



UNIVERSIDAD NACIONAL DE COLOMBIA

**MÉTODOS DE EQUIPARACIÓN DE PUNTUACIONES: LOS
EXÁMENES DE ESTADO EN POBLACIÓN CON Y SIN
LIMITACIÓN VISUAL**

LADY CATHERYNE LANCHEROS FLORIAN

Universidad Nacional de Colombia

Facultad de Ciencias Humanas

Departamento Psicología

Bogotá D.C., Colombia

2013



UNIVERSIDAD NACIONAL DE COLOMBIA

**MÉTODOS DE EQUIPARACIÓN DE PUNTUACIONES: LOS EXÁMENES DE
ESTADO EN POBLACIÓN CON Y SIN LIMITACIÓN VISUAL**

LADY CATHERYNE LANCHEROS FLORIAN

Tesis de grado para optar al título de:

Magister en Psicología

Director (a):

Ph.D., Aura Nidia Herrera Rojas

Línea de Investigación:

Métodos e Instrumentos para la investigación en Ciencias del Comportamiento

Universidad Nacional de Colombia

Facultad de Ciencias Humanas

Departamento Psicología

Bogotá D.C., Colombia

2013



UNIVERSIDAD NACIONAL DE COLOMBIA

Agradecimientos

Al culminar esta etapa de mi vida que me permitió vivir diferentes experiencias de aprendizaje, profesionales y personales quiero agradecer a las personas que participaron en este proceso y contribuyeron a la finalización del mismo en los diferentes momentos del camino recorrido...

Agradezco a mi madre Clementina Florián, a mi padre Luis Lancheros y a mi hermana Marcela Lancheros, quienes han sido mi motor y mi fuerza en este camino. Ustedes son la razón y la satisfacción de cada uno de los logros que he alcanzado en mi vida.

A la docente Aura Nidia Herrera Rojas quien confió en mí esta tarea y con sus orientaciones, palabras, sugerencias, sabiduría e infinita paciencia me apoyó, acompañó e impulsó para seguir adelante. En este recorrido me permitió descubrir que los límites los impone uno mismo y que el conocimiento es el pasaporte que permite descubrir otras realidades.

A las docentes Olga Rosalba Rodríguez Jiménez y Erika Margarita Arias Patiño quienes me guiaron por este camino y me generaron un interés especial por ésta área de la Psicología, así como al docente José Muñiz por acogerme durante mi pasantía en España y brindarme grandes oportunidades de conocimiento y crecimiento.

A mis amigas de siempre Mónica Manrique, Laura Marconi, Natalia DiazdelCastillo y Magda Mendivelso quienes en momentos de dificultad me brindaron una ayuda, consejo o apoyo para continuar. Y a sí mismo a mis compañeros que se hicieron amigos Jenny Cardenas, Javier Jiménez y Javier Villalba quienes además de sus orientaciones profesionales, me brindaron su escucha y acompañamiento durante el proceso.

A las entidades COLCIENCIAS e ICFES por el apoyo económico y asistencial brindado mediante la Convocatoria 518 de 2010.

“Uno puede devolver un préstamo de oro, pero esta en deuda de por vida con aquellos que son amables” (Proverbio).

TABLA DE CONTENIDOS

TABLA DE CONTENIDOS	4
LISTA DE TABLAS	6
LISTA DE FIGURAS	7
RESUMEN	9
ABSTRACT.....	10
INTRODUCCIÓN.....	11
OBJETIVOS	15
REVISION BIBLIOGRAFICA	17
Las personas con limitación visual	17
Equiparación de Puntuaciones.....	22
Diseños de equiparación	28
De un solo grupo	28
De grupos equivalentes	28
De grupos no equivalentes con ítems comunes o diseño de anclaje	29
Métodos de Equiparación.....	33
Métodos basados en la TCT	34

Métodos basados en la Teoría de Respuesta al Ítem	46
Error de equiparación.....	52
MÉTODO	59
Diseño	59
Instrumento.....	61
Procedimiento.....	62
Diseño de equiparación.	62
RESULTADOS.....	65
DISCUSIÓN	78
REFERENCIAS	85
ANEXO A.....	91



UNIVERSIDAD NACIONAL DE COLOMBIA

LISTA DE TABLAS

Tabla 1. Tabla de equivalencia de los Métodos de la Teoría Clásica de los Test, Base de Datos 2008 -2.....	65
Tabla 2. Dificultad real y estimada para el método media / sigma.	68
Tabla 3. Estadísticos descriptivos de los errores de equiparación para cada uno de los métodos.....	70



UNIVERSIDAD NACIONAL DE COLOMBIA

LISTA DE FIGURAS

Figura 1. Aplicación en la investigación del Diseño de anclaje para grupos no equivalentes. .	59
Figura 2. Gráfica de comparación del Error estándar de Equiparación para los Métodos Tucker y Levine.....	62
Figura 3. Magnitud de error de equiparación para cada método de equiparación en las réplicas.....	63
Figura 4. Distribución de frecuencia en matriz de respuesta que arrojó el mínimo de error de equiparación en el Método de Tucker.....	66
Figura 5. Distribución de frecuencia en matriz de respuesta que arrojó el mínimo de error de equiparación en el Método de Levine.....	67
Figura 6. Distribución de frecuencia en matriz de respuesta que arrojó el mínimo de error de equiparación en el Método de Media/Sigma.....	67
Figura 7. Distribución de frecuencia en matriz de respuesta que arrojó el máximo de error de equiparación en el Método de Tucker.....	68

Figura 8. Distribución de frecuencia en matriz de respuesta que arrojó el máximo de error de equiparación en el Método de Levine.....69

Figura 9. Distribución de frecuencia en matriz de respuesta que arrojó el máximo de error de equiparación en el Método de Media/Sigma.....69

Figura 10. Distribución del error máximo de equiparación en toda la simulación con respecto al puntaje.....71

RESUMEN

El propósito de la equiparación consiste poner en una misma escala los resultados de una o más pruebas que evalúan el mismo constructo y que cuentan con especificaciones técnicas similares (Angof, 1984). La importancia de este procedimiento radica en el uso de métodos estadísticos aplicados a las puntuaciones obtenidas en pruebas que evalúan atributos psicológicos, con el fin de garantizar la equidad en la evaluación en condiciones ajenas a las características que se pretenden medir, tal como ocurre con la población invidente, minimizando la posibilidad de beneficiar o perjudicar a un grupo particular en la interpretación que se haga de los resultados. En este trabajo se comparan tres métodos de equiparación dos de ellos basados en la Teoría Clásica de los Test: Tucker (Gulliksen, 1950)) y Levine (1955) y uno basado en la Teoría de Respuesta al Ítem: Media/Sigma de Cohen & Kim (1998). Con base en la aplicación de las pruebas de estado en Colombia SABER 11 año 2008 y utilizando el programa R, se simularon 1000 matrices de respuesta iguales a las aplicaciones de SABER 11 en longitud de la prueba, tamaños de muestra y estimaciones de los parámetros de habilidad de los examinados y parámetros de dificultad de los ítems tanto para población vidente como invidente. Usando los procedimientos de equiparación ya mencionados, para cada matriz de respuesta se construyó la tabla de conversión, se estimó el error de equiparación por medio del método bootstrap y finalmente se obtuvo el coeficiente de variación de los errores a lo largo de las simulaciones. Se encontró que el método de Media/Sigma introduce menor error de equiparación en cada matriz simulada y el método de Tucker más estable a través de las réplicas. Finalmente se realizan conclusiones en torno al uso de estos procedimientos para apoyar la equidad en la interpretación de resultados en poblaciones minoritarias como las caracterizadas por limitación visual.

Palabras Clave: Equiparación, Limitación visual, Métodos Tucker, Levine y Media/ Sigma



UNIVERSIDAD NACIONAL DE COLOMBIA

TEST EQUATING METHODS: STATE EXAMINATIONS IN POPULATIONS WITH AND WITHOUT VISUAL IMPAIRMENT

ABSTRACT

Equating is defined as the process of expressing on the same scale the results from one or more tests that measure the same construct and have similar specifications (Angof, 1984). By using statistical methods applied to the tests scores, the equating procedures pretend to ensure fairness in the evaluation, minimizing the possibility of benefit a particular group in the interpretation of the results. Three methods were compared, two of them based on Classical Test Theory: Tucker (Gulliksen, 1950)) y Levine (1955) and one based in the Item Response Theory: Media/Sigma de Cohen & Kim (1998). A simulation study was carried out based on the application of National Tests in Colombia and using the R program. 1000 response matrices were generated like the length of the test, the ability parameters of the examinees and parameters of item difficulty for both populations: with and without visual impairment. For each response matrix, a conversion table was constructed, the equating error was estimated by bootstrap method for each equating procedure and finally the coefficient of variation of the errors was calculated. The results suggest that Media/Sigma procedure shows less equating error and the Tucker procedure was more stable. Conclusions are made about the use of these procedures to support equity in the interpretation of tests results in minority populations like the visual impairment population.

Key Words: equating, visual impairment population, Tucker, Levine, Mean / Sigma Methods

INTRODUCCIÓN

La evaluación con pruebas objetivas y estandarizadas incrementa día tras día su uso y desarrollo, debido a la imparcialidad que la caracteriza en situaciones que requieren la clasificación y selección de personas en entornos educativos y laborales.

En estos escenarios, los instrumentos de evaluación adquieren un papel fundamental, debido a que las decisiones soportadas en sus resultados generan un impacto importante en el proyecto de vida de las personas. Los profesionales en Psicometría dirigen sus acciones a la búsqueda de mejoras que permitan obtener resultados válidos y confiables en la medición de los atributos psicológicos, ya que de estas mediciones, como las realizadas en el aula, el examen de estado, el de ingreso a la universidad o el concurso de méritos, depende la validez de la evidencia recolectada en un proceso de toma de decisiones que afectará positiva o negativamente a los examinados. Para obtener resultados válidos y confiables es necesario garantizar que los instrumentos utilizados, además de tener adecuadas propiedades psicométricas, permitan que las puntuaciones obtenidas sean comparables en términos de la magnitud de la habilidad de las personas, ya se trate de resultados de dos pruebas que evalúen lo mismo, de puntuaciones obtenidas en diferentes momentos a partir de un mismo instrumento o de un mismo instrumento aplicado en diferentes poblaciones.

Comparar la magnitud de la habilidad de las personas es fundamental en exámenes de aplicación masiva, en los cuales se preparan diferentes versiones de la prueba, que permitirán garantizar su seguridad o evaluar en varias oportunidades a un sujeto o a un grupo con el fin de observar logro académico o cambio en algún atributo (Navas, 2000). Se espera que además

de la rigurosidad con que son construidas las pruebas y de la calidad de las mismas, se enfatice en la importancia de la generación de puntuaciones susceptibles a la interpretación en términos de la habilidad del evaluado; en otras palabras, que los resultados obtenidos sean intercambiables en diferentes poblaciones, estén libres de sesgo y que permitan garantizar una medida precisa del atributo, independientemente de las variables propias del sujeto o de la institución.

En Colombia, el Instituto Colombiano para la Evaluación de la Educación (ICFES) es el organismo oficial que tiene como misión ofrecer el servicio de evaluación de la educación en todos sus niveles, y adelantar investigación sobre los factores que inciden en la calidad educativa, con la finalidad de ofrecer información para mejorarla. Una de estas evaluaciones es la que se realiza a los estudiantes que culminan su proceso de formación secundaria. Estas pruebas estandarizadas (SABER 11), se han aplicado año tras año desde 1980, constituyéndose en un requisito para el ingreso a la educación superior en la actualidad (Ley 30, 1992).

De acuerdo con el Decreto 869 de marzo de 2010, este examen tiene como objetivo principal comprobar el grado de desarrollo de las competencias en los estudiantes. La información obtenida de esta evaluación externa cumple diferentes propósitos en cada uno de los niveles del sector educativo: para los estudiantes se convierte en una autoevaluación que brinda herramientas que permite planear su proyecto de vida; para los establecimientos educativos es un diagnóstico que les permite revisar procesos pedagógicos y retroalimentar la calidad de sus prácticas; para las instituciones de educación superior es una herramienta que permite seleccionar los mejores candidatos para cada programa con base en su desempeño, y;

para las autoridades educativas tanto locales como nacionales es un insumo en la construcción de indicadores que permiten reflexionar sobre la calidad de la educación y generar políticas para su mejora continua.

El Examen de Estado de la educación media SABER 11 se aplica en poblaciones con diferentes características, entre las que se encuentran las personas con algún tipo de limitación ya sea física, motora, auditiva o visual, lo cual plantea la necesidad de reflexionar sobre la forma de evaluar y la calificar estos exámenes, que sirven como herramienta en la toma de decisiones educativas, determinantes para el futuro de diversos grupos poblacionales.

El ICFES realiza procedimientos de equiparación de puntuaciones de pruebas de aplicaciones diferentes del examen SABER 11, los cuales se enmarcan en un diseño de grupos no equivalentes con preguntas comunes y con estimaciones a partir de los modelos TRI, en particular bajo el modelo de Rasch empleando los momentos de las estimaciones de los parámetros. Sin embargo, no se realiza equiparación de puntuaciones entre grupos dentro de las aplicaciones (V. Cervantes, comunicación personal, 10 de Agosto, 2012).

Uno de los derechos fundamentales contemplados en la Constitución Política de Colombia (1991), es el derecho a la educación y específicamente en el artículo 68 hace referencia a la educación y la erradicación del analfabetismo en las personas con limitaciones físicas o mentales. De igual forma en la Ley 115 de 1994 y en el Artículo 1º de la Ley 361 de 1997, se señala que nadie podrá ser discriminado para el acceso a la educación y que el Estado Colombiano lo garantizará, así como la capacitación en los diferentes niveles educativos a las personas con limitación, contando con los recursos necesarios para su formación integral.

Adicionalmente, el Decreto 2082 de 1996, en su Artículo 9, indica que tanto el Servicio Nacional de Pruebas como las instituciones que realizan pruebas de validación deben tomar las medidas necesarias, los apoyos y los recursos para que las personas con limitaciones puedan presentar las pruebas considerando sus códigos, lenguaje y necesidades específicas.

Aunque estos derechos fundamentales han sido regulados por las entidades competentes, el censo realizado por el DANE en el 2005 presenta cifras que señalan debilidades en los sistemas de acceso y evaluación en el contexto educativo de las personas con limitación visual, en donde los exámenes de estado juegan un papel fundamental para garantizar los mismos. Los resultados indican que el 6,3% de la población, es decir cerca de tres millones de personas, presentan algún tipo de discapacidad y, de estos, sólo el 2,34% ingresó a algún nivel de educación superior y el 1% culminó este nivel educativo. La prevalencia de estas deficiencias refiere que de cada 100 colombianos con limitaciones, el 43,5% tienen limitaciones permanentes para ver (DANE, 2005).

En resumen, en el país uno de los exámenes que tiene mayor trascendencia al finalizar el primer ciclo de la vida escolar es el Examen de Estado SABER 11°, por lo que es necesario realizar actividades para garantizar que las diferentes formas de las pruebas sean equivalentes entre sí, que los puntajes en las pruebas sean igualmente interpretables para todos los subgrupos poblacionales y que en la aplicación año tras año se mantenga la medida del atributo a evaluar.

Esta investigación pretende evaluar a través de la simulación algunos procedimientos de equiparación para comparar su precisión y estabilidad, con el fin de proponer un método que

permita la equivalencia de puntuaciones en pruebas de aplicación masiva para población con y sin limitación visual.

Por lo anterior, la pregunta de investigación que orientó la realización de este estudio es: ¿Cuál método de equiparación permitirá estimar la equivalencia de puntuaciones con menor error, en pruebas de aplicación masiva para población con y sin limitación visual?

Con el fin de responder esta pregunta se plantearon los siguientes objetivos

OBJETIVOS

Para esta investigación se definió el siguiente objetivo general:

Seleccionar el método que permita realizar mejor estimación de la equivalencia de puntuaciones de prueba para personas con y sin limitación visual.

Para alcanzar este objetivo se formularon los siguientes objetivos específicos:

1. Resaltar la importancia de los procedimientos de equiparación en la evaluación de población con y sin limitación visual.
2. Simular procedimientos de equiparación bajo la Teoría Clásica de los Test utilizando los métodos de Tucker y Levine.
3. Simular un procedimiento de equiparación bajo la Teoría de Respuesta al Ítem utilizando el método Media/Sigma.
4. Establecer las tablas de conversión de acuerdo con la población.
5. Evaluar la precisión y la estabilidad de los métodos de equiparación seleccionados.

6. Identificar el mejor método de equiparación dadas las características de la población con y sin limitación visual.



UNIVERSIDAD NACIONAL DE COLOMBIA

REVISION BIBLIOGRAFICA

Las personas con limitación visual

La limitación visual hace referencia a la privación del código universal de gestos, el cual es vital para la comunicación humana, desde ahí existe la dificultad para interpretar a otros así como al mundo exterior (Castellanos, Álvarez, Pérez, Carrión & Ladino, 2008). En Colombia se toma como criterio para la clasificación de las personas ciegas y con baja visión, lo estipulado por la Organización Mundial de la Salud (OMS), quienes las clasifican en persona ciega total, ceguera legal y persona con baja visión. La referencia a personas con limitación visual incluye tanto las personas ciegas como aquellas con baja visión que en ambos casos requieren apoyo de un lector para resolver pruebas de aplicación masiva, (INCI & UPN, 2009).

Diferentes autores señalan una demora en la adquisición de ciertos procesos cognitivos en las personas con limitación visual, lo que sugiere que no se encuentran en igualdad de condiciones al enfrentarse a este tipo de pruebas. Especialmente se identifican demoras en la consolidación de algunos aprendizajes, como la inexistencia del objeto para los niños ciegos de nacimiento (Bigelow, 1986), en la función simbólica por la imposibilidad de imitar y formar imágenes de sí mismo y de otros (Ochaita, 1993), aún en la edad adulta para las

personas con limitación visual se dificulta usar términos referenciales, personales (yo, tú, él, ellos) y espaciales (ir, venir, subir, venir).

Ruiz (1994), indica que para atender las necesidades educativas especiales de las personas con limitación visual se deben considerar los canales de transmisión predominantes, la falta de interiorización de conocimientos espaciales y de orientación, así como un tiempo de trabajo más prolongado en comparación con las personas sin este tipo de limitación. En cuanto a las habilidades de las personas con limitación visual, aparecen entre otras, las investigaciones de Bach & Bach (2005) quienes señalan que los estímulos táctiles son representados en el cerebro durante más tiempo, adicionalmente cuando se recibe información por varios sentidos como en el caso de las personas con limitación visual, se consolida la información en áreas de asociación cortical, lo que permite que las representaciones de los objetos sean más duraderas y por tanto tengan una consolidación permanente del aprendizaje.

Las características mencionadas generan una forma diferente de aproximarse al conocimiento y a la evaluación, en ese sentido el ICFES en los Exámenes de Estado proporciona apoyos, también conocidos como acomodaciones a las personas con limitación visual para su adecuada aplicación. En el momento de la inscripción, los futuros examinados deben señalar si presentan algún tipo de discapacidad y en el momento de la aplicación a las personas con limitación visual se les asigna un lector previamente entrenado y capacitado por el INCI (Instituto Nacional para Ciegos), se recomienda dos lectores por cada evaluado para minimizar la influencia de la fatiga apoyando la mitad de la prueba, además el examen se realiza en un lugar distinto para no causar interferencia con la lectura de las preguntas y se proporcionan al lector una serie de elementos tiflológicos para facilitar la presentación del

examen, entre éstos se encuentran, una tabla de dibujo negativa y rodachina, un kit de geometría, pizarra y punzón, acetatos, papel bond base 28, hojas y marcadores de punta delgada; en el caso de las personas con baja visión se permite el ingreso al examen con ayudas tales como magnificadores o guías de escritura.

En esta situación de evaluación, los lectores cumplen un papel esencial en la inclusión educativa y social de una persona con limitación visual (INCI & UPN, 2009). A partir de diversas investigaciones Sierra & Vanegas (2001), han definido al lector como la persona que realiza un proceso de interacción con el texto con el fin de promover la comprensión, al elaborar un significado de las ideas relevantes y relacionándolas con aquellas ya interiorizadas.

La lectura oral le otorga al texto la relevancia necesaria para transmitir el mensaje y la comprensión del mismo. Entre los elementos comunicativos se encuentran el control de la fluidez, el ritmo, el control de la voz para favorecer inflexiones, el énfasis, la entonación, los hábitos posturales y elementos de significación para reconocer las palabras escritas (INCI & UPN, 2009). El lector debe manejar a la perfección estos elementos y realizar adecuadas descripciones, ya que la descripción es el recurso esencial en la lectura de las preguntas durante la prueba realizada por el lector, para así favorecer la interpretación de la información. Adicionalmente, es importante que el lector conozca el origen y edad de la limitación para establecer el tipo de descripciones o referentes que puede emplear al realizar la lectura de la prueba, debe además tener tolerancia para repetir la lectura de las preguntas las veces que sea necesario, manejar adecuadamente el tiempo por cada pregunta y ser creativo con el fin de acoplar los diferentes recursos para garantizar la comprensión de las situaciones planteadas

(INCI & UPN, 2009). El lector debe proporcionar al evaluado los detalles del texto de una manera organizada, en especial aquellos inaccesibles para las personas con limitación visual, donde el papel mediador entre el lector y el texto se requiere para realizar el análisis de éste; es así como el lector recibe la información para asimilar el significado del texto, con el fin de comunicarlo y trasmitirlo adecuadamente a la persona con limitación visual (Sierra & Vanegas, 2001). Finalmente, el lector logrará que la persona con limitación visual explore y realice sus propias hipótesis para encontrar y comprender el significado del texto, favoreciendo la conclusión e interpretación que llevará a la elección de una opción de respuesta (Allende & Condemarín, 1992). El lector debe realizar una lectura personal de la pregunta para conocer la intención antes de enunciarla, de la misma manera utilizará diferentes estrategias que permitan describir detalladamente aspectos del texto, gráficos o tablas (Correa, 2005). Sin embargo, a pesar del apoyo del lector no se debe olvidar que cuando se realiza la lectura, de un texto o una imagen a través de otra persona, la comprensión de la información puede verse afectada por las características de la lectura y de las condiciones ambientales en que se realiza.

Además de las acomodaciones realizadas para la aplicación de la prueba, está contemplada la posibilidad de realizar simulacros de las pruebas SABER para familiarizar a los futuros evaluados con el formato de la prueba y con la situación particular de evaluación (INCI & UPN, 2009). Estos apoyos, han permitido el acceso a la presentación de exámenes de selección masiva, ya sean las diseñadas por el ICFES o por las Universidades estatales que generan examen de admisión. Sin embargo, es importante identificar el efecto de estas modificaciones en las condiciones de aplicación y sus implicaciones en términos de la

comparabilidad de los resultados y las posibles diferencias que podrían existir con aquellas personas que no presentan este tipo de limitación.

En la investigación inédita realizada por Herrera, Soler, Espinosa, Lancheros y Jiménez (2012), se realizaron análisis de los puntajes obtenidos en tres aplicaciones del examen de estado, por personas con y sin limitación visual. Los resultados evidenciaron que las personas sin limitación visual se localizaban a lo largo del atributo medido, es decir en niveles bajos, medios y altos de habilidad, mientras que las personas con limitación visual se ubicaban en los niveles bajos de habilidad. Adicionalmente, se detectó Funcionamiento Diferencial de los Ítems al comparar las dos poblaciones evaluadas, indicando que existen preguntas acertadas por las personas sin limitación visual que no logran ser respondidas por las personas con limitación visual. Estas diferencias en la medición, sugieren una revisión de las metodologías de evaluación utilizadas en población minoritaria, ya que al parecer, éstas son insuficientes para estimar la magnitud real del atributo medido, donde una variable que puede afectar los resultados es la forma de aplicación, auditiva o visual. .

Las limitaciones, habilidades y apoyos que requieren las personas con limitación visual para la aplicación del Examen de Estado, promueven la reflexión acerca de la comparación de las puntuaciones resultantes del examen, dadas las condiciones diferenciales de aplicación entre la población con limitación visual y la población que no tiene este tipo de limitación, planteando la posibilidad de referirse a formas diferentes de la prueba que evalúan el mismo constructo. Es decir, que a pesar de ser aplicado el mismo contenido de la prueba a personas con y sin limitación visual, su forma de administración es diferente (auditiva vs escrita) lo que

implica que para las personas con limitación visual, se requiere más tiempo en cada pregunta y se incluyen otros materiales para el reconocimiento de gráficas y figuras.

Estas variaciones en la aplicación de la prueba pueden afectar la validez de los resultados y, por tanto, la interpretación de los puntajes. Por ejemplo, sería inadecuado afirmar con certeza que un puntaje de 50, es igual para personas con y sin limitación visual. Estas diferencias plantean la necesidad de garantizar la equivalencia entre las pruebas aplicadas a diferentes poblaciones y asegurar que las puntuaciones de uno y otro test sean intercambiables entre sí, en otros términos transformar la escala de puntuaciones de un test a la escala de puntuaciones de otro, para poder realizar comparaciones y determinar con precisión el nivel de atributo de una persona independientemente de su limitación.

Para lograr determinar con precisión del nivel de atributo independientemente del grupo poblacional al que se corresponda, se han creado diferentes métodos enmarcados en el concepto de equiparación.

Equiparación de Puntuaciones

De acuerdo con Muñiz (1997), cuando se mide una variable con pruebas diferentes, las puntuaciones que se obtienen no son directamente comparables. Con el fin de lograr puntuaciones equivalentes en evaluaciones de aplicación masiva se han utilizado diferentes métodos estadísticos agrupados bajo el rótulo de equiparación o equating. Según Kolen & Brennan (2004) la equiparación es un proceso estadístico usado con el fin de ajustar las puntuaciones de las formas de una prueba, para que los puntajes en cada una puedan ser intercambiables.

Lograr que dos formas de una prueba puedan ser equivalentes tanto en el rango como en el nivel de dificultad, es una labor difícil por lo que se hace necesario equiparar, es decir, “convertir el sistema de unidades de una forma a el sistema de unidades de la otra, y así poder afirmar que los puntajes que se derivan de las dos formas, luego de realizar la conversión puedan ser directamente equivalentes” (Angof, 1984 p. 85). A lo cual agrega Muñiz (1998), “... si efectivamente la mentada equiparación se ha hecho adecuadamente” (p. 275). Cuando se habla de puntuaciones equivalentes existen dos puntajes, uno para la forma A y otro para la forma B, y considerando que tanto A como B miden la misma variable con el mismo grado de confiabilidad (Muñiz, 1998), se considera equivalente siempre y cuando su rango percentil en cualquier grupo sea igual (Flanagam, 1951; Lord 1950, citado en Angof, 1984).

Los métodos de equiparación son utilizados para ajustar las diferencias en cuanto a la dificultad de formas alternativas de prueba, con el objetivo de hacer comparables los puntajes y así proporcionar mayor confiabilidad a la estimación de la habilidad de los evaluados (Albano, 2011). Holland y Dorans (2006) señalan que la idea de transformar los puntajes de un test a los puntajes de otro, se ha denominado como enlace (linking) y dentro de esta categoría se diferencian tres tipos: predicción (predicting), escalamiento (scaling) and equating (equiparación).

De acuerdo con lo descrito por Navas (1996), la Teoría Clásica de los Tests mantuvo el liderato en el área de Psicometría durante gran parte del siglo XX, y se centraba en crear tests cuyos resultados posteriormente eran analizados a la luz del grupo de referencia con el cual era estandarizado, denominado como medición con referencia a la norma. Por otra parte, la inversión educativa realizada en Estados Unidos requería evaluar con un alto nivel de

precisión los resultados de la misma, así que se empezó a hablar de estándares educativos, es decir, requisitos mínimos que debían adquirir los estudiantes al cursar los niveles educativos, denominado como medición con referencia a criterio. La importancia otorgada a la evaluación requirió aumentar el uso de test así como la aplicación a diferentes grupos, generando la necesidad de comparar resultados, lo que implicaría expresar las diferentes puntuaciones en una escala común, es decir, se necesitaba establecer un procedimiento que generara puntuaciones equivalentes de las pruebas utilizadas.

Debido a las implicaciones a nivel personal, educativo y político que pueden tener los resultados en pruebas de logro educativo, la legislación ha obligado a diferentes países a hacer públicos no solamente los resultados sino también las pruebas y, como afirma Muñiz (1997), “la continua crítica y discusión social de este sistema generalizado de test ha obligado a los constructores a justificar y explicar públicamente sus métodos de equiparación” (p.144).

La importancia de la equiparación surge de diferentes situaciones, una de ellas hace referencia a la garantía de equidad en un proceso de selección, clasificación o admisión, al generarse formas alternativas de una prueba porque se puede beneficiar o perjudicar a un grupo particular debido a la presentación de un test más fácil o más difícil, de modo tal que la equiparación se convierte en un mecanismo fundamental para garantizar comparaciones adecuadas de las puntuaciones obtenidas en cada prueba (Navas, 1996). Otra situación se refiere a evaluar el progreso educativo, lo cual requiere una continuidad en el proceso de evaluación que de no contar con una escala constante impediría “observar cambios reales y variaciones en las habilidades de los estudiantes evaluados” (Navas, 1996 p.298).

Este procedimiento es también importante al construir los bancos de ítems, debido a que se constituye en conocimiento fundamental para elaborar test ya sean distintos, formas alternas o paralelas. La introducción de ítems nuevos a un banco y la actualización de las estimaciones requiere utilizar estos métodos, tanto para los ítems existentes, como para el momento de la inclusión de los nuevos ítems, los cuales deben prever la forma en que se pondrán en la misma escala y la estimación de los parámetros cuando los ítems sean aplicados a diferentes grupos. En síntesis, cada vez que se actualice el banco se debe garantizar la conservación de las escalas en que se expresarán los resultados.

Finalmente, y como se menciona en repetidas oportunidades en la literatura, la equiparación es necesaria al realizar investigaciones sobre un posible funcionamiento diferencial de los ítems o sesgo, especialmente en la TRI donde se debe garantizar que las estimaciones de los parámetros se encuentran en la misma escala antes de comparar las Curvas Características de los Ítems que se obtienen de la aplicación a diferentes grupos (Navas, 1996). De esta manera, se ha encontrado que la incorporación de métodos de equiparación mejora la determinación sobre la existencia de ítems sesgados (Fidalgo, 1996).

Existen entonces diferentes usos de los procedimientos de equiparación que cobran gran importancia en los procesos evaluativos y adquiere especial relevancia cuando se evalúan grupos que aunque se enfrentan a formas paralelas de prueba, pueden diferir en su nivel de dificultad, lo que impedirá la comparación de los grupos.

Al realizar un procedimiento de equiparación se establece una población objetivo, ya sea un grupo real o hipotético y cada procedimiento es un intento por estimar la puntuación ajustada

para la población de interés. La generalización de esos resultados dependerá de la similitud entre las formas de la prueba; entre menor diferencia de dificultad y contenido los resultados de equiparación se estimarán y serán generalizados con mayor precisión a otros grupos de examinados (Livingston, 2004).

Otra limitación es que las escalas de puntajes vienen expresadas en números enteros y cuando se realiza el procedimiento, la puntuación equiparada puede ser un número decimal que no se ajusta a la forma de expresión de puntajes. Livingston (2004), señala que en estos casos se redondea el número obtenido al número entero más próximo, lo que puede llevar a que los puntajes se vean afectados por errores en el redondeo, donde una solución puede ser hacer una escala más fina aunque esto no sea lo más utilizado.

El propósito de la equiparación es finalmente otorgar una misma escala a los resultados de una o más pruebas que evalúan el mismo constructo y que cuentan con especificaciones técnicas similares. Este procedimiento es necesario porque es imposible desarrollar formas exactamente iguales en términos de dificultad a lo largo del continuo de habilidad. Lord (1980) y Navas(1996) sugieren cuatro criterios básicos para poder hablar de equiparación o para derivar puntuaciones equivalentes: (a) los test deben medir el mismo constructo, (b) debe existir simetría en la transformación de las puntuaciones de tal forma que sea indiferente el test del que se parta para realizar la equivalencia, (c) luego de la equiparación, las puntuaciones deben ser totalmente intercambiables, y (d) la tabla de equivalencia o ecuación de equiparación debe ser independiente de cualquier población. En resumen, este procedimiento debe caracterizarse por invarianza en la población, simetría y equidad (Angoff,

1984). Adicional a estos requisitos Dorans & Holland (2000) agregan que los test deben tener el mismo nivel de confiabilidad para que sean susceptibles de ser equiparados.

El proceso de equiparación de puntuaciones se desarrolla en cinco etapas básicas: (a) evaluación de la viabilidad de la equiparación y selección de un diseño para la recolección de los datos, (b) recolección de los datos realizando la aplicación de las pruebas de acuerdo con el diseño seleccionado (c) elección del método que permita equiparar las puntuaciones, es decir, un procedimiento estadístico adecuado, según el diseño escogido, para realizar la transformación de las puntuaciones, (d) estimación de la ecuación de equivalencia o la tabla de conversión y (e) evaluación del procedimiento de equiparación de puntuaciones, con el fin de identificar la magnitud del error que arroja el procedimiento, en este caso está determinado por el error típico de equiparación.

La equiparación de puntuaciones de un test es un procedimiento estadístico que finalmente está basado en el análisis de los datos. Por tanto, antes de entrar en la revisión metodológica es necesario distinguir los diseños de los estudios de equiparación de los métodos estadísticos utilizados para lograr la equivalencia entre las puntuaciones de los grupos o formas de la prueba. Los primeros hacen referencia a la estrategia o el plan para recolectar los datos que van a permitir desarrollar el procedimiento de equiparación, mientras que los segundos se refieren a la dirección estadística para analizar los datos y establecer una relación de equiparación entre éstos.

Diseños de equiparación

De acuerdo con diferentes autores (Angoff, 1984; Navas, 1996; Kolen & Brennan, 2004) dependiendo del propósito que se tenga, los diseños de equiparación se pueden clasificar en tres tipos principalmente: (a) de un solo grupo, (b) de grupos equivalentes; y (d) de grupos no equivalentes con ítems comunes o de anclaje.

De un solo grupo

En el diseño de un solo grupo se elige una sola muestra de participantes y las personas examinadas responden las dos versiones de la prueba. En este caso, se asume que la relación de equiparación observada en este grupo puede generalizarse a la población objetivo, en este caso no es necesario que el grupo de examinados sea una muestra representativa de la población. Una ventaja del uso de este diseño es que al ser la misma muestra de examinados quienes realizan el test, la potencia estadística aumenta y en general, se requiere pocos examinados para obtener un gran nivel de precisión. En este diseño es importante considerar la fatiga como variable que puede influir en el desempeño, para contrarrestar este efecto se usa el contrabalanceo, dividiendo la muestra en dos partes aleatorias para aplicar las pruebas a equiparar en orden diferente. Otra desventaja es que probablemente los resultados en la segunda prueba se vean afectados por la experiencia previa, es decir por efectos de la práctica.

De grupos equivalentes

En el diseño de grupos equivalentes se seleccionan dos muestras aleatorias representativas de la población y se asigna a cada grupo una de las pruebas, asumiendo que el azar conformó

muestras de individuos equivalentes en términos de habilidades y conocimientos en los que se puede llevar a cabo el procedimiento de equiparación. En el caso de que sea una aplicación masiva, es difícil garantizar la representatividad, por lo que se realiza una aplicación en espiral con respecto a la entrega de las formas de la prueba, para que al intercalarlas se logre una distribución equitativa de éstas. En este diseño se asume que la relación de equiparación que se observa en dos grupos de examinados puede generalizarse a la población objetivo, los dos grupos pueden ser diferentes de la población objetivo, siempre y cuando ambos difieran de la misma manera.

La principal ventaja de este diseño es que es muy práctico de aplicar, en el sentido de distribuir los cuadernillos en el orden alterno. Las limitaciones en el uso de este diseño, son en primer lugar, que para lograr precisión en el procedimiento de equiparación se requiere una muestra amplia de examinados, hasta 5 o 15 veces más para lograr el mismo grado de precisión que con el diseño de un solo grupo; la segunda limitación, se refiere a la seguridad de la prueba, ya que con frecuencia, la forma de referencia ha sido administrada con anterioridad y existe la posibilidad de un escape de información con los sujetos que la presentaron la primera vez, por haberlas visto y estudiado.

De grupos no equivalentes con ítems comunes o diseño de anclaje

El diseño de grupos no equivalentes con ítems comunes, también conocido como diseño de anclaje, se caracterizan por utilizar dos grupos de muestras diferentes en los que cada uno realiza una prueba, sin embargo ambos grupos contestan un conjunto de ítems iguales que permitirán establecer las equivalencias entre las pruebas a comparar. Aunque no hay un

acuerdo estándar sobre la cantidad de ítems denominados de anclaje, algunos autores señalan que debe corresponder al 20% de la longitud de la prueba original (Kolen & Brennan, 1995; Gau,2004)

Este conjunto de ítems puede ubicarse como parte de cada prueba, es decir como -test de anclaje interno que se refiere a una serie de preguntas que se incluyen de la prueba de referencia a la nueva forma, o como una mini prueba adicional denominada test de anclaje externo (Holland & Dorans, 2006).

Test de Anclaje Interno vs Test de Anclaje Externo

Cuando se utiliza el test de anclaje interno, la principal ventaja es la facilidad en la aplicación ya que se encuentra inmerso en la misma prueba, sin embargo puede tener dificultades de seguridad por exponer las preguntas en repetidas ocasiones. De esta exposición constante, pueden surgir valores atípicos en el análisis pues suelen ser los datos de aquellas preguntas que se han hecho más fáciles o más difíciles en la nueva forma que lo que eran en el grupo de referencia.

La ventaja de utilizar un test de anclaje interno es que el significado del puntaje en este test permanece, pues indica el mismo nivel de habilidad con la misma dificultad en ambas formas de prueba. En el momento del enlace se trata como si fuera una nueva pregunta y no como una pregunta repetida. Sin embargo, existe la posibilidad de que las preguntas cambien de dificultad, y además, este diseño puede ser inadecuado cuando se pretende equiparar una prueba con una escasa cantidad de ítems. En algunas ocasiones es necesario utilizar un test de

anclaje interno que no mida todas las habilidades que pretende la prueba para poder equilibrar la cantidad de ítems que conformarán el test.

La contraparte del test de anclaje interno es el test de anclaje externo, que consiste en una medida común, adicional a la prueba permitiendo comparar los dos grupos que toman las dos versiones de la prueba. Idealmente éste debe medir los mismos conocimientos y habilidades que la prueba a equiparar, por medio de preguntas en el mismo formato y aplicado en las mismas condiciones, aunque esto es difícil de cumplir. Debido a que el test de anclaje no es tomado por todos los examinados, las puntuaciones en éste no se incluyen para calcular las puntuaciones individuales en la prueba, solamente son usadas para el proceso de equiparación. Este plan de equiparación es poco práctico en la realidad para la mayoría de las pruebas. El uso de muestras para estas aplicaciones recuerda que los resultados de la equiparación pretenden generalizarse a otros grupos de individuos, aunque no se puede olvidar que las muestras utilizadas con frecuencia no son representativas de la población. La justificación clave de utilizar un test de anclaje externo es que los grupos se comportarán de la misma manera en el test de anclaje y en la prueba equiparada; una adecuada medida de esta relación es la correlación entre las puntuaciones de la prueba y las del test de anclaje, si esas correlaciones son fuertes, se asume que el test de anclaje es un indicador de las diferencias dentro de los grupos individuales de examinados, así se tendrá mayor confianza en el test de anclaje como indicador de las muestras que se equiparan. La principal dificultad del test de anclaje externo, es precisamente encontrar uno que permita una adecuada medición. Si el test mide habilidades diferentes a las de la prueba que se equipara, las muestras pueden diferir más

por lo que mide el test de anclaje que por las magnitudes del atributo en la prueba a equiparar; en este caso el diseño se verá afectado y la equiparación será inexacta.

Como se mencionaba anteriormente, aunque no existe un acuerdo en la literatura sobre las características del test de anclaje, se realizan algunas consideraciones. La principal es que las preguntas no hayan sido aplicadas a futuros examinados. Livingston (2004) proporciona las siguientes directrices para seleccionar los temas comunes que se incorporan en el test de anclaje interno:

(a) Incluir suficientes preguntas contenidas en la prueba de referencia, es decir, por lo menos 20 ítems comunes si la prueba es superior a 100 preguntas, y si no es el caso, se debe incluir un conjunto de preguntas que se asemeje a la prueba completa en contenido y formato; (b) idealmente debe ser una versión corta que incluya una gama completa de la dificultad, si esto no se cumple la equiparación sería imprecisa para los extremos inferiores o superiores; (c) no incluir preguntas que hayan sufrido modificaciones, estos cambios pueden hacer que la prueba resulte más fácil o más difícil, si es necesario cambiar la pregunta, ésta debe salir del grupo de ítems comunes para realizar el procedimiento de equiparación; (d) no romper una configuración de ítems, es decir, un conjunto de ítems que están vinculadas a un estímulo común, como una lectura, una imagen, un mapa, etc. y (e) si alguno de los ítems se van a incluir en el anclaje, se recomienda incluir todo el conjunto de ítems.

Otras recomendaciones que han sido mencionadas en la literatura sin mayor profundidad son: (a) evitar los últimos ítems como los de anclaje a no ser que el tiempo sea muy generoso, porque pueden quedar sin contestar o ser contestada al azar; (b) cada ítem de anclaje debe

encontrarse aproximadamente en la misma posición en ambas formas para evitar que cambie la forma de abordar el ítem por los examinados de acuerdo con la posición dentro de la prueba; (c) en igualdad de condiciones, se debe preferir aquellos ítems que correlacionen adecuadamente con el puntaje total, porque estos ítems proporcionarán más información sobre la exposición de los examinados a las dos formas de prueba (Livingston, 2004), y (d) los ítems que van a ser utilizados para el anclaje serán los que evidencien adecuadas propiedades de dificultad y discriminación con base en el modelo IRT (Gau,2004).

Realizar el procedimiento de equiparación utilizando un diseño de anclaje es una de las formas más complejas, se necesita el test de anclaje porque es inadecuado asumir que los grupos que toman las dos formas son iguales en conocimientos y habilidades, por esto es necesario utilizar esa información para ajustar las diferencias entre los grupos. La diferencia en los métodos utilizados con test de anclaje son los supuestos de la relación estadística que se generalizará a la población objetivo.

Métodos de Equiparación

Otro elemento fundamental en la equiparación de puntuaciones hace referencia al método estadístico que permite establecer equivalencia entre las puntuaciones generadas por los participantes de acuerdo con el diseño seleccionado. Navas (1996) y Kolen & Brennan (2004) clasifican los métodos de equiparación de acuerdo con la teoría psicométrica en la que se sustenten, los basados en la Teoría Clásica de los Tests (TCT) y los basados en la Teoría de Respuesta al Ítem (TRI).

Métodos basados en la TCT

Los Métodos de Equiparación fundamentados en la TCT se dividen a su vez en método lineal y método equipercentil.

1. Método Lineal

El método lineal se sustenta en que aquellas puntuaciones directas que tienen puntuaciones típicas iguales son equivalentes, es decir, que las puntuaciones de un test A equivalentes a las de un test B están dadas por la transformación lineal de las puntuaciones B; en ese sentido, se ajustan las diferencias que existen entre las dos formas de un test con referencia a sus medias y varianzas. Si X y Y representan los puntajes en los test A y B respectivamente, $\mu(X)$ y $\mu(Y)$ son las medias y $\sigma(X)$ y $\sigma(Y)$ son las varianzas, la transformación puede representarse en la siguiente ecuación:

$$\frac{X - \mu(X)}{\sigma(X)} = \frac{Y - \mu(Y)}{\sigma(Y)} \quad (1)$$

Al despejar el valor de Y se obtiene:

$$Y \equiv Y^* = \frac{\sigma(Y)}{\sigma(X)} (X - \mu(X)) + \mu(Y) \quad (2)$$

En este caso Y^* corresponde a una puntuación del test X, siendo función de la media y la desviación típica con respecto a las puntuaciones de ambos test, ajustando las diferencias en

medias y varianzas. Luego de ajustar los puntajes, el conjunto de puntajes equiparados tendrán la misma media y desviación estándar que el grupo de referencia (Livingston, 2004).

Asumiendo que el grupo 1 realiza el test X , el grupo 2 el test Y y ambos grupos el test de anclaje V . Los parámetros de la población, medias (μ) y desviaciones estándar (σ), se pueden expresar en términos de las subpoblaciones 1 y 2 que están representadas por los subíndices, adicionalmente se incluye el peso de los grupos (w) que corresponde a la cantidad de sujetos en cada grupo si se conoce o se le asigna un valor arbitrario cuando es simulado, en estos casos se recomienda $w_1=1$ y $w_2=0$. Los subíndices (s) serán posteriormente cambiados por el número correspondiente a la subpoblación (1 y 2) de que sean estimados los parámetros, con el fin de poder remplazar estos valores en la ecuación de equiparación (2) se realiza de la siguiente manera:

$$\mu_s(X) = w_1\mu_1(X) + w_2\mu_2(X) \quad (3)$$

$$\mu_s(Y) = w_1\mu_1(Y) + w_2\mu_2(Y) \quad (4)$$

$$\sigma_s^2(X) = w_1\sigma_1^2(X) - w_2\sigma_1^2(X) + w_1w_2 [\mu_1(X) - \mu_2(X)]^2 \quad (5)$$

$$\sigma_s^2(Y) = w_1\sigma_1^2(Y) - w_2\sigma_1^2(Y) + w_1w_2 [\mu_1(Y) - \mu_2(Y)]^2 \quad (6)$$

De acuerdo con el diseño y el método seleccionado, se realizan variaciones a las fórmulas que permiten realizar el procedimiento de equiparación. Para ampliar el tema véase Kolen & Brennan (2004); Navas (1996). Cuando se ha seleccionado el diseño de grupos no

equivalentes o diseño de anclaje, el método lineal puede ser basado en los datos o basado en un grupo sintético (Braun & Holland; citados en Kolen & Brennan, 1995).

Los métodos basados en los datos hacen referencia a los procedimientos aplicados a un conjunto de datos recolectados de forma empírica, e incluyen: (a) el anclaje externo cuando se aplican dos test a la población, el test principal y un test adicional llamado de anclaje, o (b) el anclaje interno que hace referencia a un conjunto de preguntas iguales inmersas en dos test diferentes. Para el anclaje externo se pueden utilizar los métodos de Lord (1955), doble equiparación, anclaje predictor y anclaje predicho; mientras que para el anclaje interno se pueden utilizar los métodos de Lord (1955), Thurstone (1925) y Swineford & Fan (1957). Una revisión detallada de estos métodos se encuentra en Navas (1996).

Los métodos basados en el grupo sintético estiman el comportamiento de los datos asumiendo que existen dos sub-poblaciones con medias, varianzas y distribución de frecuencias iguales; e incluyen: (a) *el método de Tucker*, (b) *el método de Levine* para tests con la misma y con diferente confiabilidad, y (c) *la estimación de frecuencias* (Kolen & Brennan, 2004; Navas, 1996). Este último asume que las distribuciones condicionales de las puntuaciones de las pruebas X y Y, dada una puntuación en el test de anclaje, son idénticas en las dos subpoblaciones. Con este supuesto se logra construir la distribución de frecuencias de un test a partir de las puntuaciones de ambos grupos en el test de anclaje.

La equiparación con el método lineal, utilizando el diseño de grupos no equivalentes con test de anclaje, es relativamente simple porque los parámetros necesarios para realizarlo pueden ser estimados directamente utilizando los datos disponibles. Sin embargo, este

procedimiento no es factible en grandes aplicaciones, ya que los grupos no pueden considerarse equivalentes al azar (pueden ser similares, más no equivalentes), por lo que es necesario realizar estimaciones indirectas de algunos parámetros (Kane y cols., 2009).

De acuerdo con Kane y cols. (2009), cuando se carece de todos los datos, se realiza la sustitución de parámetros, en la que se asume que las relaciones entre las puntuaciones del test de anclaje y los puntajes totales en las dos formas son invariantes en los grupos. Por medio de la sustitución de parámetros se estiman las medias y las varianzas de las dos pruebas. En una población se calculan los parámetros para una ecuación de equiparación de la siguiente manera: se estima la media y la varianza para el grupo (G1), en la forma de prueba que no presentaron, basándose los ítems comunes de la prueba que presentaron. Esto mismo se hace para el segundo grupo (G2), que presenta la forma alternativa de la prueba y para el que se estima los parámetros de la prueba X. En este procedimiento es importante determinar la población sintética que es definida como la combinación de los grupos, y cuyo peso combinado es igual a 1 (w_1 (peso grupo 1) + w_2 (peso grupo 2) = 1). En este modelo la relación de equiparación se determina remplazando los parámetros en la ecuación, por los de la población sintética.

A pesar de la sencillez de este método, existen algunas restricciones importantes de considerar para su uso. Una de ellas, es que puntajes muy bajos o muy altos para la nueva forma de prueba pueden equipararse con puntajes fuera del rango de posibles puntajes de la forma de referencia. Por ejemplo, tenemos dos formas de una prueba que contiene 100 preguntas, si la nueva forma es más difícil que la de referencia, el procedimiento de equiparación puede mostrar que un puntaje bruto de 99 de la nueva forma es comparable con

un puntaje de 103 de la forma de referencia. En ese caso un puntaje de 103 en una prueba que su máximo puntaje es 100, puede ser difícil de explicar.

Otra de las dificultades de este método es que los resultados pueden depender del grupo de examinados, si las formas difieren en dificultad. La equiparación puede asumir diferentes valores al realizar el procedimiento con individuos de alto o bajo nivel de habilidad ya sea en ambas formas, o porque una forma tenga predominantemente individuos con alto nivel de habilidad, mientras que la otra forma la resuelvan individuos con predominancia en niveles bajos de habilidad. Por esta razón, Livingston (2004) sugiere hablar de una posición relativa de los puntajes de acuerdo con la habilidad de los grupos que realicen las dos formas de prueba.

Entre los métodos que hacen parte de la equiparación lineal predominan dos ampliamente utilizados en la literatura y de fácil realización por la simplicidad de sus fórmulas, el método de Tucker y el método de Levine para test con confiabilidad igual o diferente. La primera aproximación al uso de las puntuaciones del test de anclaje es la equiparación encadenada, que consiste en equiparar las puntuaciones en la nueva forma de calificaciones en el test de anclaje y luego equiparar estas puntuaciones con el test de referencia. La cadena que se forma por estos dos procedimientos de equiparación, asume que la relación estadística que se generaliza a partir de cada muestra a la población objetivo, es una relación de equiparación.

La segunda aproximación se denomina acondicionamiento del test de anclaje o post – estratificación. En este enfoque, se utiliza la puntuación de anclaje como si fuera una variable

predictora para cada posible puntuación, se estima la distribución, generalmente solo la media y la desviación estándar de las puntuaciones en las dos formas, y así se utilizan para realizar la equiparación como si se observara en la población objetivo, siendo una relación condicional.

Los métodos que pueden ser utilizados con test de anclaje son el de estimación de frecuencias y los métodos de Tucker y de Levine, estos métodos se denominaron así por el nombre de quienes los propusieron Ledyard Tucker (Gulliksen, 1950) y Richard Levine (Levine, 1955).

Método de Tucker

El método de Tucker se basa en la regresión de los puntajes totales a partir de los resultados obtenidos en los ítems comunes y en la varianza condicional de estos, por lo que se recomienda aplicar este método con muestras pequeñas y puntajes observados. Para llevar a cabo este método, se utiliza la ecuación de regresión lineal para expresar los puntajes de la prueba X en función de la prueba Y (o viceversa) y se calcula el error estándar, definido como la desviación estándar de los puntajes igualados con resultados hipotéticos de un procedimiento de equiparación. Este error se puede utilizar como indicador de la precisión de la equiparación. Éste método tiene como supuestos: (a) que los parámetros de la ecuación de regresión tanto de X como Y, sobre el test de anclaje son idénticos en las dos subpoblaciones, y (b) la varianza de los errores de estimación tanto en X como en Y es idéntica en los dos grupos. La ventaja de utilizar este método es que las ecuaciones para realizar el procedimiento de equiparación se pueden calcular directamente a partir de los datos únicamente reemplazando

los valores. Para mayores detalles sobre este método puede verse Rodríguez (2007); Kolen & Brennan (2004); Navas (1996).

Cuando se utiliza este método para realizar equiparación lineal utilizando un diseño de test de anclaje, se debe estimar las medias y las desviaciones estándar. En este caso, se puede suponer que la puntuación media condicional en la forma nueva aumenta en forma constante, en otras palabras, es lineal con respecto a las puntuaciones del test de anclaje. En segundo lugar, se asume que la desviación estándar condicional es la misma en todos los niveles de la puntuación del test de anclaje. Por lo tanto, se puede utilizar una fórmula sencilla para calcular la media condicional de la población objetivo y adicionalmente, implica que se puede calcular un valor único para la desviación estándar condicional. Este método supone que se puede generalizar a la población objetivo tanto la ecuación lineal para estimar la media condicional en la nueva forma para otorgar un puntaje en el test de anclaje, como la estimación de la desviación estándar condicional de los puntajes sobre la nueva forma para cualquier puntaje del test de anclaje. Para lograr estas estimaciones se requieren medias y desviaciones estándar de las puntuaciones en el test de anclaje de la población objetivo. Con esta información ya es posible derivar las fórmulas para estimar la media y desviación típica de las puntuaciones en la nueva forma y en la forma de referencia para la población objetivo, para ser utilizadas en la ecuación de equiparación lineal. El resultado será una fórmula o ecuación de equiparación. Por lo general, se espera que la equiparación arroje una determinada puntuación de la nueva forma que corresponde a la una puntuación mayor en la forma de referencia.

Las ecuaciones que permiten despejar las incógnitas de las ecuaciones (3) a (6) y realizar el cálculo directamente, en la ecuación general de equiparación (2) para este método

son las que se muestran en las ecuaciones (7) a (10) donde (V) se refiere al test de anclaje y (α) se refiere a la pendiente de la regresión, los demás términos son iguales a los relacionados en las ecuaciones (3) a (6):

$$\mu_2(X) = \mu_1(X) - \alpha_1(X/V)(\mu_1(V) - \mu_2(V)) \quad (7)$$

$$\mu_1(Y) = \mu_2(Y) - \alpha_2(Y/V)(\mu_1(V) - \mu_2(V)) \quad (8)$$

$$\sigma_2^2(X) = \sigma_1^2(X) - \alpha_1^2(X/V) (\sigma_1^2(V) - \sigma_2^2(V)) \quad (9)$$

$$\sigma_1^2(Y) = \sigma_2^2(Y) - \alpha_2^2(Y/V) (\sigma_1^2(V) - \sigma_2^2(V)) \quad (10)$$

De la misma manera se calcula α , conservando el valor de los anteriores términos:

$$\alpha_1(X/V) = \alpha_2(X/V) = \frac{\sigma_1(X,V)}{\sigma_1^2(V)} = \frac{\sigma_2(X,V)}{\sigma_2^2(V)} \quad (11)$$

$$\alpha_2(Y/V) = \alpha_1(Y/V) = \frac{\sigma_1(Y,V)}{\sigma_1^2(V)} = \frac{\sigma_2(Y,V)}{\sigma_2^2(V)} \quad (12)$$

Método de Levine

El método de Levine (1955) supone que: (a) la correlación entre las puntuaciones verdaderas de las pruebas X y Y en conjunto con el test de anclaje, es igual en los grupos, (b) los parámetros de la ecuación de regresión tanto de X como Y, sobre el test de anclaje son idénticos en las dos subpoblaciones, y (c) la varianza del error de medida de los test X, Y y el

de anclaje V es idéntica para los dos grupos. A diferencia del método de Tucker el cálculo en la ecuación de equiparación no puede ser estimado de forma directa, por lo que es necesario recurrir a correcciones como las realizadas por Angoff (1984) para realizar la estimación. Este método pondera las diferencias entre los grupos en un nivel superior al método de Tucker. El método de Levine genera un procedimiento adicional en el caso que los test a equiparar presenten diferente nivel de confiabilidad.

En este método, una forma de evitar el sesgo de equiparación que se puede generar con el método de Tucker pero siguiendo con su misma lógica, consiste en realizar el procedimiento de equiparación utilizando los puntajes verdaderos en la nueva forma, la forma de referencia y el test de anclaje. Se considera como puntaje verdadero a la puntuación del examinado si la prueba fuera perfectamente fiable.

Estos puntajes nunca pueden ser conocidos, pero si es posible estimar la relación estadística que involucra puntajes verdaderos en el test y el de anclaje. Si se asume que esa relación de puntajes verdaderos es generalizable de la muestra a la población objetivo, se pueden estimar las medias y desviaciones estándar de los puntajes de la nueva forma y la forma de referencia en la población objetivo. Aunque el procedimiento estándar utiliza para sus cálculos los puntajes observados, existen dos versiones del mismo y ambas requieren adecuadas estimaciones de la confiabilidad de los test y el de anclaje en las dos muestras equiparadas.

De acuerdo con los supuestos que asume genera las siguientes fórmulas que proporcionan los valores de las incógnitas faltantes en la ecuación general de equiparación (2)

y son equivalentes a las fórmulas (7) a (10) en la mayoría de sus términos, sin embargo aparecen (\widetilde{X}) , (\widetilde{Y}) y (\widetilde{V}) que corresponden a las puntuaciones verdaderas del test correspondiente, si embargo aún no es posible remplazar los valores, por lo que se debe estimar el resultado de dividir las desviaciones típicas de los test.

$$\mu_2(X) = \mu_1(X) - \frac{\sigma_1(\widetilde{X})}{\sigma_1(\widetilde{V})} (\mu_1(V) - \mu_2(V)) \quad (13)$$

$$\mu_1(Y) = \mu_2(Y) - \frac{\sigma_2(\widetilde{Y})}{\sigma_2(\widetilde{V})} (\mu_1(V) - \mu_2(V)) \quad (14)$$

$$\sigma_2^2(X) = \sigma_1^2(X) - \frac{\sigma_1^2(\widetilde{X})}{\sigma_1^2(\widetilde{V})} (\sigma_1^2(V) - \sigma_2^2(V)) \quad (15)$$

$$\sigma_1^2(Y) = \sigma_2^2(Y) - \frac{\sigma_2^2(\widetilde{Y})}{\sigma_2^2(\widetilde{V})} (\sigma_1^2(V) - \sigma_2^2(V)) \quad (16)$$

Para estimar este resultado, se requiere el resultado en las desviaciones típicas de las puntuaciones verdaderas entre el test a equiparar y el de anclaje en las dos subpoblaciones, con este objetivo se propone las fórmulas de acuerdo con Angoff (1984) (17) y (18) si el test de anclaje es interno; mientras que (19) y (20) si el test de anclaje es externo:

$$\frac{\sigma_1(\widetilde{X})}{\sigma_1(\widetilde{V})} = \frac{1}{\alpha_1(V/X)} \quad (17)$$

$$\frac{\sigma_2(\widetilde{Y})}{\sigma_2(\widetilde{V})} = \frac{1}{\alpha_2(V/Y)} \quad (18)$$

$$\frac{\sigma_1(\widetilde{X})}{\sigma_1(\widetilde{V})} = \frac{\sigma_1^2(X) + \sigma_1(X,V)}{\sigma_1^2(V) + \sigma_1(X,V)} \quad (19)$$

$$\frac{\sigma_2(\widetilde{Y})}{\sigma_2(\widetilde{V})} = \frac{\sigma_2^2(Y) + \sigma_2(Y,V)}{\sigma_2^2(V) + \sigma_2(Y,V)} \quad (20)$$

2. Método Equipercantil

El método de equiparación equipercantil hace corresponder o hace equiparables las puntuaciones de ambos test en los que los percentiles son iguales. Al igual que en el método lineal, el cálculo que se realiza de las medias y las varianzas depende del diseño de recolección de datos (Navas, 1996; Kolen & Brennan, 2004). Así mismo, en este método existen diferentes procedimientos dependiendo si se utiliza el grupo sintético o los datos, o si el diseño es test de anclaje; el primer procedimiento se realiza mediante la estimación de frecuencias y en el segundo mediante el método directo, anclaje predictor o anclaje predicho (Navas, 1996). Este método propone que cada puntaje de la nueva forma debe ser ajustado de acuerdo con el puntaje de referencia que tiene el mismo rango percentilar en ese grupo (Livingston, 2004). Este método permitirá que las puntuaciones ajustadas generen casi la misma distribución del test de referencia en la población objetivo, con respecto a los puntajes brutos. Si la forma de la distribución de los puntajes brutos de ambas formas es igual o muy similar, se espera que los resultados del procedimiento de equiparación utilizando el método lineal o equipercantil producirán los mismos resultados.

La principal dificultad de este método radica en que las distribuciones que se observan en el test real de acuerdo con las respuestas de los examinados son irregulares, es decir que el

porcentaje de examinados no cambia gradualmente de acuerdo con el incremento de los puntajes sino que son fluctuantes. Esas irregularidades en las distribuciones de los puntajes generan problemas en este método porque no son generalizables a otros grupos de examinados que aplican el mismo test. En general, la localización, extensión y la forma de la distribución tenderá a generalizarse a otros grupos de examinados, mientras que esto no ocurre con las irregularidades en la distribución (Livingston, 2004). Como solución a esto, se puede utilizar la técnica denominada suavizado o “*smoothing*” que consiste en sustituir el resultado de la distribución inicial con una que tenga la misma ubicación, extensión y forma, pero no las irregularidades. Incluso se habla de presuavizado o “*presmoothing*” cuando se aplica a distribuciones para establecer previamente si hay una posible relación de equiparación. Para realizar estos procedimientos se han propuesto diferentes estrategias. Incluso si no se cuenta con la forma de suavizar las distribuciones, se puede realizar la equiparación con las puntuaciones observadas y luego suavizar la relación de equiparación, que también se conoce como post- suavizado o “*postsMOOTHING*” (Kolen & Brennan, 2004).

Surge la pregunta sobre cuál de los dos métodos de equiparación es más adecuado en un diseño de anclaje, debido a que el sesgo que puede surgir durante el proceso es real, previsible y explicable (Livingston, Dorans, & Wright, 1990). Esta pregunta es difícil de responder porque no existe un consenso general entre los investigadores, es importante comparar los métodos utilizando los mismos datos. Con frecuencia, la elección de un método de equiparación se reduce a lo que es creíble, dado lo que se conoce de la prueba y la población de examinados.

Kane, Mroch, Suh & Ripkey (2010) detectaron que cuando los grupos difieren en habilidad, el método de Levine hace ajustes en los puntajes mayores, el método de Tucker hace ajustes en los puntajes menores y el método lineal encadenado hace ajustes en los dos extremos. La evidencia muestra que las diferencias están dadas por efectos de regresión, en especial por la regresión a la media, que afecta diferentes modelos, dependiendo de la relación del puntaje observado entre el del test de anclaje y la prueba total asumiendo invarianza entre los grupos, así las diferencias en las magnitudes de los métodos dependen de las magnitudes de las correlaciones entre la prueba de referencia y la de anclaje.

En términos generales, los métodos basados en la TCT buscan construir una tabla de conversiones de uno a otro test que permita expresar la equivalencia entre las puntuaciones en términos de magnitud del atributo; sin embargo, con el avance de las teorías psicométricas, han realizado su incursión los métodos basados en la Teoría de Respuesta al Ítem (TRI).

Métodos basados en la Teoría de Respuesta al Ítem

En el marco de la TRI inicialmente se asumía que no era necesario equiparar puntuaciones de dos test, dado que en teoría se ha cumplido el supuesto de invarianza, es decir, que la estimación de habilidad de los examinados es independiente del test que se aplique. Sin embargo, según Pacheco (2006) las consideraciones deben realizarse con respecto a la necesidad de colocar en la misma métrica las estimaciones del atributo con diferentes medidas.

Navas (1996) señala que en la TRI, la equiparación se ha visto como un proceso que permite determinar las constantes de la transformación lineal, relacionando las estimaciones

obtenidas para los parámetros, por lo tanto, para Zhu (2001) la equiparación con este modelo se considera más exacta en los puntajes extremos de la escala, con mayores facilidades a lo largo del proceso y con menor margen de error.

Se ha descrito que el procedimiento de equiparación con base en la TRI comprende dos etapas: estimar los parámetros y colocar los parámetros en una escala común (Cook & Eignor, 1991). Keller, Keller & Parker (2011), señalan que las etapas corresponden a colocar los parámetros de los ítems en una misma escala y crear una tabla de conversión que proporcione la correspondencia entre los puntajes observados para las dos medidas. Con el fin de equiparar los puntajes, se realiza la transformación de puntuaciones de un test a otro considerando el nivel de habilidad que se estima mediante el modelo TRI, en ese sentido, la literatura sugiere que los métodos basados en esta teoría funcionan adecuadamente en grandes tamaños de muestra, por cuanto se requiere obtener una gran precisión de las estimaciones a lo largo de la escala (Navas, 2000).

Una vez que los parámetros de los ítems son llevados a una escala común, los resultados que se derivan de esos parámetros de los ítems se pueden colocar también en una escala común (Keller, Keller & Parker, 2011).

Entre las ventajas de utilizar los procedimientos de equiparación basados en la IRT se encuentran: (a) permite estimar los puntajes verdaderos equiparados para todos los valores del parámetro de habilidad (θ); (b) permite una mayor flexibilidad en la elección de formas anteriores de la prueba, debido a que los parámetros de los ítems se encuentran en una escala común; (c) si se requiere eliminar un ítem es más sencillo volver a realizar la equiparación

porque solamente deben sumarse las curvas características de los ítems finales; y (d) si los ítems están calibrados, se puede realizar un procedimiento de pre-equiparación o derivar la relación entre las formas del test antes de ser aplicado (Cook & Eignor, 1991).

Los métodos para realizar los procedimientos de equiparación bajo la TRI se han dividido en tres grandes grupos: (a) Métodos basados en los momentos, en los que se deriva el valor de las constantes considerando los dos primeros momentos en la estimación de los parámetros; para lograrlo existen diferentes procedimientos entre los que se encuentran el método de la media y la desviación típica (media/sigma), la media y la desviación típica robusta y el método iterativo de la media y la desviación típica robusta y ponderadas; (b) métodos basados en la curva característica, en donde el valor de las constantes se derivan minimizando las diferencias entre las curvas características de los test, y (c) otros métodos, entre los que se encuentran aunque con un menor uso, el método de las b 's fijas y el método de calibración concurrente. En el presente estudio se utilizó uno de los métodos basados en los momentos, conocido como media/sigma.

Método Media / Sigma

El método de la media y la desviación típica denominado media/sigma (Cohen & Kim, 1998), consiste en estimar la dificultad de los ítems del test de anclaje en cada grupo de forma independiente, es decir que para cada ítem del test de anclaje se obtendrán dos puntuaciones distintas de su dificultad. A este método se le ha criticado el no considerar la precisión con qué se estiman los parámetros, así como la existencia de valores atípicos, sin embargo para

solucionar esto se han desarrollado diferentes correcciones de este procedimiento buscando la robustez de la medida (Navas, 1996).

Se han realizado numerosos estudios que buscan evidenciar las bondades de los diferentes métodos, por una parte los estudios de simulación han proporcionado una orientación con respecto a la precisión de la estimación, mientras que los estudios empíricos han favorecido la comprensión de cómo funcionan los métodos en situaciones reales con condiciones no tan ideales (Keller, Keller & Parker, 2011). Sin embargo, no se puede olvidar que los supuestos de la Teoría de Respuesta al Ítem son difíciles de satisfacer, en especial cuando el tamaño de la muestra es pequeño (Kane, Mroch, Suh & Ripkey, 2010).

Luego de realizar la equiparación de las estimaciones de los parámetros y del nivel de habilidad (θ), puede usarse la métrica de las puntuaciones verdaderas o la métrica de las puntuaciones observadas, ya que la métrica de theta es más difícil de interpretar. En el primer caso, se consideran equivalentes las puntuaciones verdaderas de dos test cuando corresponden al mismo nivel estimado de habilidad (θ). En el segundo caso, la equiparación de puntuaciones observadas, se aplica el método equipercentil a las distribuciones que se generan teóricamente y no a las observadas (Navas, 1996; Kolen & Brennan, 2004).

Actualmente el modelo más utilizado en el marco de la TRI dada la estabilidad de las escalas generadas y su robustez a la violación del supuesto de unidimensionalidad, es el modelo de Rasch (Kolen & Brennan, 1995). En consecuencia, un procedimiento frecuente para equiparar puntuaciones es calibrar los ítems con modelo Rasch para luego realizar el

escalamiento utilizando el método media/sigma con la media y la desviación estándar de la dificultad del ítem, o con la media de la habilidad de los ítems comunes.

Si b_x es el valor estimado para la dificultad de un ítem de anclaje en el grupo 1, $\mu_x(b)$ y $\sigma_x(b)$ son la media y desviación típica de las dificultades estimadas para los ítems de anclaje en el grupo 1, b_y es la dificultad del ítem el grupo 2 y $\mu_y(b)$ y $\sigma_y(b)$ son su media y desviación estándar en el test de anclaje del grupo 2 respectivamente, las estimaciones bajo este modelo estarían dadas por la siguiente expresión:

$$\frac{b_x - \mu_x(b)}{\sigma_x(b)} = \frac{b_y - \mu_y(b)}{\sigma_y(b)} \quad 21$$

Al despejar se evidencia de la siguiente manera:

$$b_y \equiv b_y^* = \frac{\sigma_y(b)}{\sigma_x(b)} b_x + \left(\mu_y(b) - \frac{\sigma_y(b)}{\sigma_x(b)} \mu_x(b) \right) \quad (22)$$

Las constantes de la ecuación de equiparación α y β toman los siguientes valores en el método media/sigma de acuerdo con el diseño de test de anclaje:

$$\alpha = \frac{\sigma_y(b)}{\sigma_x(b)} \quad (23)$$

$$\beta = \mu_y(b) - \frac{\sigma_y(\theta)}{\sigma_x(\theta)} \mu_x(\theta) \quad (24)$$

De acuerdo con Keller, Keller & Parker (2011), se puede utilizar un criterio para examinar las diferencias entre métodos de equiparación para diferentes métodos de escalamiento. En primer lugar, se registra la correlación de los puntajes equiparados con los métodos de escalamiento dentro de cada método de equiparación; en segundo lugar, se registra la correlación de los puntajes equiparados para un método en particular dentro de los métodos de equiparación. Estos análisis proporcionan información sobre la precisión de los puntajes con el mismo método de equiparación cuando difieren los métodos de escalamiento.

En general, existen diferentes diseños y procedimientos al momento de realizar un procedimiento de equiparación, sin embargo, es importante tener en cuenta algunas consideraciones para elegir el más adecuado. Navas (2000), señala que existe consenso en que si un test se construyó a partir de algún modelo teórico, lo más conveniente será adoptar un método de equiparación desprendido del mismo modelo; los métodos de la teoría clásica funcionan bien cuando ni los test difieren significativamente en dificultad, ni los grupos en nivel de habilidad, sin embargo si los datos se ajustan a la TRI, estos métodos funcionan bien en todas las combinaciones. En general, se recomienda el método lineal cuando se trabaja con muestras pequeñas, los tests que se equiparan no son muy distintos y se requiere una precisión en una zona de la escala de habilidad mientras que los métodos como el equipercentil y los basados en la TRI funcionan mejor cuando se trabaja con una muestra grande y cuando se requiere precisión a lo largo de toda la escala. (Lawrence & Dorans, 1990).

Sin embargo, es importante señalar que independiente del método que se escoja para realizar la equiparación, cuanto más diferentes sean los grupos, las formas del test, y las especificaciones de los ítems mayor es el riesgo de hacer equiparaciones imprecisas (Navas,

2000). En consecuencia, es necesario evaluar la calidad de la equiparación, es decir, valorar que efectivamente resulten puntuaciones equivalentes a partir de la conversión de puntuaciones.

Error de equiparación

Independientemente del método que se escoja para realizar el procedimiento de equiparación, existe el riesgo de que la equivalencia realizada entre los puntajes difiera en el grado de precisión con que se estime ya sea por diferencias entre los grupos, las formas del test o las especificaciones de los ítems, tanto los comunes como los únicos (Navas, 2000).

Al finalizar el proceso de equiparación es necesario evaluar la calidad de esta, es decir, estimar el grado de precisión de la equivalencia entre las puntuaciones a partir de la conversión de las mismas.

Si bien no se puede asegurar previo al procedimiento de equiparación, cuál es mejor método a emplear en distintos momentos y considerando los diferentes tipos de diseño, es importante analizar los tipos de error que se pueden cometer. Es posible que los resultados de la equiparación se vean afectados por la variabilidad de las muestras. El error aleatorio debe ser examinado con cuidado mediante la utilización de métodos como el “jackknife” o método de las pseudo-reiteraciones con semimuestras y el método “bootstrap”. El error total del procedimiento de equiparación puede estimarse mediante el error cuadrático medio (ECM) calculado como la diferencia cuadrática entre la puntuación equiparada correspondiente a una determinada puntuación de una prueba y la puntuación que hubiera obtenido el evaluado de haberle aplicado el test contrario. El error total que afecta a la equiparación de una puntuación

se puede definir como la diferencia que existe entre la puntuación equiparada Y^* correspondiente a una determinada puntuación en el test X y la puntuación Y que habría obtenido el sujeto de haberle aplicado el test Y en lugar de X , es decir. $E = Y^* - Y$. El error cuadrático medio se representa mediante la siguiente fórmula, donde k es la cantidad de puntuaciones del test X y el subíndice j indica considerar todas las puntuaciones del test X ; el procedimiento consiste en calcular para todas las puntuaciones el error de equiparación, elevarla al cuadrado y hallar la media así:

$$ECM = \frac{\sum_i^k (Y_i^* - Y_i)^2}{k} \quad (25)$$

Entre los métodos mencionados anteriormente, el método que es utilizado con mayor frecuencia para evaluar la calidad de la evaluación es el método “bootstrap” (Albano, 2011). Consiste en un procedimiento de remuestreo intensivo que permite estimar el error típico de numerosos estadísticos. El procedimiento extrae múltiples muestras con reposición también llamadas muestras “bootstrap” para calcular repetidamente las puntuaciones equiparadas y estimar el error típico de equiparación.

En ese sentido, es indispensable realizar un procedimiento arbitrario que permita además de equiparar dos tests, determinar cuál sería la puntuación que obtendrían las personas si en lugar de contestar cierto test hubieran contestado otro; para ello se han promulgado dos tipos de formas: el paradigma circular y la simulación. La primera forma tiene como supuesto que es viable equiparar un test con él mismo, dando como conversión resultante una identidad; mientras que en el segundo se cuenta con un criterio absoluto para evaluar el error de

equiparación conociendo la conversión verdadera entre las puntuaciones de los tests (Navas, 1996).

Según Han & cols. (1997) la estabilidad de los resultados para los métodos de equiparación es evaluada equiparando una forma de prueba con ésta misma usando índices perdidos, lo cual puede lograrse equiparando los grupos completos o dividiendo el grupo en partes iguales. Para equiparar una prueba dividida en dos grupos, los criterios de equiparación son conocidos, pueden calcularse los índices de las discrepancias entre los puntajes equivalentes de un procedimiento de equiparación y los puntajes de criterio, las diferencias resultantes son conocidas como pérdida de equiparación. Estos índices que se pierden permiten evaluar la estabilidad de los métodos de equiparación al existir una relación inversa entre la estabilidad del método y la magnitud de los índices, por lo tanto, cuanto más cercanos a cero, mayor la estabilidad del método.

Livingston (2004) afirma que si se repitiera el proceso de equiparación varias veces, con diferentes examinados conservando su cantidad, y se comparan algunos de los resultados de la equiparación, éstos deberían ser similares aunque no exactamente los mismos, por lo que se debería calcular la distribución muestral de un estadístico, el puntaje equiparado para ese puntaje bruto en particular. La desviación estándar de la distribución del muestreo evidenciará hasta qué punto los resultados de la equiparación variarían de una muestra de examinados a otra. Esta desviación es denominada el error estándar de equiparación, que aunque no se puede calcular si es posible estimarla, entre más grande la muestra de examinados, menor será el error estándar de equiparación. Incluso podría llamarse el error estándar condicional de equiparación, porque es diferente para cada puntaje bruto. En el medio de la distribución

donde se encuentran la mayoría de puntuaciones de los examinados el error estándar de equiparación tiende a ser pequeño. Por el contrario, en los extremos superiores e inferiores de la distribución de los puntajes, el error estándar de equiparación tiende a ser mayor (Livingston, 2004; Von Davier & Kong, 2003).

Von Davier & Kong (2003) señalan que aunque existen diferentes métodos de equiparación, cada uno con sus propios supuestos, comparten importantes similitudes. Una de las características que comparten los métodos lineales es que se puede generar una fórmula general para el error estándar de equiparación (SEE) utilizando cualquiera de ellos. En su investigación postulan el error estándar de la diferencia de equiparación (SEED), que permite identificar la existencia de diferencias estadísticamente significativas entre los métodos de equiparación.

En la misma dirección los estudios de Harris y Crouse (1993) indican que es importante identificar si existen diferencias estadísticamente significativas, más allá de observar los diferentes errores de equiparación, sin la posibilidad de establecer que los resultados solamente reflejen errores aleatorios. En un experimento donde se comparaban tres métodos de equiparación, Von Davier & Kong (2003) al comparar los métodos de Tucker y de Levine en cuanto a la magnitud del error de equiparación evidencia que para la función de Levine es más grande que para el método de Tucker, que tiene los valores más pequeños en casi todo el rango de puntuaciones. Sin embargo, los valores obtenidos son muy cercanos para poder realizar una clara distinción sobre el menor error de equiparación.

Se han analizado otros procedimientos adicionales, desarrollados para evaluar la magnitud de error arrojado por el procedimiento de equiparación. Algunos de ellos son a) investigar acerca de la sensibilidad de las funciones de equiparación especialmente para comprobar el supuesto de invarianza de la población (Von Davier, 2003b); b) realizar un análisis de equidad con base en el proceso de equiparación, aunque también se dirige a sustentar el supuesto de invarianza, se centra en aspectos como especificar el número de subpoblaciones, comprobar si las distribuciones de las subpoblaciones son similares, calcular la diferencia estandarizada entre las medias, entre otras (Dorans & Holland, 2000); c) comparar los diferentes primeros momentos de la distribución a través de la equiparación de distribuciones antiguas (Von Davier, Holland & Thayer, 2004).

Parshall, Houghton & Kromrey (1995) realizaron una investigación para determinar el efecto del tamaño de la muestra sobre la estabilidad y el sesgo en procedimientos de equiparación lineal en dos formas de prueba bajo el diseño de grupos no equivalentes con ítems de anclaje, observando las diferencias con muestras entre 15 y 100 individuos. Se encontró que aunque no existían diferencias significativas que evidenciaran sesgo, el error estándar de equiparación (SEE) incrementaba a medida que el tamaño de muestra disminuía. En un caso similar, donde se realizaron procedimientos de equiparación lineales y no lineales utilizando el mismo diseño, con muestras entre 25 y 200 individuos, el error (SEE) era tan grande que el procedimiento de equiparación no eliminaba las dificultades que el sesgo introducía (Skaggs, 2005).

En muchas ocasiones, la calidad de la equiparación puede depender del tamaño de las muestras en relación con la población objetivo, la probabilidad de tener una muestra

representativa disminuye cuando sólo están disponibles pequeñas muestras por las características de la población. Como solución a estas dificultades se han planteado diversos procedimientos como el de la función sintética (véase, Kim, Von Davier & Haberman, 2008), sin embargo, éstos aún no han comprobado su eficacia en diferentes situaciones. Por esta razón, cuando las muestras pequeñas son inevitables se sugiere la inclusión de un número razonablemente grande de ítems comunes en un diseño de anclaje, ya que en pruebas de selección múltiple con un alto porcentaje de superposición y correlaciones altas entre los puntajes del test de anclaje y la prueba se suele reducir el error de equiparación y el sesgo (Kim, Von Davier & Haberman, 2011).

Cuando se encuentran diferencias entre los métodos la importancia de identificar en qué radican las diferencias permiten enriquecer la toma de decisiones al momento de usar uno u otro método. En el estudio de Kim, Von Davier & Haberman (2011), al realizar la comparación entre diferentes métodos de escalamiento y de equiparación en la TRI, encontraron que los participantes se clasifican de manera diferente dependiendo del método utilizado; cuando se compararon los métodos de equiparación y los métodos de escalamiento por separado, se evidenciaba una diferenciación por la elección del método de escalamiento, este tenía más impacto sobre las puntuaciones verdaderas que sobre las observadas. Esto puede deberse al hecho de construir tablas de conversión independientes para examinados diferentes con la misma puntuación en bruto, que es generada a partir de la CCT de las dos evaluaciones.

La descripción realizada de los diferentes Métodos de Equiparación evidencia que en primer lugar es preciso determinar el diseño de equiparación con el cuál se va a establecer la

relación de enlace entre las pruebas y de esta manera seguir una serie de procedimientos estadísticos que se requieren para encontrar las constantes de transformación o para establecer la tabla de equivalencia correspondiente. De los diseños descritos el que se ajusta en gran medida a la realidad de la evaluación colombiana y que introduce menor nivel de error en la selección y conformación de los grupos es el Diseño de grupos No Equivalentes con Test de Anclaje ya que es el de más fácil aplicación y es difícil garantizar grupos de poblaciones que se consideren equivalentes, especialmente cuando pertenecen a una minoría caracterizada por diferencias reales que pueden introducir diferencias en el atributo medido. Posterior a definir el Diseño de Equiparación se debe seleccionar un procedimiento de equiparación, la literatura resalta la facilidad en las estimaciones de los métodos lineales desarrollados bajo la Teoría Clásica de los Tests en donde el insumo principal es el cálculo de las medias y desviaciones estándar de las puntuaciones, tanto sus términos como procedimientos son relativamente sencillos de calcular y permiten una estimación adecuada con tamaños poblacionales pequeños. Con las mismas bondades se encuentran los Métodos bajo la Teoría de Respuesta al Ítem que han sido desarrollados asumiendo únicamente el parámetro de dificultad. En resumen la mejor forma de identificar la precisión y la estabilidad de un método de equiparación es mediante la evaluación de los errores que arroja luego de realizado el procedimiento, en esto la estadística ha tenido grandes desarrollos donde los programas computacionales permiten realizar éstos cálculos de una manera ágil y precisa.



UNIVERSIDAD NACIONAL DE COLOMBIA

MÉTODO

Diseño

Esta investigación puede caracterizarse como un experimento de MonteCarlo en el cual se definieron algunas constantes, variables independientes y una variable dependiente y se simularon las diferentes condiciones experimentales para evaluar la precisión de la equivalencia resultante de tres diferentes métodos de equiparación. Las constantes fueron: (a) la longitud de la prueba que se fijó en 24 ítems por ser la longitud de las subpruebas de lenguaje del examen de estado y (b) la razón del tamaño de muestra se fijó en 1:25 considerando los resultados de la investigación de Herrera (2005); Berrío (2008); Santana (2009). Para fijar esta condición se tomó la población de personas que reportaron alguna limitación visual compuesta por 87 examinados y se tomó una muestra aleatoria de 2175 personas sin limitación visual. Esta razón de tamaño 1:25, corresponde a 1 persona con limitación visual por cada 25 personas sin limitación visual, utilizando el muestreo aleatorio para seleccionar la población sin limitación visual. De esta manera, para los diferentes procedimientos se consideró una muestra de 87: 2175 de personas con y sin limitación visual respectivamente.

Las variables independientes fueron los métodos de equiparación a comparar. En la Teoría Clásica de los Test se seleccionó el Método de Tucker (Gulliksen, 1950) y el Método de Levine (1955), mientras que en la Teoría de Respuesta al Ítem se seleccionó el Método Media / Sigma (Cohen & Kim, 1998). Estos procedimientos fueron elegidos por ser de fácil aplicación, amplio estudio y uso en otros contextos, además de las bondades expresadas en la aplicación al diseño seleccionado y al tamaño de la muestra.

Las variables dependientes establecidas para cada método fueron el tamaño del error de equiparación con el fin de identificar la precisión del procedimiento al momento de establecer la tabla de equivalencia y el coeficiente de variación que permitió observar la estabilidad del método a lo largo de las réplicas dadas diferentes matrices de respuesta simuladas..

La investigación se realizó utilizando como referencia una de las bases de datos proporcionada por el ICFES, a través de la Convocatoria Nacional Colciencias - Icfes para la Conformación del Banco de Proyectos de Investigación en el Área de la Calidad de la Educación año 2010. La entidad proporcionó las bases de datos correspondientes a las aplicaciones del examen de estado SABER 11 en el segundo periodo de 2008.

Esta base contenía los datos sociodemográficos y las respuestas a la prueba de Lenguaje de 237.102 examinados, diferenciados en personas sin limitación visual 237.015 y con limitación visual 87. Sin embargo, como la población con limitación visual representa menos del 1% de la población total en todas las aplicaciones, situación que podía afectar la potencia estadística de las pruebas utilizadas, y con el fin de mantener la razón de

tamaños antes fijada, se seleccionó una muestra aleatoria de la población vidente, como se explicó anteriormente.

Instrumento

Aunque se trata de un estudio con datos simulados, la generación de los mismos se basó en las características y las estimaciones de los parámetros de la prueba de Lenguaje del Examen de Estado Saber 11 conformada por 24 ítems, en formato de selección múltiple con única respuesta y 4 opciones. La prueba evalúa tres componentes: función semántica de los elementos locales, sentido del texto hacia otros textos y configuración del sentido global del texto, a través de las competencias argumentativa, interpretativa y propositiva. Para este estudio, se asumió que existen dos formas de prueba de acuerdo con su modalidad de presentación (auditiva o visual). La forma de referencia es la que se aplica a las personas sin limitación visual y la forma nueva es la que se aplica a las personas con limitación visual en la modalidad auditiva. En este sentido, resulta fácil suponer que se trata de dos formas paralelas de la misma prueba que conserva igual longitud, contenido y confiabilidad.

En consecuencia, se seleccionó el diseño de equiparación de test de anclaje interno para grupos no equivalentes. Los ítems de anclaje se seleccionaron luego de identificar aquellos que no evidenciaran diferencias estadísticamente significativas de sus parámetros entre la población con y sin limitación visual y finalmente se eligieron cinco distribuidos a lo largo del continuo de dificultad que corresponden al 20% de la prueba.

Procedimiento

Generación de los datos

En primer lugar se realizó la estimación de los parámetros de habilidad para las personas y de dificultad para los ítems en la aplicación 2008 – 2, forma B utilizando el programa Winsteps versión 3.73, siguiendo el modelo de Rasch (1960). Con base en los vectores de los parámetros de dificultad y de habilidad de los individuos, obtenidos para la población con limitación visual (n=87) y la muestra aleatoria de la población sin limitación visual (n=2175) se simularon las matrices de respuestas (valores 0 y 1) asumiendo una distribución Bernoulli y se replicó 1000 veces de manera que se obtuvieron 1000 pares de matrices de respuesta. Para la generación de estos datos se escribió una rutina en el programa estadístico *R- estudio* implantando una semilla que en R se da bajo la función `set.seed()`, la semilla implantada fue `set.seed(123456789)`, esta simulación se corrió en un computador con procesador Intel® Core™ i7 CPU Q720 @1.60 GHz 1.60GHz, con RAM instalada de 6.00 GB y Sistema operativo Windows 7 Ultimate de 64 bits.

Diseño de equiparación.

Como se mencionó antes, se eligió un diseño de test de anclaje interno para grupos no equivalentes, es decir, se supone que la población con limitación visual respondió una forma de la prueba (forma oral) y las personas sin limitación visual respondió otra. El test de anclaje interno estaba compuesto por 5 ítems en cada una de las formas como se muestra en la Figura 1.

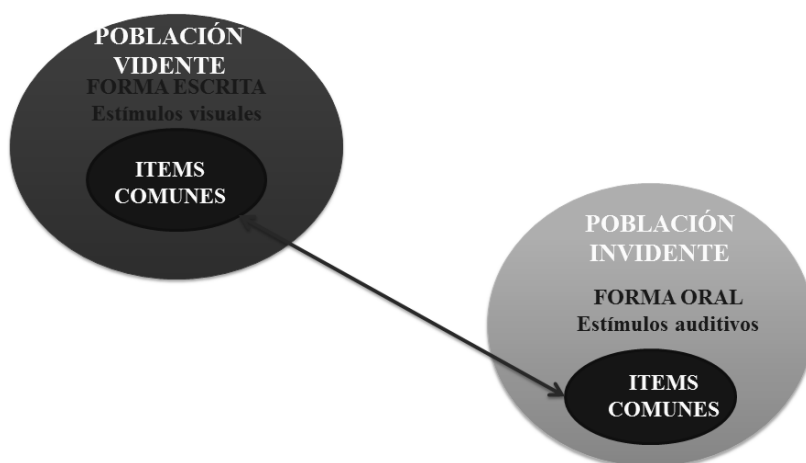


Figura 1. Aplicación en la investigación del Diseño de anclaje para grupos no equivalentes.

Tablas de equivalencias y error de equiparación

Para cada par de matrices se realizaron los tres procedimientos de equiparación antes seleccionados: el método de Tucker, el de Levine y el Media/Sigma. De la aplicación de los tres métodos se obtuvieron tablas de equivalencia de las puntuaciones de las personas con limitación visual, a partir de las personas sin limitación visual, de esta manera se obtuvieron tres tablas de equivalencias para cada par de matrices.

Se utilizó un paquete para R llamado “equate” para los métodos clásicos el cual contiene funciones para ecuaciones no-TRI bajo diseño de grupos aleatorios y no equivalentes. Para el método media-sigma se utilizó el paquete “eRm” calculando los vectores de dificultades de cada nueva prueba simulada.

Una vez calculados los pares de puntuaciones con cada uno de los métodos de equiparación que conformaban la tabla de conversión, se obtuvo el error típico para cada

puntaje equiparado, por el método bootstrap, luego se promediaron los errores de cada estimación a través de las réplicas, obteniendo un error general de equiparación para cada método en cada matriz.

Análisis de datos

Dado el objetivo de la presente investigación, comparar la magnitud de error y observar la estabilidad de cada uno de los métodos a través de las diferentes aplicaciones, el análisis de datos se realizó mediante los estadísticos descriptivos de los errores de equiparación de cada uno de los métodos. Se calcularon los errores mínimos y máximos de todas las réplicas, sus Medias y Desviaciones Estándar, además se generaron gráficos de las distribuciones de frecuencia de las matrices que arrojó el máximo y el mínimo error de equiparación en cada método, finalmente se estimó el Coeficiente de Variación del error de cada método con el objetivo de identificar aquel que generara menor error en cada matriz de probabilidad de respuesta, así como que conservara sus resultados consistentes en las 1000 matrices.



UNIVERSIDAD NACIONAL DE COLOMBIA

RESULTADOS

En la tabla 1, se observan los resultados posteriores a aplicar los procedimientos de equiparación seleccionados de la Teoría Clásica de los Test, así como el error estándar de equiparación para cada puntaje estimado con el Método Bootstrap. La columna denominada puntaje corresponde al de una persona con limitación visual y los denominados con los métodos es el puntaje que le correspondería en la forma de las personas sin limitación visual.

Tabla 1. Tabla de equivalencia de los Métodos de la Teoría Clásica de los Test, Base de Datos 2008 -2

Puntaje	Tucker	Levine	Error_Tucker	Error_Levine
0	0,892	1,100	0,526	1,237
1	1,862	1,963	0,470	1,097
2	2,831	2,826	0,417	0,961
3	3,800	3,690	0,368	0,827
4	4,770	4,553	0,323	0,699
5	5,739	5,416	0,286	0,579

Equiparación de Puntuaciones: Limitación Visual 66

Puntaje	Tucker	Levine	Error_Tucker	Error_Levine
6	6,709	6,279	0,259	0,475
7	7,678	7,142	0,247	0,399
8	8,647	8,005	0,250	0,368
9	9,617	8,869	0,269	0,394
10	10,58	9,732	0,301	0,466
11	11,55	10,59	0,341	0,568
12	12,52	11,45	0,388	0,686
13	13,49	12,32	0,439	0,814
14	14,46	13,18	0,493	0,947
15	15,43	14,04	0,550	1,084
16	16,409	14,91	0,607	1,223
17	17,373	15,77	0,666	1,364
18	18,34	16,63	0,726	1,505
19	19,31	17,50	0,786	1,648
20	20,28	18,36	0,847	1,791
21	21,24	19,22	0,908	1,935
22	22,21	20,08	0,970	2,079

Puntaje	Tucker	Levine	Error_Tucker	Error_Levine
23	23,18	20,95	1,032	2,224
24	24,15	21,81	1,094	2,369

En la figura 2 se observa la forma de la distribución de los errores para ambos métodos, en general el método de Tucker tiene menor error de equiparación en relación con el puntaje de la prueba, sin embargo, se evidencia que ambos métodos tienen un menor error en puntuaciones medias totales de la prueba que en los extremos donde el error es elevado.

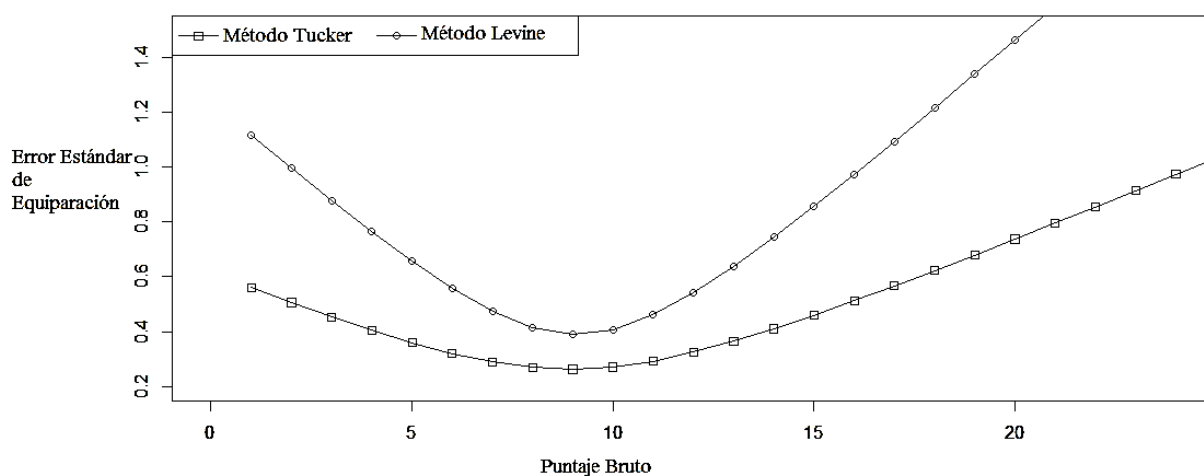


Figura 2. Gráfica de comparación del Error estándar de Equiparación para los Métodos Tucker y Levine.

En la tabla 2 se observan la dificultad estimada para la Base 2008 – 2, en la primera columna tenemos dificultad de los 24 ítems para esa simulación en la población con limitación visual, la segunda la dificultad de esa simulación para la población sin limitación visual y en

la tercera columna se encuentra la dificultad estimada con el método Media/Sigma si las personas con limitación visual resolvieran la prueba de referencia.

Tabla 2. Dificultad real y estimada para el método media / sigma.

Ítem	Dificultad		
	Con limitación visual	Sin limitación visual	Estimada
1	0.694	0.707	0.774
2	0.200	0.067	0.223
3	0.148	0.068	0.166
4	0.838	0.890	0.934
5	0.597	0.597	0.666
6	0.148	0.011	0.166
7	0.123	0.112	0.137
8	0.948	0.669	0.010
9	0.200	0.262	0.223
10	0.492	0.520	0.549
11	0.451	0.644	0.503
12	0.066	0.537	0.074
13	0.239	0.656	0.267
14	0.790	0.010	0.881
15	0.011	0.840	0.011

Ítem	Dificultad		
	Con limitación visual	Sin limitación visual	Estimada
16	0.200	0.064	0.223
17	0.948	0.439	0.010
18	0.492	0.402	0.549
19	0.362	0.161	0.404
20	0.011	0.409	0.012
21	0.492	0.437	0.549
22	0.705	0.661	0.786
23	0.180	0.719	0.201
24	0.239	0.592	0.267

En relación con los errores estándar de la aplicación 2008-2 forma B arrojados por cada método de equiparación se encontró que en el Método de Tucker fue de 0,54, en el Método de Levine 1,11 y en el de Media Sigma 0,00065. Es decir que en la base de datos real se evidencia una mejor estimación en éste último Método, seguido del de Tucker y con un mayor error general de equiparación el de Levine.

En el momento de realizar los procedimientos de equiparación y estimar los errores se observó que el método Levine arrojaba error con la imposibilidad de calcular los puntajes de equiparación en algunas matrices, por lo que se decidió aumentar la cantidad de matrices hasta 1050, con el fin de reemplazar aquellas que no tenían estimación dadas las características de

su distribución y así analizar igual cantidad de datos en todos los métodos de equiparación (Ver Anexo A).

En la Tabla 3 se encuentran los estadísticos descriptivos de los errores generales de equiparación para cada método. Se evidencia que el Método Media/ Sigma presenta menor rango de error de acuerdo con los mínimos y los máximos a lo largo de todas las réplicas, seguido del Método de Tucker y finalmente el Método de Levine que evidencia los errores de equiparación más elevados. El coeficiente de variación es menor en el Método de Tucker.

Tabla 3. Estadísticos descriptivos de los errores de equiparación para cada uno de los métodos.

Método	Mínimo	Máximo.	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%

El método que mejor se comporta a lo largo de las simulaciones es el método lineal de Tucker, sin embargo, el método Media-Sigma a pesar de reflejar el menor rango de error, tiene el coeficiente de variación se ve afectado por tener una desviación estándar tendiente a cero. El método Levine además refleja que posee mayor error que los otros 2 métodos, y su cadena parece no converger. El coeficiente de variación fue de 11,9 para el método Tucker, 31,82 para el Método Levine y 19,04 para el de Media-Sigma. En la figura 3 se evidencia la magnitud del error para los tres métodos en las matrices de respuesta.

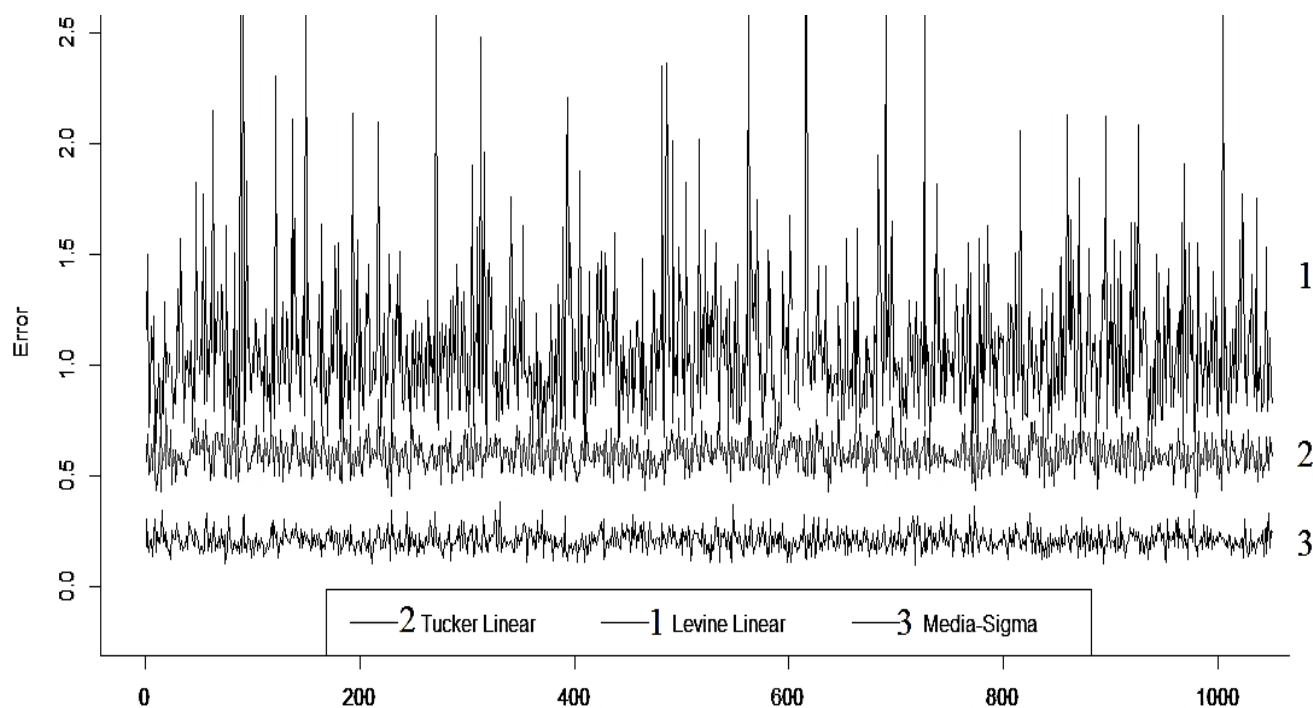


Figura 3. Magnitud de error de equiparación para cada método de equiparación en las réplicas

Estos resultados indican que el método de Levine arrojó mayor error de equiparación en comparación con los métodos de Tucker y Media – Sigma, de igual manera fue el método más inestable a lo largo de la simulación de aplicaciones. Por su parte, el método Media- Sigma arrojó menor error de equiparación en cada matriz de respuestas en relación con los métodos clásicos.

Se obtuvo la distribución de frecuencias para los errores mínimos y máximos de cada uno de los métodos con el fin de identificar las posibles causas de las diferencias.

En las figuras 4, 5 y 6 se presentan las distribuciones de las frecuencias de los puntajes en las matrices de respuesta que evidenciaron el error mínimo de equiparación para cada uno de los métodos. Solamente se presentan las relacionadas con los puntajes de las personas con

limitación debido a que tanto para los errores mínimos y máximos las distribuciones de las personas sin limitación visual presentan el comportamiento similar al de una curva normal.

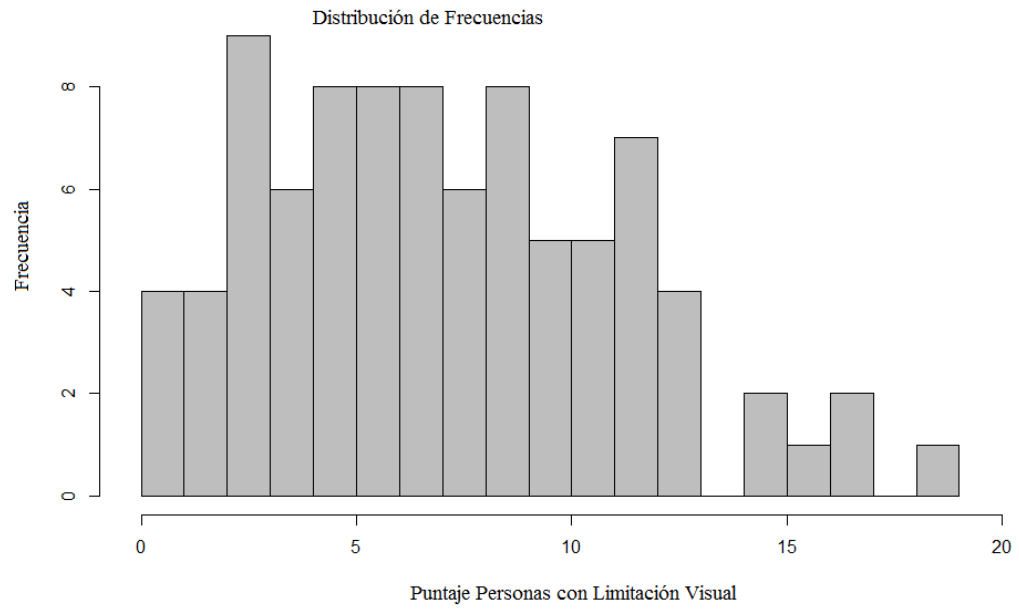


Figura 4. Distribución de frecuencia en matriz de respuesta que arrojó el mínimo de error de equiparación en el Método de Tucker.

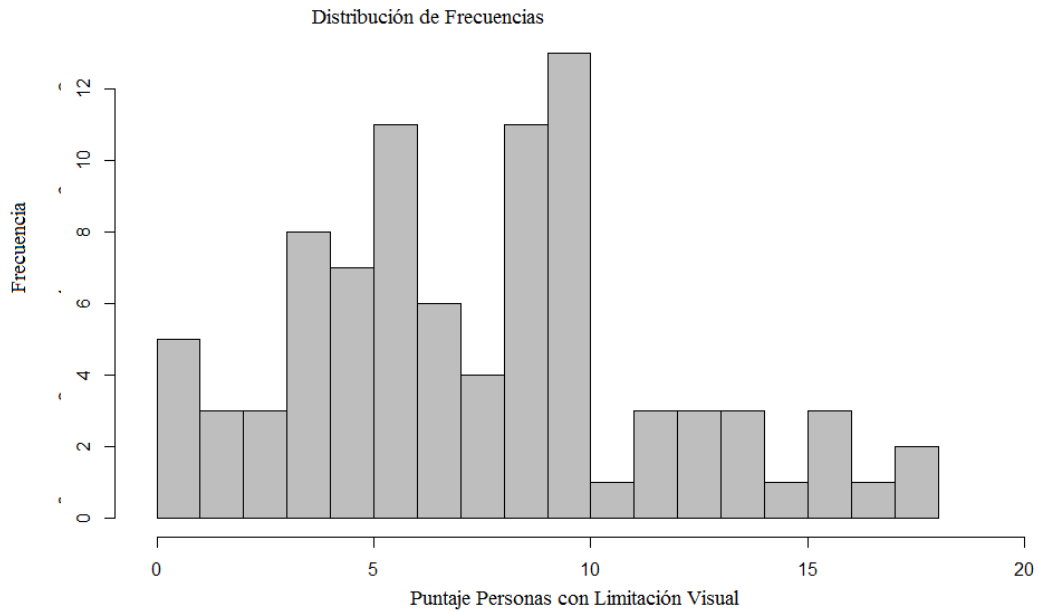


Figura 5. Distribución de frecuencia en matriz de respuesta que arrojó el mínimo de error de equiparación en el Método de Levine.

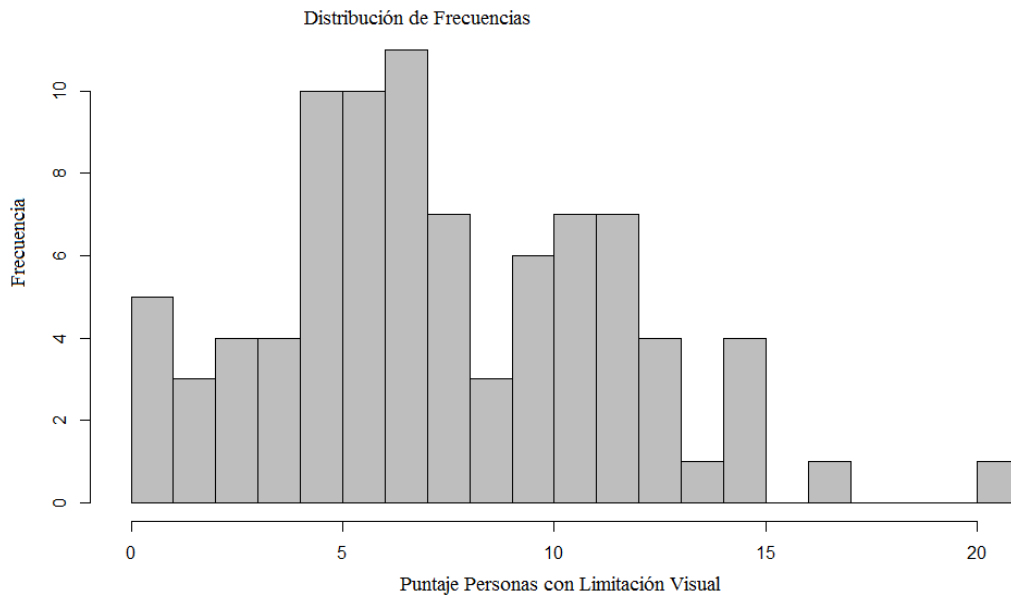


Figura 6. Distribución de frecuencia en matriz de respuesta que arrojó el mínimo de error de equiparación en el Método de Media/Sigma.

En las figuras 7, 8 y 9 se presentan las distribuciones de las frecuencias de los puntajes en las matrices de respuesta que evidenciaron un mayor error de equiparación para cada uno de los métodos.

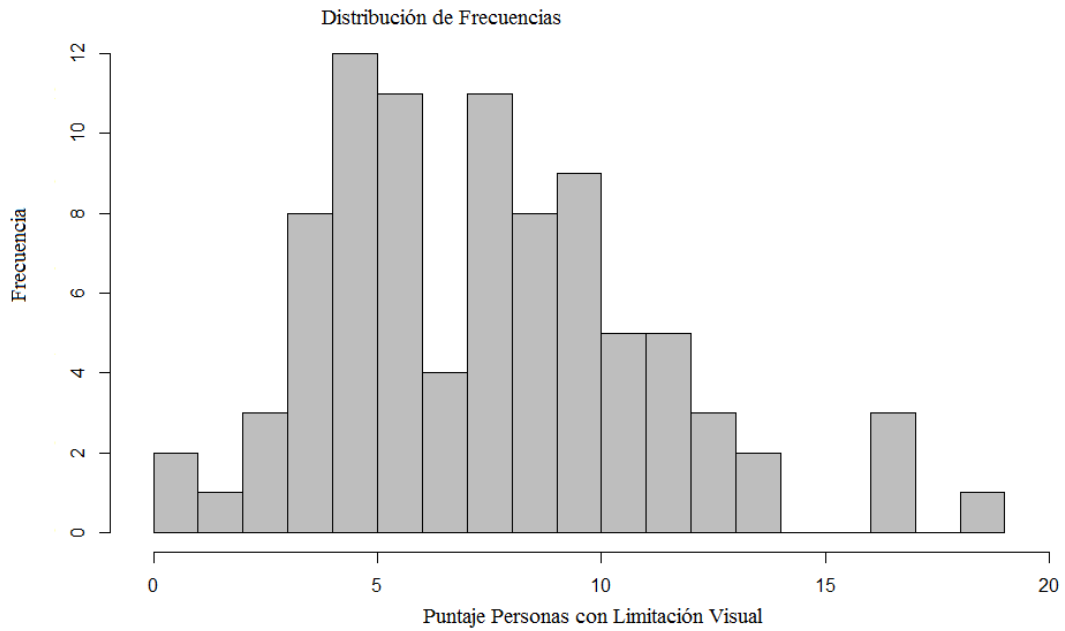


Figura 7. Distribución de frecuencia en matriz de respuesta que arrojó el máximo de error de equiparación en el Método de Tucker.

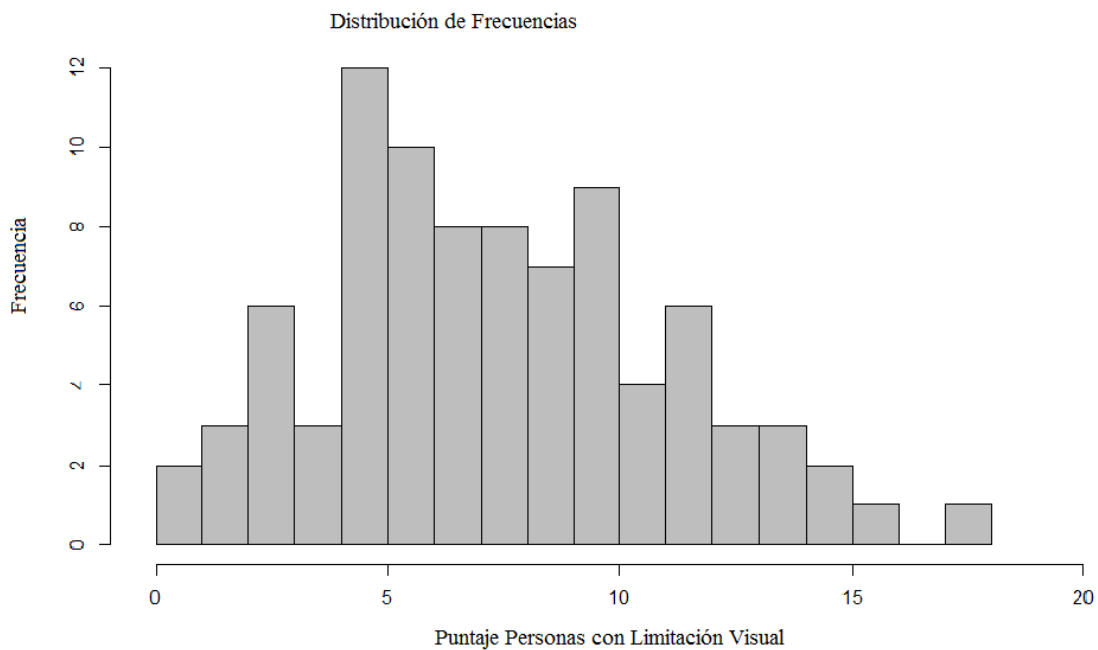


Figura 8. Distribución de frecuencia en matriz de respuesta que arrojó el máximo de error de equiparación en el Método de Levine.

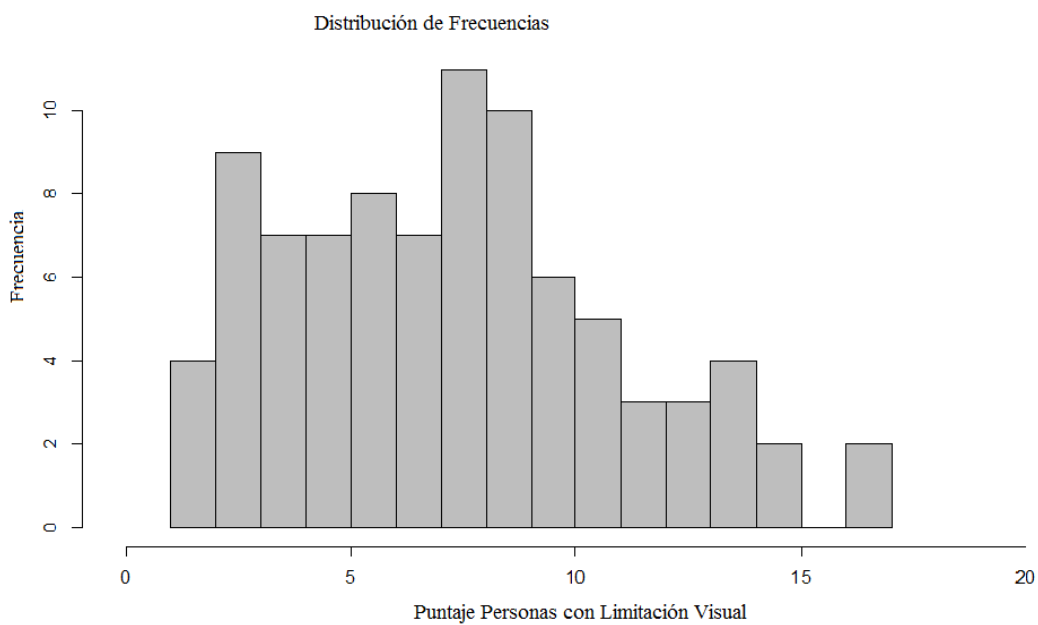


Figura 9. Distribución de frecuencia en matriz de respuesta que arrojó el máximo de error de equiparación en el Método de Media/Sigma.

En general, se observa que las distribuciones de las matrices de respuesta caracterizadas por arrojar errores mínimos en los procedimientos de equiparación contienen datos en todos los posibles puntajes que se pueden obtener, mientras que las distribuciones de las matrices caracterizadas por arrojar errores de magnitud alta en los procedimientos de equiparación evidencian ausencia de puntajes a lo largo del continuo, especialmente en los referentes a la obtención de puntuaciones elevadas en las pruebas. En todo caso, se observa que la mayoría de las distribuciones llegan hasta una puntuación máxima de 20 cuando el puntaje máximo posible es de 24. Lo anteriormente descrito se puede aclarar con la figura 10, la cual representa la distribución de los errores más alta en toda la simulación correspondiente al método de Levine.

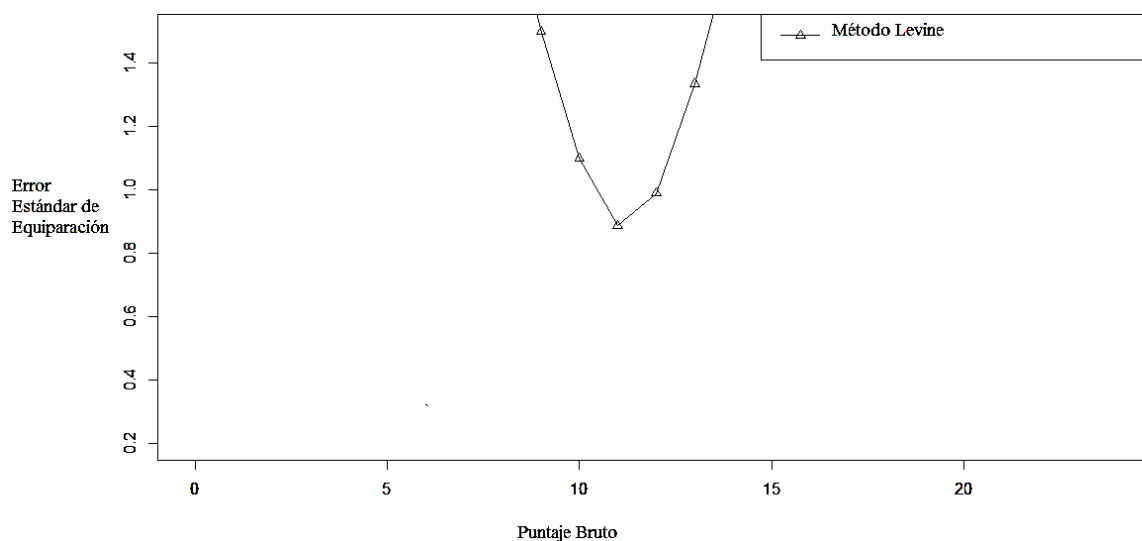


Figura 10. Distribución del error máximo de equiparación en toda la simulación con respecto al puntaje.

En esta gráfica se confirman los resultados evidenciados en las distribuciones de frecuencia de los puntajes en cuanto a los errores mínimos y máximos, en este caso el Método de Levine

Equiparación de Puntuaciones: Limitación Visual 77

arrojó el error de equiparación más alto de toda la simulación debido a que la mayoría de los puntajes se agrupan aproximadamente entre 8 y 15 cuando el máximo es 24.



UNIVERSIDAD NACIONAL DE COLOMBIA

DISCUSIÓN

Desde que se inició la medición y evaluación de características humanas, se han planteado diferentes interrogantes con respecto a la forma de realizarla, su fiabilidad y validez. Las discusiones han girado en torno a la posibilidad de acceder a aquellos atributos psicológicos por medio de preguntas que evidenciarían la magnitud del atributo en una persona. Paralelamente, los esfuerzos por garantizar evaluaciones equitativas han tomado diferentes direcciones y se han analizado aspectos que abarcan la construcción, aplicación, corrección e interpretación de las pruebas.

El enlace de puntuaciones y en especial el procedimiento denominado equiparación es uno de esos intentos por garantizar la comparabilidad de los puntajes obtenidos por las personas, más aún cuando con base en éstas se toman decisiones para su futuro educativo o laboral. Este tema ha sido de amplia divulgación desde que Angoff (1984) lo propusiera e iniciara una serie de estudios con respecto al tratamiento que debían tener esos resultados de las evaluaciones, en cuanto a su equivalencia dadas las necesidades y exigencias que se genera en situaciones de evaluación permanente. El incremento de literatura sobre el tema se ha dirigido a refinar tanto los procedimientos y ajustar las fórmulas que permitirán poner un test en términos de las unidades de otro cuando existen varias formas de prueba, pero también al descubrimiento de

nuevos diseños y procedimientos que se adapten a las diferentes situaciones de evaluación como lo exponen ampliamente Kolen & Brennan (2004). Por otra parte, también se encuentra una compilación de trabajos dedicados a evaluar los diferentes procedimientos que se han generado Kane y cols. (2010), así como las variaciones de los procedimientos en función del diseño de recolección de datos utilizado. Sin embargo, los esfuerzos por ampliar el tema no se ha dirigido únicamente al detalle técnico de los procedimientos, sino a reflejar la importancia de estos procedimientos en el uso de pruebas (Navas, 2000).

La investigación documental realizada en esta investigación pone de manifiesto cómo a pesar de esta amplia difusión sobre este tipo de procedimientos, aún es mínimo el uso de éstos que se realiza en el contexto colombiano (Cervantes, 2012). Esto se evidencia, en el tradicional uso de la interpretación de las puntuaciones de exámenes, a pesar de las acomodaciones que se realizan para las personas que por determinadas características o condiciones no pertenecen al grupo mayoritario de la población.

En Colombia, los procedimientos de equiparación se han empezado a implementar desde hace menos de una década en el Instituto Colombiano para la Evaluación de la Educación (ICFES), con el propósito de lograr comparabilidad entre las diferentes aplicaciones de los exámenes de estado que se realizan año tras año y así observar el cambio en el logro educativo. Sin embargo, aún no se han establecido procedimientos que garanticen una interpretación equitativa de los puntajes en aquellas poblaciones que reflejan resultados por debajo del promedio a lo largo de la historia evaluativa (Cervantes, 2012).

El anterior es el caso de la población con limitación visual, donde se evidencia un desempeño significativamente inferior en los exámenes de estado, así como DIF en algunos de los ítems, donde los Exámenes de Estado no contemplan las características de sus subpoblaciones (Herrera y cols., 2012). La forma de apoyar a estas personas se han basado en acomodaciones, principalmente en el apoyo de un lector en el momento de la evaluación, el cual dispone de una serie de competencias para realizar una aplicación lo más equitativa posible (INCI & UPN, 2009). Sin embargo, con respecto a los datos o a las puntuaciones que se desprenden de estas evaluaciones, aún falta asumir una postura que permita ajustar las diferencias en los puntajes con respecto la población mayoritaria.

En relación con los objetivos de ésta investigación, en la revisión teórica se expusieron las dificultades y necesidades referentes a las evaluaciones realizadas a las personas con limitación visual, así como la importancia de utilizar procedimientos de equiparación con el fin de ajustar y situar en una misma escala comparativa las puntuaciones de estos dos grupos poblacionales denominados con y sin limitación visual. El marco conceptual elaborado permitió el desarrollo de elementos fundamentales sobre el diseño, métodos y evaluación de la equiparación, que brindaron los elementos necesarios para seleccionar tanto el diseño como los procedimientos que serían susceptibles de comprobación, así como que cumplieran con las características del problema planteado.

Al aplicar los procedimientos de equiparación se generaron las tablas de equivalencia tanto para la base de datos real como para las matrices simuladas, de las cuales se evidenció que al hacer el ajuste de equiparación, los puntajes incrementan en una o dos unidades para las personas con limitación visual con referencia a las personas sin limitación visual. Sin

embargo, se evidencia que hay una inversión en los puntajes principalmente al utilizar los métodos de la Teoría Clásica de los Tests: *Tucker y Levine* a partir de un puntaje de 10, esto es explicado por la imposibilidad de contar con puntajes superiores a 12 en las personas con limitación visual, para esta prueba en la que el puntaje bruto máximo es de 24. Esta ausencia de puntajes hace que la relación de equiparación se invierta a partir de un punto en la distribución. Esto sustenta los hallazgos de Herrera y cols. (2012), donde se evidencia que las personas con limitación responden menos de la mitad de la prueba y se ubican en niveles bajos de habilidad.

Este inconveniente, es superado por el Método Media/Sigma, debido a que utiliza una medida más estable que es la dificultad, entendida como la probabilidad de acierto al ítem dada la habilidad de la persona, disminuyendo el cambio en los resultados del procedimiento de equiparación por la influencia de la cantidad de respuestas correctas que puede tener la persona.

Luego de la generación de las tablas de equivalencia correspondientes, se realizó el cálculo de los errores de cada una de los procedimientos de equiparación. El cálculo del error general de equiparación en cada matriz de probabilidad de respuesta, evidenció la magnitud de error que introducía cada procedimiento de equiparación dadas las condiciones particulares de la matriz de respuesta, para la población con limitación visual la característica fundamental consistía en acertar menos de la mitad de las preguntas abordadas.

Al analizar la magnitud del error de equiparación se evidencia que el Método Media/Sigma es el que arroja menor error de equiparación probablemente por los cálculos internos que se

evidencian invariantes de la población y los patrones de respuesta, de acuerdo con lo señalado por Zhu (2001). El método Levine arroja magnitudes de error elevadas que también se evidencian en la tabla de equivalencia, llegando incluso a invertirse completamente la relación de equiparación, es decir que este método es altamente susceptible a las variaciones en las distribuciones de frecuencia de las respuestas propiciadas por los examinados, lo cual puede indicar una inestabilidad en diferentes poblaciones, contextos y condiciones de evaluación, en concordancia con lo señalado por Kane, Mroch, Suh & Ripkey (2010).

De la misma manera, el comportamiento del coeficiente de covariación que tenía como propósito mostrar la estabilidad de los métodos dadas las diferentes posibilidades de matrices de respuesta, evidenció una mayor estabilidad en el Método de Tucker, relacionado con lo sustentado por Kolen & Brennan (2004) en relación con la pertinencia del uso de este método en muestras pequeñas; sin embargo, en los tres métodos se observa concordancia con los hallazgos de Livingston (2004) & Von Davier & Kong (2003), al confirmarse que en los extremos superiores e inferiores de la distribución de los puntajes, el error estándar de equiparación tiende a ser mayor.

Estos resultados sugieren que los métodos 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. En ese sentido, los resultados de esta investigación indican que bajo las condiciones establecidas para generar equivalencias entre personas con y sin limitación visual, el método que permite hacer intercambiables los puntajes con mayor precisión es el Método

Media/ Sigma, debido a que al realizar la conversión arroja menor error de equiparación para cada puntaje, apoyando las investigaciones de Zhu (2001).

Sin embargo, el uso de éstos procedimientos puede verse afectado por la cantidad de aciertos de las personas con limitación visual que suele ser muy bajo en relación con las puntuaciones máximas. Por lo anterior, pueden ser utilizados a mediano y largo plazo como herramientas de ajuste que permitan una interpretación más equitativa de las puntuaciones; aunque, los esfuerzos deben dirigirse a la elaboración de pruebas objetivas diseñadas de acuerdo a las necesidades del evaluado, en este caso de personas con limitación visual. Este estudio evidencia la necesidad de generar mayor investigación sobre la calidad de la evaluación en población con limitación visual, así como a otros grupos minoritarios, debido a que los instrumentos de evaluación que existen actualmente no permiten obtener una medida precisa del atributo de estos individuos, así como una medida de todo el continuo de su habilidad, lo que se evidencia en la imposibilidad de encontrar sujetos con puntajes altos en este tipo de evaluaciones.

En ese sentido, el reto en la evaluación está en utilizar esta información para identificar aquellas características que hacen de la evaluación actual una herramienta insuficiente para acceder a la información de este tipo de poblaciones; en el caso de las personas con limitación visual cómo las adaptaciones de la prueba y las diferentes acomodaciones, inciden en la calidad de las respuestas frente a las situaciones problema. La decisión de asumir la forma de presentación del examen como una forma alterna de la prueba, pone de manifiesto que el diseño de la evaluación no lo contempla como tal a pesar de que el mecanismo de recepción de la información es claramente diferente y que aún es desconocido el efecto sobre los

procesos cognoscitivos involucrados, así como en el énfasis que hacen estas personas en recursos propios como la atención y la memoria.

Se resalta la pertinencia de los estudios tipo MonteCarlo basado en simulaciones, que permiten establecer las condiciones experimentales necesarias. En el caso de la presente investigación las condiciones experimentales se establecieron de acuerdo con la realidad de la aplicación, y que por medio de la manipulación de diferentes variables permite observar el comportamiento de la variable de interés, en ese sentido se propone que para futuras investigaciones sean manipuladas otras condiciones como los razones de tamaños de muestra, así como la posibilidad de contar con datos de los sujetos a lo largo del continuo de habilidad, con el objetivo de probar la confiabilidad y estabilidad de los métodos de equiparación y así confirmar la bondad de su uso, en diferentes condiciones de probabilidad de respuesta.

Finalmente, la continuidad de este tipo de investigaciones se hace fundamental en vías de mejorar la calidad de la evaluación, ya sea porque se identifican falencias en los instrumentos diseñados y utilizados, así como para probar metodologías que permitan atenuar los efectos que genera el hecho de considerar como iguales los diferentes sectores de la población, en especial en un país que está atravesado por una amplia diversidad económica, política y social que genera características diferenciales en los individuos, lo que lleva a pensar que la evaluación debe enfocarse en el individuo más que en el grupo.



UNIVERSIDAD NACIONAL DE COLOMBIA

REFERENCIAS

- Albano, A. (2011). *Statistical Equating Methods*. Recuperado de <http://cran.r-project.org/web/packages/equate/vignettes/equatevignette.pdf>
- Allende, F. & Condemarín, M. (1992). *La Lectura: teoría evaluación y desarrollo*. Santiago de Chile: Andrés Bello.
- Angoff, W. H. (1984). *Scales, norms, and equivalent scores*. Princeton, NJ: Educational Testing Service
- Berrío, A. I. (2008). *La razón de tamaños de muestra y desajustes del modelo en la detección de ítems con funcionamiento diferencial mediante el procedimiento diferencia de la dificultad*. Tesis de Maestría. Universidad Nacional de Colombia, Bogotá.
- Bigelow, a.e. (1986). The development of reaching in blind children. *British Journal of Developmental Psychology* , 4, 355-366
- Bach, P. & Bach, R. (2005) Mecanismos de sustitución cerebral. *Revista Plasticidad y restauración neurológica*, 4, 1 - 2.
- Bogoya, D. (2003). *La llave: estándares, evaluación, mejoramiento*. En Al tablero N° 19. On line

Castellanos, W., Álvarez, S., Pèrez, C., Carriòn, C. & Ladino, A. (2008). *Desarrollo humano de la población con limitación visual por departamentos*.

Cervantes, V. Comunicación personal, 10 de Agosto, 2012.

Cohen, A.S., & Kim, S.H. (1998). An investigation of linking methods under the graded response model. *Applied Psychological Measurement*, 22,2, 116-130.

Constitución(1991). Constitución Política de Colombia. Gaceta Constitucional.

Correa, L. (2005). *Más allá de la lectura*. Instituto Nacional para ciegos INCI. Bogotá.

Cook, L. & Eignor, D. (1991). *IRT Equating Methods*. Princeton, NJ: Educational Testing Service.

Departamento Administrativo Nacional de Estadística (2005). Censo Nacional.

Dorans, N. & Holland, P. (2000). Population invariance and the equatability of tests: Basic theory and the linear case. *Journal of Educational Measurement*, 37,4, 281–306.

Gau, H. (2004). *The effect of different anchor tests on the accuracy of test equating for test adaptation*. Disertation. Ohio: The Faculty of de College of Education.

Gulliksen, H. (1950). *Theory of mental tests*. Wiley, New York.

Han, T., Kolen, M.,Pohlmann, J. (1997).A Comparison Among IRT True – and Observed – Score Equatings and Traditional Equipercentile Equating. *Applied Measurement in Education*, 10,2,105-121.

Harris, D. J., & Crouse, J. D. (1993). A study of criteria used in equating. *Applied Measurement in Education*, 6,3, 195–240.

Herrera, A. N. (2005). *Efecto del tamaño de muestra y la razón de tamaños de muestra en la detección de funcionamiento diferencial de los ítems*. Tesis de Doctorado en Evaluación y Tecnología y Tecnología Informática en Ciencias del Comportamiento. Universidad de Barcelona, España.

Herrera, A., Soler, M., Espinosa, A., Lancheros, L. & Jiménez, G. (2012). *Procedimiento para establecer equivalencia en las puntuaciones de pruebas de aplicación masiva, en personas con y sin limitación visual*. Inédito.

Holland, P & Dorans, N. (2006). *Linking and Equating*. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 187- 220). Westport, CT: Greenwood.

Instituto Nacional para Ciegos -INCI & Universidad Pedagógica Nacional - UPN. (2009). *“Miradas Valiosas” Lectores para personas con limitación visual más que una oportunidad*. Bogotá. Recuperado de: http://www.inci.gov.co/centro_documentacion/cartillas/Miradas_valiosas.pdf

Kane, M., Mroch, A., Suh, Y. & Ripkey, D. (2010). Linear Equating for the Neat Design: A Rejoinder and Some Further Comments. *Measurement*, 8, 27–37

Kane, M., Mroch, A., Suh, Y., Ripkey, D. (2009). Linear Equating for the NEAT Design: Parameter Substitution Models and Chained Linear Relationship Models. *Measurement*, 7.

Keller, L., Keller, R. & Parker, P. (2011). The Examination of the Classification of Students into Performance Categories by Two Different Equating Methods. *The Journal of Experimental Education*, 2011, 79, 30–52

- Kim, S., Von Davier, A. & Haberman, S. (2008). Small sample equating using a synthetic linking function. *Journal of Educational Measurement*, 45, 325–342.
- Kim, S., Von Davier, A., & Haberman, S. (2011). Practical Application of a Synthetic Linking Function on Small-Sample Equating. *Applied Measurement in Education*, 24, 95–114
- Kolen, M. & Brennan, R. (2004). *Test equating, scaling, and linking: methods and practices*. New York: Springer.
- Lawrence, I. & Dorans, N. (1990). Effect on Equating Results of Matching Samples on an Anchor Test. *Applied Measurement in Education*, 3,1.
- Ley 30 (1992). Ley General de Educación Superior. Diario Oficial de la República de Colombia.
- Ley 115 (1994). Ley General de Educación. Diario Oficial de la República de Colombia.
- Levine, R. (1955). *Equating the score scales of alternate forms administered to samples of different ability* (Research Bulletin 55-23). Princeton, NJ: Educational Testing Service.
- Livingston, S. (2004). *Equating Test Scores* (Without IRT). Princeton, NJ: Educational Testing Service.
- Livingston, S. A., Dorans, N. J., & Wright, N. K. (1990). What combination of sampling and equating methods works best? *Applied Measurement in Education*, 3,1, 73-95.
- Lord, F.M. (1980). *Applications of Item Response Theory to practical testing problems*. On line.
- Muñiz Fernández J. (1997). *Introducción a la Teoría de Respuesta a los Items*. Madrid: Pirámide.

- Muñiz, J (1998). *Teoría clásica de los tests*. Madrid: Pirámide.
- Navas, M. J. (1996). *Equiparación de puntuaciones*. En Muñiz, J. (Ed.), *Psicometría* (pp. 293-369). Madrid: Editorial Universitas, S. A.
- Navas, M. J. (2000). Equiparación de puntuaciones: exigencias y retos de cara al futuro. *Metodología de las Ciencias del Comportamiento* 2,2,151-165
- Ochaita, E. (1993) *Ceguera y desarrollo psicológico*. Madrid: Alianza
- Pacheco, S. (2006). *Equiparación de puntuaciones. Revisión conceptual y metodológica*. Tesis de pregrado.
- Parshall, C., Houghton, D., & Kromrey, J. (1995). Equating error and statistical bias in small sample linear equating. *Journal of Educational Measurement*, 32, 37–54
- Rasch, G. (1960). *Probabilistic Models for some Intelligence and Attainment Tests*. Danish Institute for Educational Research, Copenhagen.
- Rodríguez, O. (2007). Equiparación de puntuaciones con TRI y TCT en una prueba de ingeniería. Memorias *Segunda reunión Regional Norte, Centroamérica y Caribe de Evaluación Educativa*.
- Santana, A. C. (2009). *Efecto de la Razón de Tamaños sobre la Detección del Funcionamiento Diferencial del ítem mediante Regresión Logística*. Tesis de grado de maestría, Universidad Nacional de Colombia, Bogotá.
- Servicio Nacional de Pruebas. (1992). *El Servicio Nacional de Pruebas y sus programas de Evaluación educativa*. Documento S.N.P. Santa fe de Bogotá: Documento institucional.

- Skaggs, G. (2005). Accuracy of random groups equating with very small samples. *Journal of Educational Measurement*, 42, 309–330.
- Sierra, G. & Vanegas, N. (2001) *Construcción del discurso. Hacia la formación de un profesional competente en el uso del lenguaje*. Bogotá: EAN.
- Von Davier, A. & Kong, N.(2003a). *A Unified Approach to Linear Equating for the Non-Equivalent Groups Design*. Princeton, NJ: Educational Testing Service.
- Von Davier, A. (2003b). *Notes on linear equating methods for the Non-Equivalent Groups design* (ETS RR-03-24). Princeton, NJ: Educational Testing Service.
- Von Davier, A. A., Holland, P. W., & Thayer, D. T. (2004). *The kernel method of test equating*. New York: Springer Verlag.
- Zhu, W. (2001). An emprirical investigation of Rasch equating of motor function task. *Adapted Physical Activity Quartely*, 72-89.



UNIVERSIDAD NACIONAL DE COLOMBIA

ANEXO A

Tabla 4.

Errores en cada matriz para las 1050 réplicas.

	Tucker	Levine	Media-Sigma
[1,]	0.5675257	1.3194207	0.1852858
[2,]	0.5968180	1.1620824	0.3054535
[3,]	0.6910229	1.4998410	0.1964790
[4,]	0.5079835	0.7239741	0.1545391
[5,]	0.5210378	0.8458251	0.2075427
[6,]	0.6449163	1.1720910	0.2155911
[7,]	0.5283926	1.0254483	0.1345500
[8,]	0.7307380	1.2223856	0.2263250
[9,]	0.4889366	0.6043144	0.3060636
[10,]	0.4747605	0.6906776	0.1975836
[11,]	0.4348210	0.8435092	0.1571296
[12,]	0.4741348	0.7874669	0.2501781

Equiparación de Puntuaciones: Limitación Visual 92

[13,]	0.5383737	1.0052337	0.2016856
[14,]	0.6744460	0.8815038	0.1932385
[15,]	0.4284840	0.6300467	0.2465563
[16,]	0.6230536	0.9316978	0.3418843
[17,]	0.5706351	0.7772957	0.1996138
[18,]	0.6399504	1.2881362	0.2340209
[19,]	0.5345375	0.9202440	0.1822463
[20,]	0.8076960	0.9852721	0.2450561
[21,]	0.6649465	1.0588276	0.1847072
[22,]	0.5571066	0.9454366	0.1659883
[23,]	0.5475549	1.0495114	0.1633633
[24,]	0.7085518	NA	0.1218584
[25,]	0.4589885	0.9960735	0.2520016
[26,]	0.6194748	0.7573002	0.2273427
[27,]	0.5569126	0.8963385	0.2133506
[28,]	0.6151577	0.8263156	0.2156834
[29,]	0.4764312	0.8649708	0.2242707
[30,]	0.5297463	1.0119697	0.2856736

Tucker Levine Media-Sigma

[31,]	0.5991985	1.3424975	0.2281270
[32,]	0.5918257	0.9699019	0.2361526

Equiparación de Puntuaciones: Limitación Visual 93

[33,]	0.5511994	1.5719842	0.1665142
[34,]	0.6064331	1.2878915	0.2084999
[35,]	0.5510077	1.0983009	0.1635552
[36,]	0.5647322	0.7412305	0.1791143
[37,]	0.5365218	0.9532267	0.1820274
[38,]	0.5158650	1.0371231	0.1848165
[39,]	0.5014052	0.9807043	0.2256265
[40,]	0.5654428	1.0625377	0.1712531
[41,]	0.5709573	0.9001482	0.2926824
[42,]	0.5711416	1.0116791	0.2574383
[43,]	0.6636264	1.1136072	0.2505995
[44,]	0.6521316	0.7000093	0.2200349
[45,]	0.6048678	0.9880459	0.1946861
[46,]	0.7410790	0.8443091	0.1803745
[47,]	0.7571052	0.8409890	0.2618219
[48,]	0.5954453	1.8241379	0.2406459
[49,]	0.5953837	1.0264406	0.2254948
[50,]	0.7135561	NA	0.1950877
[51,]	0.5969489	0.9643151	0.1468171
[52,]	0.6364942	1.1156849	0.1840032
[53,]	0.6246595	0.9801806	0.1844791

Equiparación de Puntuaciones: Limitación Visual 94

[54,]	0.7030154	1.7695596	0.1531939
[55,]	0.5248981	0.8218555	0.1686873
[56,]	0.5712422	0.8781342	0.2364627
[57,]	0.7559026	1.5345369	0.2853923
[58,]	0.6247086	0.8382833	0.3321345
[59,]	0.6113304	1.0682580	0.1622171
[60,]	0.6387493	0.9505171	0.1891935
[61,]	0.4812514	1.2789313	0.2418689
[62,]	0.5316908	1.0615609	0.1581604
[63,]	0.6554058	2.1497244	0.2494466
[64,]	0.5384670	1.1371449	0.2904325
[65,]	0.5022482	0.7921648	0.2952479
[66,]	0.6800758	1.2077455	0.1622169
[67,]	0.6951167	1.0590755	0.1880552

Tucker

Levine

Media-Sigma

[68,]	0.6937234	1.3323379	0.1788038
[69,]	0.6263731	NA	0.2287810
[70,]	0.6152812	1.0440237	0.1742685
[71,]	0.6884657	1.3608114	0.2591884
[72,]	0.6880723	1.3112748	0.2111809
[73,]	0.5837032	0.8172672	0.1850946

Equiparación de Puntuaciones: Limitación Visual 95

[74,]	0.6957453	1.0415706	0.2346952
[75,]	0.5066495	0.6903230	0.1023003
[76,]	0.4970511	1.6275484	0.1311825
[77,]	0.6117817	0.8557467	0.1737426
[78,]	0.6706810	1.0058342	0.3206341
[79,]	0.6096621	0.8051995	0.2197655
[80,]	0.4863200	1.0653374	0.2047641
[81,]	0.6959789	0.6785544	0.1492875
[82,]	0.6218289	0.8228773	0.1840437
[83,]	0.6709124	1.5051164	0.2059218
[84,]	0.7174576	1.1489045	0.1895166
[85,]	0.4978111	0.7441913	0.1900184
[86,]	0.7955347	1.1584595	0.1615907
[87,]	0.4724369	0.6401610	0.2051171
[88,]	0.5615922	0.7831572	0.1516902
[89,]	0.6048186	2.7321390	0.2120241
[90,]	0.6152915	0.9709411	0.1559173
[91,]	0.6519482	3.1828758	0.2960848
[92,]	0.7039968	1.0827871	0.3220508
[93,]	0.6805055	0.7363384	0.1759696
[94,]	0.6589259	0.9264001	0.1953185

Equiparación de Puntuaciones: Limitación Visual 96

[95,]	0.6121711	1.8288818	0.1596678
[96,]	0.5216378	1.0772169	0.2284441
[97,]	0.5308666	1.1224423	0.2130105
[98,]	0.5341244	0.9015018	0.2062457
[99,]	0.5889796	0.8789485	0.1294343
[100,]	0.5884142	1.0524015	0.2088366
[101,]	0.6336240	NA	0.1510281
[102,]	0.6069036	0.9680934	0.1409240
[103,]	0.6130485	1.2089924	0.2086805
[104,]	0.7299472	1.1748279	0.2407393
[105,]	0.5585268	0.9506434	0.1614180

Tucker Levine Media-Sigma

[106,]	0.6848358	1.1187099	0.2100066
[107,]	0.6134444	1.0511118	0.1752913
[108,]	0.5635634	0.9123864	0.2183695
[109,]	0.5668794	0.9441785	0.1925933
[110,]	0.6196972	0.9926808	0.1866918
[111,]	0.5195687	0.6996360	0.1342876
[112,]	0.6714166	1.0167507	0.1501214
[113,]	0.5928216	1.2523105	0.1546199
[114,]	0.6528252	0.8373007	0.1976961

Equiparación de Puntuaciones: Limitación Visual 97

[115,]	0.5439401	0.8617130	0.1672667
[116,]	0.5589595	0.9599297	0.1848580
[117,]	0.6767430	0.6208690	0.2560043
[118,]	0.6806781	1.2159457	0.2851076
[119,]	0.6769076	0.8688519	0.1756765
[120,]	0.6280002	1.1821419	0.2998053
[121,]	0.5029633	0.7622881	0.2583069
[122,]	0.6772288	2.3072362	0.2226492
[123,]	0.5970272	1.0624419	0.2452468
[124,]	0.6320930	1.0203647	0.1314810
[125,]	0.5031674	0.8251499	0.1766270
[126,]	0.7081981	0.9945919	0.2184723
[127,]	0.6109525	1.1434319	0.2026445
[128,]	0.5899986	1.2858350	0.1891683
[129,]	0.4728492	1.0991063	0.1921647
[130,]	0.5902129	0.9304314	0.3044297
[131,]	0.6089700	0.9169307	0.2301489
[132,]	0.6497771	1.2404047	0.2315591
[133,]	0.6182611	1.1890020	0.2105673
[134,]	0.5146192	0.8742592	0.2027598
[135,]	0.5380892	0.7952734	0.1981506

Equiparación de Puntuaciones: Limitación Visual 98

[136,]	0.5783154	1.0359212	0.2366110
[137,]	0.7095784	2.1102819	0.2613928
[138,]	0.5957791	1.0729297	0.2394966
[139,]	0.6707461	1.5626911	0.2231554
[140,]	0.7296594	1.6628076	0.2321077
[141,]	0.6674698	0.9620647	0.2880060
[142,]	0.6109690	1.1370617	0.2244161
[143,]	0.5453256	1.1094736	0.2211724
[144,]	0.6010817	0.8853059	0.1749174
[145,]	0.4978886	0.8590162	0.2609800
[146,]	0.6461701	1.2229510	0.2342431
[147,]	0.6926283	1.0762068	0.1782408
[148,]	0.5661552	1.3140213	0.2071631
[149,]	0.5569085	0.7746585	0.2456449
[150,]	0.6277482	3.2931632	0.2099924
[151,]	0.5721867	1.1202767	0.2318968
[152,]	0.5736201	1.3696241	0.2039492
[153,]	0.5482442	1.0408179	0.2060735
[154,]	0.6774995	1.0460942	0.1604606
[155,]	0.5306440	0.7720817	0.1734956
[156,]	0.4878946	0.6238694	0.2611744

Equiparación de Puntuaciones: Limitación Visual 99

[157,]	0.5065622	NA	0.1870613
[158,]	0.7492476	0.9220730	0.2330680
[159,]	0.5409882	0.9495609	0.1806900
[160,]	0.6543343	0.9054097	0.2334455
[161,]	0.5903315	1.0726423	0.1917814
[162,]	0.6003815	1.3177808	0.2429431
[163,]	0.5785496	1.0213627	0.2043339
[164,]	0.5317877	0.7264141	0.1348851
[165,]	0.5516266	1.6329997	0.2103617
[166,]	0.7230428	1.1272928	0.1731035
[167,]	0.6393994	0.9804637	0.1861686
[168,]	0.6139269	0.8228158	0.1704547
[169,]	0.5800745	0.9720926	0.2388942
[170,]	0.5070456	0.6833308	0.2381577
[171,]	0.6353894	1.0319038	0.2593676
[172,]	0.4804428	0.9183407	0.2112159
[173,]	0.5715774	1.0808756	0.1999856
[174,]	0.5638207	1.2207510	0.2703919
[175,]	0.6333422	0.9659955	0.2379455
[176,]	0.5750964	0.9756565	0.1726979
[177,]	0.5115591	1.5403200	0.2222553

Equiparación de Puntuaciones: Limitación Visual 100

[178,]	0.6213698	1.3424045	0.2494508
[179,]	0.5708621	0.9414659	0.2433059
[180,]	0.6579970	1.5494590	0.1432235
[181,]	0.4873084	0.7672275	0.2405550
[182,]	0.4777655	0.6592767	0.1506266

Tucker

Levine

Media-Sigma

[183,]	0.5470111	1.3403739	0.1825876
[184,]	0.4705987	0.8594813	0.1714401
[185,]	0.7437863	0.6572923	0.1857227
[186,]	0.5998517	1.0482029	0.2519636
[187,]	0.6412564	0.9360570	0.1959304
[188,]	0.6654739	1.1911870	0.2137834
[189,]	0.5597159	0.8053088	0.2554461
[190,]	0.5661846	0.9257021	0.1638239
[191,]	0.7273655	0.7618811	0.2040852
[192,]	0.4786589	0.8589309	0.2821338
[193,]	0.5187536	1.2199510	0.1985995
[194,]	0.6006751	2.1352037	0.1908138
[195,]	0.5019052	0.8876146	0.1826696
[196,]	0.6119985	0.9891412	0.1427451
[197,]	0.5485307	1.2093132	0.2041603

Equiparación de Puntuaciones: Limitación Visual 101

[198,]	0.6296942	1.5646038	0.2038146
[199,]	0.5441392	0.7056169	0.2093393
[200,]	0.6765659	1.2502350	0.1606334
[201,]	0.5919142	0.8863751	0.1855518
[202,]	0.5189725	1.1470922	0.1507767
[203,]	0.5011723	0.8326688	0.2456585
[204,]	0.6367500	0.7964691	0.2465869
[205,]	0.6206010	1.0219279	0.2103664
[206,]	0.7109741	1.3220126	0.1611805
[207,]	0.6042086	1.1866245	0.1977458
[208,]	0.7356794	1.4517944	0.1577017
[209,]	0.6192720	0.9690005	0.2624839
[210,]	0.5382642	0.8655150	0.2249397
[211,]	0.5631689	0.8797737	0.1528088
[212,]	0.6256220	0.8932906	0.1031431
[213,]	0.5779978	1.0566402	0.1935095
[214,]	0.6146109	0.9145622	0.2122983
[215,]	0.5963366	0.9460403	0.1979324
[216,]	0.5761510	1.1748922	0.2337709
[217,]	0.8454757	2.0954001	0.2809658
[218,]	0.6044624	1.3376867	0.2687953

Equiparación de Puntuaciones: Limitación Visual 102

[219,]	0.5880823	0.8350244	0.1780571
[220,]	0.5957696	0.6792337	0.1705389
	Tucker	Levine	Media-Sigma
[221,]	0.5476264	0.8368346	0.1765612
[222,]	0.6051282	1.0579763	0.1778212
[223,]	0.7184199	1.2341114	0.2437964
[224,]	0.5484276	0.8575753	0.1762353
[225,]	0.5286012	0.9595232	0.2349604
[226,]	0.4496101	0.9364274	0.2879341
[227,]	0.6496926	1.5027300	0.1971073
[228,]	0.5461792	1.2386261	0.1984924
[229,]	0.6222971	NA	0.1164101
[230,]	0.4106422	0.6327101	0.3462681
[231,]	0.6974784	1.2081855	0.2249225
[232,]	0.5871073	0.7587743	0.2403770
[233,]	0.6740856	1.2012020	0.1888958
[234,]	0.5723364	1.1346059	0.1480777
[235,]	0.6046460	1.4070187	0.2017803
[236,]	0.4994655	1.0169456	0.2299975
[237,]	0.5262141	1.5134981	0.2542212
[238,]	0.5986212	0.9695219	0.1606173

Equiparación de Puntuaciones: Limitación Visual 103

[239,]	0.6122641	1.2344996	0.1980602
[240,]	0.6248788	0.8549568	0.1959419
[241,]	0.5375845	0.7972991	0.1654874
[242,]	0.6297196	0.8758210	0.2040188
[243,]	0.6592461	1.0413018	0.2286991
[244,]	0.5475270	1.1373803	0.3376556
[245,]	0.5133728	0.8530089	0.1421626
[246,]	0.5043618	0.9718863	0.1673419
[247,]	0.4396367	0.9165644	0.2695537
[248,]	0.5876525	0.6685450	0.1819645
[249,]	0.7268635	1.1354843	0.2443577
[250,]	0.6933107	1.1582945	0.1952839
[251,]	0.5815512	0.8602235	0.1795677
[252,]	0.6411620	1.2035179	0.2323384
[253,]	0.6210721	0.8086179	0.1394789
[254,]	0.5549576	1.0008799	0.1718369
[255,]	0.6814405	1.1142539	0.1554526
[256,]	0.5845483	1.1481288	0.2380906
[257,]	0.6380549	0.9014710	0.1807993
[258,]	0.6517582	1.1880600	0.1750339

Tucker

Levine

Media-Sigma

Equiparación de Puntuaciones: Limitación Visual 104

[259,]	0.5670245	0.8377143	0.2027933
[260,]	0.5701529	0.7928354	0.1490691
[261,]	0.5081594	1.0788162	0.2198389
[262,]	0.5588172	0.9154107	0.2283584
[263,]	0.5266926	1.2932399	0.1899473
[264,]	0.5842771	1.1488164	0.2100303
[265,]	0.5811755	0.8662220	0.2383791
[266,]	0.5836647	1.1319482	0.2975604
[267,]	0.6362939	1.0566725	0.2602678
[268,]	0.6409880	0.9212627	0.1880929
[269,]	0.5345599	1.1301792	0.1810182
[270,]	0.5888304	0.8016217	0.3368165
[271,]	0.5246781	3.2379286	0.2003095
[272,]	0.5518776	0.7703155	0.1922599
[273,]	0.6296306	0.7065677	0.1745597
[274,]	0.6099060	0.7477354	0.2059916
[275,]	0.6384561	1.1472305	0.2376398
[276,]	0.4635103	0.9252424	0.1757192
[277,]	0.6371091	1.1859644	0.1480493
[278,]	0.6274142	0.9885340	0.1276931
[279,]	0.6082389	1.0750777	0.2078098

Equiparación de Puntuaciones: Limitación Visual 105

[280,]	0.6697855	1.1794900	0.1924153
[281,]	0.6046897	1.0017584	0.1891842
[282,]	0.5214401	0.7586386	0.2425934
[283,]	0.6196262	0.8520576	0.2139869
[284,]	0.6039902	1.0328559	0.2775204
[285,]	0.7025278	1.2931373	0.1190379
[286,]	0.4650398	0.7987474	0.1927691
[287,]	0.5457764	1.3139125	0.1860154
[288,]	0.5371930	0.9686638	0.1982431
[289,]	0.5190445	1.0755821	0.1985753
[290,]	0.5174570	1.4533763	0.2299889
[291,]	0.5710339	1.2566247	0.2100357
[292,]	0.5962585	0.9959678	0.2453450
[293,]	0.5467083	0.7998685	0.2729445
[294,]	0.5699646	0.8001253	0.1606269
[295,]	0.6351443	1.2862391	0.3024885
[296,]	0.4821674	0.9678458	0.1745760

Tucker

Levine

Media-Sigma

[297,]	0.6598687	1.3333271	0.2917949
[298,]	0.7059999	1.0235107	0.1967710
[299,]	0.6289415	0.7909524	0.1863866

Equiparación de Puntuaciones: Limitación Visual 106

[300,]	0.6879471	0.9310006	0.1529381
[301,]	0.5292958	1.0523141	0.2239089
[302,]	0.5659808	0.7193160	0.2066872
[303,]	0.5300643	0.9188478	0.2509127
[304,]	0.7166677	1.1717357	0.2463533
[305,]	0.5577839	1.9022961	0.2790795
[306,]	0.5015169	0.7280892	0.1362311
[307,]	0.7315650	1.0682612	0.1502580
[308,]	0.5636990	1.1425909	0.2713005
[309,]	0.6464617	1.1008882	0.2533127
[310,]	0.6021657	1.6235447	0.2075971
[311,]	0.5741362	0.9265997	0.1821026
[312,]	0.6157652	0.8639072	0.1933623
[313,]	0.6775511	2.4817157	0.1783886
[314,]	0.4931034	0.8334861	0.1594512
[315,]	0.5898369	1.0866789	0.3181450
[316,]	0.6005681	1.9581151	0.2505626
[317,]	0.5623800	0.8421368	0.2006442
[318,]	0.6948615	0.6238931	0.1597479
[319,]	0.6397406	0.7841807	0.2110605
[320,]	0.5837108	1.4014701	0.1590054

Equiparación de Puntuaciones: Limitación Visual 107

[321,]	0.6071408	1.4612945	0.2313344
[322,]	0.6722946	1.1624964	0.2058913
[323,]	0.5960124	1.3965974	0.2374667
[324,]	0.5852058	0.9277077	0.2266026
[325,]	0.6199228	0.9040011	0.2440273
[326,]	0.6879720	0.9145263	0.3425356
[327,]	0.5867835	1.0180316	0.2613055
[328,]	0.6668715	0.7553193	0.2031992
[329,]	0.6454287	0.9695801	0.1642984
[330,]	0.6654257	0.7476793	0.2208159
[331,]	0.6709268	1.0091787	0.3814286
[332,]	0.5143355	0.8162776	0.1578462
[333,]	0.5167473	0.7825551	0.1793793
[334,]	0.6327344	1.0691938	0.2170212

Tucker

Levine

Media-Sigma

[335,]	0.6164907	1.0155934	0.2266800
[336,]	0.5859973	1.2640601	0.1818527
[337,]	0.6096186	0.9519523	0.1831897
[338,]	0.6303971	0.8356246	0.1989358
[339,]	0.5834123	0.8067518	0.1881967
[340,]	0.5209650	1.2200427	0.1965438

Equiparación de Puntuaciones: Limitación Visual 108

[341,]	0.6229465	1.7618203	0.2076151
[342,]	0.6079787	1.3358596	0.1899154
[343,]	0.5155517	1.0823727	0.1555761
[344,]	0.5553137	0.9158298	0.1548594
[345,]	0.6785178	1.2560921	0.2745409
[346,]	0.5850504	1.1338378	0.2761961
[347,]	0.7270431	0.9232147	0.1763831
[348,]	0.6231430	0.7971858	0.2350328
[349,]	0.6994859	1.3079537	0.1870341
[350,]	0.5804318	1.0913131	0.2244230
[351,]	0.5003017	0.9413870	0.2052986
[352,]	0.6262539	1.6274133	0.2776985
[353,]	0.6694950	0.9638078	0.2883974
[354,]	0.5435207	0.9764700	0.2731843
[355,]	0.6280556	0.9233439	0.2148749
[356,]	0.6311619	NA	0.1085584
[357,]	0.5270159	0.8605389	0.1703447
[358,]	0.6177124	1.0413996	0.2684105
[359,]	0.7639243	0.9583040	0.2554556
[360,]	0.5106422	0.9679731	0.2631988
[361,]	0.5967260	0.6334350	0.1879720

Equiparación de Puntuaciones: Limitación Visual 109

[362,]	0.5048414	1.0041798	0.2282137
[363,]	0.5475761	0.8477430	0.1685593
[364,]	0.6489666	0.9586337	0.2505112
[365,]	0.6439485	1.2363819	0.1727357
[366,]	0.6646905	0.8922857	0.2980527
[367,]	0.6046856	0.6281499	0.2033888
[368,]	0.5471359	1.0171001	0.2298949
[369,]	0.7964315	1.0052605	0.2133195
[370,]	0.5748652	0.6748189	0.3415395
[371,]	0.5146679	0.8819863	0.1303000
[372,]	0.6157016	0.9258858	0.2531607

Tucker

Levine

Media-Sigma

[373,]	0.5754125	0.6971656	0.2110470
[374,]	0.6083629	0.7681264	0.2146952
[375,]	0.6115255	0.9039796	0.2474140
[376,]	0.6641266	0.9455288	0.2144507
[377,]	0.5052142	0.8476593	0.2006169
[378,]	0.4824865	1.2742007	0.2370454
[379,]	0.7571967	0.9527792	0.1877456
[380,]	0.5065672	0.7187757	0.1157096
[381,]	0.6237630	1.0129119	0.2299074

Equiparación de Puntuaciones: Limitación Visual 110

[382,]	0.7699334	1.0657241	0.1810960
[383,]	0.5737934	1.1185088	0.2199375
[384,]	0.6544390	0.7294500	0.1829762
[385,]	0.6643045	1.3630694	0.1783013
[386,]	0.6113920	1.1859544	0.1973371
[387,]	0.5441594	0.7150856	0.2034819
[388,]	0.5196853	0.7923453	0.2205874
[389,]	0.6944263	1.6252656	0.1357809
[390,]	0.6125984	0.8216114	0.1625594
[391,]	0.6040457	0.9574912	0.3171640
[392,]	0.4822598	0.7063306	0.1687392
[393,]	0.6299318	1.8751529	0.1562956
[394,]	0.5938527	2.2058159	0.1160211
[395,]	0.6528835	1.2279100	0.1875920
[396,]	0.6769754	1.5493848	0.1749064
[397,]	0.6334287	1.3356052	0.1981865
[398,]	0.5771978	1.0901906	0.1509576
[399,]	0.5553642	0.7703939	0.1779533
[400,]	0.5216132	0.7613576	0.1981094
[401,]	0.5659350	1.1649294	0.1910883
[402,]	0.4792878	0.9660495	0.2066917

Equiparación de Puntuaciones: Limitación Visual 111

[403,]	0.4742125	0.8619713	0.1118263
[404,]	0.5138519	1.0010930	0.1910128
[405,]	0.5148839	1.8763966	0.1695408
[406,]	0.6310024	1.0172489	0.1863360
[407,]	0.6565586	1.2283422	0.2433754
[408,]	0.5728799	0.7319483	0.1804336
[409,]	0.6829181	1.0753912	0.1540645
[410,]	0.5434072	0.6809982	0.1122171

Tucker Levine Media-Sigma

[411,]	0.5922253	0.6372509	0.2264936
[412,]	0.6961750	0.9213496	0.1427763
[413,]	0.6080873	1.0226876	0.1495996
[414,]	0.5711751	1.4200037	0.2330460
[415,]	0.5663723	0.8251980	0.1641323
[416,]	0.5626710	0.8806613	0.2488074
[417,]	0.5869652	0.9659831	0.1980759
[418,]	0.6447270	0.9295600	0.2058625
[419,]	0.5393137	0.9610595	0.2370666
[420,]	0.5774878	0.8194797	0.2121702
[421,]	0.6529208	1.3612721	0.1897667
[422,]	0.6165606	1.4626956	0.2207381

Equiparación de Puntuaciones: Limitación Visual 112

[423,]	0.5649168	1.0401788	0.2343914
[424,]	0.6396940	1.3412222	0.2716442
[425,]	0.5959593	1.5117913	0.2944870
[426,]	0.5929606	0.9539979	0.2137405
[427,]	0.6424885	1.1268731	0.3067328
[428,]	0.5000307	0.9029447	0.1270836
[429,]	0.6134182	1.5044702	0.1458173
[430,]	0.7090294	1.3949695	0.1978145
[431,]	0.5093649	1.0004510	0.2060530
[432,]	0.5369021	1.1474585	0.1747031
[433,]	0.4912153	0.9592229	0.2126734
[434,]	0.7121225	1.0663721	0.1857836
[435,]	0.5505022	1.1493412	0.2698516
[436,]	0.6434864	1.3050577	0.1829081
[437,]	0.4700546	1.0645860	0.1837972
[438,]	0.6661868	1.5955104	0.2699240
[439,]	0.6207838	0.8181391	0.1854501
[440,]	0.6568677	1.3411989	0.2492483
[441,]	0.5605509	0.6586019	0.1967771
[442,]	0.5490391	0.6789422	0.1820123
[443,]	0.5566068	1.0911370	0.1666827

Equiparación de Puntuaciones: Limitación Visual 113

[444,]	0.6474349	0.9432545	0.2831527
[445,]	0.4908643	0.7969136	0.2188191
[446,]	0.5980404	1.1302546	0.2519924
[447,]	0.6370660	0.9606454	0.2436703
[448,]	0.5511846	0.9246007	0.2134726

Tucker **Levine** **Media-Sigma**

[449,]	0.6245799	0.9822765	0.2788950
[450,]	0.4810988	0.7415300	0.2480850
[451,]	0.5396440	0.9856883	0.2524287
[452,]	0.6730792	1.1590697	0.1560550
[453,]	0.5388035	1.0028418	0.2123695
[454,]	0.6651853	0.7853113	0.3267518
[455,]	0.6181200	1.0789327	0.2215837
[456,]	0.5858988	0.8148615	0.2161608
[457,]	0.5771109	0.9645779	0.2946972
[458,]	0.6039609	1.1868144	0.1736718
[459,]	0.6829359	1.2036463	0.2333767
[460,]	0.5327810	1.0582936	0.2109156
[461,]	0.5992532	1.0179063	0.2753146
[462,]	0.5322798	0.7282146	0.2776745
[463,]	0.5394230	1.4785206	0.1561234

Equiparación de Puntuaciones: Limitación Visual 114

[464,]	0.4921975	0.7964792	0.2142513
[465,]	0.6115673	1.0995792	0.2910426
[466,]	0.4382106	0.6886475	0.2164163
[467,]	0.5725472	0.9931895	0.1992597
[468,]	0.5729943	0.7132527	0.1465701
[469,]	0.6605258	NA	0.2117042
[470,]	0.4898327	0.7740970	0.2905329
[471,]	0.6101745	1.0886957	0.2250121
[472,]	0.5133902	0.7511648	0.2004228
[473,]	0.5709796	0.9477203	0.2043038
[474,]	0.5746017	1.3344787	0.2142234
[475,]	0.5908727	1.3173903	0.1708287
[476,]	0.4933581	0.8011221	0.2065907
[477,]	0.4996451	1.1645736	0.2238952
[478,]	0.5565763	0.9763423	0.2151102
[479,]	0.5507420	0.9905038	0.2215746
[480,]	0.5536202	1.2254257	0.1502649
[481,]	0.6113837	1.0738015	0.2054854
[482,]	0.5686755	2.3473021	0.2840761
[483,]	0.5998994	0.8318418	0.2146110
[484,]	0.4600752	0.6785238	0.2342545

Equiparación de Puntuaciones: Limitación Visual 115

[485,] 0.5611489 1.5254532 0.2087251

[486,] 0.6779042 2.3643541 0.1772478

Tucker Levine Media-Sigma

[487,] 0.5703245 1.8162288 0.2166034

[488,] 0.6146265 0.9965064 0.2749307

[489,] 0.6014829 1.1240311 0.2593838

[490,] 0.7053081 NA 0.1978625

[491,] 0.6853450 0.7383313 0.2394690

[492,] 0.6556598 2.0125691 0.2789595

[493,] 0.6416303 0.8741069 0.2302772

[494,] 0.5729300 1.0204312 0.2442695

[495,] 0.6613223 1.0447634 0.1751409

[496,] 0.5781498 0.8616975 0.1961319

[497,] 0.7607775 1.5333594 0.1732415

[498,] 0.6004660 1.2919954 0.1968390

[499,] 0.6376615 1.3337885 0.2399740

[500,] 0.5905278 1.2756785 0.1415961

[501,] 0.6382859 0.9417775 0.2573973

[502,] 0.6621518 1.0257838 0.2503133

[503,] 0.4888548 0.7110544 0.2060207

[504,] 0.5858322 1.8243797 0.2478597

Equiparación de Puntuaciones: Limitación Visual 116

[505,]	0.6662505	0.9710582	0.1794563
[506,]	0.5720123	1.0204351	0.2711049
[507,]	0.4946177	1.0908550	0.2386846
[508,]	0.6024941	0.9348400	0.1200879
[509,]	0.6491348	0.6900839	0.1533079
[510,]	0.5181158	NA	0.2161386
[511,]	0.5982902	0.7758995	0.2181072
[512,]	0.6554083	1.2371932	0.2547377
[513,]	0.6900593	NA	0.2378155
[514,]	0.5936557	0.9530673	0.2412267
[515,]	0.5968315	0.8782155	0.2386595
[516,]	0.5234651	2.0200938	0.2033474
[517,]	0.5288736	0.8083029	0.2094175
[518,]	0.7478164	1.1592291	0.1845711
[519,]	0.5943964	0.9479593	0.2316682
[520,]	0.6067458	1.2293686	0.2893028
[521,]	0.5659466	0.9088926	0.1529748
[522,]	0.6989697	1.6124811	0.2334178
[523,]	0.6808399	0.8730367	0.1478023
[524,]	0.6085896	1.3325194	0.1846050

Tucker

Levine

Media-Sigma

Equiparación de Puntuaciones: Limitación Visual 117

[525,]	0.6100753	1.0402673	0.1943788
[526,]	0.5719998	1.2187874	0.1101919
[527,]	0.5572900	0.9767712	0.2359316
[528,]	0.6844019	1.1907460	0.1255186
[529,]	0.5497270	1.3104604	0.1895550
[530,]	0.5935030	1.0363631	0.1670760
[531,]	0.6215154	1.3706937	0.2340573
[532,]	0.5797092	1.5520447	0.3114283
[533,]	0.6631567	0.9960990	0.2260224
[534,]	0.7409292	0.9622511	0.2687011
[535,]	0.6354720	0.7105268	0.1082994
[536,]	0.6009094	1.3570420	0.2408848
[537,]	0.6204543	1.1034241	0.2328692
[538,]	0.6493692	1.1018355	0.2258900
[539,]	0.5717745	1.1585709	0.2223198
[540,]	0.6293697	0.9975502	0.2615002
[541,]	0.5865746	1.1840202	0.1924783
[542,]	0.5113797	1.2540567	0.1674716
[543,]	0.7133310	0.7087898	0.2490210
[544,]	0.5990014	1.3000961	0.1668656
[545,]	0.6137651	0.8101722	0.2277598

Equiparación de Puntuaciones: Limitación Visual 118

[546,]	0.4723492	0.8359295	0.1526638
[547,]	0.6761070	0.9842600	0.1103828
[548,]	0.5968861	0.7598128	0.3715491
[549,]	0.5570476	1.0180469	0.2372773
[550,]	0.6616901	1.3751953	0.2472607
[551,]	0.6116135	0.7431052	0.2011313
[552,]	0.6578447	1.4513910	0.1936530
[553,]	0.5598919	0.7342630	0.2485949
[554,]	0.5778755	0.7773810	0.2228182
[555,]	0.5264081	0.7473047	0.1983485
[556,]	0.6499414	0.8810961	0.1910900
[557,]	0.5288485	1.0458218	0.1425433
[558,]	0.5262715	0.9783718	0.1400122
[559,]	0.6682239	0.8372957	0.2804143
[560,]	0.5826747	1.2290920	0.1899135
[561,]	0.5899691	1.0008303	0.1733939
[562,]	0.7573086	3.2303239	0.2278367

Tucker

Levine

Media-Sigma

[563,]	0.5274642	0.6851353	0.1889657
[564,]	0.7847661	1.3907880	0.2399347
[565,]	0.7144778	1.6008997	0.3089523

Equiparación de Puntuaciones: Limitación Visual 119

[566,]	0.6069543	1.2632915	0.2140019
[567,]	0.5643079	1.1578021	0.1488224
[568,]	0.5475656	1.4115720	0.1765712
[569,]	0.5570040	1.1184012	0.1965166
[570,]	0.6291955	1.7433850	0.2094582
[571,]	0.5445097	1.1962513	0.2662436
[572,]	0.5296907	0.8899055	0.2244537
[573,]	0.5465237	1.0105096	0.1855966
[574,]	0.6575590	1.0983093	0.2070956
[575,]	0.6080210	0.8236757	0.2424225
[576,]	0.4711799	0.7085260	0.2954871
[577,]	0.5658199	0.9621993	0.1901298
[578,]	0.6144864	0.8658552	0.2474799
[579,]	0.6871927	1.0598286	0.1596614
[580,]	0.6097680	1.5168828	0.1357123
[581,]	0.4607497	1.4254809	0.2208181
[582,]	0.4839762	1.0783806	0.1484763
[583,]	0.7115268	1.3445956	0.3072276
[584,]	0.5771101	0.9266999	0.1583327
[585,]	0.6019418	0.8445433	0.2267360
[586,]	0.7034573	0.9541912	0.1808312

Equiparación de Puntuaciones: Limitación Visual 120

[587,]	0.5768120	0.7486175	0.2456674
[588,]	0.5354257	0.6325876	0.1938621
[589,]	0.4911764	0.6995158	0.2326180
[590,]	0.5171578	0.8311646	0.1552989
[591,]	0.6098735	0.7760041	0.2311898
[592,]	0.5357203	0.7171340	0.2473663
[593,]	0.5952922	0.7283029	0.1766132
[594,]	0.5984699	1.4244548	0.2861375
[595,]	0.6640628	1.2230268	0.1743836
[596,]	0.6227666	0.9629422	0.1689785
[597,]	0.5550563	1.1707694	0.2150250
[598,]	0.6107480	1.1716532	0.1119022
[599,]	0.6423625	0.8931631	0.1522116
[600,]	0.6591022	0.8675613	0.1987175

Tucker

Levine

Media-Sigma

[601,]	0.6186748	1.6777257	0.2046499
[602,]	0.6927970	1.3577739	0.1114835
[603,]	0.6314889	1.2528959	0.1786000
[604,]	0.6779911	NA	0.2197302
[605,]	0.6759508	0.8301252	0.1803038
[606,]	0.5622183	1.0262896	0.2109011

Equiparación de Puntuaciones: Limitación Visual 121

[607,]	0.6609708	1.0864560	0.1579788
[608,]	0.5210927	1.1598180	0.1665762
[609,]	0.5292086	0.8159018	0.1785098
[610,]	0.6835547	0.7978810	0.1755305
[611,]	0.6719722	NA	0.2332912
[612,]	0.6504618	NA	0.2266473
[613,]	0.5291468	0.9082382	0.2041498
[614,]	0.5425009	1.1456915	0.3243775
[615,]	0.4974605	1.3454186	0.1374083
[616,]	0.6567737	3.7139315	0.1258789
[617,]	0.5020960	0.9877848	0.2149855
[618,]	0.6041095	NA	0.2363038
[619,]	0.6014782	1.0556722	0.1741917
[620,]	0.7135414	0.9372432	0.2401323
[621,]	0.5127724	0.8894975	0.1352629
[622,]	0.6791019	1.1461276	0.2541609
[623,]	0.4898437	0.9385219	0.3092781
[624,]	0.5303364	0.9192173	0.1285182
[625,]	0.7526369	1.2140759	0.1640898
[626,]	0.7058660	1.2311471	0.2385426
[627,]	0.6136833	1.4446750	0.3114801

Equiparación de Puntuaciones: Limitación Visual 122

[628,]	0.6176012	1.0504685	0.2248210
[629,]	0.6627010	1.1230186	0.1671569
[630,]	0.6251838	1.1540925	0.2033401
[631,]	0.6972554	0.9349653	0.1602397
[632,]	0.5715728	1.0754076	0.2319757
[633,]	0.5181779	0.9619668	0.1897679
[634,]	0.5764169	1.4460635	0.2168021
[635,]	0.5144757	0.7727599	0.3034374
[636,]	0.6331732	0.9800453	0.1705727
[637,]	0.4322162	0.8741791	0.2352094
[638,]	0.5532927	0.9086717	0.2156321

Tucker

Levine

Media-Sigma

[639,]	0.4684552	0.9973448	0.2484736
[640,]	0.7003632	0.8408555	0.2456316
[641,]	0.6110916	0.8837285	0.1833794
[642,]	0.5699655	0.9482210	0.2009084
[643,]	0.5524455	0.9583354	0.1613495
[644,]	0.6157270	0.9628252	0.1863910
[645,]	0.6716000	0.8616078	0.2757063
[646,]	0.8404293	1.2642736	0.2099569
[647,]	0.6784510	1.0933620	0.1734932

Equiparación de Puntuaciones: Limitación Visual 123

[648,]	0.5736484	1.1147512	0.2107076
[649,]	0.5210381	0.7554256	0.2590196
[650,]	0.6311546	0.9794516	0.2353224
[651,]	0.5734810	0.7532310	0.2469035
[652,]	0.5572357	0.9926632	0.1437718
[653,]	0.5780164	1.5731777	0.2727474
[654,]	0.5616545	1.1838078	0.2003207
[655,]	0.5866280	1.0074287	0.1424592
[656,]	0.5678526	0.8031919	0.2153696
[657,]	0.6627455	0.9737837	0.2132710
[658,]	0.5384017	0.9808407	0.1589835
[659,]	0.6019743	0.8174109	0.2536200
[660,]	0.5817721	1.2859251	0.2299837
[661,]	0.6876591	1.0388496	0.1969809
[662,]	0.5394352	0.6964258	0.2227723
[663,]	0.6501172	1.6186445	0.2506373
[664,]	0.5813652	1.0043708	0.2237292
[665,]	0.5808599	1.3064292	0.1626843
[666,]	0.7349121	0.7695353	0.2087132
[667,]	0.6615609	0.7959633	0.2075906
[668,]	0.6010156	0.8101747	0.1639346

Equiparación de Puntuaciones: Limitación Visual 124

[669,]	0.5159903	0.9303828	0.2328192
[670,]	0.5356542	0.8070038	0.2694247
[671,]	0.6245686	1.1173373	0.2034634
[672,]	0.4532570	0.8096587	0.2237589
[673,]	0.6793667	1.1306005	0.2565635
[674,]	0.5435618	1.0274219	0.1813579
[675,]	0.5571004	1.0304907	0.1338290
[676,]	0.6009519	0.9043266	0.1945155

Tucker

Levine

Media-Sigma

[677,]	0.6620415	0.9720641	0.1766252
[678,]	0.5964296	0.7326768	0.2390462
[679,]	0.5212056	1.1756341	0.1690404
[680,]	0.6901641	1.1107385	0.1879958
[681,]	0.5578594	1.0889780	0.1468330
[682,]	0.5399513	0.6768484	0.1865534
[683,]	0.7001573	1.9470297	0.2551488
[684,]	0.6990310	1.6418590	0.1668141
[685,]	0.6545347	1.4407730	0.1734158
[686,]	0.6503757	1.2168125	0.1818160
[687,]	0.6079067	0.8197224	0.2080171
[688,]	0.6270071	1.1105670	0.2767732

Equiparación de Puntuaciones: Limitación Visual 125

[689,]	0.5738284	1.4682803	0.2210332
[690,]	0.6699581	2.8845137	0.1501899
[691,]	0.6111390	0.9813289	0.1634500
[692,]	0.5446055	1.0851695	0.1656180
[693,]	0.6809002	1.4097053	0.1868263
[694,]	0.5194853	0.8440699	0.1949458
[695,]	0.6235297	1.5104016	0.2529603
[696,]	0.8105969	1.6461380	0.1448180
[697,]	0.7134080	0.9048879	0.2173967
[698,]	0.6393608	1.0324594	0.2416783
[699,]	0.4876181	0.9895903	0.1771672
[700,]	0.5660984	0.9765782	0.1715603
[701,]	0.5819090	0.9356508	0.2100315
[702,]	0.5407529	1.0141273	0.1910986
[703,]	0.6227499	0.9658058	0.3129607
[704,]	0.6501599	0.6520196	0.1642594
[705,]	0.6564100	1.0307033	0.1818736
[706,]	0.6043088	0.7708862	0.2175536
[707,]	0.5352695	0.8414213	0.1450109
[708,]	0.7047704	0.7858299	0.1700392
[709,]	0.7142414	0.9629978	0.2179090

Equiparación de Puntuaciones: Limitación Visual 126

[710,]	0.5057662	0.7587071	0.1858949
[711,]	0.6229385	1.0433580	0.2703684
[712,]	0.5395226	1.2936975	0.2254034
[713,]	0.6099516	1.0939499	0.2276226
[714,]	0.5681269	0.9495935	0.2566756

	Tucker	Levine	Media-Sigma
[715,]	0.5421390	1.1377497	0.3155663
[716,]	0.6460879	0.8092633	0.2935850
[717,]	0.6563604	1.1186158	0.1009889
[718,]	0.5954321	1.1507807	0.2012008
[719,]	0.5629479	1.2857351	0.2383134
[720,]	0.6673090	0.9906302	0.3190593
[721,]	0.4863625	0.7599196	0.2323810
[722,]	0.5868897	1.0479157	0.2753473
[723,]	0.6135825	1.1961969	0.1389961
[724,]	0.6169466	0.9884835	0.1515774
[725,]	0.6506135	1.0139766	0.2032226
[726,]	0.7015138	2.6435756	0.1637019
[727,]	0.5712991	0.7249655	0.2851482
[728,]	0.5293705	1.1770858	0.1592372
[729,]	0.6683660	1.0200888	0.1744254

Equiparación de Puntuaciones: Limitación Visual 127

[730,]	0.6681387	1.0580519	0.2648086
[731,]	0.5146844	0.5785527	0.2342025
[732,]	0.5687315	0.9010636	0.2761245
[733,]	0.5530413	1.1070782	0.1503355
[734,]	0.6280432	0.8974124	0.2421520
[735,]	0.6414181	0.8185684	0.2077370
[736,]	0.5590003	1.4374297	0.1911229
[737,]	0.5966194	1.1899649	0.2421652
[738,]	0.6061773	1.8209258	0.1689434
[739,]	0.5404496	1.0122380	0.1841238
[740,]	0.6371825	1.3051205	0.2214307
[741,]	0.5582797	0.8385250	0.2943163
[742,]	0.6192928	0.8888781	0.2447758
[743,]	0.6316816	1.1176597	0.2000347
[744,]	0.5505128	0.8314078	0.3093604
[745,]	0.5228965	1.4354810	0.2360268
[746,]	0.5968967	0.8612298	0.1896688
[747,]	0.5580848	0.9889726	0.2208560
[748,]	0.5702875	1.1190564	0.1943300
[749,]	0.5623764	1.0288695	0.1828546
[750,]	0.5612905	0.9913905	0.2337098

Equiparación de Puntuaciones: Limitación Visual 128

[751,]	0.5732228	0.9530791	0.3245714
[752,]	0.5563380	1.0576392	0.1353911
	Tucker	Levine	Media-Sigma
[753,]	0.5489014	0.8299094	0.1795869
[754,]	0.5465003	0.9386832	0.1733996
[755,]	0.6167949	1.1140819	0.2452168
[756,]	0.5449108	1.3645024	0.2601259
[757,]	0.6131977	1.1781138	0.2189545
[758,]	0.5970377	1.1254678	0.2188229
[759,]	0.5407290	0.8062712	0.2119741
[760,]	0.5930073	0.7792854	0.2242935
[761,]	0.6953581	1.1642254	0.1798790
[762,]	0.6745369	1.2716831	0.1383010
[763,]	0.5242312	0.8175611	0.2025596
[764,]	0.7343655	0.9573200	0.1597004
[765,]	0.5785422	0.8754535	0.2408172
[766,]	0.5695405	1.1436383	0.2390512
[767,]	0.5938260	1.5486222	0.1509557
[768,]	0.5467350	0.9168281	0.3269402
[769,]	0.7041932	1.4175540	0.2372751
[770,]	0.4844154	0.4695651	0.1791678

Equiparación de Puntuaciones: Limitación Visual 129

[771,]	0.5306571	0.9130886	0.1380955
[772,]	0.6449184	0.9941308	0.3615610
[773,]	0.5122494	0.9661898	0.2349044
[774,]	0.4339577	0.6946425	0.2955449
[775,]	0.6848331	1.0955771	0.2336173
[776,]	0.4968280	0.6124895	0.1919508
[777,]	0.5659040	1.5742648	0.1586016
[778,]	0.6517331	0.9704484	0.2282193
[779,]	0.4880017	1.1383292	0.2062301
[780,]	0.5746235	0.7483560	0.2648522
[781,]	0.6113191	1.4530065	0.2058127
[782,]	0.5404527	0.9195801	0.2083668
[783,]	0.5636734	1.0305778	0.2226510
[784,]	0.5804053	1.1147931	0.1972583
[785,]	0.7118213	1.6318162	0.1688189
[786,]	0.5299620	1.2856349	0.1625602
[787,]	0.4973026	0.9119993	0.2065926
[788,]	0.6836251	1.2327657	0.2587221
[789,]	0.6955883	1.0042327	0.1495065
[790,]	0.7530783	1.0474834	0.1902316

Tucker

Levine

Media-Sigma

Equiparación de Puntuaciones: Limitación Visual 130

[791,]	0.5558171	0.8213812	0.1789369
[792,]	0.8262478	1.1479361	0.1978673
[793,]	0.6010973	0.9200572	0.1736570
[794,]	0.5732838	0.8633411	0.1686638
[795,]	0.6595100	1.1780261	0.1989983
[796,]	0.5758157	1.0754745	0.2035177
[797,]	0.6647546	1.1416731	0.2159137
[798,]	0.5086606	1.1060202	0.1862928
[799,]	0.6110420	1.1273127	0.2418143
[800,]	0.6242001	0.9748489	0.2525700
[801,]	0.6632422	0.9266927	0.2283801
[802,]	0.5320283	0.7862711	0.1954233
[803,]	0.7915990	0.9340160	0.1668343
[804,]	0.5483167	1.2759255	0.1264664
[805,]	0.6550800	0.9618602	0.2654465
[806,]	0.7591754	1.2741368	0.2141649
[807,]	0.6802252	1.2617957	0.1823825
[808,]	0.5807438	1.1320766	0.2561531
[809,]	0.6050239	0.9185228	0.2145250
[810,]	0.6824176	0.8026030	0.1852582
[811,]	0.5572349	1.5051159	0.2857826

Equiparación de Puntuaciones: Limitación Visual 131

[812,]	0.5580080	1.0511756	0.2402033
[813,]	0.5504507	0.8849690	0.2501662
[814,]	0.6792095	0.9733320	0.2422776
[815,]	0.5362112	2.0603940	0.1853218
[816,]	0.5735873	1.1543156	0.1725713
[817,]	0.5807995	0.7523509	0.2161284
[818,]	0.5596862	0.9062884	0.2163117
[819,]	0.5051015	1.0904736	0.1937824
[820,]	0.6205521	0.7763973	0.2056146
[821,]	0.4928507	0.7778537	0.1661048
[822,]	0.5099119	0.7653143	0.2955617
[823,]	0.6789464	1.1972423	0.2309393
[824,]	0.5704867	1.2397147	0.3285320
[825,]	0.7388865	0.9924914	0.2449499
[826,]	0.6185253	1.1697745	0.1482748
[827,]	0.7430536	1.0628576	0.1784332
[828,]	0.6474922	0.8505833	0.2410818

Tucker

Levine

Media-Sigma

[829,]	0.7441001	0.9373248	0.1582565
[830,]	0.6417663	1.0900988	0.1488753
[831,]	0.6548852	1.0033166	0.2640727

Equiparación de Puntuaciones: Limitación Visual 132

[832,]	0.6732242	0.9490663	0.2181754
[833,]	0.4915467	0.7932219	0.1543408
[834,]	0.6776886	0.9837247	0.2695585
[835,]	0.6101261	1.3439956	0.1249345
[836,]	0.5428399	0.9158314	0.1945934
[837,]	0.8062431	1.1123992	0.1838899
[838,]	0.4465573	1.0881579	0.1351397
[839,]	0.6667988	1.1279452	0.1999920
[840,]	0.5693770	1.2645178	0.1331813
[841,]	0.6040888	0.6692969	0.2228082
[842,]	0.4987835	0.8871705	0.1383846
[843,]	0.6277012	0.9379802	0.2401649
[844,]	0.5008659	0.6927095	0.1619458
[845,]	0.5805440	1.1408158	0.1558935
[846,]	0.5750093	1.3235065	0.2146660
[847,]	0.4526624	0.8871133	0.1834816
[848,]	0.5763806	0.8327630	0.2195905
[849,]	0.6865294	1.0504452	0.2091615
[850,]	0.5243854	0.9431044	0.1706527
[851,]	0.5478769	0.7969879	0.2234039
[852,]	0.6793071	1.4496695	0.1580227

Equiparación de Puntuaciones: Limitación Visual 133

[853,]	0.5653224	0.8921827	0.2252306
[854,]	0.5757219	1.4869217	0.2720602
[855,]	0.6682798	0.8423564	0.1810552
[856,]	0.5767605	1.2180242	0.1954758
[857,]	0.7334122	0.8710061	0.1910247
[858,]	0.5493297	1.0241159	0.2778966
[859,]	0.5885045	2.1274010	0.2434897
[860,]	0.6555668	1.2302959	0.2280884
[861,]	0.5503003	1.0953391	0.1767338
[862,]	0.5562656	1.6527412	0.1551045
[863,]	0.5580633	1.0742787	0.2122257
[864,]	0.6233733	1.1575993	0.2027055
[865,]	0.6999689	1.4764775	0.2250717
[866,]	0.4876033	1.0388011	0.2121545

Tucker

Levine

Media-Sigma

[867,]	0.6659042	0.7689836	0.2433091
[868,]	0.6015345	0.8987290	0.2492739
[869,]	0.6848942	1.2368775	0.1285868
[870,]	0.6946656	1.8440921	0.2185437
[871,]	0.5558220	0.6246621	0.2342879
[872,]	0.6596693	1.3230874	0.1697739

Equiparación de Puntuaciones: Limitación Visual 134

[873,]	0.5932690	0.8838484	0.3139680
[874,]	0.7014617	0.8756452	0.1843507
[875,]	0.6924213	0.7968357	0.2542660
[876,]	0.6198571	0.7674736	0.1371744
[877,]	0.5965470	1.1495072	0.2711921
[878,]	0.6631521	1.0896452	0.2251386
[879,]	0.6935092	1.5277318	0.1657724
[880,]	0.6533313	1.1210020	0.1651262
[881,]	0.5751259	1.1661085	0.2079629
[882,]	0.5524440	0.8860467	0.2599149
[883,]	0.5816789	1.0991093	0.1667005
[884,]	0.5390739	1.0630150	0.1797471
[885,]	0.5100711	1.1369208	0.2188078
[886,]	0.8866930	0.8906439	0.1990982
[887,]	0.4410697	0.9878789	0.1479728
[888,]	0.5283710	0.7376812	0.2823744
[889,]	0.6652317	0.9500159	0.2210685
[890,]	0.6287871	1.1117481	0.2000496
[891,]	0.5630165	0.9461111	0.1770030
[892,]	0.5956566	1.3473385	0.2547704
[893,]	0.6919287	1.3854548	0.1063109

Equiparación de Puntuaciones: Limitación Visual 135

[894,]	0.5844862	1.2080789	0.1822140
[895,]	0.5392347	2.1254777	0.1866714
[896,]	0.5716152	0.7115744	0.1823348
[897,]	0.6020239	1.1084613	0.2819224
[898,]	0.6256583	1.0323748	0.2369166
[899,]	0.6627517	1.3631720	0.1747503
[900,]	0.6055764	1.0996388	0.1971482
[901,]	0.5505950	1.1921047	0.2087558
[902,]	0.6278024	0.8164327	0.2044972
[903,]	0.6397441	1.5650336	0.2077432
[904,]	0.5171668	0.8059798	0.1539296

Tucker

Levine

Media-Sigma

[905,]	0.5442425	0.7681938	0.2915867
[906,]	0.6654451	1.3885349	0.1608350
[907,]	0.5169577	0.8046759	0.1831766
[908,]	0.6082065	0.9860494	0.1753206
[909,]	0.5094740	1.5109642	0.3126360
[910,]	0.7667494	1.2225623	0.1863197
[911,]	0.6229787	0.8689343	0.2332466
[912,]	0.5416517	0.8392598	0.1951890
[913,]	0.6574903	1.1354415	0.1309102

Equiparación de Puntuaciones: Limitación Visual 136

[914,]	0.4936267	0.9200545	0.1730274
[915,]	0.6712636	0.9582791	0.1996545
[916,]	0.6930635	0.8036378	0.2056723
[917,]	0.5461797	1.3287164	0.2663759
[918,]	0.5601360	0.7597488	0.2393062
[919,]	0.6594087	1.6400307	0.2483429
[920,]	0.6940702	0.8858620	0.2249220
[921,]	0.6184837	1.0323083	0.2634440
[922,]	0.4709806	1.6428356	0.2103306
[923,]	0.5886624	1.2699415	0.2466466
[924,]	0.7184759	1.1677552	0.2064165
[925,]	0.7113976	2.0869485	0.2448124
[926,]	0.5416990	1.0142553	0.1910396
[927,]	0.5287755	0.9512461	0.1487328
[928,]	0.6031573	1.1603482	0.1677833
[929,]	0.5861198	1.0002745	0.2332934
[930,]	0.4820170	1.2227622	0.1630786
[931,]	0.5678720	1.2566071	0.1488515
[932,]	0.5401634	1.1576062	0.2147311
[933,]	0.6717128	1.0312254	0.2284261
[934,]	0.5682276	0.6836355	0.1250835

Equiparación de Puntuaciones: Limitación Visual 137

[935,]	0.6591172	0.8105644	0.1986851
[936,]	0.5998996	0.7882986	0.1852049
[937,]	0.5273163	0.9908356	0.1701397
[938,]	0.5327827	0.5667110	0.2717694
[939,]	0.5910453	0.9512926	0.2088086
[940,]	0.5530562	1.0266633	0.1797479
[941,]	0.6299123	0.7301815	0.2296738
[942,]	0.6088463	1.5021567	0.2092833

Tucker

Levine

Media-Sigma

[943,]	0.5575091	0.8218872	0.2781211
[944,]	0.5517522	1.3688283	0.2608568
[945,]	0.6237259	1.4167003	0.2520610
[946,]	0.5760667	0.8813128	0.2833302
[947,]	0.5523465	0.7537724	0.2097746
[948,]	0.5849353	0.7547847	0.2260767
[949,]	0.5962831	1.3035183	0.2385376
[950,]	0.6016450	1.0519503	0.1345294
[951,]	0.5830008	0.8321521	0.1109831
[952,]	0.5547039	1.3577142	0.3052323
[953,]	0.5468479	1.1260756	0.1758398
[954,]	0.5867323	1.4375034	0.2473619

Equiparación de Puntuaciones: Limitación Visual 138

[955,]	0.6403105	1.0861043	0.2225680
[956,]	0.6267990	0.8116586	0.2109049
[957,]	0.6458946	1.0818226	0.2166980
[958,]	0.4629583	0.8138581	0.2448722
[959,]	0.5575999	0.9763964	0.1852552
[960,]	0.5875858	1.0246359	0.2999380
[961,]	0.5882845	1.1327115	0.2799003
[962,]	0.7113774	1.1879033	0.1730005
[963,]	0.7102953	0.8784106	0.3102366
[964,]	0.6176739	1.2416439	0.2237322
[965,]	0.5305297	0.8553964	0.2193021
[966,]	0.6055335	1.6451864	0.2376635
[967,]	0.5366030	0.6406505	0.1780187
[968,]	0.6458657	1.9083984	0.2331534
[969,]	0.4480978	1.1415318	0.2033148
[970,]	0.5664015	0.7843958	0.1739434
[971,]	0.6383978	1.0916789	0.2150849
[972,]	0.5747542	1.0634070	0.1211925
[973,]	0.5214115	1.5546208	0.3082026
[974,]	0.5896930	1.0783839	0.2241492
[975,]	0.5826432	1.0282835	0.2329262

Equiparación de Puntuaciones: Limitación Visual 139

[976,]	0.6068416	1.2459222	0.2007315
[977,]	0.6067973	0.7877358	0.3429612
[978,]	0.4558912	0.6098014	0.1811006
[979,]	0.4011705	1.0453639	0.1849682
[980,]	0.4515425	0.7912462	0.1364860

Tucker

Levine

Media-Sigma

[981,]	0.6025911	1.5538652	0.1746193
[982,]	0.7033264	1.3454058	0.1916698
[983,]	0.5014342	1.0067040	0.1984162
[984,]	0.5973419	1.0944645	0.2288050
[985,]	0.5640076	0.8152489	0.1676760
[986,]	0.5535833	0.9819598	0.2769997
[987,]	0.6961473	0.9800483	0.2229541
[988,]	0.6109187	1.2369052	0.1832413
[989,]	0.5266714	0.7633720	0.1909426
[990,]	0.6637859	0.8455023	0.2737610
[991,]	0.5633177	1.0629973	0.1752906
[992,]	0.5958439	0.8225145	0.2295400
[993,]	0.5370020	1.2028444	0.2582917
[994,]	0.6926635	0.9699148	0.1669656
[995,]	0.5545327	1.4230422	0.2299379

Equiparación de Puntuaciones: Limitación Visual 140

[996,]	0.5896873	0.9032114	0.1881035
[997,]	0.6632013	1.3026223	0.2147531
[998,]	0.6132600	1.1510937	0.2059397
[999,]	0.5074094	0.7948512	0.2267357
[1000,]	0.5931364	0.7700120	0.2023932
[1001,]	0.4863870	1.0201975	0.2448044
[1002,]	0.5857911	1.0336683	0.2416670
[1003,]	0.4363378	0.6348196	0.2119447
[1004,]	0.6611785	3.4579435	0.2207127
[1005,]	0.6343724	0.8904332	0.2163652
[1006,]	0.6192943	1.2619118	0.1903078
[1007,]	0.7069022	0.9157369	0.2087599
[1008,]	0.6339980	0.9968281	0.1875756
[1009,]	0.5302399	0.7867321	0.1767460
[1010,]	0.5119724	0.7814770	0.2312138
[1011,]	0.5673607	1.1058415	0.2268425
[1012,]	0.7012956	0.8313588	0.1457596
[1013,]	0.6242228	1.1637709	0.1928894
[1014,]	0.6110408	0.9423714	0.1605033
[1015,]	0.6274188	1.0660183	0.2224239
[1016,]	0.5611897	1.2674886	0.2003886

Equiparación de Puntuaciones: Limitación Visual 141

[1017,] 0.5412501 0.8331817 0.1857590

[1018,] 0.5726158 0.9295828 0.1941084

Tucker Levine Media-Sigma

[1019,] 0.5548567 1.0189660 0.1843391

[1020,] 0.5796454 1.5651588 0.1757751

[1021,] 0.5852474 1.0117254 0.2121636

[1022,] 0.6432291 1.7724106 0.2226333

[1023,] 0.6858272 1.5232999 0.1332737

[1024,] 0.5105836 0.9446608 0.3034802

[1025,] 0.6265819 1.0455016 0.2290809

[1026,] 0.5257125 0.9317825 0.2086541

[1027,] 0.5390223 0.8173957 0.1704393

[1028,] 0.7054359 1.0752846 0.1376029

[1029,] 0.5333698 1.2016076 0.1505355

[1030,] 0.6136186 1.1068474 0.2583647

[1031,] 0.6568769 1.4103191 0.1601402

[1032,] 0.6265369 1.0560657 0.1871503

[1033,] 0.5506623 0.8552067 0.1992595

[1034,] 0.5196992 0.9385043 0.1511120

[1035,] 0.6135087 1.7544707 0.1951374

[1036,] 0.5905160 0.7911310 0.1747050

Equiparación de Puntuaciones: Limitación Visual 142

[1037,]	0.4745929	0.9708370	0.2138996
[1038,]	0.6742870	0.8904534	0.2227905
[1039,]	0.5293343	1.0834885	0.1937208
[1040,]	0.5954009	0.7944604	0.2164022
[1041,]	0.4926111	0.8221884	0.2506438
[1042,]	0.5456870	0.8321889	0.1386049
[1043,]	0.5449223	1.1754734	0.2612705
[1044,]	0.5765049	1.3037913	0.2691740
[1045,]	0.6741415	1.5346048	0.1482768
[1046,]	0.5944479	0.7940983	0.2553184
[1047,]	0.5241194	0.8923749	0.3310723
[1048,]	0.6737503	1.1243103	0.1743789
[1049,]	0.6279915	0.8740011	0.2534320
[1050,]	0.5935343	0.8293195	0.2483598
