



Métodos de equiparación en pruebas SABER 11^o: Comparación de procedimientos en personas videntes e invidentes

Aura Nidia Herrera Rojas

Investigador Principal

Catheryne Lancheros, MSc (c)

Javier Jiménez, MSc (c)

Autores ponencia

Universidad Nacional de Colombia

Proyecto financiado por ICFES-COLCIENCIAS CONV. 518-2010

Universidad Nacional de Colombia – DIB Sede Bogotá



Prosperidad para todos

Antecedentes



- En pruebas de aplicación masiva el desempeño de los invidentes es significativamente inferior.
- Se aplica la misma prueba para videntes e invidentes pero por canales sensoriales diferentes.

(Rojas y cols., Inédito)

EQUIPARACIÓN DE PUNTUACIONES

Es un proceso estadístico usado con el fin ajustar las puntuaciones de las formas de un test para que los puntajes en cada una de las formas puedan ser intercambiables.

(Kolen & Brennan, 2004)

Proceso



Selección de un Diseño



Recolección de datos



Selección de un Método



Establecimiento ecuación de equivalencia o tabla de conversión

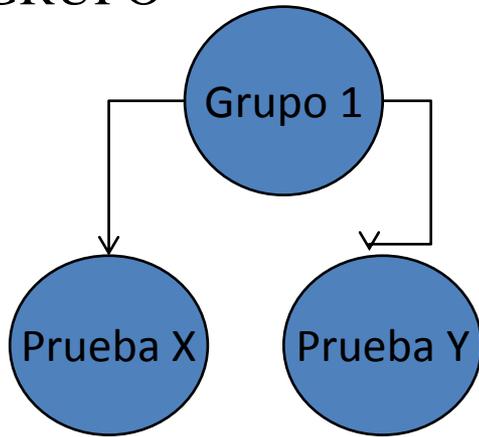


Evaluación de la equiparación

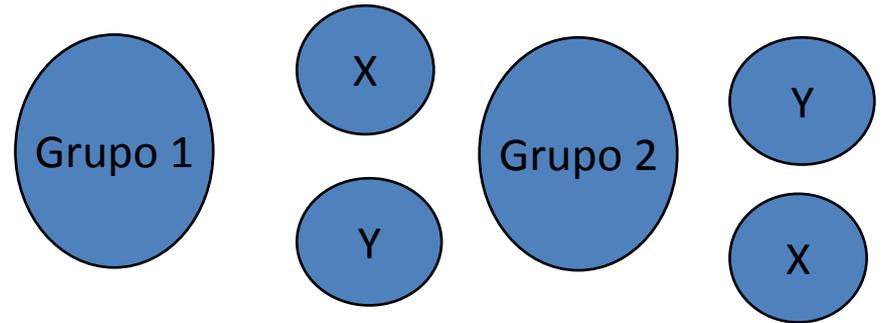
Diseños



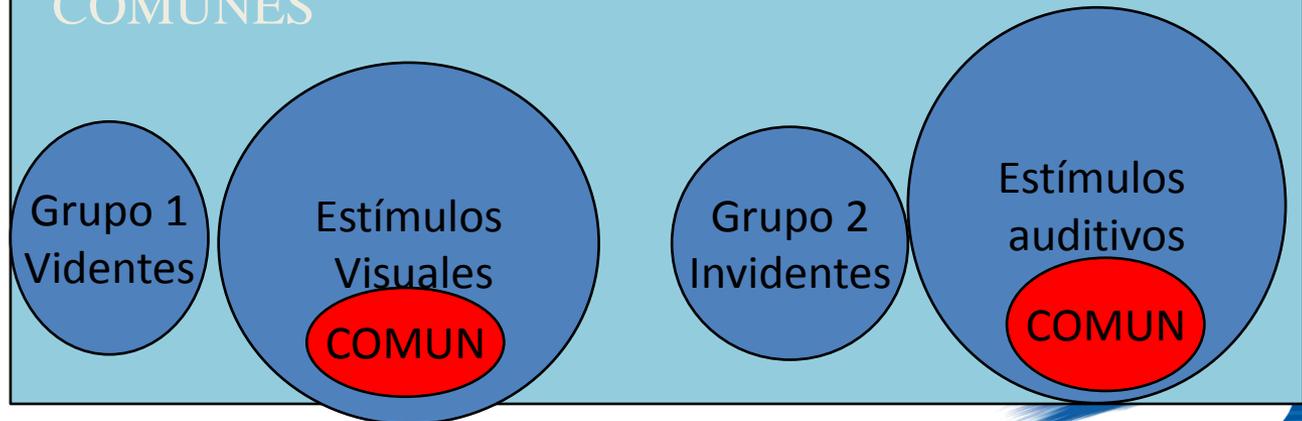
DISEÑO DE UN SOLO GRUPO



DOS GRUPOS EQUIVALENTES AL AZAR



DOS GRUPOS NO EQUIVALENTES CON ITEMS COMUNES



Métodos



Métodos de Equiparación

IRT

TCT

Lineal

Equipercantil

Métodos basados en los momentos

Métodos basados en la CCT

Otros

Estimación de frecuencias

Tucker
Levine

Media /
Sigma



Pregunta de Investigación

¿Cuál método de equiparación permitirá estimar la equivalencia de puntuaciones con menor error, en pruebas de aplicación masiva para población vidente e invidente?

Método



Bases de Datos

- Aplicaciones del examen de estado SABER 11 en los años 2008 -2, diferenciada por población vidente e invidente.
- 237.102 Videntes
- 87 Invidentes

Diseño

- Cuantitativo
- Experimental
- Simulación (Condiciones estables, Variables dependientes e independientes)



Instrumento

- Subprueba de *Lenguaje* SABER 11 aplicación 2008 II.
- 36 ítems (de selección múltiple con única respuesta) los cuales conformaban dos formas de prueba cada una de 24 ítems.

Condiciones Estables

- **Tamaño de la prueba:** 24 ítems.
- **Razón del tamaño de la muestra:** Ajustada a 1:25
Invidentes (87) - Videntes (2175)

Variables Independientes

Métodos de equiparación



Modelo TCT

- Método lineal de Tucker
- Método lineal de Levine

Modelo TRI

- Métodos Media/Sigma

Variables Dependientes

- Error de equiparación para cada método
- Coeficiente de Variación

Procedimiento



Se selecciona el **Diseño de ítems comunes y grupos no equivalentes** ó **diseño de anclaje**



Se eligieron cinco ítems de anclaje distribuidos a lo largo del continuo de dificultad y correspondiendo al 20% de la prueba.

Simulación de Datos



Parámetros de ítems y personas modelo de Rasch (Winsteps)

Vector de Dificultad

Vector de habilidad

1000 pares de matrices de probabilidad de respuesta utilizando distribución Bernoulli.

Procedimiento



- Se utilizó el programa estadístico R
- Desarrollo de métodos de equiparación a cada matriz de métodos de probabilidad Tucker, Levine y Media/Sigma.
- Obtención tablas de equivalencias de las puntuaciones de los invidentes con respecto a los videntes.
- Error de equiparación por cada puntuación equiparada utilizando el método bootstrap.
- Promedio de los errores de equiparación para cada matriz simulada.
- El coeficiente de varianza de cada método de equiparación a través de las replicas.
- Comparación de la estabilidad de los métodos

RESULTADOS

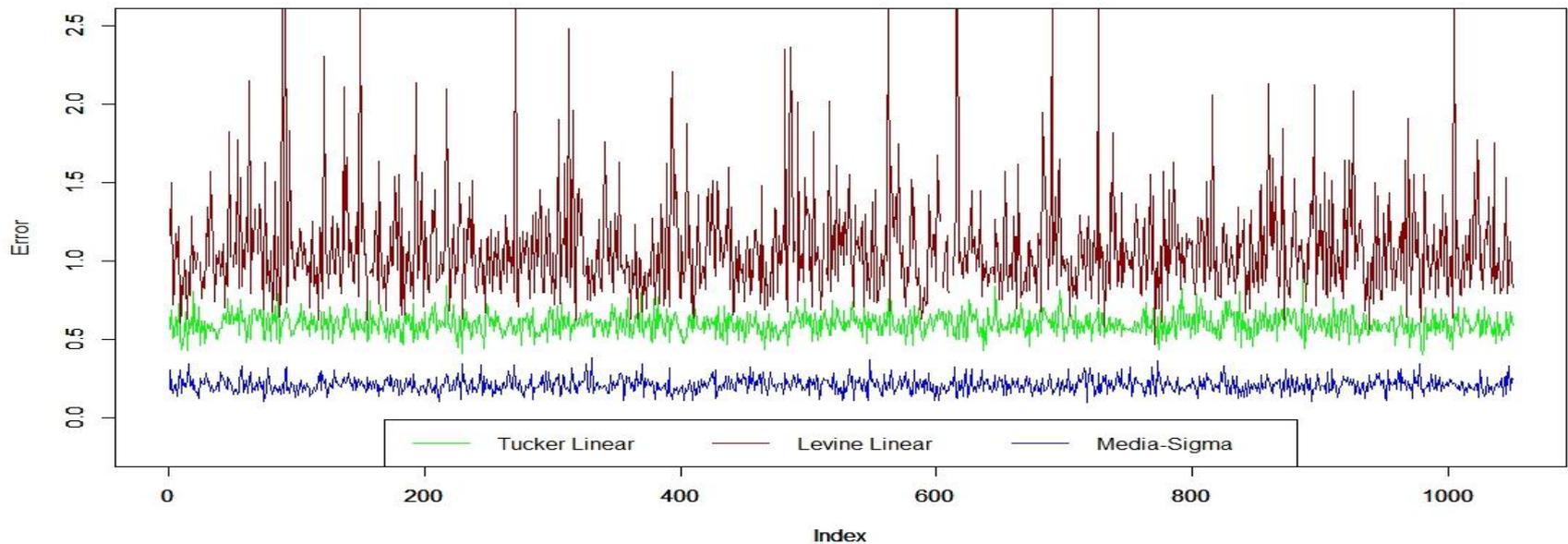


Estadísticos descriptivos de los errores de equiparación para cada método.

Método	Mín	Máx	Media	Desv. típica	Coefficiente variación
Tucker	0,40	0,89	0,60	0,07	11,9 %
Levine	0,47	3,71	1,08	0,34	31,82%
Media/Sigma	0,10	0,38	0,21	0,04	19,04%

Magnitud de error de equiparación para cada método de equiparación en las réplicas

Cadenas de los Errores para cada Método



CONCLUSIONES



- Los procedimientos de equiparación son una herramienta viable para obtener puntuaciones intercambiables entre la población invidente y la vidente cuando se mide el mismo constructo.
- Estos procedimientos favorecen procesos de evaluación equitativos, al contribuir en interpretaciones validas sobre el nivel de habilidad de una población determinada.
- El método Media/Sigma aporta un menor error de equiparación en cada matriz simulada.
- El método lineal de Tucker presenta una adecuada estabilidad a lo largo de las réplicas.

CONCLUSIONES



- Es necesario elaborar pruebas objetivas que permitan evaluar el constructo medido a lo largo de la dimensión de la habilidad con el fin de tener información en sus diferentes niveles.
- Estos estudios ponen de relevancia la importancia de generar mayor conocimiento sobre la calidad de la evaluación dirigida a población minoritaria.

GRACIAS



Proyecto de Investigación CONV COLC/UNAL RC 255-2011: **Procedimiento para establecer equivalencia en las puntuaciones de pruebas de aplicación masiva en personas con y sin limitación visual.** Fondo Nacional de financiamiento para la Ciencia, la Tecnología y la Innovación Francisco José de Caldas y la Universidad Nacional de Colombia. Convocatoria 518. Convocatoria Nacional Colciencias - Icfes para la Conformación del Banco de Proyectos de Investigación en el Área de la Calidad de la Educación año 2010.

Convocatoria Nacional de Investigación y de Creación Artística de la Universidad Nacional de Colombia - DIB. Modalidad 1. Fortalecimiento a grupos de investigación y creación artística con proyección internacional: **Hacia una integración entre estadística y Psicología generación del nuevo conocimiento y proyección social**

Grupo de Métodos e Instrumentos para la Investigación en Ciencias del Comportamiento UNIVERSIDAD NACIONAL DE COLOMBIA



Prosperidad para todos