



# INEVA en acción

Boletín informativo

Volumen 1, Número 2, 2005

## INTRODUCCIÓN

Este boletín informativo del Programa de **IN**vestigación y **EVA**luación Educativa (conocido como **INEVA**) del Departamento de Estudios Graduados de la Facultad de Educación en la Universidad de Puerto Rico, Recinto de Río Piedras tiene como propósito principal exponer comentarios, noticias, reseñas y síntesis de trabajos relacionados con las áreas de investigación, estadística, evaluación y medición educativa. Además, es un medio para establecer vínculos de comunicación e intercambio con otras personas interesadas, en y fuera de Puerto Rico.

Este segundo número incluye un artículo breve acerca del tema de *Test Equating*, las reseñas de dos organizaciones vinculadas con la evaluación, enlaces relacionados con la investigación y evaluación educativa, dos noticias acerca de presentaciones y publicaciones de la facultad del Programa y de la publicación reciente del *Handbook of Qualitative Research*.

## TEST EQUATING

El conjunto de procedimientos estadísticos conocidos como *Test Equating* (equivalencia o equiparación de las puntuaciones de dos o más pruebas) se encuentra en la literatura relacionada al campo de la psicometría desde el

1950. Sin embargo, su importancia comenzó a ser reconocida a principios de la década de los 80. Este giro en importancia se puede atribuir a tres factores: (1) el aumento de pruebas estandarizadas que utilizan más de una versión o formulario, (2) la atención de las críticas a las pruebas estandarizadas y (3) el aumento en importancia del movimiento de *accountability* (rendir cuentas) y *fairness in testing* (justicia en la administración de pruebas y el informe de resultados). Estos procedimientos son utilizados para establecer equivalencia entre puntuaciones obtenidas cuando se administran diferentes formularios de una prueba, los cuales han sido construidos siguiendo especificaciones estadísticas preestablecidas y cuyos ítems provienen de la misma área de contenido. Cuando los formularios cumplen con estos dos requisitos se dice que son paralelos.

Las personas encargadas de desarrollar pruebas estandarizadas crean diferentes formularios de una misma prueba para evitar que las personas tengan conocimiento sobre los ítems antes de ser administradas (efecto de difusión de tratamiento). Otro propósito es evitar el efecto de la práctica, ya que en muchas ocasiones un individuo puede tomar una prueba más de una vez. Sin embargo, al crear una prueba, en esencia, se está seleccionando una muestra de ítems de un área de contenido. Al crear diferentes formularios, las pruebas resultantes poseen diferentes ítems por lo que

inevitablemente puede variar el nivel de dificultad de un formulario a otro.

En una situación ideal, en la cual se construyan formularios paralelos, un individuo obtendría aproximadamente la misma puntuación si le fuese administrado uno u otro formulario, o sea, serían intercambiables. Lamentablemente, éste no es el caso ya que siempre estará presente en la puntuación el efecto del nivel de dificultad de los ítems que incluye el formulario administrado. Mediante el uso de los procedimientos de *Test Equating* se establecen equivalencias y se informa una puntuación estándar.

Existen dos clasificaciones dentro de los procedimientos de *Test Equating*. En primer lugar, están aquellos que surgen de la teoría clásica de medición donde la puntuación total de una persona se visualiza como la suma de la “puntuación real” y una constante relacionada al error. En segundo lugar, se encuentra la visión contemporánea de *Test Equating* donde se trata de separar la habilidad del individuo del nivel de dificultad de cada ítem. Esta última corresponde a la Teoría de Respuesta del Ítem (*IRT* por sus siglas del inglés, *Item Response Theory*).

Entre los procedimientos de la teoría clásica de medición se encuentran: *Mean Equating*, *Linear Equating* y *Equipercentile Equating*. *Mean Equating* asume que las distribuciones difieren sólo en promedio. *Linear Equating* asume que las puntuaciones obtenidas difieren tanto en promedio como en desviación estándar. *Equipercentile Equating* se utiliza cuando las distribuciones, además de diferir en promedio y desviación estándar, difieren en la forma de la distribución.

Los procedimientos de *IRT* difieren de aquellos basados en los métodos clásicos pues producen parámetros en función de las puntuaciones en los ítems en vez de en la puntuación total. Entre los procedimientos de *IRT* más usados se

encuentran el de un parámetro (conocido como el Modelo de *Rasch*) y el de tres parámetros.

La forma en que se construye y se asignan puntuaciones en una prueba influye en la confiabilidad de los resultados, la validez de las inferencias y la precisión con la que se pueden establecer equivalencias. Ninguna técnica estadística substituye la necesidad de crear y seleccionar ítems y administrar una prueba cuidadosamente. Si la prueba se desarrolla con cuidado y los formularios tienen ítems que provienen del mismo contenido, el procedimiento de *Test Equating* se puede utilizar para establecer equivalencias cuando existen diferentes niveles de dificultad. Cuando este procedimiento se realiza apropiadamente nos permite medir crecimiento, establecer tendencias y combinar datos de diferentes formularios de una misma prueba.

Para información específica acerca de estos procedimientos refiérase al libro de Kolen y Brennan (1995) titulado *Test Equating: Methods and Practices*.

Dr. Víctor E. Bonilla Rodríguez

## PÁGINAS ELECTRÓNICAS

***African Evaluation Association***  
(<http://www.afrea.org/home/index.cfm>)

La Asociación Africana de Evaluación es una organización sombrilla para las asociaciones nacionales de evaluación y las redes en África. La página está dirigida a profesionales del área de evaluación. La misma tiene como propósito promover, fortalecer y apoyar la evaluación en África mediante el intercambio de información y de los avances en la evaluación.

Esta página presenta un documento que establece las pautas para la evaluación en África. Entre los recursos disponibles se incluye información y enlaces a revistas

profesionales reconocidas del área de evaluación, y documentos e información que son producto de las conferencias y actividades de esta asociación. Entre estos documentos se incluye un manual sobre el enfoque utilitario de Michael Quinn Patton en África.

Del mismo modo, provee materiales para adiestramiento en evaluación y modelos o enfoques para la evaluación. Además, ofrece una lista de cotejo para ayudar a planificar evaluaciones, negociar contratos, revisar el progreso y asegurar que la evaluación se complete adecuadamente. Se puede acceder también a bibliografías anotadas sobre evaluaciones realizadas en África y a los *African Evaluation Guidelines*. Esta asociación pone a disposición de otras organizaciones los estándares para la evaluación de programas en África, los cuales fueron creados utilizando como base los estándares del *National Joint Committee on Standards for Educational Evaluation*.

Connie F. Walker Egea

**United Kingdom Evaluation Society**  
(<http://www.evaluation.org.uk/>)

La *United Kingdom Evaluation Society* (UKES) se fundó en 1994 con el propósito de promover la teoría, práctica, entendimiento y utilización de la evaluación y su contribución al conocimiento público. Además, provee información acerca de las actividades de la UKES, noticias de actividades y trabajos de evaluación, contratos, eventos y recursos de la UKES y de otras organizaciones internacionales.

Esta página está dirigida a profesionales de la evaluación, practicantes y delegados de evaluación del gobierno nacional y local, comunidad investigativa y consultores independientes. Se incluye información adecuada para el campo de la evaluación, enlaces de interés detallados y recientes (Tabla

de “Last Updated” en la página principal). Ayuda tanto al evaluador como a todo aquel que necesite la información y sea ajeno al campo.

Entre los recursos que posee la página se encuentran: un glosario de conceptos y una bibliografía de textos y documentos relacionados al campo de la evaluación. Además, contiene las guías que especifican los elementos que se deben considerar en una evaluación por parte de los evaluadores, participantes y los delegados de evaluación, entre otros. Estas guías son propias de la UKES y de otras organizaciones las cuales se pueden acceder a través de enlaces disponibles.

Juan P. Vázquez Pérez

## RECURSOS EN LA RED

**Research Randomizer**  
([www.randomizer.org/index.htm](http://www.randomizer.org/index.htm))

Posee un programado que permite la selección y la asignación aleatoria de sujetos, además de asignar participantes a condiciones experimentales.

**Practical Assessment, Research and Evaluation**  
(<http://PAREonline.net>)

Revista electrónica dirigida a los profesionales en la educación. Provee acceso gratuito a artículos sobre problemas y prácticas en la educación.

**The Qualitative Report**  
(<http://www.nova.edu/ssss/QR/index.html>)

Revista electrónica dirigida a la investigación cualitativa. Contiene artículos en texto completo publicados desde 1990 y enlaces a otros recursos. Es publicada por la Nova Southeastern University.

## PRESENTACIONES DE LA FACULTAD

La Dra. María del R. Medina y la Dra. Ada L. Verdejo presentarán una mesa redonda con el título de *Teaching Program Evaluation Across Countries: Communalities and Differences* en la Conferencia Anual de la *American Evaluation Association* y de la *Canadian Evaluation Society*. Cada 10 años estas dos organizaciones y otras organizaciones internacionales dedicadas a la evaluación de programas se unen para llevar a cabo este evento. La misma se llevará a cabo en Toronto, Canada del 24 al 30 de octubre de 2005. Puede encontrar información adicional acerca de este evento en la página electrónica: <http://www.eval.org/eval2005>

## PUBLICACIONES DE LA FACULTAD

La más reciente edición del *Mental Measurement Yearbook* (Spies & Plake, 2005) incluye dos críticas de pruebas escritas por la Dra. María del R. Medina: *New Standards Reference Examinations: Mathematics* y *Receptive One-Word Picture Vocabulary Test-Spanish-Bilingual Edition*. El *Mental Measurement Yearbook* es la publicación más antigua y prestigiosa dedicada a la crítica de pruebas en el mundo. Fue creada en 1938 por Oscar Buros en el instituto que lleva su nombre en la Universidad de Nebraska-Lincoln. Las personas que critican las pruebas para esta publicación se seleccionan de manera cuidadosa de un grupo destacado de profesionales en diversos campos. La versión electrónica de las críticas se puede acceder en la página electrónica <http://buros.unl.edu>

## PUBLICACIONES NUEVAS

La tercera edición del *Handbook of Qualitative Research* se publicó este año. Como en las pasadas ediciones Norman K. Denzin e Yvonna S. Lincoln son los editores. Esta edición incluye un total de 44 capítulos, de los cuales 16 son de temas distintos a los de la segunda edición. Además, incorpora nuevos autores y coautores que representan, precisamente, una muestra de la gama de expresiones e interpretaciones que alberga el campo de la investigación cualitativa. Como sus ediciones anteriores, recomendamos su lectura para continuar con la discusión y reflexión sobre los asuntos teóricos, éticos, metodológicos y políticos de la investigación cualitativa.

Las opiniones vertidas en esta publicación son de los (las) autores(as) y no representan las del Programa de INEVA. Las personas interesadas en escribir para esta publicación deben comunicarse con la Junta Editora a nuestra dirección electrónica ([ineva@uprrp.edu](mailto:ineva@uprrp.edu)). De la misma manera pueden enviarnos sus comentarios y sugerencias acerca de esta publicación.

## JUNTA EDITORA

Dr. Víctor E. Bonilla Rodríguez  
Madelyn E. Cintrón Rodríguez  
Chamary Fuentes Vergara  
Dra. María del R. Medina Díaz  
Rosa L. Román Oyola  
Juan P. Vázquez Pérez  
Connie F. Walker Egea

Conceptuación Gráfica  
Dr. Víctor E. Bonilla Rodríguez