



**ANÁLISIS PSICOMÉTRICO DE UNA PRUEBA DE LECTO
ESCRITURA APLICADA EN GRADO TERCERO DE UN
COLEGIO DE LA LOCALIDAD DE SUBA**

**PSYCHOMETRIC ANALYSIS OF A READING TEST
APPLIED IN THE THIRD GRADE OF A SCHOOL OF A
TOWN IN BOGOTÁ**

Tiberio Andrés Liz Motta, Licenciado en Matemáticas y Física, Universidad de los Llanos, Villavicencio-Meta. Contacto: liz2091@gmail.com.

RESUMEN

El presente artículo contiene los resultados de la investigación para medir la confiabilidad de una prueba de lecto escritura, basada en la prueba TALE y aplicada en un colegio de la localidad de Suba en la ciudad de Bogotá; el análisis de este trabajo se hizo a través de la perspectiva de la teoría clásica de los test y se complementa con la teoría de respuesta al ítem. Lo anterior, como fundamento práctico que permita utilizar estos resultados como herramienta de apoyo en el momento de decidir cómo fortalecer los elementos de evaluación del estudiantado.



La prueba se aplicó a 248 estudiantes de grado tercero, los cuales correspondían al 100% del público objetivo al que se podía aplicar dicho test; los resultados de las teorías mencionadas, muestran que el test es confiable y que para mejorar el índice de confiabilidad es necesario eliminar tres ítems, por lo tanto el 80% de los ítems evaluados cumplen con un nivel alto de confiabilidad y se recomienda que la prueba pueda seguir siendo utilizada.

Palabras Clave: confiabilidad, TRI, TCT, psicometría. Lecto escritura.

ABSTRACT

This article contains the results of the research to measure the reliability of a reading and writing test, based on the TALE test and applied in a school in the town of Suba in the city of Bogotá; The analysis of this work was done through the perspective of the classical theory of tests and is complemented with the theory of response to the item. The above, as a practical basis that allows using these results as a support tool when deciding how to strengthen the elements of student evaluation.

The test was applied to 248 third grade students, which corresponded to 100% of the target audience to whom the test could be applied; The results of the aforementioned theories show that the test is reliable and that to improve the reliability index it is necessary to eliminate three items, therefore 80% of the evaluated items meet a high level of reliability and it is recommended that the test can continue to be used.

Keywords: reliability, TRI, TCT, psychometry. I read writing.



INTRODUCCIÓN

La preocupación por la calidad de la educación en Bogotá, ha sido uno de los pilares fundamentales en las políticas distritales y nacionales, lo cual se evidencia en el artículo 15, en el programa de formación integral: más y mejor tiempo en los colegios, dentro del plan de desarrollo económico, social, ambiental y de obras públicas del distrito capital 2020-2024, un nuevo contrato social y ambiental para la bogotá del siglo XXI (Alcaldía de Bogota, 2020) del gobierno distrital, donde menciona como resultado aumentar la calidad de la educación con la pertinencia de la jornada única y completa, a la vez disminuir brechas de inequidad que existen en el sistema educativo, donde implementarán estrategias pedagógicas innovadoras dirigidas al desarrollo de habilidades y competencias del Siglo XXI para los niños, niñas, adolescentes y jóvenes de la ciudad en el ámbito urbano y rural.

Dentro de los procesos evaluativos, es importante brindar elementos para el análisis sobre los efectos de los procesos de aprendizaje en los estudiantes, por lo tanto es necesario que las evaluaciones aplicadas a los mismos sean confiables, permitiendo aportes y conclusiones fidedignos, que de igual manera servirán como referentes al momento de tomar decisiones relacionadas con los esquemas de enseñanza del distrito. Existe un amplio consenso en todos los países de la región en que la evaluación sistemática y correctamente orientada constituye una de las estrategias principales para la mejora de la calidad de la enseñanza, y es por eso que no hay que perder el foco en la importancia al momento de abordar una evaluación educativa y la utilización de los resultados que se obtienen al aplicar la misma. (VANEGAS & JIMENEZ, 2016).

A veces se considera que se ha evaluado el aprendizaje, cuando solamente se ha medido; evaluar exige comparar el resultado de la medición con un punto de referencia que



establezca el resultado que se debería obtener de esta manera, llegar a un juicio sobre lo adecuado o inadecuado de lo medido. Un test es confiable cuando permite hacer mediciones precisas sobre lo que se está evaluando; las mediciones aplicadas a una prueba y específicamente la medición de la confiabilidad permite generar elementos esenciales para producir resultados que describan, con un alto nivel de precisión, (VANEGAS & JIMENEZ, 2016) las competencias de las personas evaluadas, en este caso, los estudiantes de un colegio de la localidad de Suba.

Por lo anterior, se hace necesario determinar la confiabilidad de un test de lecto escritura que se aplicó a los estudiantes, del colegio de la localidad en mención, y para ello, se tendrán en cuenta los fundamentos de la Teoría Clásica (TCT) y la Teoría de Respuesta al Ítem (TRI).

REFERENTES TEÓRICOS

Estamos en la era del conocimiento, donde cada decisión que se toma, requiere que se realice basados en evidencia, que esté sustentada en datos reales e interpretada de acuerdo al contexto; bajo este paradigma en el contexto educativo, se pide ahora que los programas educativos que se ponen a prueba en las instituciones escolares, en forma de planes curriculares o PEI, demuestren su efectividad (Páramo & Hederich, 2014) y esto requiere de una pedagogía basada en la evidencia, donde el programa cuente con un proceso riguroso y metódico de recolección de información que tenga como objetivo evidenciar los retos y avances del mismo y permita mejorar la calidad educativa.



En este sentido, la evaluación es un instrumento para el mejoramiento que permite obtener información válida y confiable sobre las consecuencias de acciones específicas, para así optimizar los esfuerzos (Ministerio de Educación Nacional, 2006). Sin embargo, El hecho de que haya una evaluación no resuelve el problema del mejoramiento de la calidad de la enseñanza, pero sí permite ayudar a mejorarlo, ya que es una ventana para conocer de manera práctica, que aspectos de los esquemas de aprendizaje pueden ser mejorados, (VANEGAS & JIMENEZ, 2016). Es por eso que, más allá de los objetivos prácticos como lo es la organización de alianzas productivas de trabajo entre los miembros del equipo de la institución educativa y el acompañamiento de la evolución del desempeño de los estudiantes, la realización del diagnóstico de entrada supone un aspecto que favorece el conocimiento que tiene el docente sobre el estudiante. Eso significa que debe ser un proceso de indagación científica, apoyado en una base epistemológica y cuyo objeto lo constituye la totalidad de los sujetos (individuos o grupos) o entidades (instituciones, organizaciones, programas, contextos familiar, socio-ambiental, etc.) (Hernández, 2015).

Bajo este planteamiento, la aplicación de pruebas estandarizadas y ajustadas al contexto, se presentan como herramientas para conocer el nivel en que se encuentra un grupo de estudiantes respecto a la apropiación de los contenidos, así como de otras microhabilidades relacionadas con los Derechos Básicos de Aprendizaje (DBA) y las Mallas de Aprendizaje, que deberían haber sido consolidadas. Por lo tanto es necesario que las evaluaciones aplicadas a los mismos sean confiables, permitiendo aportes y conclusiones fidedignos (VANEGAS & JIMENEZ, 2016).



De acuerdo a lo anterior, para diagnosticar las habilidades fundantes de la lectoescritura existen varias pruebas estandarizadas a nivel mundial, como la Prueba de Articulación de Fonemas (PAF) (Vallés Arándiga, 1995). Esta prueba evalúa la existencia de alteraciones disléxicas en la población infantil. Su aplicación es individual, y va dirigida a niños de 5 a 8 años, incluso en edades superiores a niños que presentan problemas de pronunciación (Manso & Méndez, 2003), o la prueba de ECOMPLEC-SEC, que permite evaluar la comprensión lectora de los individuos en tres tipologías textuales (narrativo, expositivo y discontinuo) (Bojacá, 2017).

Sin embargo, una de las más utilizada y que tiene un alto grado de confiabilidad es la “Early Grade Reading Assessment (EGRA)” versión Gran Canarias, que esta diseñada para documentar el desempeño estudiantil en las habilidades de lectura en los grados iniciales, (USAID, 2009). Esta batería se ha aplicado en más de 11 países y 19 idiomas, Se ha adoptado y utilizado por otros asociados en la ejecución en más de 30 países y más de 60 idiomas. Los datos de EGRA se han utilizado para la retroalimentación sobre la práctica docente. Muchos países han mostrado interés en utilizarlo como un trampolín para mejorar la lectura, y se han ido a rediseñar su formación del profesorado en torno al tema (USAID, 2009).

Por otro lado, existe la TALE, que es un Test de Análisis de Lectura y Escritura, definido como "una prueba destinada a determinar los niveles generales y las características específicas de la lectura y escritura de cualquier niño en un momento dado del proceso de adquisición de tales conductas (Cervera, 2015, pág. 3). Este Test fue utilizado en el año 2017 para establecer el nivel de comprensión lectora y su influencia en el rendimiento académico de los estudiantes del Primer Año de Bachillerato de la Unidad Educativa Juan de Velasco de la ciudad de Riobamba (Aguayo &



Carrasco, 2016). Con esto se determinó que en la comprensión de oraciones, los estudiantes tenían un nivel deficiente en un 37%, aceptable en un 53% y excelente en un 11% y la comprensión específica de textos, los estudiantes tenían en su gran mayoría un nivel deficiente en un 51%, aceptable en un 42% y excelente en un 7% de comprensión específica de textos (palabras) (Aguayo & Carrasco, 2016). Sin embargo, durante la revisión del estudio, no fue posible establecer el nivel de confiabilidad del instrumento.

En el año 2016, se aplicó la misma batería de pruebas en el estudio sobre la relación entre atención sostenida y procesos lectoescritores en niños de edades entre 7 y 8 años. Para esto, se tomó una muestra de 40 niños del grado segundo de primaria de un colegio distrital de Bogotá. Los resultados obtenidos dejan entre ver que si existe una correlación entre atención sostenida y procesos de lectoescritura, específicamente en comprensión lectora, dictado y copia, además, permitió establecer que en lectoescritura, los participantes presentaron tiempo y errores muy superiores a los promedios para el nivel II, reflejando niveles de atención bajo (Rojas, 2016). Sin embargo, al igual que la anterior investigación, ésta también careció de un proceso de confiabilidad.

Los estudios anteriores demuestran que el Test de Análisis de Lectura y Escritura TALE se puede ajustar al contexto de los estudiantes o establecer una prueba con base a estas, sin embargo, como se mencionó en ambos estudios igual que en muchos otros, carecen de un análisis psicométrico, que les permita determinar la confiabilidad del instrumento, y con esto, poder hacer mediciones precisas sobre lo que se está evaluando; las mediciones aplicadas a una prueba y específicamente la medición de la confiabilidad permite generar elementos esenciales para



producir resultados que describen, con un alto nivel de precisión, las competencias de las personas evaluadas (VANEGAS & JIMENEZ, 2016).

- Psicometría

Existen varias referencias sobre la historia de la psicometría, una de ellas se remonta a las dinastías del antiguo imperio chino, quienes generaron los primeros sistemas de evaluación de los individuos en función de su habilidad (Meneses, 2017, pág. 26). Estos avances lograron poner en marcha los sistemas de clasificación oficial mediante la expedición de títulos formales que acreditaban el acceso a los diferentes niveles de responsabilidad. (Meneses, 2017, pág. 26). Sin embargo, no fue sino hasta la invención y la incorporación del papel a partir del siglo XVI que fue trasladando la confianza en el juicio personal basado en impresiones, hacia la administración de pruebas institucionales basadas en una autoridad imparcial, que objetivaba las destrezas y los conocimientos requeridos en los ámbitos educativo y administrativo. (Meneses, 2017, pág. 26).

Por otro lado, las primeras pruebas de evaluación psicológica se dieron a mediados del siglo XIX, teniendo en cuenta, el esfuerzo de los profesionales orientados al estudio de los problemas mentales y las lesiones cerebrales. (Meneses, 2017, pág. 27). Estos antecedentes y con el estudio sistemático de las diferencias humanas durante el mismo siglo, que gracias a los trabajos de Friedrich W. Bessel (1784-1846) y Carl F. Gauss (1777-1855) y las contribuciones de Gustav T. Fechner (1801-1887) y Hermann von Helmholtz (1821-1894), constituyeron un avance importante en la sistematización de la medida de las sensaciones psicológicas producidas por la estimulación física (Meneses, 2017, pág. 28).



Sin embargo, Sir Francis Galton (1822-1911), James McKeen Cattell (1860- 1944) y Alfred Binet (1857-1911) son considerados como padres de la psicometria moderna, pues los dos primeros autores se aventuraron a postular relaciones entre sus medidas y el intelecto, llegando a conclusiones a veces controvertidas y otras ampliamente criticadas con posterioridad por su simplismo. Por su parte, Binet adoptó un enfoque innovador y fue responsable de lo que se considera el primer test de aplicación general para la medida de las habilidades cognitivas. (Meneses, 2017, pág. 28).

No fue sino hasta 1905, que la psicometria ingresó en las aulas de clase, gracias, a que Binet y su colega Théodore Simon desarrollaron el primer test para cuantificar la inteligencia, teniendo un grupo de referencia en niños de 3 a 13 años. Esto condujo a la definición de lo que se conoce, hasta la actualidad, como edad mental (Meneses, 2017, pág. 29). Los siguientes avances se dieron durante la primera guerra mundial, gracias al apoyo del Gobierno de Estados Unidos, pues los tests deberían servir como instrumentos de evaluación psicológica para el reclutamiento de los nuevos soldados. (Meneses, 2017, pág. 29).

- Teoría Clasica de los Test (TCT).

Un test psicologico es un procedimiento de medida objetiva y estandarizada de una muestra de comportamientos. (Meneses, 2017, pág. 39), que contiene tres elementos fundamentales, como: ser objetivo, estandarizado y enfrentarse a fenomenos psicologicos no observables, mediante una muestra de comportamientos, la cual permite establecer inferencias adecuadas sobre el comportamiento general de las personas más allá de los elementos específicos evaluados mediante tests. (Meneses, 2017, pág. 40).



Esta teoría se basa en el modelo lineal clásico propuesto por Spearman, que articula el proceso de medida definiendo tres conceptos fundamentales: la puntuación verdadera, la puntuación empírica y el error de medida. Así, partiendo del modelo operacional de medida psicométrica, esta teoría de los tests no está interesada en el sistema de relaciones empíricas, sino que centra su atención en el análisis de las puntuaciones obtenidas para valorar los errores cometidos en el proceso de medida indirecta de los fenómenos psicológicos. (Meneses, 2017, pág. 51).

De acuerdo a lo anterior se puede establecer que una teoría clásica de los tests es una teoría que proporciona modelos para las puntuaciones de los tests; por tanto, el problema central de una teoría de los tests es la relación que existe entre el nivel del sujeto en la variable no observable que se desea estudiar y su puntuación observada en el test (Salvador Chacón Moscoso, 2008, pág. 37).

Aunque la TCT resulta útil para establecer una aproximación para la estimación de las puntuaciones verdaderas, a partir de las puntuaciones empíricas. Tiene unas limitaciones que no son fáciles de resolver desde el modelo lineal clásico. Tal y como discutieron Lord y Novick (1968) en su reformulación, esta teoría tiene dos dependencias importantes en relación con los instrumentos y las muestras utilizadas. Por un lado, una dependencia de las puntuaciones empíricas obtenidas respecto a los instrumentos. Por otro lado, una dependencia de las propiedades psicométricas de los tests respecto a las muestras utilizadas para desarrollarlos. (Meneses, 2017, pág. 53).



- Teoría de respuesa al ítem (TRI)

Para solucionar lo anterior, la Teoría de la Respuesta a los Items, supone un cambio de planteamientos respecto a la TCT, aunque ambas teorías no pueden considerarse contrapuestas sino más bien complementarias (Salvador Chacón Moscoso, 2008, pág. 38). Pues mientras que la primera establece una relación lineal entre la puntuación del sujeto y su valor en el rasgo medio, la segunda, establece esta relación no lineal.

la teoría TRI plantea la existencia de una relación entre las puntuaciones en la variable latente y la probabilidad de acertar cada ítem introduciendo diferentes funciones matemáticas para modelar adecuadamente esta relación. Así, una vez seleccionada la función más adecuada se construyen lo que se denominan las curvas características de los ítems, una modelización de las respuestas que puede tener en cuenta diferentes parámetros, como la capacidad de discriminación de los ítems, su dificultad o la probabilidad de que sean acertados al azar. (Meneses, 2017, pág. 54).

El término Teoría de la Respuesta a los Items se debe a Lord (1980) y tienen sus orígenes en trabajos publicados en las décadas de los años 30 y 40 (Brodgen, 1946; Ferguson, 1942, Richardson, 1936; Tucker, 1946; Lawley, 1943, 1944), pero que se desarrollaron fundamentalmente a partir de las publicaciones de Lord en la década de los 50 (Lord, 1952, 1953a, 1953b). En la década siguiente se desarrollan modelos logísticos más tratables en un nivel matemático que los modelos de ojiva normal. Rasch (1960), presenta un modelo logística de un parámetro, y Birnbaum (1968) propone en la obra de Lord y Novick (1968) *Statistical theories of mental test scores*, otros dos modelos de dos y tres parámetros. Estos



tres modelos logísticos de uno, dos y tres parámetros, constituyen actualmente el tronco básico de la TRI. (Salvador Chacón Moscoso, 2008, pág. 38).

- TALE

El Test de Analisis de Lecto Escritura TALE es una prueba destinada a determinar los niveles generales de la lectura y escritura de cualquier niño en un momento dado del proceso de adquisición de tales conductas (Muñoz Carpio Karen Estefanía, 2017).

Fue diseñada con fines de asistencia, enseñanza e investigación. Frente a los test de prelectura que analizan los requisitos para la lectura, este test trata de analizar la lectura establecida. Comprende dos partes (Lectura y Escritura) cada una de las cuales está integrada por varias pruebas. (Josep Toro, 2020).

Lectura: esta parte de la prueba consiste en realizar una actividad que permita verificar la fluidez lectora y el proceso de comprensión.

Fluidez lectora: El Ministerio de Educación Nacional expone en las Mallas de Aprendizaje de Lenguaje que en el grado tercero es un imperativo mejorar la fluidez lectora de los estudiantes, pues ésta permite que el lector se enfoque en la comprensión, ya que puede decodificar e identificar palabras de forma precisa, lo que se evidencia en una lectura con entonación adecuada y que respeta los signos de puntuación. También explican que la fluidez posee tres componentes principales: leer palabras y textos con precisión, la automaticidad o identificación rápida y sin esfuerzo de las palabras y la lectura expresiva y a un ritmo adecuado (Ministerio de Educación, 2018, pág. 16). En este caso se debe escoger



un texto entre 90 y 110 palabras, de estructura sencilla y con vocabulario conocido (no se deben superar las 6 palabras nuevas por texto) que permitan determinar el nivel en que se encuentra el estudiante.

Comprensión de lectura: En el grado tercero se espera que los estudiantes fortalezcan los procesos de comprensión lectora, para ello es esencial su contacto con diversidad de textos para que reconozca sus contenidos, estructuras, y propósitos comunicativos. Por tal razón la prueba centra las preguntas de comprensión en tres niveles de lectura:

La lectura literal Es el nivel básico de lectura centrado en las ideas y la información que está explícitamente expuesta en el texto. La Lectura literal es reconocimiento de detalles (nombres, personajes, tiempos y lugar del relato), reconocimiento de la idea principal de un párrafo o del texto, identificación de secuencias de los hechos o acciones, y identificación de relaciones de causa o efecto (identificación de razones explícitas relacionadas con los hechos o sucesos del texto). (Herazo, 2020)

La lectura inferencial: se constituye la lectura implícita del texto y requiere un alto grado de abstracción por parte del lector. El objetivo de la lectura inferencial es la elaboración de conclusiones y se reconoce por inferir detalles adicionales, inferir ideas principales no explícitas en el texto, inferir secuencias de acciones relacionadas con la temática del texto, inferir relaciones de causa y efecto (partiendo de formulación de conjeturas e hipótesis acerca de ideas o razones),



predecir acontecimientos sobre la lectura y interpretar el lenguaje figurativo a partir de la significación literal del texto. (Herazo, 2020)

La lectura crítica: es la lectura de carácter evaluativo donde intervienen los saberes previos del lector, su criterio y el conocimiento de lo leído, tomando distancia del contenido del texto para lograr emitir juicios valorativos desde una posición documentada y sustentada. Los juicios deben centrarse en la exactitud, aceptabilidad y probabilidad; pueden ser: de adecuación y validez (compara lo escrito con otras fuentes de información), de apropiación (requiere de la evaluación relativa de las partes) y de rechazo o aceptación (depende del código moral y del sistema de valores del lector). (Herazo, 2020).

Dictado de palabras y oraciones: El propósito del dictado es verificar el dominio de la conciencia fonológica (correspondencia fonema – grafema) y la automatización del proceso de codificación de palabras y oraciones de uso cotidiano de los estudiantes y la apropiación del código escrito.

Escritura de un texto corto: El propósito del ejercicio es comprender el nivel de apropiación del código escrito para producir un texto con una intención comunicativa, con elementos mínimos del texto narrativo y que éste posea adecuación a nivel global y local y convencionalismos como la segmentación, ortografía y uso de signos de puntuación.

METODOLOGÍA



El tipo de estudio que se resulto más pertinente para el desarrollo del presente proyecto es el estudio observacional, de tipo descriptivo, donde a principal característica de los estudios descriptivos es que se limitan simplemente a dibujar el fenómeno estudiado (Jorge Veiga de Cabo, 2008).

- Variables de estudio.

Comprende los 10 ítem que componen la prueba de lecto escritura aplicada.

- Participantes.

Se contó con la participación de docientos cuarenta y ocho estudiantes que hacen parte del grado tercero de un colegio de la localidad de Suba de la ciudad de Bogotá, elegido para realizar la prueba. La prueba se aplicó al 100% de la población, en el primer trimestre academico del año 2019. Docientos cuarenta y ocho estudiantes, 124 corresponden al género femenino y 124 al género masculino, con una edad aproximada entre 8 a 9 años.

- Instrumentos de recolección de información.

La recolección de la información para determinar la confiabilidad de un test aplicado a grado tercero, de un colegio de la localidad de Suba en la ciudad de Bogotá, se hizo a través de un test de lecto escritura, basado en la prueba TALE, cuyas habilidades a evaluar eran las de lectura y escritura, a través de cinco microhabilidades, tales como: fluidez de la lectura, comprensión lectora, dictado de palabras y oraciones, escritura de oraciones y escritura de textos cortos. La prueba se aplicó durante el primer trimestre del año 2019. El análisis se realizó utilizando la teoría clásica de los test y la teoría de respuesta al ítem.

- Estructura del Test.

El Test consta de 10 ítems, distribuidos de la siguiente forma.

| HABILIDADES | MICRO HABILIDADES | ITEMS |
|-------------|-------------------|-------|
|-------------|-------------------|-------|



| | | |
|-----------|---------------------------------|---------------|
| LECTURA | Fluidez de la lectura | 1 |
| | Comprensión lectora | 2, 3, 4, 5, 6 |
| ESCRITURA | Dictado de palabras y oraciones | 7 |
| | Escritura de oraciones | 8 |
| | Escritura de textos cortos | 9, 10 |

RESULTADOS

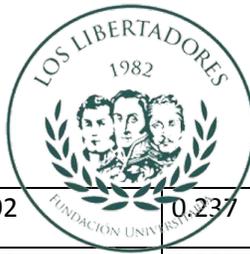
- Análisis de modelo Rasch

El modelo Rasch es un recurso que permite analizar la calidad técnica de los ítems que componen la prueba de acuerdo al nivel de dificultad, que se establece desde valores negativos, hasta los positivos, indicando los primeros valores una dificultad menor y los segundos, un nivel de dificultad mayor, se sugiere que aquellos valores que se encuentren menores o cercanos a -3, son ítems muy fáciles y aquellos mayores o cercanos a +3, son muy difíciles.

El cero representa el nivel de dificultad promedio, por tanto, aquellos valores que se acerquen a 0, ya sea positivos o negativos, indican este nivel de dificultad. De acuerdo a la base dicotómica construida a partir de los resultados de la prueba, se construyó el modelo permitiendo generar la siguiente tabla 1. Para este caso el nivel de dificultad se encuentra en la columna (Estimate std). También permite ver el error de medida (Error), como el límite inferior (lower CI) y superior (upper CI) de dificultad.

Table 1. Resultados modelo RASCH

| Item Easiness Parameters (beta) with 0.95 CI: | | | | |
|---|--------------|-------|----------|----------|
| | Estimate std | Error | lower CI | upper CI |



| | | | | |
|---------------|--------|-------|--------|--------|
| beta 1 | -1.392 | 0.237 | 0.928 | 1.857 |
| beta 2 | -1.120 | 0.235 | -1.581 | -0.660 |
| beta 3 | -2.194 | 0.248 | -2.681 | -1.708 |
| beta 4 | -0.905 | 0.234 | -1.363 | -0.446 |
| beta 5 | -2.701 | 0.260 | -3.211 | -2.191 |
| beta 6 | -1.927 | 0.243 | -2.404 | -1.450 |
| beta 7 | -2.701 | 0.260 | -3.211 | -2.191 |
| beta 8a | 6.261 | 1.120 | 4.066 | 8.456 |
| beta 8b | 6.261 | 1.120 | 4.066 | 8.456 |
| beta 9 Animal | -1.096 | 0.235 | -1.556 | -0.636 |
| beta 9 Grande | -1.678 | 0.240 | -2.148 | -1.208 |
| beta 10 | 3.193 | 0.401 | 2.407 | 3.979 |

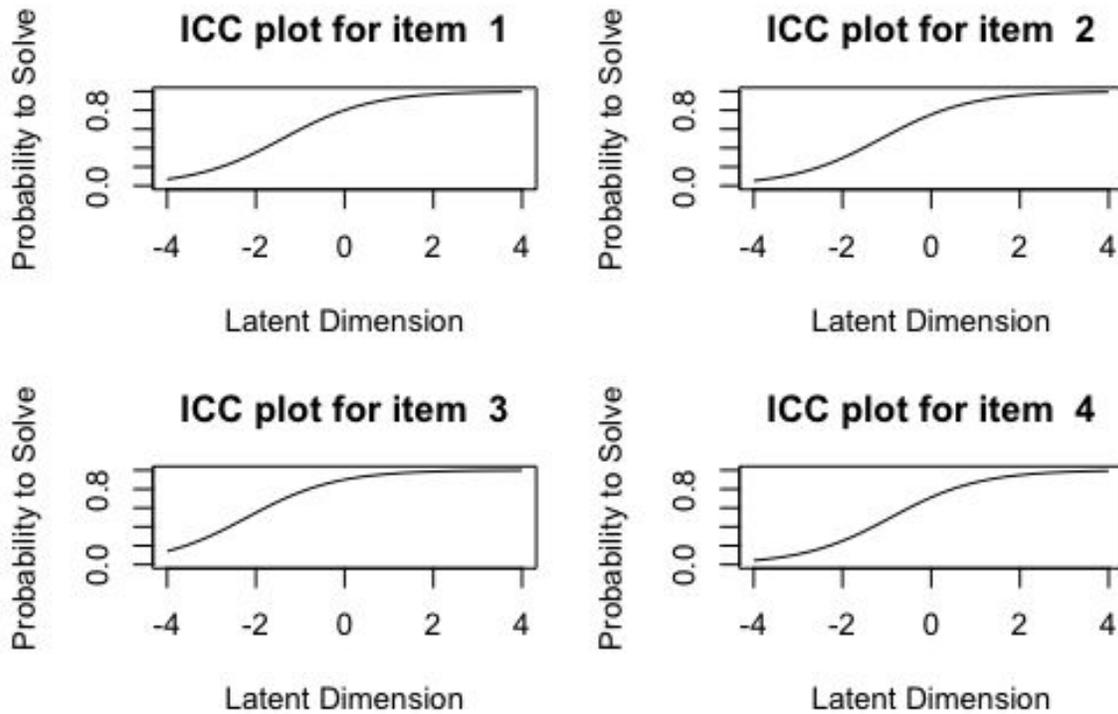
De acuerdo a la tabla 1, se puede evidenciar algunos valores positivos muy grandes, los cuales se encuentran en los ítems 8a, 8b y 10, demostrando que éstos tienen un mayor nivel de dificultad. Por otro lado, se puede evidenciar que en los ítems 1, 2, 3, 4, 5, 6, 7 y 9 donde se le pedía al estudiante realizar una oración con el sustantivo “Animal” y adjetivo “Grande” su nivel de dificultad se ajusta al modelo, ya que están dentro del rango de -3 a +3 lo que indica que, por ahora, estos tres ítems se encuentran acorde a los estudiantes.

- Gráficas por Ítem

Las gráficas por ítem se presentan por cada uno, con su debida interpretación, teniendo en cuenta que en el eje de las x indica el nivel de habilidad de los individuos, y en el eje y la probabilidad de resolver el ítem.



Item 1 - 4



Bajo este esquema, se puede ver que en el ítem 1, los estudiantes que tienen una habilidad menor, tienen cerca del 0% de probabilidad de resolver el ítem, pero aquellos estudiantes que tienen una habilidad promedio o mayor, tienen más del 70% de probabilidad de resolverlo. Todo lo anterior indica que el ítem tiene una dificultad menor, razón de esto es que en la tabla 1, tenemos que el nivel de dificultad de este ítem es de -1.392.

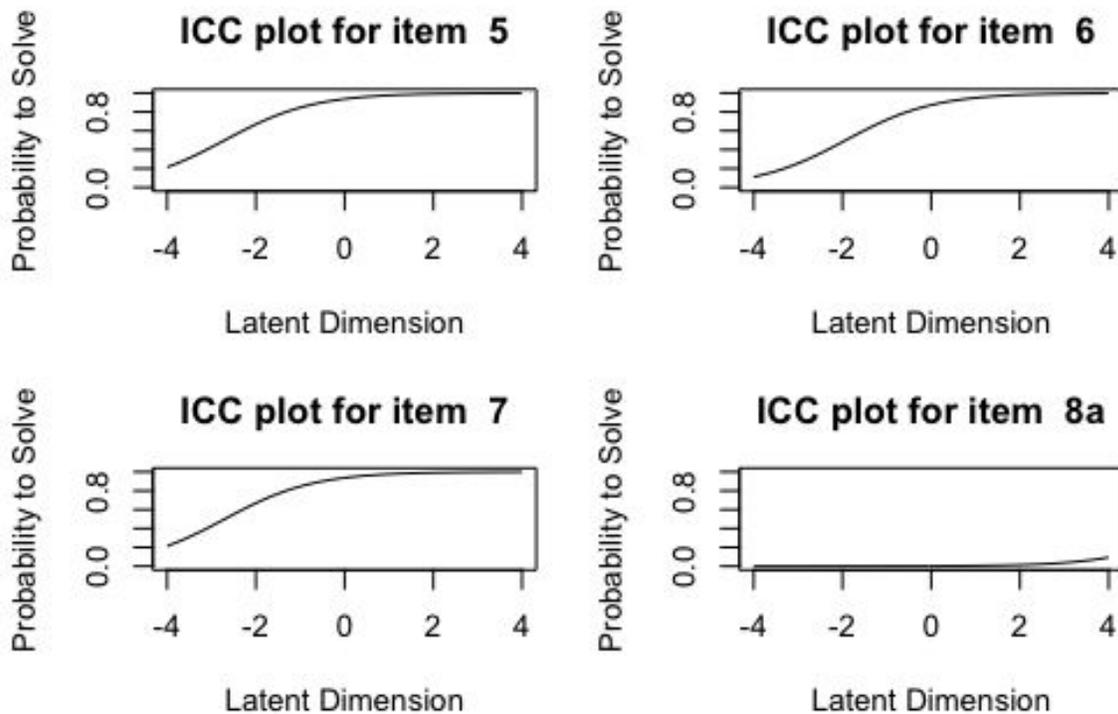
Observando el ítem 2, se puede evidenciar que los estudiantes con un nivel de habilidad menor, tienen cerca de 10% de probabilidad de responderlo, mientras que los que tienen una habilidad media o mayor, tienen cerca de un 60% de solucionarlo, se podría decir que este ítem se encuentra ajustado, además, observando la tabla 2, podemos ver que el nivel de dificultad del ítem es -1.120. Lo que nos da a entender que tiene un nivel de dificultad promedio.



En el ítem 3, se observa que los estudiantes con un nivel de habilidad promedio, tienen cerca de 80% de probabilidad de solucionar el ítem, mientras que aquellos que tienen una habilidad menor tienen cerca de un 20% de probabilidad de solucionarlo. Este ítem podría ser candidato a ser ajustado, pues si el nivel de dificultad es menor y la tabla uno lo confirma con - 2.194.

Si observamos el ítem 4, se puede ver que los estudiantes con una habilidad promedio, pueden solucionar el ítem con una probabilidad de cerca de un 60%, mientras que aquellos que cuentan con una habilidad menor, tienen una probabilidad por debajo de 10%, lo que nos indica que este ítem se encuentra acorde a la habilidad de los estudiantes, conclusión que nos reafirma el nivel de dificultad del ítem que se presenta en la tabla uno, que es de -0.95.

Item 5 – 8a



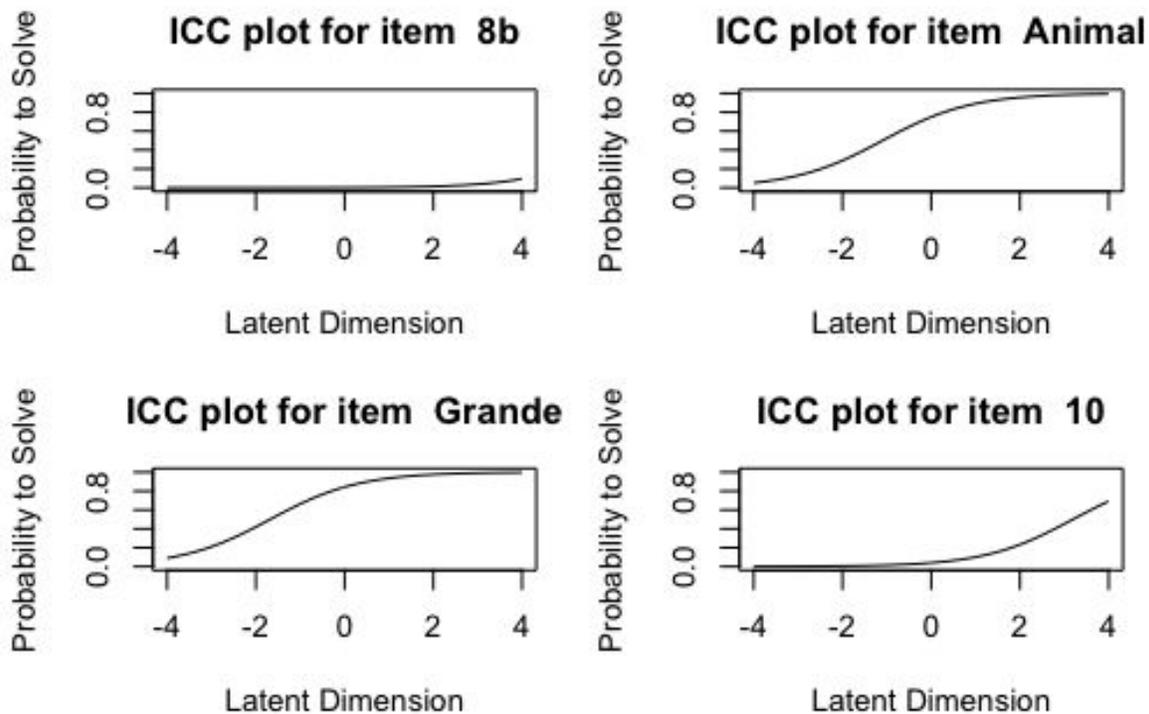


El ítem 5, evidencia que los estudiantes con un nivel de habilidad promedio, tienen cerca de un 80% de probabilidad de solucionarlo, mientras que aquellos que tienen una habilidad menor, tienen cerca de un 20%, esto nos puede indicar que el ítem no tiene mucha dificultad para su solución y cuenta con un nivel de dificultad de -2.701.

En el ítem 6, se puede ver que los estudiantes que tienen una habilidad menor, tienen cerca de 10% de probabilidad de responder el ítem, mientras que los estudiantes que tienen una habilidad promedio, tienen una probabilidad arriba del 80% para responder el ítem, lo que se puede inferir que el ítem tiene una menor dificultad y se confirma con el nivel de dificultad que es de - 1.927.

En el ítem 7, se evidencia que los estudiantes que tiene un nivel de habilidad menor, tienen una probabilidad de solucionarlo de cerca del 20%, mientras que aquellos que tienen una habilidad promedio o mayor a este, la probabilidad de solución es por arriba del 85%, determinando que es fácil de solucionar, esto se confirma cuando se observa el nivel de dificultad, que es de -2.701.

En el ítem 8a, se evidencia que los estudiantes que podrían solucionarlo, son aquellos que tienen una habilidad alta, pues tendría cerca de un 10% de realizarlo. Esto indica que el ítem tiene una dificultad muy alta, lo que se comprueba con la tabla uno, pues este ítem tiene un nivel de dificultad de 6.261.



El ítem 8b presenta el mismo comportamiento que el anterior, por tanto, se puede inferir que su dificultad es alta, lo que se comprueba en la tabla, que evidencia que el nivel de dificultad es de 6.261.

El ítem 9, donde el estudiante tenía que escribir una oración con el sustantivo animal, se puede ver que los estudiantes que tienen una menor habilidad, tienen una probabilidad de responder cercana a 5%, mientras que aquellos que tienen una habilidad promedio, tienen una probabilidad de respuesta cercana a 50%. Lo que se puede inferir que este ítem está ajustado y cuenta con un nivel de dificultad de -1.096.

En el ítem 9, donde el estudiante tenía que escribir una oración con el adjetivo grande, se puede ver que los estudiantes que tienen una menor habilidad, tienen una probabilidad de



responder cercana a 5%, mientras que aquellos que tienen una habilidad promedio, tienen una probabilidad cercana a 80%, este ítem podría ser sujeto a ajuste, pues su dificultad es menor, además que su nivel es de -1.678.

En el ítem 10, se puede ver que los estudiantes que tienen una habilidad promedio, tienen una probabilidad de responder cercana a 5%, mientras que aquellos que tienen una habilidad mayor, tienen una probabilidad cercana de 70%. De acuerdo a esto, se puede inferir que el ítem tiene una dificultad mayor, lo cual se confirma con en la tabla, que presenta un nivel de dificultad de 3.193.

- Análisis de modelo Cronbach.

El alpha de Cronbach permite medir la fiabilidad de la consistencia interna del instrumento, el método permite estimar la fiabilidad de un instrumento de medida a través de un conjunto de ítems, teniendo en cuenta los resultados de esto.

El alpha puede ser calculado a partir de la correlación entre los ítems como una función del número de ítems del instrumento y la correlación media de estos, bajo estos aspectos, el alpha arroja valores entre 0 a 1, siendo los valores cercanos a 0, son valores inaceptables, mientras los que se acercan a 1, son valores aceptados: dentro de un análisis exploratorio estándar, el valor de consistencia interna en torno a 0.7 es adecuado y es el nivel mínimo aceptable (Frías-Navarro, 2020).

De acuerdo con lo anterior, se calculó el alpha obteniendo los siguientes resultados.



Tabla 1. Valor del alpha

| raw_alpha |
|-----------|
| 0.8 |

Raw_alpha, indica el valor del coeficiente, a nivel general, el cual es 0.8, lo que nos permite inferir que se considera aceptable la confiabilidad de los resultados de la prueba en forma general.

Tabla 2. Valor promedio de correlación

| average_r |
|-----------|
| 0.22 |

Por otro lado, el average_r, es el valor promedio de correlación entre los ítems. Entre más cercano sea este valor a 1, los ítems han tenido más asociación entre ellos. El valor de 0.22, nos indica que los ítems en promedio tienen una baja asociación entre ellos.

Esto evidencia que es posible obtener valores altos del alpha, aunque los ítems no tengan correlaciones fuertes entre ellos. Por tanto no podemos asumir que tenemos ítems con una fuerte asociación entre ellos únicamente calculando el alpha.

Tabla 3. Intervalo de confianza

| 95% confidence boundaries | | |
|---------------------------|-------|-------|
| lower | alpha | upper |
| 0.76 | 0.8 | 0.83 |



El anterior cuadro, muestra el intervalo de confianza del alpha con un 95%. Intervalos altos, indican que el valor del alpha tiene un mayor error de medida. Para este caso, se podría decir que es un intervalo moderado.

Tabla 4. Valor del alpha, dependiendo de la ausencia de cada ítem.

| Reliability if an item is dropped: | | | | | | |
|------------------------------------|-----------|-----------|-----------|----------|-------|-------|
| items | raw_alpha | std.alpha | average_r | alpha se | var.r | med.r |
| 1 | 0.76 | 0.74 | 0.21 | 0.020 | 0.033 | 0.23 |
| 2 | 0.78 | 0.75 | 0.22 | 0.018 | 0.035 | 0.23 |
| 3 | 0.77 | 0.75 | 0.21 | 0.019 | 0.035 | 0.23 |
| 4 | 0.79 | 0.76 | 0.23 | 0.017 | 0.037 | 0.25 |
| 5 | 0.78 | 0.76 | 0.22 | 0.018 | 0.037 | 0.23 |
| 6 | 0.78 | 0.75 | 0.22 | 0.018 | 0.036 | 0.23 |
| 7 | 0.77 | 0.75 | 0.21 | 0.019 | 0.034 | 0.23 |
| 8a | 0.80 | 0.78 | 0.24 | 0.017 | 0.021 | 0.26 |
| 8b | 0.80 | 0.78 | 0.24 | 0.017 | 0.021 | 0.26 |
| 9a Animal | 0.77 | 0.75 | 0.21 | 0.019 | 0.034 | 0.23 |
| 9b Grande | 0.77 | 0.75 | 0.21 | 0.019 | 0.033 | 0.23 |
| 10 | 0.80 | 0.79 | 0.25 | 0.017 | 0.033 | 0.28 |

Reliability if an item is dropped, refiere las puntuaciones del alpha como variarían dependiendo de la ausencia de algún ítem. Bajo este aspecto, el alpha cae cerca o más de 0.03 puntos, si se quitaran los ítems, 1 o 3, 7, 9a animal, y 9b grande.

Resulta importante resaltar que si falta los ítems, 8a, o 8b, o 10, no afectarían la confiabilidad de los resultados.



Tabla 5. Correlación del ítem Vs puntuación total

| Item statistics | | | | | | | |
|-----------------|-----|-------|-------|-------|--------|-------|-------|
| | n | raw.r | std.r | r.cor | r.drop | mean | sd |
| 1 | 248 | 0.71 | 0.67 | 0.66 | 0.607 | 0.601 | 0.491 |
| 2 | 248 | 0.63 | 0.59 | 0.55 | 0.502 | 0.556 | 0.498 |
| 3 | 248 | 0.65 | 0.62 | 0.58 | 0.539 | 0.718 | 0.451 |
| 4 | 248 | 0.54 | 0.50 | 0.42 | 0.387 | 0.520 | 0.501 |
| 5 | 248 | 0.56 | 0.54 | 0.48 | 0.440 | 0.778 | 0.416 |
| 6 | 248 | 0.62 | 0.59 | 0.55 | 0.499 | 0.681 | 0.467 |
| 7 | 248 | 0.66 | 0.63 | 0.60 | 0.566 | 0.778 | 0.416 |
| 8a | 248 | 0.12 | 0.37 | 0.25 | 0.096 | 0.004 | 0.064 |
| 8b | 248 | 0.12 | 0.37 | 0.25 | 0.096 | 0.004 | 0.064 |
| Animal | 248 | 0.67 | 0.62 | 0.60 | 0.548 | 0.552 | 0.498 |
| Grande | 248 | 0.68 | 0.64 | 0.62 | 0.570 | 0.645 | 0.479 |
| 10 | 248 | 0.22 | 0.28 | 0.15 | 0.143 | 0.044 | 0.206 |

Item statistics, nos presenta las correlaciones (raw.r) de los ítems con la puntuación total de la escala. Valores más cercanos a 1 indican una asociación más fuerte entre ítem y puntuación. En este caso, los ítems 8a y 8b nos presentan una correlación de (0.12) indicando que es baja, al igual que el ítem 10. Valdría la pena analizarlos más detalladamente, pues se espera que los ítems tengan una buena asociación con la puntuación.

Tabla 6. Proporción de respuestas

| Non missing response frequency for each item | | | |
|--|------|------|------|
| items | 0 | 1 | miss |
| 1 | 0.40 | 0.60 | 0 |



| | | | |
|--------|------|------|---|
| 2 | 0.44 | 0.56 | 0 |
| 3 | 0.28 | 0.72 | 0 |
| 4 | 0.48 | 0.52 | 0 |
| 5 | 0.22 | 0.78 | 0 |
| 6 | 0.32 | 0.68 | 0 |
| 7 | 0.22 | 0.78 | 0 |
| 8a | 1.00 | 0.00 | 0 |
| 8b | 1.00 | 0.00 | 0 |
| Animal | 0.45 | 0.55 | 0 |
| Grande | 0.35 | 0.65 | 0 |
| 10 | 0.96 | 0.04 | 0 |

Con respecto a Non missing response frequency for each item, nos muestra la proporción de respuestas para cada opción posible para cada uno de los items, incluidas las respuestas perdidas. En este orden de ideas, la tabla evidencia que no existe respuestas perdidas en la prueba y que el ítem 8a y 8b, tienen el 100% de respuestas en la opción 0, que para nosotros, era equivalente a haber fallado en la solución del ítem, al igual que el ítem 10, quién presenta un 96% de respuesta en la misma opción de respuesta.

DISCUSIÓN DE LOS RESULTADOS.

La aplicación del modelo Rasch permitió ver el nivel de dificultad de cada ítem y con ello, visualizar en las curvas, la probabilidad de respuesta que necesitaba un estudiante con habilidad promedio para resolver cada ítem.



De lo anterior se pudo establecer que los ítem que más se ajustaban al nivel de habilidad promedio de los estudiantes, son aquellos que se encuentran en un nivel de dificultad entre -1 y 1 . Con respecto a los demás, es necesario realizar un ajuste para, bien sea bajar el nivel de dificultad, o aumentarlo.

Por otro lado, se puede ver que los ítems 8a, 8b y 10, son aquellos que presentan mayor dificultad, sin embargo, los dos primeros, tienen un error mucho más alto de los demás, como también un intervalo más amplio, lo que se puede inferir que algo sucedió en la aplicación en la prueba, ya que estos comportamientos en los datos no es normal.

En el caso de la aplicación del modelo de Cronbach, se pudo establecer que la prueba cuenta con un nivel de confiabilidad aceptable, ya que el α calculado fue de 0.8 , sin embargo, el valor promedio de correlación entre los ítems es bajo, pues su resultado fue de 0.22 , por tanto, es necesario realizar un análisis más profundo entre los ítems más adelante.

Observando los ítems, se puede evidenciar que el ítem que más puede disminuir la confiabilidad de los resultados, es el ítem 1, ya que si se elimina, el resultado descendería a 0.76 . Por otro lado, los ítems que no afectarían la confiabilidad si se llega a eliminar alguno de ellos son los 8a, 8b, y 10, esto coincide con que estos tres ítems, son los que presentan una menor asociación con la puntuación. Esto puede deberse a que los ítems mencionados, cerca del 100% de los datos fueron categorizados con (0), que para este caso, indicaría como respuestas erróneas, o sin respuesta.

CONCLUSIONES



De acuerdo al análisis realizado, los ítems se pueden agrupar en tres grupos, el primero, son aquellos que presentan un nivel de dificultad entre -2 y 2 y que son significativos para la confiabilidad de la prueba, pues éstos mismos harían variar la confiabilidad entre 0.01 a 0.03 puntos en forma negativa (1, 2, 4, 6, 9 Animal y 9 Grande). Los segundos ítems, son aquellos que tienen un nivel alto de dificultad y que coincide que si falta alguno de estos ítems, no afectaría la confiabilidad de los resultados, (8a, 8b y 10) Y por último, se encuentran aquellos ítems que tienen una dificultad mucho menor, o sea de a -2, hasta -3, los cuales, si faltaran, afectarían la confiabilidad de los resultados de una forma no tan certera, como los del primer grupo (3, 5 y 7).

En el ítem 8 se le pedía al estudiante que escribiera dos oraciones que el profesor les dictaba y correspondían a las siguientes: “a. Los elefantes son tan inteligentes como las ballenas y los simios” y “b. Los colmillos del elefante les sirven para atacar y defenderse”. Sin embargo, cuando se calificó la prueba, se pudo evidenciar que en el 100% de las pruebas, los estudiantes habían acertado las dos oraciones y sus respuestas correspondían a los siguientes párrafos “Los elefantes son inteligentes y sociales, y los colmillos de los elefantes son blancos y grandes”. En vista que en la mayoría de pruebas se presentaba resultados cercanos a estos dos párrafos, se puede inferir que la indicación del facilitador cambió al momento de aplicar el instrumento y por esta razón se origina la discrepancia en los resultados, generando que por un lado, el nivel de dificultad sea mayor, como se evidenció con el modelo Rasch, como también que la ausencia de éstos, podría no afectar la confiabilidad de los resultados.

El análisis de los dos modelos, permitió evidenciar que los ítems fallados, o no resueltos que tuvieron en su mayoría estas calificaciones, hace que el nivel de dificultad aumente y que esto no



genere cambios en la confiabilidad de los resultados. Este hecho se visualiza en el ítem 10, donde cerca del 96% de los datos obtuvieron esta calificación y no presenta una variación en el α de Cronbach.

Los ítems 2, 4, 6, son aquellos que podrían estar ajustados al nivel de habilidad promedio de los estudiantes y que pueden afectar muy poco la confiabilidad de los resultados, por tanto, serían los que se pueden desarrollar.

Por otro lado, en el ítem 10, que correspondía a plantear una historia, donde el personaje iba a ser un elefante, lo cual se debería tener personajes, tiempo, espacio y se coherente, se encontró en la mayoría de resultados que las historias planteadas carecían si bien de una o varias de estas características, por tanto en su mayoría los resultados no fueron muy satisfactorios.



REFERENCIAS BIBLIOGRÁFICAS

- Olga Lucía Uribe-Enciso, S. C.-G. (Diciembre de 2014). Relación entre la lecto-escritura, el desempeño académico y la deserción estudiantil. *Relación entre la lecto-escritura, el desempeño académico y la deserción estudiantil*. Bucaramanga, Colombia.
- Esperanza Guarneros Reyes, L. V. (Febrero de 2014). *Habilidades lingüísticas orales y escritas para la lectura y escritura en niños preescolares*.
- Páramo, P., & Hederich, C. (2014). Revista Colombiana de Educación. *Educación basada en la evidencia*. Bogota: Universidad Pedagógica Nacional.
- Ministerio de Educación Nacional. (Marzo de 2006). Evaluar y promover el mejoramiento. *Al tablero*. Bogota: Ministerio de Educación Nacional.
- Hernández, M. A. (31 de Julio de 2015). EL DIAGNÓSTICO EDUCATIVO, UNA IMPORTANTE HERRAMIENTA PARA ELEVAR LA CALIDAD DE LA EDUCACIÓN EN MANOS DE LOS DOCENTES. *Atenas*. Matanzas, Cuba: Universida de Matanzas, Camilo Cienfuegos.
- USAID. (2009). *Manual para la evaluación inicial de la lectura en niños de educación primaria*. RTI International.
- Aguayo, E. D., & Carrasco, Á. S. (2016). TESIS PREVIA A LA OBTENCIÓN DEL GRADO DE MAGISTER EN DOCENCIA, MENCIÓN INTERVENCIÓN PSICOPEDAGÓGICA. *La comprensión lectora y su influencia en el rendimiento académico de los estudiantes del Primer Año de Bachillerato de la Unidad Educativa Juan de Velasco de la ciudad de Riobamba, período Septiembre 2014 - Marzo 2015*. Riobamba, Ecuador: UNIVERSIDAD NACIONAL DE CHIMBORAZO .
- Rojas, C. M. (Diciembre de 2016). *Estudio sobre la relación entre atención sostenida y procesos lectoescritores en niños de edades entre 7 y 8 años*. Bogotá, Colombia: Universidad Internacional de la Rioja.
- VANEGAS, E. A., & JIMENEZ, L. O. (2016). ANÁLISIS PSICOMÉTRICO DE UNA PRUEBA DE FÍSICA PARA GRADO DECIMO DEL COLEGIO NICOLÁS GÓMEZ DÁVILA - IED. Bogota: Fundación Universitaria los Libertadores.
- Manso, J. M., & Méndez, M. J. (2003). Valoración de las Dificultades de Lectoescritura en Niños, Adolescentes y Adultos. *Puertas a la Lectura*. Merida, España: Universida de Extremadura.
- Bojacá, C. P. (Julio de 2017). Aprender a Pensar y Reflexionar: propuesta de estrategias cognitivas y metacognitivas en la comprensión lectora. *Maestría en Educación con énfasis en el Aprendizaje de la Lectoescritura y las Matemáticas*. Bogotá, Colombia: Universida de externado.
- Salvador Chacón Moscoso, J. A.-G. (2008). La Teoría Clásica de los tests. En J. A.-G. Salvador Chacón Moscoso, *Fundamentos de la medición psicológica*. Sevilla.
- Cervera, J. T. (2015). Sinopsis. En J. T. Cervera, *T.A.L.E. Test de Análisis de lectoescritura*.
- Meneses, J. (2017). La evaluación psicometrica. En J. Meneses, *La evaluación psicometrica*. Cooperativa Editorial Magisterio.



- Muñoz Carpio Karen Estefanía, S. L. (Junio de 2017). UNIVERSIDAD CENTRAL DEL ECUADOR. *La Coordinación Visomotriz y su incidencia en la Escritura en los niños/as de tercero y cuarto año de básica de la Unidad Educativa San José la Salle, ubicada al sur de la Ciudad de Quito, en la Parroquia la Magdalena, en el período 2016-2017*. Quito, Ecuador: UNIVERSIDAD CENTRAL DEL ECUADOR.
- Josep Toro, M. C. (2020). *TALE: TEST DE ANALISIS DE LECTOESCRITURA (2ª ED.)*. Obtenido de casa del libro:
<https://www.casadellibro.com/libro-tale-test-de-analisis-de-lectoescritura-2-ed/9788477740261/337503>
- Herazo, Z. D. (2020). *La lectura y sus tipos* . Obtenido de Corporación Universitaria Rafael Nuñez:
<https://www.curn.edu.co/clye/108-portal-palabras/lectura/896-lectura-y-sus-tipos.html#:~:text=Para%20esto%20se%20abordan%20tres,predominante%20en%20el%20%C3%A1mbito%20acad%C3%A9mico.&text=La%20Lectura%20inferencial%20se%20constituye,abstracci%C3%B3n%20por>
- Ministerio de Educación. (2018). *Mallas de aprendizaje, lenguaje, grado 3º. Mallas de aprendizaje, lenguaje, grado 3º*. Bogota, Colombia.
- Jorge Veiga de Cabo, E. d. (2008). *Modelos de estudios en investigación aplicada: conceptos y criterios para el diseño. Aula Abierta*.
- Frías-Navarro, D. (2020). *APUNTES DE CONSISTENCIA INTERNA DE LAS PUNTUACIONES DE UN INSTRUMENTO DE MEDIDA*. España: Universidad de Valencia.