítems de Medición I

	Media Puntaje Logit	Error del Modelo	Ajuste a l a Media MNSQ	Ajuste a los Extremos MNSQ
Media	12.4	.53	1.00	1.00
Desviación Estándar	5.1	.10	0.12	0.25
Puntaje Máximo	3.00	.75	1.23	1.57
Puntaje Mínimo	-2.18	.46	0.80	0.68
Índice de Confiabilidad	Índice Real	.85		

En el análisis de los ítems se determinó que 11 de los 67, quedaron ubicados por encima de dos desviaciones logit ; 4 de competencia interpretativa y 7 de competencia argumentativa. Debajo de 2 desviaciones logit quedaron 5 ítems; 3 de competencia interpretativa y 2 de argumentativa.

Discusión

El objetivo general del presente estudio fue diseñar una prueba para evaluar competencias en estudiantes de Psicología en el área de Medición y Evaluación. La prueba diseñada cumplió con los requisitos psicométricos básicos exigidos por la TRI; la confiabilidad más alta se encontró en los ítems de Medición y Evaluación I y II, con un valor de .85 y 0.80 respectivamente, la confiabilidad más baja se encontró en los ítems de Medición IV con un valor de .59. Bajo los lineamientos de la TRI, se puede concluir que se dio respuesta a la pregunta planteada en la presente investigación, ya que el instrumento elaborado resultó adecuadamente ajustado, de acuerdo con el método de mínimos cuadrados estandarizados (MNSQ); los puntajes de los ítems se encontraron dentro del rango de ajuste que está entre 0.7 y 1.3 en valores logit; esto indica que las preguntas se distribuyeron según los parámetros que plantea el modelo.

A continuación, se presentan los temas que resultaron ser más fáciles y más difíciles en cada una de las asignaturas evaluadas. En Medición y Evaluación I, los ítems más fáciles para los estudiantes estaban relacionados con los temas de confiabilidad e historia de la medición. Probablemente, el énfasis que el programa de medición hace en confiabilidad ha permitido la aprehensión de dicho concepto por parte de los estudiantes. Así también, la historia de la medición, puede ser un tema de fácil manejo, en la medida en que no está relacionado con niveles complejos de cognición.

Por otro lado, los ítems más difíciles hicieron alusión a la parte teórica y práctica de las escalas de medición, tema que en los diferentes programas de Psicología no solamente se estudia en Medición y Evaluación, sino también en Metodología y Estadística. Estos resultados sugieren que en los programas, se debe seguir haciendo énfasis en estos contenidos básicos.

En Medición y Evaluación II, el tema más fácil fue el de intereses vocacionales, tópico más utilizado en Psicología vocacional, tal vez porque posee un carácter motivador que la mayoría de los autores le confiere y al cual se le da bastante relevancia en la asignatura. El ítem más difícil se relacionó con el concepto de validez que en la práctica resulta complicada su aplicación.

En Medición y Evaluación III, la pregunta más fácil, al igual que en Medición y Evaluación I, se relacionó con la historia de la medición. La más difícil, tuvo que ver con los tipos de inteligencia planteados por Gardner; probablemente éste sea un tema poco trabajado, ya que en esta asignatura, los estudiantes se dividen para exponer en clase los diferentes autores y esto puede dificultar el aprendizaje de cada temática por parte de la totalidad del grupo de estudiantes.

En lo relacionado con Medición y Evaluación IV, la pregunta más fácil hacía referencia a la prueba de personalidad MMPI, es posible que el texto dado en el ítem, especificara claramente la respuesta solicitada o que los docentes hayan hecho énfasis en la aplicación de esta prueba. La pregunta más difícil se relacionó con selección de personal; se recomienda por tal razón, dar mayor relevancia a este tema.

Basados en el análisis realizado a las respuestas dadas a los diferentes ítems, se recomendó al programa de psicología que sirvió de base para la elaboración de los ítems profundizar en los siguientes temas: escalas de medición, validez, teoría de inteligencia de Gardner y selección de personal. También se sugirió revisar las diferentes estrategias pedagógicas y didácticas utilizadas por los diferentes docentes, ya que en el caso específico del tema de la teoría de la inteligencia de Gardner que puede no ser de gran dificultad, es posible que no se haya dado una respuesta satisfactoria por parte de los estudiantes, debido a que en esta asignatura, los estudiantes exponen las diversas teorías de inteligencia por subgrupos, de forma que solo algunos manejan en profundidad determinado tema. Es necesario revisar si dentro de este método, el docente profundiza sobre el tema y lo evalúa.

Respecto a las competencias evaluadas por los ítems, el desempeño evidenciado en el análisis psicométrico de la prueba no estableció diferencias entre las diferentes acciones (interpretativa, argumentativa y propositiva), ya que a lo largo de la variable los ítems se distribuyeron equitativamente.

Teniendo en cuenta que la prueba diseñada reunió los requisitos exigidos por la TRI, se sugirió aplicar el instrumento a los estudiantes que en el año 2004 debían presentar los exámenes ECAES, como preparación para un mejor desempeño. Se recomendó además, aplicar el instrumento a otro grupo de estudiantes para corroborar si el nivel de dificultad y facilidad de los temas anteriormente presentados se mantiene.

Referencias

- Aldana, E., Chaparro, L., García, G., Gutiérrez, R., Llinás, R., Palacios, M., Patarroyo, M., Posada E., Restrepo A. & Vasco E. (1996). *Colombia: al filo de la Oportunidad*. Santafé de Bogotá: COLCIENCIAS.
- Avendaño, B. L. & Medellín, E. (2001). Consideraciones sobre las teorías de Medición: un ejemplo de aplicación bajo la Teoría de Respuesta al Ítem (TRI). *ACTA Colombiana de Psicología*. 5: 87-99. Bogotá: Universidad Católica de Colombia.
- Avendaño, B. L. (2003). Teoría de Respuesta al Ítem (TRI): Otra Alternativa para la Medición y Evaluación. *Suma Psicológica*. Volumen 10, # 2. paginas 235 - 245.
- Bogoya, M., Torrado, M., Vinent, M., Jurado, F., Pérez, M., García, G., Acevedo, M., Cárdenas, F., Granés, J. & Sarmiento F. (1999). *Hacia una Cultura de la Evaluación para el siglo XXI*. Bogotá: Universidad Nacional de Colombia.
- Bogoya D., Vinent, M., Restrepo, G., Torrado, M., Jurado, F., Pérez, M., Acevedo, M., García, G., Sarmiento., Cárdenas F., Granés, J. & Díaz, L. (2000). *Competencias y Proyecto Pedagógico*. Bogotá: Universidad Nacional de Colombia.
- Hambledon, R.K. & Swaminathan, H. (1985). *Item Response Theory. Principle and Applications*. Kluwer-Nijhoff Publishing: Boston.
- Martínez, R. (1996). *Psicometría: Teoría de los Test Psicológicos y Educativos*. Madrid: Síntesis.
- Muñiz, J. (1990). *Teoría Clásica de los Test*. Madrid: Pirámide.
- Pardo, C. (2001). El Modelo de Rasch: Una Alternativa para la Evaluación Educativa en Colombia, *Revista Acta Colombiana de Psicología*, 9-21.
- República de Colombia. Ministerio de Educación Nacional. *Decreto 0917 de 2001*. Recuperado en Mayo 14 de 2003 disponible en <http://www.mineduccion.gov.co/>
- República de Colombia. Ministerio de Educación Nacional. *Decreto 1527 de 2002*. Recuperado en Mayo 14 de 2003 disponible en <http://www.mineduccion.gov.co/>
- República de Colombia, Ministerio de Educación Nacional. Decreto Número 1716 de 2001. ICFES. Recuperado Noviembre 14 de 2001 disponible en: <http://www.icfes.gov.co/es/direccion/polcal/decreto1716MEDICINA.doc>
- República de Colombia, Ministerio de Educación Nacional. Decreto Número 2233 de 2001. ICFES. Recuperado Noviembre 14 de 2001 disponible en: <http://www.icfes.gov.co/es/direccion/polcal/decreto2233INGENIERIA MECANICA.doc>
- República de Colombia. Ministerio de Educación Nacional. Decreto 2790 de 1994. Recuperado en Mayo 14 de 2003 disponible en <http://www.mineduccion.gov.co/>
- República de Colombia. Ministerio de Educación Nacional. Ley 115, Ley General de Educación de 1994. Recuperado en Mayo 14 de 2003 disponible en <http://www.mineduccion.gov.co/>
- República de Colombia. Ministerio de Educación Nacional. Ley 30 de 1992. Recuperado en Mayo 14 de 2003 disponible en <http://www.mineduccion.gov.co/>
- Torrado, M. (1998). *De la evaluación de Aptitudes a la evaluación de competencias*. ICFES: Bogotá
- Tristán, A. (2001). *Modelo para el análisis de reactivos objetivos por computadora*. México: Instituto Tecnológico de San Luis Potosí.
- Vargas, N., Scopetta, O., Cristancho, C., Cubides, Y., Flórez, N., Gamboa, C., Montenegro, R., Ortega, M., Pérez, J. & Romero, J. (1995). *Evaluación Sumativa del Currículo de la Facultad de Psicología de la Universidad Católica de Colombia, primera etapa: Construcción del marco general y Proceso de Ajuste y*

Efectividad del Instrumento de Conocimientos para el Ciclo Básico de Formación. Tesis de Grado. Bogotá: Universidad Católica de Colombia.

Autoras

Bertha Lucía Avendaño, María del Pilar Fresneda, Aida Ingrith Martínez & Ingrid Santos.
Universidad Católica de Colombia. Bogotá, Colombia. E-mail: blavendano@hotmail.com