

XIV Jornadas de Investigación y Tercer Encuentro de Investigadores en Psicología del Mercosur. Facultad de Psicología - Universidad de Buenos Aires, Buenos Aires, 2007.

Dificultades en la difusión y el desarrollo de la teoría de respuesta al ítem en Argentina.

Abal, Facundo Juan Pablo, Lozzia, Gabriela, Picón Janeiro, Jimena, Galibert, María Silvia y Aguerri, María Ester.

Cita:

Abal, Facundo Juan Pablo, Lozzia, Gabriela, Picón Janeiro, Jimena, Galibert, María Silvia y Aguerri, María Ester (2007). *Dificultades en la difusión y el desarrollo de la teoría de respuesta al ítem en Argentina. XIV Jornadas de Investigación y Tercer Encuentro de Investigadores en Psicología del Mercosur. Facultad de Psicología - Universidad de Buenos Aires, Buenos Aires.*

Dirección estable: <https://www.aacademica.org/000-073/588>

ARK: <https://n2t.net/ark:/13683/e8Ps/4ue>

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

DIFICULTADES EN LA DIFUSIÓN Y EL DESARROLLO DE LA TEORÍA DE RESPUESTA AL ÍTEM EN ARGENTINA

Abal, Facundo Juan Pablo; Lozzia, Gabriela; Picón Janeiro, Jimena; Galibert, María Silvia; Aguerri, María Ester
Facultad de Psicología, UBACyT, Universidad de Buenos Aires
Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT-PICT)

RESUMEN

A pesar de su uso extendido, la Teoría Clásica de Test (TCT) presenta limitaciones teóricas y prácticas. Un nuevo enfoque psicométrico, la Teoría de Respuesta al Ítem (TRI), ofrece soluciones y está en auge desde hace varias décadas en Estados Unidos y algunos países europeos. Sin embargo, su desarrollo en Argentina es incipiente. En el presente trabajo se sistematizaron los factores que han dificultado la difusión y aplicación de la TRI: 1) El estancamiento producido a nivel histórico por la necesidad de evolución de la informática, 2) La falta de profesionales del campo de la psicología que dominen el fuerte componente matemático de la TRI, 3) Los elevados costos operativos (software, gran tamaño muestral, recursos y tiempo) que conlleva validar una prueba desde esta perspectiva, y 4) Los riesgos que se asumen al pretender ajustar un modelo tan riguroso a datos reales. La TRI profundiza el estudio del funcionamiento de los ítems develando aspectos ocultos para la TCT que pueden dañar la validez y confiabilidad de un instrumento. Se concluye que se debería priorizar la calidad del instrumento por sobre los costos e intentar superar los factores que entorpecen la puesta en práctica de la TRI.

Palabras clave

TRI TCT Psicometría Medición

ABSTRACT

DIFFICULTIES OF DIFFUSION AND APPLICATION OF ITEM RESPONSE THEORY IN ARGENTINE

In spite of its widely spread use, Classical Test Theory (CTT) presents theoretical and practical limitations. A new psychometric approach, Item Response Theory (IRT), offers solutions and it has been on increased for several decades in the United States and some European countries. Nevertheless, its development in Argentina is still incipient. The present work systematizes the factors that have hindered the diffusion and the application of IRT: 1) The stagnation produced in the course of the history by the necessity of more developed computing facilities, 2) The lack of professionals in the psychology field capable to understand deeply the mathematical background of IRT, 3) The high operative costs (software, big sample sizes, resources and time) that are required to validate tests from this perspective, and 4) the risks that have to be taken when trying to fit such a restrictive model to real data. IRT deepens the study of item functioning and so it can reveal some aspects that are hidden for CTT, which may damage the validity and reliability of measurement instruments. Quality should be prioritized over costs and there should be made stronger efforts to overcome the obstacles to put IRT into practice.

Key words

IRT CTT Psychometric Measurement

En sus primeros trabajos Spearman (1904) formuló las bases de un modelo sencillo, robusto y parsimonioso que le permitió fundamentar las puntuaciones de los test y estimar el error de medida inherente a todo proceso de medición (Muñiz, 1994; Barbero-García, 2003). Su modelo lineal de puntuaciones supone que el puntaje observado de un sujeto en un test es el resultado de la suma de un puntaje verdadero más un error de medida. Esta es la base teórica donde se asientan los conceptos de validez y confiabilidad desarrollados durante la primera mitad del siglo XX y que han servido para garantizar la calidad de la medida psicológica. En la actualidad muchos de los tests psicométricos son validados exclusivamente según este modelo, que se conoce en la literatura como la Teoría Clásica de los Tests (TCT).

A pesar de su notable expansión, no han faltado las voces críticas respecto de los supuestos del modelo y los indicadores globales que utiliza (Gómez-Benito & Hidalgo-Montesinos, 2003). Los principales problemas de la TCT ya habían sido referenciados por Thurstone en 1928: la medición de las variables no es independiente del instrumento empleado y las propiedades de los instrumentos dependen de los sujetos utilizados para establecerlas.

Estas limitaciones encontraron soluciones en un nuevo enfoque psicométrico: la Teoría de Respuesta al Ítem (TRI). Muñiz (1997) y Barbero-García (2003) coinciden en ubicar el nacimiento formal de la TRI con la publicación de la tesis doctoral de Lord en 1952. El objetivo sustancial de esta teoría es la construcción de instrumentos de medición con propiedades invariantes entre poblaciones. Por otra parte, permite expresar las propiedades del test en función de la aditividad de las propiedades de los ítems que lo componen, es decir que se independiza de los instrumentos particulares de medición (tests) y toma a cada ítem como una unidad de análisis básica.

Los aportes de este nuevo enfoque han traído aparejadas transformaciones en la forma de diseñar y analizar tests. Asimismo, ha proporcionado soluciones válidas para viejos problemas que presentaba la TCT en: la construcción de banco de ítems, la elaboración de test adaptativos, el análisis del sesgo de los ítems, la equiparación de puntajes obtenidos en distintas formas de test, entre otros (Martínez-Arias, 1999).

La TRI está en auge desde hace varias décadas en Estados Unidos y algunos países europeos. Sin embargo, su difusión, estudio y aplicación en Argentina son muy incipientes. Si la TRI ha demostrado ser una teoría potente que supera las limitaciones teóricas y prácticas de la TCT ¿por qué razón, a más de 50 años de su introducción, ha ganado tan poco terreno en nuestro medio? El objetivo de esta presentación será intentar alcanzar una respuesta para este interrogante.

DESARROLLO

Muchos son los factores que han dificultado la difusión y aplicación de la TRI. La experiencia que han tenido los psicómetras de los países más desarrollados así lo demuestra (Hambleton & Jones, 1993; Muñiz, 1997). A medida que los profesionales se adentran en esta nueva perspectiva y pretenden ponerla en práctica deben enfrentar los siguientes obstáculos.

FACTORES TECNOLÓGICOS

Ya desde sus orígenes la TRI vio entorpecido su desarrollo por las fuertes limitaciones tecnológicas de la época. Debe recordarse que la mayoría de las ideas principales ya se habían desarrollado a nivel teórico a mediados de siglo XX, pero la complejidad de los cálculos hacía extremadamente dificultosa su puesta en práctica.

En la década del '70 comenzaron a surgir los primeros software que permitían resolver rápidamente y de forma fiable los laboriosos cálculos para las estimaciones de los parámetros. Pero sólo con el avance de la informática se vislumbraron las posibilidades de aplicación de los modelos de la TRI. Fue recién en la década del '80, a casi 30 años de su gestación, que el uso masivo de las computadoras personales tornó accesible la TRI para los investigadores de los países desarrollados.

Así como Lord dio origen a la TRI en 1952, su libro *Applications of item response theory to practical testing problems* publicado en 1980 supuso un renacimiento para la teoría. En efecto, para Muñiz (1997) este texto fue un hito en la historia de la TRI dado que marcó el inicio de su expansión.

Si bien los factores tecnológicos parecen actualmente superados debe recordarse que Argentina siempre se encuentra un paso más atrás en la evolución de la informática.

FACTORES ASOCIADOS A LA FORMACIÓN

Sin lugar a duda, el potente componente matemático de la TRI es uno de los factores que más le resta atractivo para los profesionales que buscan formarse en esta perspectiva. Esto ya había sido señalado por Hambleton y Jones (1993) y Álvaro-Page (1993) quienes sostuvieron que las nociones matemáticas ponen en desventaja a la TRI sobre todo frente a un modelo tan simple como el de la TCT. La mayoría de los especialistas no están acostumbrados a trabajar con modelos matemáticos complejos. En contraposición, parecen manejarse con mayor comodidad con los cálculos sencillos que utiliza la TCT y que pueden ser obtenidos fácilmente con los programas estadísticos comunes.

A esto debe sumarse una resistencia natural a incorporar un conjunto de nuevas nociones. Muchos de los psicómetras argentinos se han formado bajo las condiciones que rigen el modelo clásico y han confirmado su utilidad en la práctica. Esta corroboración no es un aspecto menor si se tiene en cuenta que en Argentina la formación universitaria en el área siempre ha estado más centrada en la Evaluación Psicológica que en la Psicometría. Basados en recurrentes experiencias exitosas, algunos profesionales adoptan las soluciones que la TCT ofrece a pesar de sus limitaciones teóricas. La facilidad conceptual del modelo de la TCT puede obstaculizar el aprendizaje de la TRI si no se tiene en claro cuál es el beneficio que reporta.

Otro aspecto importante de la formación abre la pregunta sobre aquello que deben aprender los nuevos profesionales. Más allá de la hegemonía alcanzada de la TRI, en la práctica la TCT aún no ha perdido vigencia a nivel mundial. Quizás sea esto lo que haya motivado a Nunnally y Bernstein (1995) a decir "cualquier estudiante de psicometría serio necesitará conocer los fundamentos de la TRI aunque los procedimientos clásicos serán suficientes en la mayoría en las aplicaciones" (p. 443). Al respecto, Argentina aún se encuentra en un período de transición. Por el momento, se están dando los primeros pasos en la enseñanza de esta teoría de manera paulatina pero auspiciosa. Sólo el tiempo permitirá determinar si la formación de los actuales estudiantes traerá los frutos de un mayor desarrollo de la TRI.

FACTORES OPERATIVOS

Un importante factor que afecta actualmente la expansión de la TRI son los costos operativos que implica realizar la validación de un instrumento desde esta perspectiva. Este motivo es privativo de los países como la Argentina, cuya economía se

encuentra en desventaja frente a las de los países desarrollados.

Los software requeridos obligan a una gran inversión ya que estos programas tienen funciones específicas y no integran los paquetes estadísticos usados comúnmente. De alguna manera, esto reduce la accesibilidad de la TRI a posibles usuarios y, simultáneamente, encarece el costo de los programas porque no tienen un amplio mercado que los consuma (Nunnally & Bernstein, 1995).

Asimismo, para conseguir estimaciones estables de los parámetros, la TRI exige muestras que superen los 500 individuos. Un tamaño muestral tan grande también multiplica los gastos de procedimiento. La cantidad de protocolos y recursos humanos se reducen si se optan por un tamaño de muestra de entre 200 y 500 sujetos como permite la TCT (Muñiz, 1997). Se podrían realizar menos administraciones y se haría más rápidamente la carga de datos. Resulta notable el ahorro de insumos, recursos y tiempo que deben ponerse en la balanza a la hora de decidirse por alguna de las teorías psicométricas.

FACTORES ASOCIADOS AL AJUSTE DEL MODELO

Una vez realizada la inversión económica, es posible que los resultados no acompañen. La aplicación de los modelos de la TRI a la medición de variables psicológicas ha generado cambios en la manera de abordar el problema de la construcción de ítems y de pruebas. Debe recordarse que la TRI está sostenida por fuertes supuestos a los que la mayoría de los datos difícilmente se acomodan. La rigurosidad de estos requisitos pone en riesgo el ajuste del modelo a los datos. Es por esto que, cuando los supuestos de aplicabilidad son rechazados, se puede recurrir a una modelización con la TCT como segunda opción. Como afirmó Martínez-Arias (1995) frente a problemas concretos que imposibiliten alcanzar el modelo de la TRI el constructor sabe que puede aplicar un modelo menos restrictivo.

COMENTARIOS FINALES

Un análisis clásico presenta las limitaciones que surgen de utilizar indicadores globales en la evaluación de la calidad de los ítems. Este análisis resulta útil pero no exhaustivo. Profundizar en el estudio con la TRI del funcionamiento de los ítems puede develar aspectos que permanecen ocultos por la TCT y que incluso pueden dañar a la validez y confiabilidad de un instrumento. También la TRI permite garantizar que la medida de una persona y los parámetros estadísticos de los ítems se mantengan constantes más allá de los cambios en la composición del instrumento y del grupo normativo.

La TRI no tiene como objetivo reemplazar a la TCT. En efecto, se recomienda iniciar el análisis de los ítems con la TCT y después de depurar los reactivos con dicha teoría, aplicar el o los modelos que correspondan de la TRI.

La evaluación psicológica necesita instrumentos válidos y confiables porque de ellos dependen muchas de las decisiones que se toman en los ámbitos de aplicación de los test. Por lo tanto se debería priorizar la calidad del instrumento por sobre los costos y los factores que entorpecen la puesta en práctica de la TRI. Hoy más que nunca, se manifiesta de gran utilidad para los profesionales del área estar atentos a los avances que la psicometría ha dado.

BIBLIOGRAFÍA

- ÁLVARO-PAGE, M. (1993). *Elementos de Psicometría*. Madrid: Eudema.
- BARBERO-GARCÍA, M.I. (2003). *Introducción a la Psicometría*. En M. I. Barbero-García; E. Vila-Abad & J.C. Suárez-Falcón (Eds.), *Psicometría*. (pp. 23-75) Madrid: Universidad Nacional de Educación a Distancia.
- GÓMEZ-BENITO, J. & HIDALGO-MONTESINOS, M.D. (2003). *Desarrollos recientes en Psicometría*. *Avances en Medición*, 1, 17-36.

- HAMBLETON, R.K. & JONES, R.W. (1993). Comparison of >Lord, F. M. (1952) A theory of test scores. Psychometric Monograph N°7. Iowa City, IA: Psychometric Society.
- LORD, F.M. (1980). Applications of item response theory to practical testing problems. Hillsdale, N. J.: Lawrence Erlbaum.
- MARTÍNEZ-ARIAS, M.R. (1995). Psicometría: Teoría de los Tests Psicológicos y Educativos. Madrid: Síntesis.
- MARTÍNEZ-ARIAS, M.R. (1999). Los avances de la psicometría y la construcción de test. En F. Silva (Ed.), Avances en Evaluación Psicológica. Valencia: Promolibro.
- MUÑIZ, J. (1994). Teoría Clásica de Test. Madrid: Pirámide.
- MUÑIZ, J. (1997). Introducción a la teoría de respuesta a los ítems. Madrid: Ediciones Pirámide.
- NUNNALLY, J.C. & BERNSTEIN, I.J. (1995). Teoría Psicométrica. Nueva York: McGraw-Hill.
- SPEARMAN, C.E. (1904). General Intelligence objectively determined and measured. American Journal of Psychology, 15, 201 -293.
- THURSTONE, L.L. (1928). Attitudes can be measured? American Journal of Sociology, 33, 529-554.