

UNIVERSIDAD DEL VALLE DE GUATEMALA

Facultad de Educación



**Determinar la confiabilidad y validez de la Prueba de Lectura Inicial
aplicada en escuelas de Quiché**

Trabajo de graduación presentado por Carmen María Escobar Ramírez,
para optar al grado académico de Magister en Medición, Evaluación
e Investigación Educativas

Guatemala

2019

UNIVERSIDAD DEL VALLE DE GUATEMALA

Facultad de Educación



**Determinar la confiabilidad y validez de la Prueba de Lectura Inicial
aplicada en escuelas de Quiché**

Trabajo de graduación presentado por Carmen María Escobar Ramírez,
para optar al grado académico de Magister en Medición, Evaluación
e Investigación Educativas

Guatemala

2019

Vo. Bo.

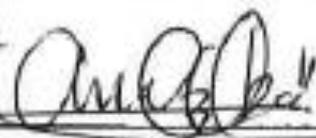


M.A. Viviane Yvette Bolaños Gramajo

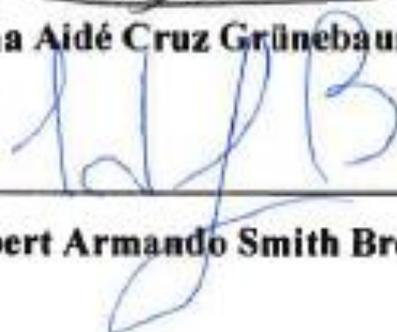
Tribunal examinador:



M.A. Viviane Yvette Bolaños Gramajo



M.A. Ana Aidé Cruz Grünebaum



M.A. Herbert Armando Smith Brolo

Fecha de aprobación del examen de graduación:

Guatemala, 7 de noviembre 2019

PRÓLOGO

Agradezco a cada una de las personas que contribuyeron con este logro, en especial a...

Mi familia porque en ella se manifiesta Dios, siempre han creído en mi potencial y cada día me animan a tomar decisiones que contribuyen a mi realización y felicidad; entre ellas continuar mis estudios en la Maestría en Medición, Evaluación e Investigación Educativa en la Universidad del Valle de Guatemala.

Todos mis catedráticos expertos en las áreas de Medición, Evaluación o Investigación, quienes me permitieron descubrir nuevos campos, en los que he tenido oportunidades laborales que me han permitido confirmar el impacto del ámbito educativo, el cual me apasiona y en el que deseo especializarme.

Mariela Zelada, decana de la Facultad de Educación, quien escuchó mi interés y me refirió con *Magda Cifuentes* quien fue el enlace para que School the World aceptara que mi trabajo de tesis, fuera a partir del análisis de la base de datos que ellos recopilaron en el 2018.

Silvia Culajay, quien me recordó que podría tener una excelente asesora, graduada de la misma maestría, con amplia experiencia y con calidad humana.

Yvette Bolaños Gramajo quien fue una asesora extraordinaria, desde el inicio siempre estuvo dispuesta a orientarme con asertividad; compartió de su tiempo para reunirse, incluso los fines de semana completos; con paciencia me escuchó y resolvió cada una de mis dudas; reflexionamos a partir de nuestras experiencias y sueños; compartió lecturas y links interesantes; fue paso a paso para que lograra los análisis en los software estadísticos y con su estilo de vida es un ejemplo que inspira a realizar todo con amor y excelencia.

Equipo de School the World quienes desde diferentes lugares se unieron a la primera reunión por Skype. Ana Joya (Honduras), Kathy López (Quiché) y Kener Us (Quiché) técnico que demostró su disponibilidad para compartir la información respecto al proceso que lideró para la adaptación y aplicación de la prueba de Lectura Inicial y con alegría siempre dio seguimiento a las consultas que en el proceso surgieron.

ÍNDICE GENERAL

PRÓLOGO.....	v
ÍNDICE GENERAL.....	vi
ÍNDICE DE GRÁFICAS.....	viii
ÍNDICE DE TABLAS.....	x
RESUMEN.....	xii
I. INTRODUCCIÓN.....	1
A. Formulación del problema.....	2
B. Justificación.....	3
C. Hipótesis de investigación.....	7
II. MARCO TEÓRICO.....	8
A. Marco conceptual.....	8
B. Marco referencial.....	12
C. Marco contextual.....	35
D. Marco normativo-jurídico.....	48
III. MARCO METODOLÓGICO.....	51
A. Alcances del estudio.....	51
B. Objetivos de la investigación.....	52
C. Participantes.....	52
D. Escenario.....	53
E. Instrumentos de recolección de información.....	58
F. Procedimiento.....	63
G. Diseño del método.....	69
H. Análisis de los datos o información.....	69
I. Consideraciones éticas.....	70
IV. RESULTADOS DE LA INVESTIGACIÓN.....	72
A. Descriptivos por cada Subprueba.....	72
B. Correlación de Pearson en cada Subprueba.....	77
C. Análisis de los ítems.....	84
D. Resultados con Rasch.....	95
E. Indicadores de ajuste del modelo.....	96
F. Análisis diferencial de los ítems.....	111
G. Confiabilidad.....	111
H. Análisis Factorial Exploratorio.....	118
I. Calibración de los ítems con TRI.....	126

J. Calificación de los ítems.....	133
V. DISCUSIÓN DE LOS RESULTADOS.....	140
A. Correlación en cada Subprueba.....	140
B. Dificultad (TCT/TRI).....	141
C. Discriminación y habilidad.....	143
D. RMSE.....	145
E. Infit (WMS) y Outfit (UMS).....	145
F. Análisis diferencial.....	146
G. Confiabilidad.....	146
H. Análisis factorial exploratorio.....	148
I. Calibración (Dificultad-Discriminación-Adivinación).....	149
J. Total de puntuaciones.....	149
K. Validez.....	150
VI. CONCLUSIONES.....	151
VII. RECOMENDACIONES.....	153
VIII. REFERENCIAS.....	155
IX. ANEXOS.....	162

ÍNDICE DE GRÁFICAS

Gráfica 1. Histograma del total de puntuaciones obtenidas por los estudiantes	73
Gráfica 2. Correlación de los ítems-total de la Subprueba de alfabeto en la población y la muestra.....	78
Gráfica 3. Correlación ítems-total de la Subprueba de alfabeto con el total de la prueba en la población y la muestra	78
Gráfica 4. Correlación de los ítems-total de la Subprueba de palabras simples en la población y la muestra	79
Gráfica 5. Correlación de los ítems-total de la Subprueba de palabras simples con el total de la prueba en la población y la muestra	79
Gráfica 6. Correlación de los ítems-total de la Subprueba de palabras inventadas en la población y la muestra	80
Gráfica 7. Correlación de los ítems-total de la Subprueba de palabras inventadas con el total de la prueba en la población y la muestra	80
Gráfica 8. Correlación de los ítems-total de la Subprueba de lectura oral en la población y la muestra (parte 1)	81
Gráfica 9. Correlación de los ítems-total de la Subprueba de lectura oral en la población y la muestra (parte 2)	81
Gráfica 10. Correlación de los ítems-total de la Subprueba de lectura oral con el total de la prueba en la población y la muestra (parte 1)	82
Gráfica 11. Correlación de los ítems-total de la Subprueba de lectura oral con el total de la prueba en la población y la muestra (parte 2)	82
Gráfica 12. Correlación de los ítems-total de la Subprueba de comprensión lectora en la población y la muestra	83
Gráfica 13. Correlación de los ítems-total de la Subprueba de comprensión lectora con el Total de la prueba en la población y la muestra.....	83
Gráfica 14. Dificultad de los ítems en la Subprueba de alfabeto	84
Gráfica 15. Discriminación de los ítems en la Subprueba de alfabeto.....	84
Gráfica 16. Desviación Estándar de los ítems en la Subprueba de alfabeto	85
Gráfica 17. Dificultad de los ítems en la Subprueba de palabras simples	86
Gráfica 18. Discriminación de los ítems en la Subprueba de palabras simples	86
Gráfica 19. Desviación Estándar de los ítems en la Subprueba de palabras simples.....	87
Gráfica 20. Dificultad de los ítems en la Subprueba de palabras inventadas.....	88
Gráfica 21. Discriminación de los ítems en la Subprueba de palabras inventadas	88
Gráfica 22. Desviación Estándar de los ítems en la Subprueba de palabras inventadas.....	89
Gráfica 23. Dificultad de los ítems en la Subprueba de lectura oral.....	90
Gráfica 24. Discriminación de los ítems en la Subprueba de lectura oral	91
Gráfica 25. Desviación Estándar de los ítems en la Subprueba de lectura oral	91
Gráfica 26. Dificultad de los ítems en la Subprueba de comprensión lectora	92
Gráfica 27. Discriminación de los ítems en la Subprueba de comprensión lectora	93
Gráfica 28. Desviación Estándar de los ítems en la Subprueba de comprensión lectora.....	93
Gráfica 29. Error Estándar de la Subprueba de alfabeto.....	96
Gráfica 30. Habilidad para responder la Subprueba de alfabeto.....	96
Gráfica 31. Error Estándar de la Subprueba de palabras simples	97
Gráfica 32. Habilidad para responder la Subprueba de palabras simples	98
Gráfica 33. Error Estándar de la Subprueba de palabras inventadas	99
Gráfica 34. Habilidad para responder la Subprueba de palabras inventadas	99
Gráfica 35. Error Estándar de la Subprueba de lectura oral.....	100

Gráfica 36. Habilidad para responder la Subprueba de lectura oral.....	101
Gráfica 37. Error Estándar de la Subprueba de comprensión lectora	102
Gráfica 38. Habilidad para responder la Subprueba de comprensión lectora	102
Gráfica 39. Dificultad en la Subprueba de alfabeto	103
Gráfica 40. Desviación Estándar en la Subprueba alfabeto	104
Gráfica 41. Dificultad en la Subprueba de palabras simples.....	105
Gráfica 42. Desviación Estándar en la Subprueba palabras simples.....	105
Gráfica 43. Dificultad en Subprueba de palabras inventadas.....	106
Gráfica 44. Desviación Estándar en la Subprueba de palabras inventadas	107
Gráfica 45. Dificultad en la Subprueba de lectura oral	108
Gráfica 46. Desviación Estándar en la Subprueba lectura oral	108
Gráfica 47. Dificultad en Subprueba comprensión lectora	109
Gráfica 48. Desviación Estándar en la Subprueba comprensión lectora.....	110
Gráfica 49. Confiabilidad si el ítem es eliminado en la Subprueba de alfabeto, palabras simples e inventadas	112
Gráfica 50. Confiabilidad si el ítem es eliminado en la Subprueba de lectura oral y comprensión lectora	113
Gráfica 51. Confiabilidad de la prueba de dos mitades sin el total.....	116
Gráfica 52. Confiabilidad de la prueba de dos mitades con el total.....	118
Gráfica 53. Dificultad de la Subprueba alfabeto en la calibración de los ítems.....	126
Gráfica 54. Discriminación de la Subprueba alfabeto en la calibración de los ítems	126
Gráfica 55. Adivinación de la Subprueba alfabeto en la calibración de los ítems	127
Gráfica 56. Dificultad de la Subprueba palabras simples en la calibración de los ítems	127
Gráfica 57. Discriminación de la Subprueba palabras simples en la calibración de los ítems	128
Gráfica 58. Adivinación de la Subprueba palabras simples en la calibración de los ítems ...	128
Gráfica 59. Dificultad de la Subprueba palabras inventadas en la calibración de los ítems ..	129
Gráfica 60. Discriminación de la Subprueba de palabras inventadas en la calibración de los ítems	129
Gráfica 61. Adivinación de la Subprueba palabras inventadas en la calibración de los ítems	130
Gráfica 62. Dificultad de la Subprueba lectura oral en la calibración de los ítems	130
Gráfica 63. Discriminación de la Subprueba lectura oral en la calibración de los ítems	131
Gráfica 64. Adivinación de la Subprueba lectura oral en la calibración de los ítems.....	131
Gráfica 65. Dificultad de la Subprueba comprensión lectora en la calibración de los ítems .	132
Gráfica 66. Discriminación de la Subprueba comprensión lectora en la calibración de los ítems	132
Gráfica 67. Adivinación de la Subprueba comprensión lectora en la calibración de los ítems	133

ÍNDICE DE TABLAS

Tabla 1. Resultados de los estudiantes en lectura	1
Tabla 2. Diferencias entre TCT y TRI	22
Tabla 3. Rango del indicador de dificultad	25
Tabla 4. Probabilidades de Rasch de respuesta correcta	26
Tabla 5. Rango del nivel de discriminación	27
Tabla 6. Valores de anclaje	30
Tabla 7. Porcentaje de los indicadores de cobertura en Guatemala y el depto. de Quiché	37
Tabla 8. Porcentaje de los indicadores de cobertura en tres de los municipios de Quiché	38
Tabla 9. Tasa de Pobreza Rural por nivel y Coeficiente de Gini en Quiché y municipios de intervención.....	38
Tabla 10. Hitos de School the World.....	40
Tabla 11. Caracterización de la comunidad educativa por cada escuela	45
Tabla 12. Subpruebas de la prueba STAR y de lectura inicial de STW	60
Tabla 13. Estadísticos de respuesta correcta de la Prueba de lectura inicial.....	73
Tabla 14. Descriptivos para la Subprueba de alfabeto	74
Tabla 15. Descriptivos para la Subprueba de palabras simples	75
Tabla 16. Descriptivos para la Subprueba de palabras inventadas	75
Tabla 17. Descriptivos para la Subprueba de lectura oral.....	76
Tabla 18. Descriptivos para la Subprueba comprensión lectora	77
Tabla 19. Promedio de la Subprueba alfabeto	85
Tabla 20. Los ítems más difíciles y que mejor discriminan.....	85
Tabla 21. Promedio de la Subprueba palabras simples.....	87
Tabla 22. Ítems más difíciles y que mejor discriminan	87
Tabla 23. Promedio de la Subprueba palabras inventadas	89
Tabla 24. Ítems más difíciles y que mejor discriminan	89
Tabla 25. Promedio de la Subprueba lectura oral	92
Tabla 26. Ítems más difíciles y que mejor discriminan	92
Tabla 27. Promedio de la Subprueba comprensión lectora.....	93
Tabla 28. Ítems más difíciles y que mejor discriminan	94
Tabla 29. Estadística de calidad de la escala.....	95
Tabla 30. Descriptivos del WMS y UMS estándar de la Subprueba de alfabeto.....	97
Tabla 31. Descriptivos del WMS y UMS estándar de la Subprueba palabras simples.....	98
Tabla 32. Descriptivos del WMS y UMS estándar de la Subprueba palabras inventadas	100
Tabla 33. Descriptivos del WMS y UMS estándar de la Subprueba lectura oral	101
Tabla 34. Descriptivos del WMS y UMS estándar de la Subprueba comprensión lectora....	103
Tabla 35. Estadísticos descriptivos del modelo de habilidad de la Subprueba alfabeto	104
Tabla 36. Estadísticos descriptivos del modelo de habilidad de la Subprueba palabras simples	106
Tabla 37. Estadísticos descriptivos del modelo de habilidad de la Subprueba palabras inventadas	107
Tabla 38. Estadísticos descriptivos del modelo de habilidad de la Subprueba lectura oral ...	109
Tabla 39. Estadísticos descriptivos del modelo de habilidad de la Subprueba comprensión lectora.....	110
Tabla 40. Confiabilidad de la prueba en la población basado en los elementos tipificados ..	111
Tabla 41. Indicadores de confiabilidad basados en diferentes métodos	111
Tabla 42. Confiabilidad de la prueba según subpruebas.....	113
Tabla 43. Confiabilidad de la prueba incluyendo los ítems y el total global	114

Tabla 44. Confiabilidad de la prueba de dos mitades sin el total.....	114
Tabla 45. Confiabilidad de la prueba de dos mitades con el total.....	117
Tabla 46. KMO y prueba de Bartlett de la prueba con 38 ítems.....	118
Tabla 47. Varianza total explicada de la prueba con 38 ítems.....	119
Tabla 48. KMO y prueba de Bartlett de la prueba con 16 ítems.....	120
Tabla 49. Varianza total explicada de la prueba con 16 ítems.....	120
Tabla 50. KMO y prueba de Bartlett de la Subprueba de Alfabeto.....	121
Tabla 51. Varianza total explicada de la Subprueba de alfabeto.....	121
Tabla 52. KMO y prueba de Bartlett de la Subprueba palabras simples.....	122
Tabla 53. Varianza total explicada de la Subprueba palabras simples en la muestra.....	122
Tabla 54. Varianza total explicada de la Subprueba palabras simples en la población.....	122
Tabla 55. KMO y prueba de Bartlett de la Subprueba de palabras inventadas.....	123
Tabla 56. Varianza total explicada de la Subprueba de palabras inventadas.....	123
Tabla 57. KMO de la Subprueba comprensión lectora.....	124
Tabla 58. Varianza total explicada de la Subprueba comprensión lectora en la muestra.....	124
Tabla 59. Varianza total explicada de la Subprueba comprensión lectora en la población.....	125
Tabla 60. Tabla de puntuaciones en la Subprueba alfabeto.....	134
Tabla 61. Tabla de puntuaciones en la Subprueba palabras simples.....	134
Tabla 62. Tabla de puntuaciones en la Subprueba palabras inventadas.....	135
Tabla 63. Tabla de puntuaciones en la Subprueba lectura oral.....	136
Tabla 64. Tabla de puntuaciones en la Subprueba comprensión lectora.....	139
Tabla 65. Análisis diferencial entre Escuelas con la Subprueba alfabeto.....	167
Tabla 66. Análisis diferencial entre edades con la Subprueba alfabeto.....	168
Tabla 67. Análisis diferencial entre idiomas con la Subprueba alfabeto.....	168
Tabla 68. Análisis diferencial entre sexo con la Subprueba alfabeto.....	168
Tabla 69. Análisis diferencial entre Escuelas con la Subprueba palabras simples.....	168
Tabla 70. Análisis diferencial entre edades con la Subprueba palabras simples.....	169
Tabla 71. Análisis diferencial entre idiomas con la Subprueba palabras simples.....	169
Tabla 72. Análisis diferencial entre sexo con la Subprueba palabras simples.....	169
Tabla 73. Análisis diferencial entre Escuelas con la Subprueba palabras inventadas.....	169
Tabla 74. Análisis diferencial entre edades con la Subprueba palabras inventadas.....	170
Tabla 75. Análisis diferencial entre idiomas con la Subprueba palabras inventadas.....	170
Tabla 76. Análisis diferencial entre sexo con la Subprueba palabras inventadas.....	170
Tabla 77. Análisis diferencial entre Escuelas con la Subprueba lectura oral.....	170
Tabla 78. Análisis diferencial entre edades con la Subprueba lectura oral.....	171
Tabla 79. Análisis diferencial entre idiomas con la Subprueba lectura oral.....	171
Tabla 80. Análisis diferencial entre sexo con la Subprueba lectura oral.....	171
Tabla 81. Análisis diferencial entre Escuelas con la Subprueba comprensión lectora.....	171
Tabla 82. Análisis diferencial entre edades con la Subprueba comprensión lectora.....	172
Tabla 83. Análisis diferencial entre idiomas con la Subprueba comprensión lectora.....	172
Tabla 84. Análisis diferencial entre sexo con la Subprueba comprensión lectora.....	172

RESUMEN

Este informe de tesis contiene los resultados de la investigación con un enfoque cuantitativo, la cual se realizó para determinar las propiedades psicométricas, confiabilidad y validez, de la Prueba de Lectura Inicial. El instrumento fue adaptado por el equipo técnico de *School the World (STW)* a partir de la prueba de lectura *School based Test About Reading (STAR)* la cual fue diseñada y socializada por Visión Mundial. La prueba contiene 206 ítems en total, distribuidos en cinco Subpruebas las cuales son: alfabeto (27 ítems), Palabras Simples (20 ítems), Palabras Inventadas (20 ítems), Lectura Oral (129 ítems) y Comprensión Lectora (10 ítems).

La aplicación se realizó en la plataforma Tangerine, el miércoles 10 de octubre 2018 y estuvo a cargo del equipo técnico de School the World con el apoyo de estudiantes de la Universidad del Valle de Guatemala del Campus Altiplano. Se aplicó en nueve Escuelas Oficiales Rurales Mixtas (EORM) que se ubican en el departamento de Quiché y se caracterizan por ser comunidades que viven en extrema pobreza y a su vez actualmente participan en los diferentes programas de intervención educativa que tiene School the World. La muestra estuvo formada por 60 estudiantes de 3ro. primaria que participaron de forma voluntaria.

A partir de la base de datos que el equipo compartió, se realizaron los análisis estadísticos en Excel, SPSS y jMetrik. El promedio de respuestas correctas brindadas por los estudiantes evaluados fue de 113 (54.85%) y la mediana de 112 (54.36%) sobre un total de 206 ítems. Se correlacionaron los ítems de cada Subprueba en la población y la muestra, con y sin el total de la prueba. Por medio del coeficiente de Pearson se identificó la correlación de los ítems, la cual fue mayor (0.70) en la Subprueba de lectura oral.

Se calculó la confiabilidad por medio del estadístico Alfa de Cronbach y se obtuvo que en total la Prueba de lectura inicial tiene 0.982, lo cual refleja una confiabilidad alta y estable. Se establecieron los indicadores de ajuste, el RMSE y el KMO para cada Subprueba y el análisis de ítems. De acuerdo con la TCT la Subprueba de lectura oral fue la más difícil y para la TRI fue la Subprueba de comprensión lectora la más difícil de responder; ambas Subpruebas evalúan los macroprocesos asociados a las operaciones de más alto nivel asociadas a la comprensión de lo que se lee o al

procesamiento semántico. La Subprueba de comprensión lectora fue la que menor discriminó entre las habilidades de los estudiantes para responder correctamente.

A partir de los análisis realizados se concluyó que la Prueba de lectura inicial, aplicada por *School the World* es confiable porque al aplicarla en diferentes escuelas se pudo comprobar la consistencia y estabilidad de las puntuaciones obtenidas y es válida en la medición del constructo de la habilidad lectora en los estudiantes de 3ro. primaria. Se recomienda para futuras aplicaciones la selección de una muestra aleatoria más amplia que permitan una mayor variabilidad de los datos analizados.

I. INTRODUCCIÓN

Al reconocer que cualquier conocimiento que se adquiriera a partir de la lectura servirá como base para añadir nuevos conocimientos, implica que se considere la relación directa que existe entre la competencia lectora con el éxito escolar o laboral y el grado de autonomía personal. La escuela es quien tiene el compromiso de desarrollar las competencias para la vida, entre ellas la competencia Lectora, la cual se debe enseñar intencionadamente, debido a que la lectura es un proceso complejo. Para la sociedad, el desarrollo de esta competencia implica mejorar y diversificar los programas educativos y las prácticas pedagógicas, para atender la adquisición paulatina de competencias comunicativas de distinto nivel, en las cuales se consideren especialmente a los grupos excluidos y vulnerables (OREALC, 2005).

Al analizar los resultados de lectura de los estudiantes de último año de ciclo de diversificado (graduandos) que realizó el Ministerio de Educación durante el 2018, los resultados de los estudiantes de 3ro básico en la Evaluación del Programa Internacional para la Evaluación de los Estudiantes (PISA) y los resultados en la Evaluación Nacional de Primaria 2014 para 3.º y 6.º primaria, se observa que a nivel Nacional, en los diferentes niveles evaluados, la mitad de la población de estudiantes guatemaltecos alcanzan un nivel de logro.

Tabla 1. *Resultados de los estudiantes en lectura*

	Graduados 2018 (Mineduc)	3ro. básico 2013 (PISA)	6to. Primaria 2014 (Mineduc)	3ro. primaria 2014 (Mineduc)
Guatemala	34.81%	3% en el nivel 2 o más	40.40%	49.93 %
Depto. de Quiché	25.02%	5.24%	20.16%	28.79%

(Mineduc, 2018)

Para acompañar el proceso escolar que tienen las escuelas en Guatemala, desde el 2009 *School the Word (STW)* trabaja con las comunidades que viven en condiciones de extrema pobreza. Por medio de la participación comunitaria STW ha generado diversos programas educativos con los cuales se busca un impacto en las comunidades para mejorar la calidad de vida. Entre las actividades que han realizado se encuentra el fortalecimiento

de las habilidades Lectoras de las niñas y los niños. Para evaluar el impacto de las acciones implementadas, en 2018 adaptaron la Prueba de Lectura *School based Test About Reading* de aquí en adelante *STAR*, la cual fue socializada por el equipo de Visión Mundial. Se aplicó la Prueba de lectura inicial en la plataforma de Tangerine en 9 escuelas ubicadas en el departamento de Quiché que participan en los programas de STW. Como se describe más adelante, algunas de estas escuelas trabajan con aula multigrado que en su mayoría poseen estudiantes que hablan K'iché como idioma materno. Para esta investigación, la muestra estuvo conformada por 60 estudiantes que estudian en 3ro. primaria.

La oportunidad de contar con una medición de la habilidad de lectura de los estudiantes que son beneficiados con acciones de STW, motivó el interés de adaptar la escala así como la consecuente responsabilidad de analizar las propiedades psicométricas de la prueba adaptada para este contexto guatemalteco. El análisis de la confiabilidad y validez de la Prueba de lectura inicial aplicada por STW complementa y enriquece el análisis que desde la Facultad de Educación se inició, al redactar el informe de los resultados obtenidos por cada una de las escuelas en las cinco Subpruebas.

Por medio de la determinación de las propiedades psicométricas, confiabilidad y validez, de la Prueba de lectura inicial y el análisis de los ítems, esta investigación se propone aportar a los esfuerzos de validación de la prueba en el contexto guatemalteco, particularmente en escuelas con recursos limitados pero que actualmente son beneficiadas por las iniciativas de desarrollo de programas de lectura. Asimismo, es un aporte de información útil para *STW*, quien a través de los programas de monitoreo puede apoyar al trabajo que los docentes realizan en el aula, al diseñar un plan de mejora que permita afianzar las habilidades lectoras de las niñas y los niños, quienes podrán tener mayor éxito en su formación académica y a lo largo de su vida.

A. Formulación del problema

El proceso de medir de forma sistemática la adquisición de las habilidades Lectoras, en los estudiantes del primer ciclo de educación primaria, permite unificar y estimular los esfuerzos más efectivos para mejorar el desempeño de las niñas y los niños durante los siguientes años de escolaridad. Actualmente las investigaciones apoyan la intervención

temprana como el mejor predictor para el futuro académico exitoso, particularmente en las áreas críticas de lectura y lenguaje (Learning Renaissance, 2015).

La alfabetización es un proceso permanente, que puede expresarse y evaluarse en un continuo de grados, y no una capacidad dicotómica, que los estudiantes tienen o no tienen. Por eso la alfabetización no se define como el alcance de un umbral básico de herramientas de lectura y escritura, sino como la posibilidad de desenvolverse socialmente, lograr metas personales, y desarrollar el conocimiento propio y de los otros a lo largo de toda la vida, a partir de información escrita (OREALC, 2005).

Al considerar la importancia de medir las habilidades Lectoras en los grados iniciales de la primaria, también surge la inquietud acerca de ¿cómo medir las habilidades Lectoras? ya que originalmente quien asume el rol de determinar si las niñas y los niños comprenden lo que leen, suele ser el docente, por medio de diferentes instrumentos, en su mayoría informales. Por ello surgió la necesidad de contar con un instrumento de medición válido y confiable para asegurar que la información es sólida y rigurosa, ya que según Gregory (2012) es absurdo y poco ético basar las decisiones importantes en resultados de pruebas que no pueden replicarse.

1. Planteamiento del problema

Determinar la confiabilidad y validez de la Prueba de Lectura Inicial aplicada en escuelas de Quiché.

2. Pregunta de investigación

Desde un enfoque cuantitativo inferencial se responde a la pregunta de investigación: ¿Cuáles son las propiedades psicométricas, confiabilidad y validez de la Prueba de Lectura Inicial aplicada en 3ro primaria, en 9 escuelas del departamento de Quiché?

B. Justificación

La consecución de una educación de calidad es la base para mejorar la vida de las personas y el desarrollo sostenible. Para lograrlo Guatemala ha priorizado el cuarto

Objetivo de Desarrollo Sostenible (ODS), a través de la meta 4.1 que propone para el 2030, velar que todas las niñas y todos los niños tengan una enseñanza primaria y secundaria completa, gratuita, equitativa y de calidad que produzca resultados de aprendizaje pertinentes y efectivos; lo cual reducirá el analfabetismo en la población (SEGEPLAN, 2018).

A su vez, por medio del Plan Nacional de Desarrollo y los Objetivos de Desarrollo Sostenible se especifican las prioridades nacionales, sobre las cuales se deben armonizar las acciones impulsadas por las instituciones del Estado, sean estas públicas o privadas, además de la cooperación internacional, de manera que se pueda organizar, coordinar y articular en función de los intereses y prioridades del desarrollo nacional a largo plazo. Las prioridades nacionales de desarrollo son producto de un proceso de integración, basado en análisis técnico, de las metas de los ODS priorizadas por Guatemala y las metas del Plan K'atún 2032. Ambas reflejan las áreas temáticas del desarrollo que deberán ser la guía para que el país pueda mejorar las condiciones de vida de la población. Desde la Política #2 de Educación Nacional, se indica el mejoramiento de la calidad del proceso educativo para asegurar que todas las personas sean sujetos de una educación pertinente y relevante. Y específicamente en el tercer objetivo estratégico se reconoce que debe fortalecerse el sistema de evaluación para garantizar la calidad educativa (Consejo Nacional de Educación, 2010).

Actualmente existen planes, políticas y objetivos trazados para mejorar la calidad educativa y a su vez se han realizado evaluaciones a nivel nacional y regional, en las cuales los resultados obtenidos reflejan que la mayoría de los estudiantes no poseen un nivel de desempeño satisfactorio en lectura. De acuerdo con los resultados del informe del *Programa Internacional para la Evaluación de Estudiantes (PISA)*, cerca del 20% de los estudiantes de los países de la *Organización para la Cooperación y el Desarrollo Económicos –OCDE-* no obtiene, de media, las competencias Lectoras básicas. Esta proporción se ha mantenido estable desde el 2009. La diferencia en lectura a favor de las niñas disminuyó 12 puntos en la prueba PISA entre 2009 y 2015; el rendimiento de los niños mejoró, particularmente entre aquellos con mejores resultados, mientras que el de las niñas empeoró, sobre todo entre aquellas con peores resultados (OCDE, 2015).

A nivel regional, en el 2013 Guatemala participó en el *Tercer Estudio Regional Comparativo y Explicativo* (TERCE), en el que se evaluaron los logros de aprendizaje en las áreas de lenguaje (lectura y escritura), entre otras, en estudiantes de 3ro. y 6to. primaria. Los resultados de aprendizaje de Guatemala se ubican por debajo del desempeño promedio a excepción de la prueba de escritura de sexto grado, en la que el país destaca por sobre la media regional lo cual evidencia el rezago que persiste en el desarrollo de la comprensión Lectora en los estudiantes que se encuentran finalizando el I y II ciclo de educación primaria (UNESCO, 2015).

Así mismo por medio de la Evaluación Nacional de Primaria, la cual es una evaluación muestral que se realiza en centros educativos públicos del país, con una muestra representativa a nivel municipal para 3ro. y 6to. primaria, los resultados de la última evaluación realizada en el 2014 muestran que 49.93% de los estudiantes de 3ro. primaria obtuvieron un resultado de logro en lectura el cual aumentó a diferencia del 46.63% que fue obtenido en el 2013 (Quim, M.; Cruz, A. y Santos, J., 2014). A nivel departamental, 28.79% de los estudiantes de 3ro. primaria en Quiché obtuvieron este resultado, el cual aumentó en comparación a la última evaluación que fue de 21.93% (Digeduca, 2018). Sin embargo, los resultados reflejan que aún hay que potencializar las destrezas que permitan a los estudiantes tener un nivel de desempeño alto, principalmente en el área rural en donde la brecha es mayor en comparación con el área urbana.

Sin un alfabetismo elemental hay pocas posibilidades de que las niñas y los niños puedan escapar del ciclo intergeneracional de la pobreza ya que la capacidad de leer y entender un texto simple es una de las habilidades más fundamentales que un estudiante puede aprender (USAID, 2009).

En Guatemala, los niveles de desempeño alcanzados por los estudiantes en lectura, están por debajo del promedio de la OCDE y PISA-D, lo que representa un desafío para fortalecer las acciones dentro de los establecimientos educativos, para que estos sean espacios que brinden oportunidades de aprendizaje que permitan, en alguna medida, reducir las brechas de desigualdad entre hombres y mujeres, áreas, niveles socioeconómicos, entre otros; logrando que, independientemente de su origen o características, los jóvenes tengan derecho a una educación en la que no solo ingresen a la escuela, si no también desarrollen las habilidades mínimas esperadas (Ministerio de Educación, 2018).

Demasiados niños están abandonando la escuela primaria con habilidades de lectura inadecuadas para el siguiente nivel de instrucción. (Torgesen, 2004) Para mejorar el rendimiento de lectura, las escuelas deben examinar los programas básicos de lectura y si estos están alineados con la investigación de lectura con base científica (Vaughn Gross Center for Reading & Language Arts., 2005).

Durante los últimos 10 años en Guatemala ha trabajado la organización *School the World* (STW), la cual apuesta a la educación de comunidades que viven en extrema pobreza. Entre sus intervenciones se ha propuesto medir las habilidades Lectoras de 60 estudiantes que estudian en 3ro. primaria en 9 escuelas, las cuales participan en los programas educativos. Para medir dichas habilidades, decidieron adaptar la prueba de lectura STAR por sus siglas en inglés (*School based Test About Reading/Prueba escolar sobre lectura*) la cual fue diseñada y compartida por *Visión Mundial (World Vision/WV)*.

En este informe se presenta el análisis que se realizó para determinar las propiedades psicométricas, confiabilidad y validez de la Prueba de lectura inicial, aplicada por STW. Considerando que una prueba puede ser confiable, pero no necesariamente válida. Por ello es necesario demostrar que la Prueba de lectura inicial es confiable y válida. Para identificar ambas propiedades Bostwick y Kyte (2005) proponen la analogía en la que hay tres dianas con las marcas que han hecho tres lanzadores. La primera diana tiene las marcas dispersas, las cuales representan que los tiros no son válidos ni confiables; la segunda diana tiene en un punto cercano todas las marcas, sin embargo, estas se encuentran lejos del blanco, lo que significa que los tiros son confiables, pero no válidos; y la tercera diana tiene todas las marcas cercanas entre sí y muy cerca del blanco, lo cual significa que los lanzamientos son confiables y válidos.

La información de la prueba fue socializada por equipo de STW, quien desde el inicio cedió la base de datos con los resultados de la aplicación que fue realizada en el año 2018. Posteriormente se sostuvieron frecuentes espacios de comunicación para brindar información respecto a la adaptación y aplicación de la prueba. Como producto de la investigación se entregó a STW insumos con el potencial de sustentar la toma de decisiones que favorezcan a la mejora de los procesos de evaluación de cada uno de los programas en los cuales se elaboren planes para reflexionar y acompañar las prácticas de los docentes.

La elaboración de este análisis a su vez, representó una oportunidad para vincular la teoría que se exploró durante la formación recibida en la Maestría de Evaluación, Medición y Evaluación Educativa, con la cual se propuso aportar al trabajo de STW. También se constituyó en un espacio para vincular la experiencia en la aplicación de pruebas y la pasión por el área de lectoescritura. Esta presentación de la tesis representó un compromiso para culminar con un ciclo de formación y crecimiento personal.

C. Hipótesis de investigación

$H_0 = p \geq 0.05$	$H_1 = p \leq 0.05$
---------------------	---------------------

La hipótesis nula establece que **p** es mayor o igual a **0.05**

La hipótesis alterna establece que **p** es menor o igual a **0.05**

Supuestos teóricos:

- a. La Prueba de lectura inicial no discrimina la habilidad de comprensión Lectora de los estudiantes de 3ro. primaria en las 9 escuelas de Quiché.
- b. La Prueba de lectura inicial discrimina la habilidad de lectura de los estudiantes de 3ro. primaria en las 9 escuelas de Quiché.

II. MARCO TEÓRICO

A. Marco conceptual

1. Variables

- a. Habilidad en lectura
- b. Alfabeto
- c. Lectura de palabras simples
- d. Descodificación de palabras inventadas (pseudopalabras)
- e. Lectura oral
- f. Comprensión lectora
- g. Instrumento de medición
- h. Propiedades psicométricas
- i. Dificultad
- j. Discriminación
- k. Adivinación
- l. Confiabilidad
- m. Validez

2. Definición conceptual

- a. Habilidad de lectura. Consiste en la obtención de información (capacidad para localizar información en un texto), la interpretación de textos (capacidad para construir significados y hacer inferencias a partir de la información escrita) y la reflexión y evaluación (capacidad para relacionar el texto con los conocimientos, ideas y experiencias) (OCDE/ Instituto Nacional de Calidad y Evaluación, 2000).
- b. Alfabeto. Habilidad que está relacionada con la capacidad de recordar las formas de las letras escritas y sus nombres, así como los sonidos (USAID, 2013).

- c. Lectura de palabras simples. Capacidad para decodificar palabras no relacionadas entre sí, por medio de una lista de palabras familiares que son seleccionadas de materiales de lectura de los primeros grados (de dificultad creciente) por su alta frecuencia (USAID, 2009).
- d. Decodificación de palabras inventadas (pseudopalabras). Capacidad para decodificar palabras que no existen y que a su vez no se pueden leer con facilidad al reconocerlas visualmente (World Vision, 2017).
- e. Lectura oral. La fluidez en la Lectura oral es una medida de la competencia global en la lectura. (Hasbrouck y Tindal, 2006). La capacidad de traducir letras en sonidos, unificar estos en palabras, procesar conexiones, relacionar el texto con significados, y hacer inferencias para deducir información no explícita en el texto (World Vision, 2017).
- f. Comprensión lectora. Aplicación de estrategias para entender, recordar y encontrar significado a lo que se ha leído, además de estar en capacidad de comunicarlo (World Vision, 2017).
- g. Instrumento de medición. Prueba cuyas respuestas se obtienen, codifican y transfieren a una matriz o base de datos y se preparan para su análisis mediante un paquete estadístico (Hernández, 2014).
- h. Propiedades psicométricas. Procedimientos de medición para determinar la validez y confiabilidad de la prueba que mide un determinado constructo (Hernández, 2014).
- i. Dificultad. El indicador de dificultad del ítem proviene de la proporción de sujetos que respondieron correctamente el ítem, entre el total de sujetos que respondieron el ítem. El indicador tiene valores entre 0 y 1 en donde 0 indica la máxima expresión de dificultad y 1 la mínima (Quim & Santos, 2015).

- j. **Discriminación.** La discriminación de elementos se refiere a la capacidad de un elemento para diferenciar entre los estudiantes en función de qué tan bien conocen el material que se está evaluando (University of Washington, 2019).
 - k. **Adivinación.** Parámetro de adivinación que representa la probabilidad de que por azar se adivine la respuesta correcta. En un reactivo de opción múltiple, incluso los examinados con muy bajo nivel de habilidad tienen una probabilidad mayor a cero de responder de manera correcta (Anastasi & Urbina, 1998).
 - l. **Confiabilidad.** Es el grado de consistencia y estabilidad de las puntuaciones obtenidas a lo largo de sucesivos procesos de medición con un mismo instrumento, lo que implica ausencia de errores de medida (Hernández, 2014).
 - m. **Validez.** Es la propiedad psicométrica que determina el grado en que un instrumento mide la variable que pretende medir. La validez es un concepto del cual pueden tenerse diferentes tipos de evidencia: evidencia relacionada con el contenido, evidencia relacionada con el criterio y evidencia relacionada con el constructo (Hernández, 2014).
3. **Definición operacional**
- a. **Habilidad de lectura.** Capacidad del estudiante de 3ro. primaria para leer y comprender un texto que se obtuvo como resultado en la Prueba de lectura inicial.
 - b. **Alfabeto.** Resultados obtenidos en la Prueba de lectura inicial en la Subprueba de alfabeto que indicó un nivel de habilidad de las niñas y los niños para reconocer las 27 letras del alfabeto, al indicar el nombre o sonido de cada letra (mayúscula o minúscula).
 - c. **Lectura de palabras simples.** Resultados obtenidos en la Prueba de lectura inicial en la Subprueba de palabras simples que indicó un nivel de habilidad de las niñas y los niños para leer una lista de 20 palabras comunes que no están relacionadas entre sí.

- d. Decodificación de palabras inventadas (pseudopalabras). Resultados obtenidos en la Prueba de lectura inicial en la Subprueba de palabras inventadas que indicó un nivel de habilidad de las niñas y los niños para decodificar los sonidos de las letras que forman las 20 palabras sin sentido.
- e. Lectura oral. Resultados obtenidos en la Prueba de lectura inicial en la Subprueba de lectura oral que indicó un nivel de habilidad de las niñas y los niños para leer en voz alta, con fluidez y precisión, una historia compuesta por 129 palabras, durante 60 segundos.
- f. Comprensión de lectura. Resultados obtenidos en la Prueba de lectura inicial en la Subprueba de comprensión lectora que indicó un nivel de habilidad de las niñas y los niños para responder correctamente, a las 10 preguntas relacionadas con la historia que leyeron previamente en voz alta y que están clasificadas en el nivel literal e inferencial de comprensión lectora.
- g. Instrumento de medición: Prueba de lectura inicial compuesta por cinco Subpruebas (Alfabeto, Palabras simples, Palabras inventadas, Lectura oral y Comprensión lectora) las cuales son una adaptación que realizó el equipo de STW según la prueba de lectura STAR.
- h. Propiedades psicométricas. Medidas que incluyeron el análisis de dificultad y discriminación de los ítems, la confiabilidad y validez de la Prueba de lectura inicial.
- i. Dificultad. Índice que estableció la proporción de estudiantes que respondieron correctamente a los ítems de la Prueba de lectura inicial.
- j. Discriminación. Índice que permitió determinar que tan bien los ítems permitieron identificar y discriminar entre los estudiantes con o sin habilidad en lectura según los resultados obtenidos por la Prueba de lectura inicial.

- k. Adivinación. Índice que confirmó la presencia o ausencia de respuesta al azar por parte de los estudiantes al responder la Prueba de lectura inicial.
- l. Confiabilidad. Índice que determinó el grado en el que el instrumento fue consistente y exacto con diferentes métodos. Se realizó por medio del análisis de correlación de mitades (Prueba de dos mitades), análisis de ítem-total y el indicador de coeficiente Alfa.
- m. Validez. Grado en el que la medida obtenida representó al constructo de lectura que se deseaba medir, de acuerdo con la capacidad de discriminación de la prueba (validez discriminante). Se realizó durante el proceso de esta investigación al analizar los resultados obtenidos en la muestra y en las diferentes unidades o escuelas por medio del análisis factorial exploratorio (AFE) y durante las pruebas de hipótesis en cada una de las Subpruebas de la Prueba de lectura inicial.

B. Marco referencial

1. Importancia de la lectura

La capacidad de leer y comprender un texto simple es una de las más fundamentales habilidades que un estudiante puede aprender. Sin una alfabetización básica, hay pocas posibilidades de que una niña o un niño puedan escapar del ciclo intergeneracional de la pobreza. La evidencia indica que aprender a leer tanto a tiempo como a una velocidad suficiente es esencial para aprender a leer bien (RTI, International, 2016).

Un sistema de educación básica de buena calidad debería proporcionar a las niñas y los niños las habilidades de lectura necesarias para aprender en el hogar, la escuela y la comunidad. Sin embargo, en muchos países, los estudiantes matriculados en la escuela por seis años no pueden leer y entender un texto simple. Los estudiantes que no saben leer con facilidad y entender lo que están leyendo cuando están en 3er. grado tienen menos probabilidades de transición de "aprender a leer" a "leer para aprender" las materias principales en 4to. grado y los siguientes grados. Esto los pone en riesgo de abandonar la

educación primaria y disminuye su potencial para participar como miembros productivos y efectivos de la sociedad (World Vision International, 2017).

La lectoescritura inicial es la base que la niñez necesita para tener éxito en todas las otras áreas de educación. Las niñas y los niños primero necesitan "aprender a leer" para que puedan "leer para aprender" (RTI, International, 2016).

2. Habilidades lectoras

La adquisición de la alfabetización se vuelve más difícil a medida que los estudiantes crecen; los estudiantes que no aprenden a leer en los primeros grados es más probable que repitan y eventualmente sean expulsados de la escuela. Es decir, si las habilidades básicas fundamentales no se adquieren desde el principio, las brechas en los resultados de aprendizaje (entre los estudiantes que han dominado las habilidades fundamentales de lectura y los que no lo hacen) crecen con el tiempo (RTI, International, 2016).

Según Chall (1996) la adquisición de la lectura es vista como un proceso de desarrollo complejo. Para ello son necesarias las habilidades del orden superior (la fluidez y la comprensión) las cuales se basan en las habilidades de orden inferior (la conciencia fonológica y decodificación), y a su vez las habilidades de orden inferior han demostrado ser un predictivo de los logros de la lectura posteriores. Por lo tanto, incluso si los niños aún no pueden leer un pasaje con comprensión, no obstante, se puede medir su progreso hacia la adquisición de las habilidades de orden inferior que son pasos necesarios a lo largo del camino hacia ese fin (RTI, International, 2016).

Los microprocesos son los que están más directamente relacionados con el acceso al léxico y el procesamiento sintáctico. La finalidad básica de los microprocesos está en el procesamiento léxico y sintáctico, para determinar si el sujeto realiza la decodificación de forma integrada y eficaz, si existen fallos en las habilidades de estos procesos hay indicios que sugieren que las limitaciones de decodificación o en el empleo inadecuado de los signos gráficos y de puntuación influyen negativamente en la comprensión (Junta de Andalucía., s.f).

Los macroprocesos son las operaciones de alto nivel asociadas a la comprensión del texto o procesamiento semántico. La finalidad básica sobre el procesamiento semántico es

determinar el grado de comprensión del mensaje. Se trata de detectar si existen fallos en las habilidades de extracción del significado del texto e identificación de la idea principal del mismo (Junta de Andalucía., s.f).

La evidencia afirma que únicamente cuando el lector decodifica las palabras que lee de manera automática, logra prestar atención a los significados y a lo que globalmente quiere decir el texto. Las niñas y los niños lectores que no tienen fluidez al leer estarán menos motivados, tendrán más dificultades para aprender los contenidos de las diferentes áreas curriculares y comprenderán menos lo que leen (Soriano, Miranda, Soriano, Nievas, & Félix, 2011).

Según OREALC (2005) en el área de Lenguaje los estudiantes deben desarrollar las siguientes competencias para la vida: la capacidad para tomar decisiones con información suficiente a la hora de expresarse y de interpretar mensajes, la habilidad para resolver problemas comunicativos, la habilidad para comunicarse afectiva y efectivamente, y la capacidad para pensar y expresar el pensamiento en forma crítica y creativa. Para lograrlo se requiere de una alfabetización oportuna y pertinente que permita desarrollar en los estudiantes los microprocesos y macroprocesos previamente descritos. Generalmente cinco componentes son aceptados como necesarios para dominar el proceso de lectura: conciencia fonológica, fonética (método de instrucción que ayuda a enseñar las relaciones entre sonido y símbolo), vocabulario, fluidez y comprensión (Vaughn & Linan-Thompson, 2004; Armbruster, et al., 2003).

a. Principio alfabético.

El principio alfabético es un concepto abstracto que se necesita enseñar de manera explícita y sistemática a los estudiantes, cuando los estudiantes entienden que los mapas de los sonidos se forman en letras, empezarán a aprender a decodificar palabras. Ya que el conocimiento del alfabeto incluye el conocimiento de los nombres de letras individuales, sus características gráficas distintivas, y qué fonema (s) cada uno representa (RTI, International, 2016). Es una habilidad crítica que las niñas y los niños deben adquirir para convertirse en lectores y es la primera Subprueba que conforma la prueba (World Vision International, 2017).

b. Palabras simples.

Las palabras familiares son palabras de alta frecuencia seleccionadas en materiales de lectura de grados iniciales (1ero. a 3er grado) las cuales son en el idioma y el contexto del estudiante. (RTI, International, 2016). La capacidad de leer una lista de palabras familiares es una habilidad crítica de lectura y se incluye en la segunda Subprueba de la prueba, es una medida del reconocimiento de palabras que impide que los estudiantes puedan adivinar cada palabra siguiente debido a que no están relacionadas (World Vision International, 2017).

c. Decodificación de palabras sin sentido (pseudopalabras).

Además de poder leer palabras familiares fuera de contexto, es igualmente importante que los estudiantes puedan decodificar palabras que no son familiares. Esta es la tercera Subprueba que conforma la prueba, sin esta habilidad, las niñas y los niños solo podrán leer las palabras que han memorizado, una táctica que les impedirá seguir aprendiendo vocabulario a medida que se vuelve cada vez más complejo en los grados superiores. Las evaluaciones de alfabetización que carecen de un componente de decodificación pueden no captar la incapacidad de un estudiante para leer palabras desconocidas (World Vision International, 2017).

d. Lectura oral.

La fluidez de la Lectura oral es una medida de la competencia general de lectura, la capacidad de traducir letras en sonidos, unificar sonidos en palabras, procesar conexiones, relacionar significado al texto y hacer inferencias para completar la información faltante (Hasbrouck y Tindal, 2006).

La fluidez puede verse como un puente entre el reconocimiento de palabras y la comprensión de textos. Si bien la decodificación es el primer paso para el reconocimiento de las palabras, los lectores deben avanzar en su capacidad de decodificación hasta el punto en que se convierta en automática; entonces su atención será libre de pasar de las letras y

palabras individuales a las ideas contenidas en el texto (Armbruster et al., 2003; Hudson, Lane, Pullen & Torgesen, 2005).

La fluidez de la Lectura oral es una medida de las habilidades avanzadas de lectura, ya que conecta los procesos necesarios para identificar con éxito las letras, las palabras más utilizadas y las palabras decodificables para probar la competencia de lectura. En esta Subprueba el estudiante lee en voz alta una historia, durante un minuto. Es probable que el hecho de no completar con éxito esta Subprueba apunte a la dificultad del estudiante para completar las tres tareas anteriores que también abordan las habilidades básicas de lectura (World Vision International, 2017).

e. Comprensión lectora.

Como respuesta a la historia que los estudiantes leyeron en voz alta, es importante evaluar a los estudiantes sobre su capacidad para retener y recordar la información que han leído. Una incapacidad para comprender y expresar esta comprensión de un pasaje apropiado para su grado sugiere que el estudiante también puede tener dificultades para descifrar o leer con fluidez (World Vision International, 2017).

La lectura es más que decodificar, pues intervienen distintos procesos de comprensión. Mabel Condemarín (1981) basada en Barret (1967) clasificó las dimensiones cognitivas y afectivas de la comprensión Lectora. Los niveles más conocidos y utilizados son comprensión literal, inferencial y crítica. Barret identifica dos más: reorganización de la información y apreciación Lectora (USAID-Mineduc, 2017).

3. Lectura en 3° primaria

La capacidad de lectura es tanto un resultado de la oportunidad de aprender como un factor crítico que determina si los estudiantes pueden continuar aprendiendo y avanzar en la escuela. Si los estudiantes no adquieren un nivel adecuado de capacidad de lectura al principio de su escolaridad, se retrasan cada vez más, por lo que su oportunidad continua de aprender depende de su nivel de capacidad de lectura (USAID, 2009).

De acuerdo con el Currículo Nacional Base CNB de 3ro. primaria, en la malla curricular del área de Comunicación y Lenguaje L-1 se establece que las niñas y los niños al finalizar el grado deben alcanzar las siguientes competencias:

- a. Interpreta mensajes orales emitidos con lenguaje cotidiano, científico y literario
- b. (Competencia 1).
- c. Expresa oralmente sus opiniones y pensamientos utilizando las estrategias adecuadas (Competencia 2).
- d. Utiliza elementos del lenguaje no verbal en el análisis de expresiones emitidas en otros lenguajes y por diferentes interlocutores (as) (Competencia 3).
- e. Aplica diversas estrategias de lectura para la asimilación de la información, la ampliación de conocimientos y como recreación (Competencia 4).
- f. Expresa sus ideas por escrito utilizando la estructura de las palabras y las modificaciones que sufren en su relación con las demás (Competencia 5).
- g. Enriquece su vocabulario básico con palabras generadas con la aplicación de diferentes estrategias (Competencia 6).
- h. Redacta textos con diferentes propósitos apegándose a las normas del idioma (Competencia 7).
- i. Aplica las herramientas del lenguaje oral y escrito además de recursos diversos en la investigación y en la generación de nuevos conocimientos (Competencia 8) (Mineduc, 2012).

4. Pruebas estandarizadas

Los procedimientos estandarizados son tan importantes que se incluyen como un criterio esencial para una evaluación válida en los estándares para la evaluación educativa y psicológica por medio de pruebas (Gregory, 2012). En las aplicaciones típicas, los encargados de administrar una prueba deben seguir de manera cuidadosa los procedimientos estandarizados para la aplicación y la evaluación que especifica el editor. Las especificaciones relacionadas con las instrucciones que deben darse al examinado, los límites de tiempo, la forma de presentación o respuesta de los reactivos, y los materiales o el equipo de prueba deben respetarse de manera estricta. Solos deben hacerse excepciones

con base en el juicio profesional, considerando de modo cuidadoso, principalmente en las aplicaciones clínicas (AERA, APA, & NCME, 1999).

Según Gregory, (2012) de manera general, las pruebas se pueden agrupar en dos campos: pruebas grupales y pruebas individuales. Las pruebas individuales son instrumentos que, por su diseño y propósito, deben aplicarse a una sola persona. En las pruebas individuales los examinadores deben familiarizarse íntimamente con los materiales y las instrucciones antes de comenzar la aplicación. Esto implica extensa práctica y anticipación de circunstancias poco comunes y de la respuesta adecuada. Un examinador bien preparado memoriza los elementos clave de las instrucciones verbales y está listo para manejar lo inesperado. Muchas de las pruebas tienen aspectos complejos de aplicación que, de no tomar en cuenta, podrían causar que el estudiante cometa errores innecesarios en los reactivos. La atención adecuada a los detalles de aplicación es esencial para obtener resultados válidos.

5. Pruebas para evaluar la lectura en grados iniciales

La evaluación es una herramienta privilegiada al servicio de la mejora continua y su valor pedagógico es brindar información sobre los resultados para tomar decisiones que mejoren la educación. Por medio de las pruebas tradicionales basadas en lápiz y papel, se requiere que las niñas y los niños ya hayan adquirido fluidez en la lectura y habilidades de comprensión para responderlas. Si no lo han hecho (es decir, si no pueden leer la pregunta o escribir la respuesta), la prueba no podrá medir con precisión lo que las niñas y los niños saben en términos técnicos. Y los resultados se verán afectados por un efecto suelo, con un alto número de estudiantes que alcanzan puntajes cero. En esos casos, la prueba indica únicamente lo que las niñas y los niños no saben, pero no lo que saben o dónde se encuentran a lo largo del camino del desarrollo de la comprensión Lectora (RTI, International, 2016).

Es por medio de una evaluación oral, que se puede obtener más información sobre lo que las niñas y los niños realmente saben y dónde están en el proceso de adquisición de lectura desde el principio. Las evaluaciones orales también pueden ayudar a detectar tempranamente las habilidades Lectoras y el crecimiento a lo largo del tiempo, es decir,

cambios que aún no son detectables en una prueba en papel pero que no obstante constituyen un avance hacia la adquisición de la lectura (RTI, International, 2016).

Específicamente en Guatemala, la evaluación de lectura, entendida como comprensión Lectora, en el Nivel de Educación Primaria a escala nacional, se inició en los años 90. Se evaluó a los estudiantes a partir de 9 años en adelante, con el objetivo de tener indicadores nacionales (Fortín, 2013). Posteriormente en el 2004, se inició la evaluación a primero primaria, con el intento de crear una evaluación con base en un criterio, el cual fue el Currículo Nacional Base. Y en el 2006, cuando surge la Dirección General de Evaluación e Investigación Educativa (Digeduca) del Ministerio de Educación, se lleva a cabo una evaluación de lectura para primero, tercero y sexto primaria con una muestra a nivel nacional, que sirve como un mecanismo de rendición de cuentas para el país (Cotto & Del Valle, 2017).

a. La evaluación de lectura en grados iniciales ELGI.

Esta prueba es una adaptación de la prueba original EGRA (Early Grade Reading Assessment). Evalúa el lenguaje oral, el principio alfabético, la conciencia fonológica, la rapidez automatizada de nombramiento, la decodificación, la fluidez de lectura, la comprensión de lectura y la escritura. Actualmente continúa en proceso de mejora para ser más eficaz al medir el desempeño lector de los estudiantes en primer grado del Nivel de Educación Primaria (Cotto & Del Valle, 2017).

b. La prueba de lectura STAR.

Por medio de la prueba de lectura STAR se proporciona información crítica sobre la capacidad de lectura fundamental de las niñas y los niños. STAR reemplaza a FLAT como la herramienta recomendada para evaluar la lectura, mide las habilidades de lectura para los estudiantes del grado 3 (2do. primaria en Guatemala) que son el foco de la alfabetización de los programas de Visión Mundial.

El objetivo de la herramienta es ayudar al personal de la Oficina Nacional a comprender las habilidades de lectura de las niñas y los niños para informar sobre la

programación educativa local. Los componentes de la evaluación incluyen: Conocimiento de las letras, leer la mayoría de las palabras usadas, leer las palabras decodificables, lectura de cuentos y comprensión Lectora. Respecto a la “Competencia mínima en comprensión Lectora” se refiere al logro de la lectura apropiada para el nivel de grado. La evaluación STAR está diseñada para ser utilizada con estudiantes del grado 3, que en Guatemala es 2º primaria o también al finalizar grado 2, el cual equivale a 1º primaria en Guatemala (World Vision International, 2017).

Es importante recopilar información de antecedentes sobre los estudiantes evaluados, para observar las tendencias que se relacionan con la lectura y el impacto de la equidad en la adquisición de las habilidades. Se incluyen preguntas sobre el sexo del estudiante, la edad, el idioma que se habla en el hogar y el entorno de alfabetización en el hogar. También hay preguntas sobre las prácticas de los maestros en el aula y la participación de los estudiantes en actividades de acción comunitaria. La inclusión de la información de antecedentes de los estudiantes y los datos relacionados con la eficacia de la escuela y el programa proporciona una imagen más completa para comprender las calificaciones de lectura de los estudiantes. También permite el análisis de acuerdo con las diferentes poblaciones marginadas para garantizar que aquellos que no están teniendo éxito puedan ser identificados adecuadamente. Es obligatorio, no opcional, recopilar esta información (World Vision International, 2017). La prueba de lectura STAR socializada por Visión Mundial incluye la lectura en el idioma ixil y español.

c. La Prueba de Lectura Inicial de STW.

Es el nombre de la prueba que School the World adaptó en base a la prueba de lectura STAR. Esta prueba se implementó en el idioma español, desde la plataforma Tangerine, la cual fue socializada por el equipo de Visión Mundial en Guatemala. En el siguiente capítulo correspondiente a la metodología se ampliará la información específica sobre la prueba, adaptación y aplicación.

6. Psicometría

La psicometría está representada etimológicamente como la teoría y técnica de medición de procesos mentales. Se aplica especialmente en los campos de la psicología y la educación. (Pasquali, 2009) El nivel más básico de información que brinda una prueba es la puntuación natural, la cual corresponde al número de respuestas que respondieron los estudiantes. Las puntuaciones naturales por sí solas carecen de significado, por lo que resulta necesario hacer comparaciones e identificar si los ítems fueron fáciles o difíciles, en especial con la edad de los estudiantes evaluados (Gregory, 2012).

Las pruebas referidas a la norma se emplean para clasificar a los estudiantes a lo largo de un continuo con fines comparativos y las pruebas referidas al criterio se usan para comparar los logros de un estudiante examinado, con un estándar de desempeño definido con anterioridad. Las pruebas referidas al criterio identifican el dominio (o la falta del dominio) en relación con competencias específicas y predeterminadas. Este tipo de pruebas es utilizado cada vez más en los sistemas educativos, para evaluar el dominio que han logrado los estudiantes de las habilidades académicas esperadas en cada grado escolar. La información que brindan proporciona la base para determinar la intervención que se dará a los estudiantes rezagados, evaluar el currículo y determinar el éxito que tiene cada escuela para implementar el currículo (Gregory, 2012) .

La psicometría moderna se remonta a dos fuentes: la teoría de la prueba clásica (TCT) y la teoría de respuesta al ítem (TRI). En un sentido general, la psicometría intenta explicar el significado de las respuestas dadas por los sujetos en una serie de tareas, típicamente nombradas ítems (Pasquali, 2009).

a. Teoría Clásica de los Test (TCT).

Charles Spearman (1904) estableció las bases para la teoría que después fue ampliada y revisada por algunos psicólogos contemporáneos (Feldt y Brennan, 1989; Lord y Novick, 1968; Kline, 1986). La teoría clásica de los test también se conoce como teoría de puntuaciones verdaderas y falsas, fue la base para el desarrollo de los instrumentos durante el siglo XX (Gregory, R., 2012).

b. Teoría de Respuesta al Ítem (TRI).

La teoría de respuesta al ítem o teoría del rasgo latente (TRI) es un conjunto de modelos matemáticos y herramientas estadísticas con usos generalizados. Ha sustituido a la teoría clásica de pruebas como modelo preferido para la elaboración de instrumentos. La TRI plantea una sola dimensión de la habilidad o el rasgo subyacente en la que se basan todos los ítems, e hipotetiza que cada estudiante evaluado posee cierta cantidad del rasgo latente que se está midiendo (Gregory, R., 2012).

La TRI, a diferencia de la TCT, se centra más en las propiedades individuales de los ítems que en las propiedades globales de la prueba, de ahí su nombre. Se puede decir que uno de sus propósitos es intentar obtener la puntuación que corresponde a una persona en una dimensión o rasgo, como por ejemplo, su inteligencia, su nivel en un cierto rasgo de personalidad, su dominio en una cierta materia, etc (Jiménez, K.; Montero, E., 2013). En la Tabla 2 se presentan las diferencias entre la TCT y TRI las cuales fueron adaptadas por Muñiz (2010).

Tabla 2. Diferencias entre TCT y TRI

Características	Teoría Clásica (TCT)	Teoría de Respuesta a los ítems (TRI)
Modelo	Lineal	Exponencial (modelo Poison, Ojiva normal, modelos Rasch, logísticos, etc.)
Asunciones	Débiles (fáciles de cumplir por los datos)	Fuertes (difíciles de cumplir por los datos)
Invarianza de las mediciones	NO	SI
Invarianza de las propiedades del test	NO	SI
Escala de las puntuaciones	Entre 0 y la puntuación máxima del test	Entre $-\infty$ y $+\infty$
Énfasis teórico	Test	Ítem
Relación Ítem-Test	Sin especificar	Curva característica del ítem
Errores de Medida	Error típico de medida común para toda la muestra	Varía según el nivel de aptitud del sujeto
Tamaño muestral	Puede funcionar bien con muestras de 200 a 500 sujetos	Se recomienda más de 500 sujetos.

(Muñiz, 2010)

Los coeficientes de correlación están presentes en los análisis generados desde la TCT y TRI. El Coeficiente de correlación expresa el grado de relación lineal entre dos conjuntos de puntuaciones obtenidas de los mismos estudiantes. Los coeficientes de correlación pueden tomar valores que van de -1.00 a +1.00. Un coeficiente de correlación

de +1.00 significa una relación lineal perfecta entre los dos conjuntos de puntuaciones, cuando dos medidas tienen una correlación de +1.00 el orden de los estudiantes es idéntico para ambos conjuntos de puntuaciones. Un coeficiente de correlación de -1.00 significa una relación igualmente fuerte, pero con una correspondencia inversa: la puntuación más alta en una variable corresponde a la puntuación más baja en la otra y viceversa (Gregory, 2012).

El coeficiente r de Pearson es el estimador por excelencia de la confiabilidad de los instrumentos en la psicometría clásica. Dos son las suposiciones básicas de confiabilidad desde este punto de vista. La primera se refiere a la obtención de puntajes idénticos, y la segunda se refiere a la obtención de idénticas posiciones relativas en las distribuciones de 10 puntajes obtenidos por los individuos en medidas sucesivas (Silva, A., 1987).

Desde la TRI se desarrolló el Modelo Rasch que permite realizar análisis psicométricos de las pruebas. En 1960 el matemático danés Georg Rasch propuso un modelo de medida que permite solventar muchas de las deficiencias de la TCT y construir pruebas más adecuadas y eficientes (Prieto & Delgado, 2003). El modelo propuesto por Rasch (1960) se fundamenta en los siguientes supuestos: El atributo que se desea medir puede representarse en una única dimensión en la que se situarían conjuntamente las personas y los ítems; el nivel de la persona en el atributo y la dificultad del ítem determinan la probabilidad de que la respuesta sea correcta. Si el control de la situación es adecuado, esta expectativa es razonable y así debe representarla el modelo matemático elegido (Prieto & Delgado, 2003).

Existen distintas escalas métricas de los valores de las personas y los reactivos. La más utilizada es la escala Lógit, que es el $\text{Ln}[\text{Pis}/(1-\text{Pis})]$, es decir $(\theta_s - \beta_i)$. La localización del punto cero de la escala es arbitraria, pero por tradición el punto cero indica la habilidad media de los sujetos (Embretson y Reise, 2000). El intervalo de probabilidad se ubica entre -5 y 5 o bien entre -3 y 3. Los procedimientos de análisis permiten detectar a los reactivos y a los estudiantes que no se ajustan al modelo. En este caso, es muy sencilla la interpretación de los parámetros de las personas (los valores de θ mayores que 0 indican que las personas tienen una probabilidad superior a 0,50 de éxito en los ítems de dificultad media). Aunque la escala lógit puede adoptar valores entre más y menos infinito, la gran mayoría de los casos se sitúa en el rango ± 5 (Prieto & Delgado, 2003).

Con la aplicación del modelo de Rasch, de la Teoría de Respuesta al Ítem (TRI), se obtiene un análisis detallado de la calidad técnica y del aporte de cada ítem, independientemente del conjunto de individuos sobre el que fue aplicada la prueba, lo cual es fundamental para la conformación de un banco de ítems. También es posible estimar el error de medida asociado a la medición realizada, así como obtener una estimación del dominio que posee el estudiante en los conocimientos evaluados. Con lo anterior es posible ir más allá de la estimación ya que se puede: Estudiar la relación entre el nivel de habilidad de los evaluados y el nivel de dificultad de cada ítem, con lo cual se podrían determinar los distintos niveles de desempeño de los estudiantes y poder realizar recomendaciones más acertadas y aplicar un enfoque psicométrico basado en TRI en cada proceso, tanto en el desarrollo de la prueba (construcción, juzgamiento, ensamblaje) como en su administración y calificación. Además, cuenta con más evidencias empíricas y validaciones científicas, que deben ser consideradas en las inferencias que se derivan de los resultados, a partir del uso de la prueba, para la toma de decisiones (Jiménez & Montero, 2013).

El modelo de Rasch postula que la probabilidad que tiene un estudiante de responder correctamente una pregunta depende de la diferencia entre el nivel de habilidad de la persona y el nivel de dificultad de la pregunta. Esto supone que ambas se deben ubicar en la misma escala con las mismas unidades, lo cual se constituye en una ventaja del modelo. El modelo de Rasch, al igual que otros modelos enmarcados en la Teoría de Respuesta al Ítem parten de una serie de hipótesis referentes a los datos que se utilizan en la medición del fenómeno latente: monotonicidad, independencia local y multidimensionalidad (Badenas, 2009).

Las ventajas del modelo de Rasch con respecto a otros modelos de la TRI es que es muy simple de aplicar. Otra característica, es que permite analizar las interacciones entre los estudiantes y los ítems. Además, las medidas que se obtienen no dependen de las condiciones específicas de cómo son obtenidas. Otra propiedad, es la unidimensionalidad del instrumento, es decir, que todos los reactivos que lo componen contribuyan a evaluar una sola característica o competencia. Finalmente, la independencia local postula que la respuesta a un reactivo no está influenciada por las respuestas a los otros reactivos del instrumento (Martín, Díaz, Córdoba, & Picquart, 2011).

Los aspectos principales que se analizan de un ítem son: dificultad, discriminación, distractores, omisiones y adivinación. En el análisis de los ítems según Rasch, los parámetros de las personas y los ítems se expresan en las mismas unidades y se localizan en el mismo continuo (Prieto & Delgado, 2003). A continuación, se describen estos indicadores.

a) Dificultad.

El indicador de dificultad del ítem proviene de la proporción de sujetos que respondieron correctamente el ítem, entre el total de sujetos que respondieron el ítem. El indicador tiene valores entre 0 y 1 en donde 0 indica la máxima expresión de dificultad y 1 la mínima. En la Tabla 3 se presentan los rangos y criterios de acuerdo con el indicador de dificultad de los ítems, los cuales fueron adaptados por Quim & Santos (2015) desde García, González & Ballesteros (2001). El índice de dificultad del ítem es una herramienta útil para identificar los ítems que deben modificarse o descartarse (Gregory, 2012).

Tabla 3. *Rango del indicador de dificultad*

Rango	Criterio
I.D. < 0.25	Muy difíciles
0.25 < 0.45	Difíciles
0.45 < I.D < 0.55	Normales
0.55 < 0.75	Fáciles
I.D. > 0.75	Muy fáciles

(Quim, M. y Santos, J., 2015)

Para los elementos con una alternativa correcta que vale un solo punto, la dificultad del elemento es simplemente el porcentaje de estudiantes que contestan un elemento correctamente. En este caso, también es igual a la media del elemento. El índice de dificultad del objeto varía de 0 a 100; cuanto mayor sea el valor, más fácil será la pregunta. La dificultad del ítem es relevante para determinar si los estudiantes han aprendido el concepto que se está probando. También juega un papel importante en la capacidad de un elemento para discriminar entre los estudiantes que conocen el material evaluado y los que no (University of Washington, 2019). En jMetrik el bparam representa la dificultad del ítem (Meyer, 2014). En la Tabla 4 se presenta el conjunto de las relaciones de probabilidad de éxito entre la capacidad (en lógit) y la dificultad (en Lógit) las cuales son parte del Modelo Rasch: independiente de cualquier contenido de la prueba (Bond & Fox, 2015).

Tabla 4. Probabilidades de Rasch de respuesta correcta

		Dificultad del ítem en aumento de lógit >>												
		-	-	-	-	-	-	0.00	0.50	1.00	1.50	2.00	2.50	3.00
Capacidad de la persona en aumento de lógit >>		3.00	2.50	2.00	1.50	1.00	0.50							
3.00		0.99	0.99	0.99	0.99	0.98	0.97	0.95	0.92	0.88	0.82	0.73	0.62	0.50
2.50		0.99	0.99	0.99	0.98	0.97	0.95	0.92	0.88	0.82	0.73	0.62	0.50	0.38
2.00		0.99	0.99	0.98	0.97	0.95	0.92	0.88	0.82	0.73	0.62	0.50	0.38	0.27
1.50		0.99	0.98	0.97	0.95	0.92	0.88	0.82	0.73	0.62	0.50	0.38	0.27	0.18
1.00		0.98	0.97	0.95	0.92	0.88	0.82	0.73	0.62	0.50	0.38	0.27	0.18	0.12
0.50		0.97	0.95	0.92	0.88	0.82	0.73	0.62	0.50	0.38	0.27	0.18	0.12	0.08
0.00		0.95	0.92	0.88	0.82	0.73	0.62	0.50	0.38	0.27	0.18	0.12	0.08	0.05
-0.50		0.92	0.88	0.82	0.73	0.62	0.50	0.38	0.27	0.18	0.12	0.08	0.05	0.03
-1.00		0.88	0.82	0.73	0.62	0.50	0.38	0.27	0.18	0.12	0.08	0.05	0.03	0.02
-1.50		0.82	0.73	0.62	0.50	0.38	0.27	0.18	0.12	0.08	0.05	0.03	0.02	0.01
-2.00		0.73	0.62	0.50	0.38	0.27	0.18	0.12	0.08	0.05	0.03	0.02	0.01	0.01
-2.50		0.62	0.50	0.38	0.27	0.18	0.12	0.08	0.05	0.03	0.02	0.01	0.01	0.01
-3.00		0.50	0.38	0.27	0.18	0.12	0.08	0.05	0.03	0.02	0.01	0.01	0.01	0.01

(Bond & Fox, 2015)

b) Discriminación y Beta.

El índice de discriminación del ítem es un indicador estadístico de la eficiencia con el que el ítem discrimina entre los estudiantes que obtienen puntuaciones altas y bajas en toda la prueba (Gregory, 2012). Es importante comprender que el índice de discriminación proviene de la correlación del resultado en el ítem con la calificación total de la prueba y se aduce que el resultado del ítem fue eliminado del total de la prueba para hacer la correlación. El elemento tendrá una baja discriminación si es tan difícil que casi todos se equivocan o adivinan, o tan fácil que casi todos lo entienden bien.

De acuerdo con la Universidad de Washington (2019) el índice de discriminación es “bueno” si está por encima de 0.30, es “justo” si está entre 0.10 y 0.30 y es “pobre” si está por debajo de 0.10. En la Tabla 5 se presentan los criterios para identificar el nivel de discriminación de los ítems los cuales fueron adaptados por Quim & Santos (2015) desde Muñiz et al (2005). El índice de discriminación del reactivo se define como una medida para determinar si las competencias o habilidades que mide la prueba también las mide el ítem.

Tabla 5. *Rango del nivel de discriminación*

Rango	Criterio
Igual o mayor que 0.40	El ítem discrimina muy bien
Entre 0.30 y 0.39	El ítem discrimina bien
Entre 0.20 y 0.29	El ítem discrimina poco
Entre 0.10 y 0.19	El ítem se debe mejorar
Menor de 0.10	El ítem carece de utilidad para discriminar

(Quim, M. y Santos, J., 2015)

c) Adivinación.

Parámetro que representa la probabilidad de que por azar se adivine la respuesta correcta. En un reactivo de opción múltiple, incluso los examinados con muy bajo nivel de habilidad tienen una probabilidad mayor a cero de responder de manera correcta (Anastasi & Urbina, 1998).

Las ventajas del modelo de Rasch solo pueden ser obtenidas si los datos empíricos se ajustan al modelo. De acuerdo con la ecuación, la probabilidad de respuesta a un ítem depende solo de los niveles de la persona y el ítem en el atributo medido. La presencia de respuestas aberrantes tales como que personas poco competentes resuelvan correctamente ítems difíciles, indicaría que los parámetros de sujetos e ítems son meros numerales carentes de significado teórico. La falta de ajuste podría deberse a diversos factores: multidimensionalidad o sesgo de los ítems, falta de precisión en el enunciado o en las opciones, respuestas al azar, falta de motivación o cooperación, errores al anotar la respuesta, copiado de la solución correcta, etc (Karabatsos, 2000).

1) Infit (WMS) y Outfit (UMS).

Existen dos medidas de bondad de ajuste en el modelo de Rasch: el Infit que se interpreta como ajuste interno, es un valor sensible al comportamiento inesperado que afecta a los reactivos cuya dificultad está cerca del nivel de habilidad de una persona y el Outfit que se interpreta como ajuste externo, es un valor sensible al comportamiento inesperado que afecta a los reactivos cuya dificultad está lejos del nivel de habilidad de una persona. Estos estadísticos se reportan como medias cuadráticas de residuales (MNSQ) y

como residuales estandarizados (ZSTD). En la práctica el criterio que se aplica es que los valores de la media cuadrática deben estar entre 0.8 y 1.3 y los valores estandarizados deben estar entre -2 y 2, lo cual indica un ajuste razonable (González, 2008). Según Meyer (2014) el Infit y Outfit significan estadísticas de ajuste residual cuadrado y tienen un valor esperado de 1 y siempre son positivos en valor. Los valores cercanos a 1 indican un buen ajuste de los datos del modelo, y los valores divergen de 1 a medida que aumenta el grado de inadaptación. Si los datos realmente se ajustan al modelo sus valores esperados son 0.0. Si es menor que 0.0 indica que es demasiado predecible y si es mayor de 0.0 indica falta de previsibilidad. Los valores estandarizados pueden ser positivos o negativos; es necesario considerar que las estadísticas estandarizadas del Infit y Outfit son sensibles al tamaño de la muestra, solo son útiles para muestras de menos de 300 examinados (Linacre, J., 2002).

Existen diferentes reglas básicas para decidir si un ítem o persona se ajusta bien. Un rango estrecho de los valores medios cuadrados de Infit y Outfit como 0.8 a 1.2 se recomiendan para apuestas (Bond & Fox, 2007), pero los valores entre 0.5 y 1.5 siguen siendo productivos para medición (Linacre, 2012). Los valores de la media cuadrática menores a 0.8 o valores estandarizados menores a -2 indican datos con demasiado determinismo o poco estocásticos, mientras que valores de la media cuadrática mayores a 1.3 o valores estandarizados mayores a 2 indican alta posibilidad de azar (Tristán, 2001).

2) Desviación estándar del modelo (Std. UMS y Std. WMS).

En el análisis del modelo Rasch en jMetrik, las dos estadísticas de equipo son etiquetadas como UMS (unweighted mean squares) y Std. UMS para indicar la estadística de ajuste cuadrático medio no ponderado y estadístico de ajuste de cuadrado medio no ponderado estandarizado, respectivamente. Del mismo modo, las estadísticas Infit están etiquetadas como WMS (weighted mean squares) y Std. WMS para indicar la estadística de ajuste cuadrado media ponderada y el ajuste de cuadrado medio ponderado estandarizado (Meyer, J., 2014). Linacre (2012), Meyer sugiere que los valores de UMS y WMS deben tener un rango de 0.5 a 1.5, con valores mayores a 2 que indiquen que el elemento en realidad degrada la medición. Los valores estandarizados de la (Std. UMS y

Std. WMS) son útiles para tamaños de muestra más pequeños y el valor absoluto de estas estadísticas no debe ser mayor que 3 (Tennant, J., 2015).

3) Raíz del error cuadrático medio (RMSE).

El Root Mean Squared Error (RMSE) o error cuadrático medio o dispersión de muestra de valores en torno a medidas calibradas (Wright, 1996) es uno de los parámetros “errores estándar más realistas” de medida. El modelo RMSE para medidas de niños es 0.48, (no muy lejos de lo real RMSE de 0.52) e implica precisión relativa en las estimaciones modelo de las medidas. La desviación estándar estimada "verdadera" del modelo para las medidas de los niños es 0.92 (Linacre, 1998, p. 57). Las estadísticas de ajuste del error cuadrático medio muestran el tamaño de la aleatoriedad, es decir, la cantidad de distorsión del sistema de medición, ya que 1.0 son sus valores esperados. Los valores inferiores a 1.0 indican que las observaciones son demasiado predecibles (redundancia, los datos se adaptan al modelo). Los valores superiores a 1.0 indican imprevisibilidad (ruido no modelado, datos que se adaptan al modelo). Estadísticamente, los cuadrados medios son estadísticas de chi cuadrado divididas por sus grados de libertad. Las medias cuadráticas son siempre positivas (Linacre, J., 2002). En general, los cuadrados medios cerca de 1.0 indican poca distorsión del sistema de medición, independientemente del valor estandarizado. Al evaluar la media de los cuadrados por encima de 1.0 antes que la media de los cuadrados, muy por debajo de 1.0, ya que la media de la media por lo general se ve obligada a estar cerca de 1.0 (Linacre, J., 2002)

Los valores de anclaje pueden no coincidir exactamente con los datos actuales. En la medida en que no lo hagan, las estadísticas ajustadas pueden ser engañosas. Los valores de anclaje que son demasiado centrales para los datos actuales tienden a hacer que los datos parezcan encajar demasiado bien. Los valores de anclaje que están demasiado dispersos para los datos actuales tienden a hacer que los datos parezcan ruidosos (Linacre, J., 2002).

Tabla 6. Valores de anclaje

Valor estandarizado	Implicación para la medición
≥ 3	Datos muy inesperados si se ajustan al modelo (perfectamente), por lo que probablemente no lo hagan. Pero, con un tamaño de muestra grande, el desajuste sustancial puede ser pequeño.
2.0 -2.9	Datos notablemente impredecibles.
-1.9-1.9	Los datos tienen una previsibilidad razonable.
≤ -2	Los datos son demasiado predecibles. Otras "dimensiones" pueden estar limitando los patrones de respuesta.

(Linacre, J., 2002)

Se dice que un ítem muestra funcionamiento diferencial (Differential Item Functioning, DIF) cuando presenta propiedades psicométricas diferentes en función del grupo en el que ha sido administrado, después que los grupos han sido igualados en el rasgo o habilidad medida por la prueba (Angoff, 1993). Que una prueba o un ítem esté sesgado implica que la prueba o el ítem favorecen injustamente a un grupo sobre otro, es decir, se encuentra disparidad en el comportamiento del mismo en función de alguna variable grupal (p.e., edad, sexo, nivel socioeconómico, bagaje cultural y dominio lingüístico) que no resulta relevante para el constructo que trata de medir el test (Gómez & Hidalgo, 2003)

Por medio del estudio de la fiabilidad de un instrumento de medida, se debe explorar hasta qué punto los resultados que se obtienen a partir de su aplicación están afectados por el error que se ha cometido al medir. Si el error es pequeño, se puede confiar en los resultados de la prueba; si el error es grande, el proceso de medición deja de tener sentido (Barrios & Cosculluela, 2013). Murphy y Davidshofer (2005) afirman que en cualquier contexto de evaluación una fiabilidad por debajo de 0,6 se consideraría baja e inaceptable. Kaplan y Saccuzo (2009) sugieren que coeficientes de fiabilidad que oscilan entre 0,7 y 0,8 son suficientemente buenos para la mayoría de las ocasiones en las que las pruebas se utilizan para fines de investigación. Y alrededor de 0.90 es un nivel elevado de confiabilidad.

El coeficiente alfa de Cronbach. El procedimiento más utilizado para evaluar la consistencia interna de una prueba es el coeficiente Alfa de Cronbach (Muñiz, 1996 y 2003). La confiabilidad es una generalización de la fórmula KR. Para el cálculo del alfa de Cronbach se emplea el promedio de todas las correlaciones existentes entre los ítems del instrumento que atributan al concepto latente que se pretende medir (Barraza, 2007). El coeficiente alfa de Cronbach (1951) expresa la consistencia interna de una prueba a partir

de la covariación entre sus ítems. Cuando más elevada sea la proporción de la covariación entre los ítems respecto a la varianza total de la prueba, más elevado será el valor del coeficiente alfa de Cronbach y más elevada su fiabilidad (Barrios & Cosculluela, 2013). El coeficiente alfa es un índice de la consistencia interna de los ítems, es decir, de su tendencia a correlacionarse de manera positiva entre sí. Por tradición se ha considerado que el coeficiente alfa es un índice de la unidimensionalidad, es decir, del grado en que una prueba o escala mide un solo factor (Gregory, R., 2012).

Otros coeficientes también exploran indirectamente la homogeneidad o la dimensionalidad de una escala, es decir, unidimensional o multidimensional, o cuantos dominios explora la escala (Cortina, 1993). Entre ellos está: el coeficiente de Guttman's lambda-2 es la estimación de confiabilidad en el cálculo de la puntuación regresiva de Kelley. Transformación no lineal de las opciones en el panel puntuación incluyen rango percentil y puntuación normalizada. Estos son ambos calculados a partir de la puntuación total. El coeficiente de Feldt-Gilmer es una variante para cuando la escala está dividida en tres o más partes iguales (Feldt, L. y Brennan, R., 1989) el coeficiente de Feldt-Brennan es el cálculo de una estimación de confiabilidad para medidas clásicamente congénicas (Meyer, 2014).

El coeficiente Raju beta propuesto por Raju, pondera la consistencia interna de una escala cuando es dividida en dos o más partes desiguales (Kristof, W, 1974). Cuando falta interrelación en los componentes de una escala, se puede subestimar la consistencia interna (Osburn, h, 2000). El coeficiente KR21 (Kuder y Richarson) antes de que Cronbach propusiera el coeficiente alfa como indicador de la consistencia interna de un test, Kuder y Richardson presentaron dos fórmulas de cálculo de este indicador, conocidas como KR20 y KR21 que son casos particulares cuando los ítems son dicotómicos y se codifican las dos alternativas de respuesta posibles como 0 y 1 (Barrios & Cosculluela, 2013).

Para la confiabilidad si los ítems de las pruebas se eliminan, el análisis de confiabilidad a través de la consistencia interna permite estudiar la propiedad en la prueba, pero el valor del análisis es el que permite observar la relación entre cada ítem y el funcionamiento total de la prueba; por lo tanto, se puede saber qué tanto mejoraría o empeoraría la consistencia de la prueba si se decide eliminar el ítem (Bogliaccini, Cardozo y Rodríguez, 2005).

La consistencia interna de un cuestionario de preguntas es la característica fundamental de la confiabilidad de un instrumento para tomar datos en la investigación, está dada por el grado de homogeneidad total y covariación que tienen los ítems entre sí, salvo que se tenga interés por conocer la consistencia interna de dos o más partes del instrumento tales como la primera mitad, los reactivos pares o impares etc., se determinaría por partes el Alpha de Cronbach. La variación del coeficiente es entre 0 y 1. Si el valor es mayor que 0.8, el instrumento se considera confiable y adecuado para investigar. El cálculo puede realizarse mediante dos fórmulas, considerando la varianza de los ítems y la matriz de correlación de los ítems (Frías, 2013). La consistencia interna de un instrumento se puede evaluar mediante el procedimiento de dos mitades (Cozby, 2005) método de las mitades o Split-half method (Briones, 2001). Se obtiene una estimación de la confiabilidad por mitades al correlacionar los pares de puntuaciones obtenidas de las mitades equivalentes de una prueba aplicada solo una vez a una muestra representativa de estudiantes. Si las puntuaciones de las dos mitades de la prueba, obtenidas a partir de una sola aplicación, reflejan una fuerte correlación, entonces las puntuaciones de las dos pruebas completas obtenidas en dos aplicaciones separadas, también debería revelar una fuerte correlación (Gregory, R., 2012).

Una prueba es válida en la medida en que las inferencias que se hagan a partir de ella sean apropiadas, significativas y útiles (Gregory, R., 2012). Diferentes instancias académicas, así como la mayoría de los autores contemporáneos contemplan a la validez como un concepto unitario en el que se considera toda la evidencia acumulada con respecto a la interpretación de las puntuaciones de un instrumento para el propósito establecido (Mertens, 2005; Bostwick y Kyte, 2005). En este sentido, la validez no puede descansar en un solo tipo de evidencia (Messick, 1995, 1989 y 1988).

En términos estadísticos la validez se define como la proporción de la varianza verdadera que es relevante para los fines del examen. El término relevante se refiere a lo que es atribuible a la variable, características o dimensión que mide la prueba. En este sentido, generalmente la validez de una prueba se define ya sea por medio de la relación entre sus puntuaciones con alguna medida de criterio externo, o bien la extensión con la que la prueba mide un rasgo subyacente específico hipotético o “constructo” (Aliaga, 2007).

La validación de una prueba es el proceso de acumular evidencia para apoyar tales inferencias. Una variedad de evidencias puede obtenerse de las puntuaciones producidas por una prueba dada, y hay muchas formas de acumular evidencia para apoyar una inferencia específica. La validez, sin embargo, es un proceso unitario. Aunque la evidencia puede ser acumulada de muchas formas, la validez se refiere siempre al grado en que esa evidencia apoya las inferencias que se hacen a partir de las puntuaciones” (American Psychological Association, American Educational Research Association, and National Council on Measurement in Education, 1985) La confiabilidad es una condición necesaria, pero no suficiente, de la validez (Gregory, R., 2012).

La validez no se puede resumir en un solo indicador o índice numérico, al igual que ocurre con la confiabilidad sino que la validez de las puntuaciones de una prueba se asegura mediante la acumulación de evidencia teórica, estadística, empírica y conceptual del uso de las puntuaciones (Aliaga, 2007). A continuación se definen los diferentes tipos de validez:

- a) Validez convergente (correlación interna). Se demuestra cuando una prueba tiene una correlación elevada con otras variables o pruebas con las cuales comparte una superposición de los constructos (Gregory, R., 2012).
- b) Validez predictiva (regresión). En la validez predictiva las medidas del criterio se recaban en el futuro, por lo general meses o años después de obtener las puntuaciones de la prueba (Gregory, R., 2012).
- c) Validez de constructo. El análisis factorial es un modelo estadístico especializado que resulta particularmente útil para investigar la validez de constructo. El propósito es identificar el número mínimo de determinantes (factores) que se requieren para explicar las intercorrelaciones entre una batería de pruebas (Gregory, R., 2012). Por medio de la validez de constructo se tiene la evidencia que apoya que las conductas observadas o las puntuaciones obtenidas en una prueba son indicadores o manifestaciones del constructo teórico que se pretendía medir (Martínez, R., 2005).

El análisis factorial exploratorio (AFE) es una técnica de la estadística multivariada que se usa en psicometría para obtener evidencias de las dimensiones subyacentes, factores o componentes que están presentes en el instrumento. Recurre al método inductivo o exploratorio, dado que el investigador analizará el conjunto de datos y variables observadas

sin tener a priori ninguna hipótesis acerca de la estructura y dejará que sea el análisis de los datos quienes proporcionen información al respecto (Ibáñez, 2016). A nivel global, las cargas factoriales de los ítems (que representan el peso o nivel de importancia del ítem en cada factor) se consideran óptimas si son iguales o mayores a 0.3, en valor absoluto. Antes de realizar un análisis psicométrico con la TCT, la TRI o Rasch es importante evidenciar, utilizando el análisis factorial exploratorio, que el instrumento mide fundamentalmente solo un rasgo o constructo, pues este es un supuesto que debe cumplirse para que la aplicación de estos modelos de medición sea válida (Jiménez, K.; Montero, E., 2013). Otras de las exigencias del análisis factorial exploratorio es que debe utilizar muestras grandes, de aproximadamente 300 participantes, para obtener resultados útiles y relativamente estables (Tabachnick & Fidell, 2001).

El Indicador de KMO es la medida de adecuación muestral de Kaiser-Meyer Olkin (KMO) 1970 que refleja la influencia de todos los factores: tamaño de las correlaciones entre los ítems, tamaño de la muestra, número de factores y número de ítems. Esta medida de adecuación indica cuán grande es la correlación entre las variables medidas. Si las correlaciones son suficientemente grandes, la matriz se considera adecuada para su factorización porque ofrecerá resultados estables, replicables en otras muestras diferentes, independientemente del tamaño de la muestra, o del número de factores, o del número de ítems. La adecuación de los datos al Análisis Factorial se considera “suficiente” cuando el resultado de la medida KMO tome valores entre .70-.79, y “satisfactoria” cuando tome valores mayores a 0.80.

- a) Una correlación próxima a cero quiere decir que el responder bien a esa pregunta no tiene que ver con estar bien en el conjunto del examen.
- b) Una correlación negativa, sobre todo si es de cierta magnitud, quiere decir que el responder bien a esa pregunta está relacionado con estar más bien mal en el conjunto de la prueba (en principio se trata de una mala pregunta, o quizás hay un error en la clave de corrección).
- c) Una correlación positiva quiere decir que el responder bien a esa pregunta está relacionada con un buen resultado en el conjunto de la prueba. Los ítems con mayores correlaciones positivas son los más discriminantes, los que mejor diferencian a los mejores y peores estudiantes (Morales, P., 2008).

Martínez (2005) señala que antes de realizar un AFE deben evaluarse tres supuestos principales: la normalidad, linealidad y multicolinealidad de las puntuaciones. Para obtener el KMO se debe de elegir la extracción con el método: Componentes principales, Mínimos cuadrados no ponderados, Mínimos cuadrados generalizados, Máxima verosimilitud, Factorización de ejes principales, Análisis alfa y Análisis de imagen; así como la rotación: Varimax, Oblim directo, Quartimax, Equamax y Promax. (SPSS, 2010). En la validación de cuestionarios los métodos de rotación más comunes son dos. El método Varimax que mediante la rotación ortogonal minimiza el número de variables que tienen saturaciones altas en cada factor, de forma que se simplifica la interpretación de los factores. Y el método Promax, que mediante la rotación oblicua permite que los factores estén correlacionados y suele ser muy útil, en conjuntos grandes de datos (Ibáñez, 2016).

En el contexto de la validación de una prueba, el objetivo que persigue la calibración de los ítems es verificar el cabal desempeño estadístico de los ítems y del instrumento en su conjunto, así como fijar un valor para los parámetros de dificultad, error de estimación, correlación ítem-prueba, ajuste próximo y lejano y discriminación, de cada uno de los ítems que finalmente son admitidos. Aquellos ítems cuyos parámetros presentan valores que no satisfacen los criterios establecidos se excluyen del procesamiento (Bogoya, Barragán, Contento, & Ocaña, 2014).

C. Marco contextual

1. Departamento de Quiché

Se encuentra situado en la región Noroccidente de Guatemala. Limita al Norte con México; al Sur con los departamentos de Chimaltenango y Sololá; al Este con los departamentos de Alta Verapaz y Baja Verapaz; y al Oeste con los departamentos de Totonicapán y Huehuetenango (Instituto Nacional de Estadística, 2010). Está organizado en 21 municipios de los cuales en tres School the World tiene programas de intervención.

a. Territorio y cultura.

El departamento de Quiché tiene una extensión territorial de 8,378 Km². Se encuentra a una altitud de 1,196 y 2,310 msnm (Anexo 1). El clima es templado – frío y la producción agrícola se basa en el maíz amarillo, maíz blanco, frijón negro y arroz. Los idiomas predominantes son el Uspanteko, Ixil, Sakapulteko, K'iché y Español (Anexo 2).

b. Número de habitantes.

Según las proyecciones de población, el número de habitantes del departamento de Quiché hasta el 2013 fue de 1, 019,290 personas representando el 6.6% de la población total, estimada para ese año en 15, 438,383. La población creció 3.41% entre 2012 y 2013, porcentaje superior a la tasa de crecimiento nacional, que fue de 2.32%. Tiene una densidad poblacional de 109 personas por Km² (Instituto Nacional de Estadística, 2010). Según los datos del último censo del INE, en el 2013 la población a nivel municipal en San Pedro Jocopilas era de 29,035 habitantes, en Santa Cruz del Quiché era de 102,782 habitantes y en Santo Tomás Chichicastenango de 148,855 habitantes. Según las proyecciones de población del departamento de Quiché para 2013, los dos municipios con mayor cantidad de población son: Chichicastenango y Santa Cruz del Quiché (Instituto Nacional de Estadística, 2014). En los cuales se encuentran ubicadas la mayoría de las escuelas que aceptaron participar en la aplicación de la Prueba de Lectura inicial.

c. Índices de cobertura.

Desde el 2009 al 2013 el índice de analfabetismo en el departamento presentó una tendencia decreciente, pasando de 37.6 en 2009 a 30.1 en 2013. En el nivel primario, para el 2013 la tasa de escolaridad alcanzó un 99.0%, en tanto que la tasa de aprobación fue de 81.8%. Por su parte, la tasa de retención, es decir los estudiantes que se inscribieron y continuaron sus estudios en el nivel primario, pasó de 96.9% a 98.1% (Mineduc, 2019).

De acuerdo con el anuario estadístico presentado por el Ministerio de Educación de Guatemala, en la versión preliminar presentada en el 2019, durante el ciclo escolar 2018 el

porcentaje de los indicadores de cobertura para el nivel de Primaria en la República de Guatemala y el departamento de Quiché fueron los siguientes:

Tabla 7. *Porcentaje de los indicadores de cobertura en Guatemala y el depto. de Quiché*

Lugar	Tasa de repitencia	Tasa de retención	Tasa de deserción	Tasa de promoción	Tasa de no promoción	Fracaso escolar
República de Guatemala	9.16%	95.32%	4.68%	88.64%	11.36%	15.51%
Depto. de Quiché	12.25%	95.17%	4.83%	84.42%	15.58%	19.66%

(Mineduc, 2019)

d. Municipios con escuelas participantes en los programas de STW.

La aplicación de la Prueba de Lectura inicial, fue realizada con el consentimiento de 9 escuelas las cuales se encuentran ubicadas en el departamento de Quiché, en tres de los municipios:

- 1) En San Pedro Jocopilas participó la escuela de Chocruz.
- 2) En el municipio de Santa Cruz del Quiché participó la escuela de Xesic I.
- 3) Y en el municipio de Santo Tomás Chichicastenango participaron las escuelas ubicadas en el Cantón Panimaché Quinto Alto, Cantón Río Mactzul, Cantón Mactzul Séptimo, Cantón Chuguexa Cuarto, Cantón Agua Viva, Cantón Chulumal Segundo y Cantón Xalbaquiej.

En los tres municipios en donde se ubican las nueve escuelas evaluadas, durante el ciclo escolar 2018 presentaron los indicadores de la tasa de repitencia, retención, deserción, promoción, de no promoción y de fracaso escolar para el nivel de Primaria con los siguientes porcentajes:

Tabla 8. *Porcentaje de los indicadores de cobertura en tres de los municipios de Quiché*

Municipio del depto. de Quiché	Tasa de repitencia	Tasa de retención	Tasa de deserción	Tasa de promoción	Tasa de no promoción	Fracaso escolar
San Pedro Jocopilas	12.06%	94.96%	5.04%	83.5%	16.5%	20.71%
Santa Cruz del Quiché	10.90%	97.05%	2.95%	87.02%	12.98%	15.55%
Santo Tomás Chichicastenango	10.20%	94.84%	5.16%	85.48%	14.52%	18.93%

(Mineduc, 2019)

e. Pobreza y pobreza extrema.

La pobreza extrema, es el nivel en el que se encuentran las personas que no alcanzan a cubrir el costo del consumo mínimo de alimentos. Y la pobreza total, son las personas que alcanzan a cubrir el costo del consumo mínimo de alimentos, pero no así, el costo mínimo adicional para otros bienes y servicios básicos (Instituto Nacional de Estadística, 2011).

Tabla 9. *Tasa de Pobreza Rural por nivel y Coeficiente de Gini en Quiché y municipios de intervención*

Lugar	Pobreza extrema rural	Pobreza total rural			Coeficiente de Gini rural
	Incidencia	Incidencia	Brecha	Severidad	
Depto. de Quiché	20.2%	76.9%	26.6%	12.2%	0.26%
Santa Cruz del Quiché	15.5%	65.2%	21.2%	9.0%	0.26%
Chichicastenango	28.4%	82.9%	31.5%	14.6%	0.24%
San Pedro Jocopilas	9.6%	55.4%	16.0%	63%	0.26%

(Instituto Nacional de Estadística, 2011)

La incidencia de la pobreza total y extrema corresponde al número de personas que se encuentran por debajo de la línea de pobreza total y extrema, respectivamente, como proporción de la población total. El Coeficiente de Gini es un índice que mide la desigualdad de una distribución, en este caso del consumo. Su valor varía entre cero y uno, donde cero corresponde a la perfecta igualdad y uno a la desigualdad total (Instituto Nacional de Estadística, 2011).

2. School the World

Es una organización sin fines de lucro, que trabaja para crear un cambio sostenible desde un nivel profundamente local para garantizar una educación de calidad y accesible para todas las niñas y los niños en las familias de las aldeas rurales de Guatemala y Honduras.¹

a. Historia.

En 2007 Kate Curran, fundadora de School the World, dejó su carrera como abogada para viajar por el mundo. Su llamada personal a la acción siguió a la repentina muerte de su hermano, padre y madre en menos de dos años. Las últimas palabras de su madre, “He tenido una gran vida” y el compromiso de toda la vida de su padre con el servicio público la inspiraron a comenzar un viaje para definir “una gran vida” que la llevó a través de los cuatro continentes. Pudo identificar que por medio de la educación se puede impactar en el cambio de vidas, especialmente de las niñas y los niños más vulnerables. Junto con los gobiernos locales, los socios y padres profundamente comprometidos brindan la oportunidad de aprender a las niñas y los niños que viven en las comunidades más pobres para que puedan trabajar hacia un futuro más brillante (School the World, 2019).

Kate también reconoció la oportunidad de construir un mundo interconectado de futuros líderes. A través del Programa de Aprendizaje de Servicio Estudiantil de School the World, los estudiantes de los Estados Unidos se comprometen a financiar la educación de una aldea y viajan para obtener una visión de primera mano del valor de sus esfuerzos y de los desafíos que aún quedan por delante. Algunos de los hitos más relevantes en School the World se presentan en la Tabla 10.

¹Para más información visitar: <http://schooltheworld.org/index.php>

Tabla 10. Hitos de School the World

Año	Hitos
2009	Reunión con el alcalde de Chichicastenango, Guatemala para acordar una asociación pública, que creó la base para el modelo efectivo de School the World para mejorar la educación. En los últimos 9 años, se han asociado con alcaldes en 75 escuelas.
2012	DOLE llevó el modelo a Honduras para llevar la educación a los vecindarios cerca de sus operaciones. Una nueva cultura y un nuevo clima, pero luchas muy similares.
2013	Presentación del Programa de Servicio de Aprendizaje, para traer a los jóvenes de EE. UU a conocer School the World. El primer viaje incluyó 19 estudiantes, en los últimos años se han superado los 600 World Changers.
2014	Se inició la construcción de los patios de recreo en las escuelas y se enfatizó el poder del juego.
2016	Lanzamiento del video Journey con Hilario y su caminata de dos horas y media a la escuela. Más de 70 millones lo han visto en las redes sociales y más de medio millón han compartido esta visión impactante de los desafíos de acceder a la educación en muchas partes del mundo.
2018	Aplicación de la Prueba de Lectura Inicial en 3° primaria.
2019	10mo. Aniversario de School the World

(School the World, 2019)

Desde septiembre de 2009, School the World ha construido 82 escuelas y 49 patios de recreo en dos países. Ha creado 401 bibliotecas en el aula, ha capacitado a 5,635 padres como "primeros educadores" y capacitado a 323 maestros y beneficiado a 9,582 niños (School the World, 2019). En el año 2019 School the World cumple 10 años de estar aportando a la educación. En la proyección a futuro se han propuesto llegar a 100 escuelas, ingresar a un tercer país y ampliar la programación a la educación de la primera infancia y la escuela secundaria inferior. Con la finalidad de asegurar que los infantes comiencen la escuela preparados para aprender y asegurar que completen su educación con habilidades fundamentales para desarrollarse en la vida (School the World, 2019).

Es una de las primeras ONG en exigir la asociación comunitaria y solicitar que el municipio contribuya con el 50% de los costos, debido a que la comunidad debe aportar la tierra y mano de obra no calificada. También requieren que los padres participen en programas de educación para padres. Es una estrategia que ha sido elogiada por USAID, los ministros de educación y los principales expertos en educación. Las operaciones del programa de educación son administradas completamente por personal local. Ya que se considera que ellos son los que están más familiarizados con los desafíos locales, las costumbres y los idiomas y pueden tener el mayor impacto, para cumplir con el objetivo de brindar acceso a una educación de calidad para quienes viven en las zonas rurales de América Central.

b. Programas de intervención educativa.

Desde la primera infancia hasta la adolescencia, los programas brindan educación de calidad a las niñas y los niños que viven en la pobreza extrema. School the World se asocia con líderes locales para invertir en un futuro mejor, una aldea a la vez. Construir escuelas es solo el comienzo, crear un suministro sostenible de libros, capacitar a los maestros y comprometer profundamente a los padres es parte de una estrategia detallada para fomentar el sentido de pertenencia, el orgullo y el valor de la educación (School the World, 2019).

El Programa de Escuelas Comunitarias muestra lo que puede suceder cuando una comunidad entera se organiza en torno al poder de la educación. Se comienza negociando con los municipios locales para proporcionar el 50% de los nuevos costos de infraestructura. El 50% restante es propiciado por los patrocinadores corporativos, donantes individuales o World Changers que participan en el Programa de Ciudadanía Global. Una vez que el Ministerio de Educación aprueba una nueva escuela comunitaria y garantiza suficientes maestros, se hace una reunión con los líderes comunitarios y los padres. Se les consulta si están dispuestos a proporcionar la tierra y el trabajo, asistir a programas bimensuales de capacitación de padres y contribuir con fondos para los libros. Cuando los líderes y los padres están de acuerdo, comienza la asociación (School the World, 2019).

1) Nueva escuela.

Una escuela de tres aulas es el primer paso. A menudo reemplaza estructuras improvisadas o estructuras de adobe en peligro de colapso. La escuela simboliza mucho más que el edificio físico, es un símbolo de la asociación entre School the World, la comunidad y el gobierno local. Es un símbolo de progreso, proporcionando una razón para la esperanza. Cada salón de clases está equipado con una pizarra blanca, un escritorio para el maestro, estantes para una biblioteca llena de libros, escritorios con el tamaño correcto para las niñas y los niños y murales pintados a mano que inspirarán a la niñez a soñar más allá del aula (School the World, 2019).

2) Empoderamiento de los padres.

La mayoría de los padres con los que trabajan son analfabetos, pero quieren que sus hijos tengan la educación y la oportunidad que nunca tuvieron. A través de capacitaciones bimensuales para padres, se les da la posibilidad a los padres de convertirse en los primeros educadores de sus hijos. Aprenden cómo apoyar en el aprendizaje de sus hijos dentro y fuera del aula. Aprenden sobre sus derechos, los derechos de sus hijos y las responsabilidades del maestro (School the World, 2019).

3) Formación de profesores.

Los docentes en áreas rurales golpeadas por la pobreza extrema comienzan su carrera docente con una capacitación mínima. A menudo se encuentran en entornos de enseñanza muy difíciles con pocos suministros y un apoyo limitado del gobierno. Incluso solo el llegar y salir de la escuela puede ser un desafío diario. Por medio de capacitaciones grupales mensuales y capacitación individual en el aula se brinda a los maestros dos años de capacitación intensiva. El plan de estudios para los grados de 1ro. a 3ro. se enfoca en las destrezas de lectura y matemáticas en los grados iniciales. Y el plan de estudios para los grados de 4to. a 6to. se centra en las técnicas de gestión de aula que mejoran el aprendizaje. Ambos currículos incluyen componentes para el aprendizaje social, emocional y habilidades básicas para la vida. También se construye una cultura de orgullo profesional, influencia en los compañeros y motivación, por medio del reconocimiento de los maestros que dominan las mejores prácticas a través de premios especiales (School the World, 2019).

4) Poder del juego.

La mayoría de los niños matriculados en las escuelas van a trabajar antes o después de la escuela para ayudar a mantener a sus familias. Por lo que es necesario proporcionar a las niñas y los niños un lugar seguro para descubrir, imaginar y explorar lo cual es un aspecto vital en su desarrollo. Por medio de la construcción de patios de recreo, y la integración de

juegos organizados y actividades de aprendizaje al aire libre para desarrollan el aprendizaje social, emocional, las habilidades cognitivas, aumentar la autoconfianza y la igualdad de género (School the World, 2019).

5) Libros y suministros.

A través de la asociación con los padres en cada comunidad y generosos donantes, se logra almacenar y reabastecer las bibliotecas en las aulas con hermosos libros que inspiran a los niños a soñar más allá del aula. Los niños se emocionan tanto al recibir libros nuevos que incluso esperan fuera de su salón de clases, corriendo hacia el camión de reparto cuando lo ven acercarse a su escuela (School the World, 2019).

6) Oportunidad básica.

Al considerar que solo un pequeño porcentaje de niños en zonas rurales de Centroamérica pueden continuar su educación después del sexto grado, School the World se ha comprometido a eliminar los obstáculos para dar a cada uno de sus estudiantes la oportunidad de continuar su educación a través de la escuela secundaria de educación básica (School the World, 2019).

7) Solo lo básico.

Incluso cuando los estudiantes tienen una escuela secundaria en su comunidad o a poca distancia a pie, a menudo no pueden pagar un uniforme, los útiles escolares y las asignaciones de tareas requeridas en ese nivel. Es por ello que desde School the World se ofrecen becas para cubrir las necesidades básicas para cada estudiante que lo necesite (School the World, 2019).

8) Modelos alternativos de aprendizaje.

Donde no hay escuelas cercanas, School the World se asegura de que cada estudiante tenga acceso a al menos un modelo alternativo. En las ubicaciones centrales, aseguran un maestro para un programa nuevo e innovador que combina el estudio independiente con la enseñanza y la evaluación un día a la semana. En las comunidades más remotas, apoyan para el acceso a programas de estudio independientes basados en radio (School the World, 2019).

9) Becas para niñas.

Enfatizan la equidad de género y la educación para las niñas a través de las capacitaciones para padres en la escuela primaria y al proporcionar becas completas para las niñas que muestran pasión por la educación y el aprendizaje para asistir a una escuela secundaria tradicional de tiempo completo. Para eliminar todas las barreras, se cubre el transporte, los útiles escolares y los suministros mensuales (School the World, 2019).

c. Escuelas evaluadas en Quiché.

Las nueve escuelas seleccionadas para aplicar la Prueba de lectura inicial actualmente también participan en los programas de intervención educativa de School the World.

- 1) Escuela Oficial Rural Mixta Agua Viva
- 2) Escuela Oficial Rural Mixta Chocruz
- 3) Escuela Oficial Rural Mixta Chulumal Segundo Jornada Matutina
- 4) Escuela Oficial Rural Mixta Chulumal Segundo Jornada Vespertina
- 5) Escuela Oficial Rural Mixta Chuguexa Cuarto
- 6) Escuela Oficial Rural Mixta Mactzul Séptimo
- 7) Escuela Oficial Rural Mixta Panimaché Quinto Alto
- 8) Escuela Oficial Rural Mixta Río Mactzul
- 9) Escuela Oficial Rural Mixta Xalbajquiej
- 10) Escuela Oficial Rural Mixta Xesic I

d. Caracterización de las escuelas.

Las deficiencias en la calidad educativa, relacionadas con la preparación deficiente de los maestros y los recursos insuficientes, llevan al 76 por ciento de todas las niñas y los niños en áreas rurales que ingresan al primer grado a abandonar la escuela antes de terminar la escuela primaria. Esta situación se ve agravada por las malas condiciones de salud, la desnutrición infantil, y trabajo infantil, ya que muchas de las niñas y los niños pobres deben contribuir a los ingresos de su familia (USAID, 2006). Este es el contexto en el cual se encuentran las escuelas con las que School the World trabaja programas de intervención educativa.

Para conocer más detalles (distancia, tiempo de intervención, proyecto de infraestructura, intervenciones y el estado actual) sobre las escuelas que aceptaron participar en la aplicación de la Prueba de lectura inicial ver el Anexo 3. A continuación en la Tabla 11 se presenta la caracterización de la comunidad educativa de las 9 escuelas que aceptaron participar en la aplicación de la Prueba de lectura inicial de STW, por medio de la información específica de cada escuela se pueden identificar los logros y desafíos en cada comunidad.

Tabla 11. *Caracterización de la comunidad educativa por cada escuela*

Escuelas	Población estudiantil	Equipo docente	Comunidad educativa	Características
EORM Agua Viva	Las niñas y los niños son participativos. Por las tardes ayudan a sus padres en el cultivo de papas. Es una comunidad muy pobre.	La mayoría de los docentes son bilingües, el docente de 3° se ha formado en el PADEP. Hay nuevos docentes y se rotan en los grados. (Programa de Académico de desarrollo Profesional Docente, - PADEP-, implementado por el MINEDUC)	Los padres de familia participaron activamente en la intervención inicial. La mayoría de los padres de familia tienen baja escolaridad o son analfabetos, sin embargo, apoyan a sus hijos.	Cuentan con libros de lectura, se realizan lecturas con los estudiantes. Hay nivel de Preprimaria.

Escuelas	Población estudiantil	Equipo docente	Comunidad educativa	Características
EORM Chocruz	A las niñas y los niños les gusta leer. Son amigables y cariñosos.	La mayoría de los docentes cuentan con formación a nivel de diversificado. El director tiene formación en Pedagogía y un docente cuenta con formación en PEM.	Los padres de familia llegan pocas veces a la escuela. Brindan apoyo desde casa.	El director exige a los docentes que cumplan con el horario de clases. Hay nivel de Preprimaria desde el 2008.
EORM Chuguxa	Las niñas y los niños se comunican con facilidad. Hay 110 estudiantes. Es una comunidad con escasos recursos y niños desnutridos. El acceso a la escuela es difícil.	La directora es maestra de 3° primaria la cual es un aula multigrado. Hay tres maestras quienes tienen dificultad para asistir todos los días.	El día de la aplicación no había clases y por medio de los padres se logró convocar a los estudiantes. La comunidad muestra disposición para gestionar otros proyectos con STW.	Cada aula cuenta con libros de lectura. Se reabrió el nivel de Preprimaria en el 2018. En junio 2018 se firmó el convenio de cierre del proyecto.
EORM Chulumal Jornada Matutina	Las niñas y los niños son colaboradores. Viven en una comunidad pobre, sin embargo, por su proximidad tienen más acceso a Chichicastenango.	Los docentes están estudiando en el PADEP. La mitad del equipo docente es bilingüe. La escuela se ha dividido en dos jornadas. Cuenta con 23 docentes quienes fueron capacitados por UVG. Los docentes cumplen con su horario de clases y receso.	La comunidad realiza visitas en la escuela. Es evidente el apoyo de los padres en las tareas de los estudiantes. En la comunidad hay varios profesionales.	En la escuela hay biblioteca. Hay oportunidad para que los padres de familia se involucren en los talleres. Hay dos secciones en el nivel de Preprimaria desde el 2009.
EORM Chulumal Jornada Vespertina	Las niñas y los niños participan motivadamente en las actividades educativas.	La docente de 3ro. Primaria tiene 12 años de trabajar en la escuela. Tiene experiencia para motivar a los estudiantes a aprender. Todos los docentes participaron en las capacitaciones de la UVG.	A los padres de familia se les dificulta participar en los talleres de formación.	Cada aula cuenta con libros de lectura.
EORM Mactzul VII	Las niñas y los niños son de familias de escasos recursos con cuadros de desnutrición. Hay participación diaria en las clases.	Hay rotación de los docentes en los diferentes grados. Han participado en las capacitaciones de la UVG sin embargo por el horario ya no han asistido.	La mayoría de los padres de familia tienen baja escolaridad o son analfabetas. El COCODE, Consejo Educativo y los padres de familia participan activamente.	En las aulas hay libros de lectura los cuales se trabajan en el horario de lectura diaria. La maestra de 6° utiliza la estrategia de parafraseo y la elaboración de gráficos a partir de la lectura. Hay nivel de Preprimaria desde el 2005.

Escuelas	Población estudiantil	Equipo docente	Comunidad educativa	Características
EORM Panimaché Quinto	Es una comunidad extremadamente pobre.	Una docente y el director atiende aula multigrado (1° a 3° y de 4° a 6°). Los docentes se han esforzado en ser constantes.	Hay participación activa de los padres, quienes en su mayoría son analfabetos o con baja escolaridad.	Las dos aulas cuentan con libros de lectura. No hay nivel de Preprimaria, está en gestión.
EORM Río Mactzul	Las niñas y niños son participativos, en su mayoría son niños. La población es de escasos recursos y se manifiesta el machismo.	El docente de 3° primaria ha gestionado proyectos para mejoras en la escuela. Los docentes participaron en las capacitaciones a maestros por la UVG, sin embargo, la mayoría ya no labora en la escuela.	Algunos padres y madres de familia ayudan a sus hijos en las tareas de la escuela y visitas escolares.	Cuentan con biblioteca, hora de lectura asignada y tres docentes son licenciados en Pedagogía. Hay nivel de Preprimaria desde el 2010.
EORM Xalbaquej	Se identifica que hay estudiantes desnutridos. Por las tardes ayudan a sus padres y hermanos a sacar arena, piedrín y piedra del río.	El docente de 3ro. Primaria trabaja desde 1998 en diferentes grados. Los otros docentes cuentan con formación en el PADEP y PEM. Todos los docentes son bilingües. Han participado en la capacitación a docentes que da UVG.	Los padres de familia colaboran con los docentes y se involucran para apoyar en las tareas de los hijos.	Los docentes cuentan con un horario de lectura diario en el que motivan a los estudiantes a leer. Hay nivel de Preprimaria desde el 2006.
EORM Xesic	Las niñas y los niños participan activamente en las actividades educativas. La mayoría de las familias son de escasos recursos y con cuadros de desnutrición.	El docente de 3° ha trabajado durante dos años en el mismo grado. Se caracteriza por ser dinámico y motiva a los estudiantes a ser responsables.	El COCODE, el alcalde comunitario, la directora y el Consejo Educativo se han esforzado por convocar a los padres de familia, sin embargo, la mayoría muestra poco interés. Los docentes y padres de familia se organizaron para hacer un horario para el uso correcto de los juegos.	Aun no cuentan con libros de lectura. La escuela cuenta con un director técnico quien no tiene a su cargo un grado en específico. La escuela no cuenta con agua para consumo diario. Hay nivel de Preprimaria desde el 2004.

(López, 2019)

D. Marco normativo-jurídico

1. Ley de Educación Nacional de Guatemala

En el capítulo I define las características, estructura, integración y funciones del Sistema en el:

- a. Artículo 3. Define que el Sistema Educativo Nacional es el conjunto ordenado e interrelacionado de elementos, procesos y sujetos a través de los cuales se desarrolla la acción educativa, de acuerdo con las características, necesidades e intereses de la realidad histórica, económica y cultural guatemalteca.
- b. Artículo 7. La función fundamental del Sistema Educativo es investigar, planificar, organizar, dirigir, ejecutar y evaluar el proceso educativo a nivel nacional en sus diferentes modalidades.

Y en el capítulo VI relacionado al planteamiento y la evaluación en el capítulo único en el:

- a. Artículo 69. La evaluación es un proceso inherente a la acción educativa y debe realizarse en forma pedagógica y permanente, a fin de determinar los logros cuantitativos y cualitativos de la educación en función de sus fines y principios.
- b. Artículo 70. La evaluación del sistema Educativo Nacional deberá efectuarse permanentemente de conformidad con el reglamento de esta ley.
- c. Artículo 71. La evaluación del rendimiento escolar debe realizarse solamente en períodos y sistemas con carácter obligatorio y permanente, en base a las necesidades socioeducativas del país y estará regulada por la reglamentación específica (Ley de Educación Nacional, 1991).

2. Acuerdo Gubernativo 225-2008

Reglamento Orgánico Interno del Ministerio de Educación en el:

a. Artículo 8. Indica que la Dirección General de Evaluación e Investigación Educativa - DIGEDUCA- es la dependencia del Ministerio de Educación responsable de evaluar el desempeño de la población estudiantil, los subsistemas escolar y extraescolar, así como divulgar la información. Dentro de sus funciones se encuentran:

1) Evaluar el Sistema Educativo Nacional y los programas de Ministerio de Educación que lo requieran, mediante la medición del desempeño de los estudiantes, de acuerdo con los estándares educativos.

2) Analizar las evaluaciones que permitan establecer la dirección y magnitud de los cambios del sistema, en términos del desempeño de los estudiantes, y que proporcionen insumos con los cuales se puedan desarrollar estrategias para mejorar el desempeño de los actores educativos.

3) Establecer los lineamientos, la reglamentación y la implementación de mecanismos para evaluar las competencias y el desempeño de docentes y directores, en las áreas curriculares y de competencias lingüísticas.

4) Desarrollar investigaciones en variables directa e indirectamente asociables con el aprendizaje (factores asociados).

5) Ejecutar evaluaciones diagnósticas del recurso humano docente
(Mineduc, 2008).

3. Acuerdo Ministerial 1171-2010

Reglamento de Evaluación de los Aprendizajes para los niveles de Educación Preprimaria, Primaria y Media de los subsistemas de educación escolar y extraescolar en todas sus modalidades en el Capítulo XV relacionado a la Evaluación Externa, en

a. Artículo 46.

Define que la evaluación externa es la que se realiza por la dependencia especializada del Ministerio de Educación, designada para el efecto y ajena al centro educativo, por medio de instrumentos de evaluación que cumplen con criterios técnicos de calidad. Estos instrumentos son administrados y calificados con lineamientos y condiciones específicas y no es vinculante con la promoción de las y los estudiantes.

b. Artículo 47.

El propósito de la evaluación externa es obtener información oportuna y pertinente que:

- 1) Permita establecer el nivel de calidad del Sistema Educativo Nacional.
- 2) Permita detectar áreas débiles del Sistema Educativo Nacional.
- 3) Facilite la toma de decisiones con respecto al proceso de enseñanza y aprendizaje.
- 4) Brinde elementos que conformen un marco de referencia para la mejora continua del Sistema Educativo Nacional.
- 5) Incida en la toma de decisiones para la formulación de políticas educativas (Mineduc, 2010).

III. MARCO METODOLÓGICO

A. Alcances del estudio

El alcance es exploratorio de carácter transversal, ya que la investigación se centra en analizar cuál es la confiabilidad y validez de la Prueba de lectura inicial aplicada una vez en 3ro. primaria en 9 escuelas del departamento de Quiché (Hernández, R, 2014).

De acuerdo con Hernández (2014) el propósito de la investigación con alcance exploratorio es examinar un tema de investigación poco estudiado, del cual se tienen muchas dudas o no se ha abordado antes. Su valor permite que se familiarice con fenómenos desconocidos al obtener información para realizar la investigación más completa en un contexto particular y establecer prioridades para investigaciones futuras.

Principalmente en esta investigación se reconoce que a nivel nacional son escasas las instituciones que se dedican a realizar este tipo de análisis, debido a que se requiere de tiempo y expertos que puedan interpretar la información recopilada. Por medio del diseño de la muestra y la elección de los sujetos, la cual fue representativa de las unidades de análisis donde se implementan los programas de STW. El alcance y generalización de las conclusiones de esta investigación se remite a las unidades de análisis identificadas por STW.

Al determinar las propiedades psicométricas de la Prueba de lectura inicial se podrá identificar si la prueba mide los constructos de forma confiable y válida, lo cual es un alcance que permitirá a STW retroalimentar el proceso de evaluación y tomar decisiones para las siguientes aplicaciones de la prueba de lectura, en base a la evidencia que se aportará. Es necesario considerar que los resultados no se pueden generalizar, ya que son específicos para el contexto en el que fueron aplicadas las pruebas de lectura inicial en las escuelas ubicadas en los tres municipios del departamento de Quiché.

B. Objetivos de la investigación

1. Objetivo general

Analizar las propiedades psicométricas, confiabilidad y validez de la Prueba de lectura inicial, aplicada en 3ro. Primaria, en nueve escuelas del departamento de Quiché.

2. Objetivos específicos

a. Indagar sobre los procedimientos que se emplearon para adaptar la Prueba de lectura inicial, la cual fue aplicada en nueve escuelas de intervención por STW.

b. Indagar sobre el proceso de campo que se realizó para aplicar la Prueba de lectura inicial en 3ro. Primaria, en las nueve escuelas de intervención por STW.

c. Describir los cinco componentes que conforman la Prueba de lectura inicial, aplicada en 3ro. Primaria, en las nueve escuelas de intervención por STW.

d. Analizar las propiedades psicométricas de la Prueba de Lectura Inicial.

e. Identificar el nivel de confiabilidad (consistencia interna) y validez de la Prueba de lectura inicial.

C. Participantes

La unidad de análisis está conformada por las nueve escuelas seleccionadas por conveniencia por el equipo de School The World, de las cuales 7 escuelas (*Panimaché Quinto Alto, Río Mactzul, Mactzul Séptimo, Chuguexa Cuarto, Agua Viva, Chulumal Segundo y Xalbaquiej*) son de post intervención, ya que han participado durante 5 años en las intervenciones de STW a través de talleres para padres, capacitación a docentes y donación de libros o se encuentran en un período de monitoreo. Y de las tres escuelas recién construidas o que participaban por primera vez, dos aceptaron participar (*Chocruz y Xesic*

I) y debido a que una de ellas no aceptó, se coordinó para aplicar la prueba en la escuela *Chulumal Segundo* en la jornada vespertina y matutina (Joya, Us, & López, 2019).

Para la aplicación de la Prueba de lectura inicial se obtuvo una muestra conformada por 60 estudiantes, niñas y niños entre 8 a 14 años quienes estudiaron en 3ro. Primaria en el año 2018, en alguna de las nueve escuelas que se ubican en el departamento de Quiché, las cuales representan un subgrupo de la población total de escuelas en donde tiene intervención STW. La selección de los estudiantes a quienes se les aplicó la prueba fue aleatoria, elegidos por sorteo (3 hombres y 3 mujeres). Todos los estudiantes estudian en 3ro. primaria y cada uno de ellos aceptó participar desde el inicio al final de la Prueba de lectura inicial. Dentro de los datos generales que los estudiantes compartieron, fueron su edad, sexo e idiomas que hablan.

D. Escenario

1. Adaptación de la prueba de lectura

El equipo de STW participó en un taller impartido por Visión Mundial, en el cual el equipo compartió la prueba de lectura STAR y la alternativa para evaluar las habilidades Lectoras de los estudiantes en los grados iniciales de la primaria, por medio de la plataforma Tangerine en la cual se registran las respuestas de cada aplicación y a su vez se genera una base de datos en Excel, con la información recopilada.

A partir de la exploración y autorización del equipo de Visión Mundial, el técnico de proyectos de STW creó su usuario y adaptó la prueba de lectura STAR en la plataforma de Tangerine, desde la cual se hicieron varias pruebas piloto en diferentes escuelas y grados que participan en los programas de intervención de STW. La Prueba de lectura inicial de STW incluye los mismos ítems y Subpruebas. Se omitió la lectura en idioma Ixil, el cual se especificaba en el documento: Literacy Bost instrument Guatemala_ESP elaborado por Visión Mundial. De igual forma se omitieron las partes III, IV, V y VI en las cuales se realizan preguntas sobre el contexto escolar.

Se hicieron aplicaciones piloto con estudiantes de 1ero. Primaria y se observó que muchos de los estudiantes aún estaban en el proceso de adquisición de la lectura y

reconocían muy pocas letras. Debido a ello se decidió aplicar en 2do. grado y la mayoría de los estudiantes no logró culminar la prueba. Así que se decidió implementar una aplicación en el grado de 3ro. primaria y los datos recopilados son los que permitieron determinar las propiedades psicométricas de la Prueba de Lectura Inicial adaptada por STW. La prueba de lectura STAR desarrollada por Visión Mundial consta de doce partes las cuales son:

- a. Parte I: Identificación de la información específica del estudiante.

- b. Parte II: Introducción y presentación. Solicitud del consentimiento del estudiante para ser evaluado.

- c. Parte III: Información sobre el récord académico del estudiante.

- d. Parte IV: Información sobre la cultura Lectora del estudiante.

- e. Parte V: Prácticas de los maestros en el aula.

- f. Parte VI: Participación en actividades de lectura después de la escuela.

- g. Parte VII: Lectura de letras minúsculas.

- h. Parte VIII: Lectura de palabras más usadas.

- i. Parte IX: Decodificación de palabras.

- j. Parte X: Lectura de una historia con tiempo límite.

- k. Parte XI: Registro de palabras leídas o no leídas.

1. Parte XII: Preguntas de comprensión sobre la historia leída (World Vision International, 2017).

La Prueba de lectura inicial que fue adaptada por el equipo de STW consta de las siguientes partes:

a. Parte I: Introducción y presentación. Solicitud del consentimiento del estudiante para ser evaluado.

b. Parte II: Fecha y hora de aplicación.

c. Parte III: Nombre del evaluador.

d. Parte IV: Información del estudiante. (Edad, si estudia, nombre de la escuela, grado, nombre y apellido, sexo e idioma que habla en su casa)

e. Parte IV: Alfabeto en español. (27 letras)

f. Parte V: Palabras simples en español. (20 palabras)

g. Parte VI: Palabras inventadas. (20 palabras)

h. Parte VII: Lectura oral de un pasaje. (129 palabras)

i. Parte VIII: Comprensión lectora. (10 preguntas) (School the World, 2019).

2. Selección de la muestra

Para la selección de la muestra se contó con el asesoramiento del equipo de la Facultad de Educación de la Universidad del Valle de Guatemala (UVG), quien por medio de comunicaciones electrónicas con el equipo de STW, propuso que se evaluaran 6 estudiantes en cada una de las diez escuelas que se seleccionaron por conveniencia siete

escuelas con post-intervención y tres escuelas que se integran por primera vez (Joya, Us, & López, 2019). Se agregó que durante la aplicación no se tomó en cuenta una escuela, debido a que no aceptaron participar en la aplicación y por ello se aplicó la Prueba de lectura inicial en una misma escuela en sus dos jornadas (Facultad de Educación, 2018).

La elección de los estudiantes que participaron en la aplicación estuvo a cargo de cada uno de los aplicadores, de acuerdo con el técnico de STW durante la visita al salón de clases los aplicadores seleccionaron al azar a tres niñas y tres niños.

3. Aplicación de la Prueba de lectura inicial

a. Previo a la aplicación.

Se realizó la comunicación con las autoridades de la Facultad de Educación de la UVG en el Campus Altiplano para que asignaran a los estudiantes que apoyarían en la aplicación de la Prueba de lectura inicial. El encargado de los proyectos técnicos junto con ayuda del equipo de STW se encargó de organizar el taller para los aplicadores, el cual se realizó en las instalaciones de la Universidad del Valle en el Campus Altiplano y tuvo una duración de 3 horas aproximadamente.

El propósito del taller fue dar la inducción necesaria a los aplicadores, para que conocieran e identificaran la prueba y el proceso de aplicación que realizarían. Entre las actividades que se realizaron fueron: asignar una Tablet (las cuales la Universidad prestó) a cada evaluador, leer cada una de las instrucciones de la prueba, acceder al usuario en Tangerine y aplicar la prueba para resolver las dudas de los aplicadores, previo a aplicar en las escuelas seleccionadas. Debido a que no todos los aplicadores viven cerca del lugar en donde se les convocó, algunos estuvieron ausentes en el taller.

Antes de la aplicación se contactó a cada una de las autoridades de las escuelas para notificarles el día y la hora en que los aplicadores los visitarían. Se les solicitó su apoyo para que dispusieran de un espacio adecuado en el cual se pudieran aplicar las pruebas a los estudiantes de 3ro. primaria.

En total fueron 19 aplicadores quienes apoyaron. Por medio de un correo electrónico se les notificó la escuela en la que evaluarían, se consideró que fueran áreas geográficas cercanas a su residencia.

b. Durante la aplicación.

Se convocó a los aplicadores para que el miércoles 10 de octubre 2018 llegaran a las escuelas asignadas, en algunos casos se reunieron en un punto para luego dirigirse al establecimiento. La aplicación de las pruebas se realizó durante el horario regular de clases. En cada escuela se asignaron dos aplicadores quienes eran responsables de llevar el material asignado, que consistía en: la Tablet y las hojas de lectura. En cada escuela estuvo un técnico del equipo de STW quien apoyó en la aplicación de la prueba. El protocolo para seguir en cada escuela fue el siguiente:

- 1) Saludo a las autoridades de la escuela y al docente de 3ro. Primaria.
- 2) Ubicación del espacio para evaluar (preferiblemente un área cómoda y libre de ruido).
- 3) Saludo y presentación con los estudiantes de 3ro. Primaria.
- 4) Elección de los estudiantes a evaluar (fue el azar tres hombres y tres mujeres).
- 5) Evaluación individual de dos estudiantes por aplicador.

Ya que cada uno necesitó una hora aproximadamente con cada estudiante. El evaluador siguió la lectura desde la Tablet en donde marcó únicamente los errores y los ítems que el estudiante no leyó. (Durante el recreo se continuó con la aplicación de la prueba).

- 6) Al culminar se dejó el espacio en orden y se agradeció al docente y autoridades por el apoyo.

c. Después de la aplicación.

Cada aplicador se encargó de entregar personalmente el material asignado. Y se agradeció a cada uno por su apoyo. Después de tener todo el material, se revisó que los datos aparecieran guardados en la base de Tangerine, ya que cada dispositivo estaba sincronizado con el correo del encargado de la aplicación. El técnico se encargó de descargar la base de datos para identificar las respuestas obtenidas y verificar que los datos estuvieran completos. Y desde la Facultad de Educación de la Universidad del Valle en el Campus Central, se apoyó con la elaboración del informe de los resultados de la aplicación de la Prueba de lectura inicial en 3ro. primaria.

E. Instrumentos de recolección de información

Para explorar la Prueba de lectura inicial y conocer el proceso de adaptación y aplicación que se realizó por el equipo de STW, se tuvo la autorización de las autoridades de School the World, así como el apoyo del equipo técnico de School the World conformado en su mayoría por profesionales jóvenes, de Guatemala y Honduras. Por medio de Skype y el correo electrónico el equipo estuvo dispuesto a aclarar las inquietudes acerca del proceso y compartieron la información que corresponde a los siguientes documentos:

1. Base de datos en Excel. La cual incluye todos los datos y respuestas que se registraron directamente en la plataforma Tangerine, durante la Prueba de lectura inicial.
2. Informe para STW. El cual fue elaborado por el equipo de la Facultad de Educación y socializado con el equipo de STW, a finales del 2018.
3. Prueba de lectura STAR. Documento de Word: Literacy Boost instrument Guatemala_ESP el cual es la plantilla que Visión Mundial utilizó para aplicar la prueba de lectura STAR y la misma que STW agregó en plataforma Tangerine, en donde creó la Prueba de lectura inicial.
4. Prueba de lectura inicial. Al tener acceso al usuario de School the World, se pudo explorar la Prueba de lectura inicial, la cual fue adaptada por STW y diseñada en la

plataforma Tangerine, desde la cual los aplicadores ingresaron las respuestas de los estudiantes. Tangerine es un software de recopilación de datos electrónicos de código abierto diseñado para su uso en Android dispositivos móviles. Su principal uso es permitir el registro de las respuestas de los estudiantes en los grados iniciales, en las evaluaciones de habilidades de lectura y matemáticas. Tangerine también se utiliza para capturar las respuestas al entrevistar estudiantes, profesores y directores. Tangerine fue desarrollado por RTI International con fondos de RTI y Google.org (s.n, 2019).

En 2009, el software de recopilación de datos electrónicos Tangerine fue concebido como una solución para abordar algunos de los desafíos encontrados en la administración basada en las pruebas de papel, particularmente para grandes empresas nacionales que requieren mucha mano de obra. RTI desarrolló el software durante el período de septiembre de 2010 a septiembre de 2011 como un código abierto. Como una encuesta basada en navegador (pero fuera de línea) utilizando la tecnología de código abierto de vanguardia optimizada para Tablet de pantalla táctil, permite a Tangerine compatibilidad con una amplia gama de dispositivos móviles, idiomas y escritura de scripts. Una característica crítica es que permite la sincronización de datos de varios asesores, trabajando simultáneamente sin conexión en diferentes dispositivos, con una base de datos central, sobre una variedad de tipos de conexión (RTI, International, 2012).

Las ventajas en el uso de Tangerine es que se tiene disponibilidad rápida de los datos de evaluación; se mejora la calidad y la consistencia de los datos y hay menos errores de medición debido a campos faltantes, errores de transcripción de datos, tipos o formatos de datos no válidos, datos ilegibles o incomprensibles, etc.; se reduce la cantidad de papel y suministros, así como una logística más manejable requerida para preparar y gestionar las evaluaciones en papel (fotocopiado, clasificación, grapado, embalaje, etc.); hay posibilidad de acceder a información y medios adicionales dentro del dispositivo portátil, es decir, audio, archivos, video tutoriales, manuales de instrucciones, sistema de posicionamiento global (GPS), etc (RTI, International, 2012).²

La prueba aplicada está conformada por 206 ítems los cuales conforman las siguientes Subpruebas que se presentan en la Tabla 12.

² <http://www.tangerinecentral.org/demo/>

Tabla 12. *Subpruebas de la prueba STAR y de lectura inicial de STW*

Subprueba	Habilidad de lectura	Habilidad demostrada por...
Identificación de letras/Alfabeto	Conocimiento del alfabeto	Proporcionar nombre/sonido de las letras
Palabras más usadas/Palabras Simples	Reconocimiento de palabras	Lectura de palabras extraídas de una lista de palabras comunes utilizadas en el libro de texto de tercer grado
Palabras decodificadas/Palabras Inventadas	Decodificación	Indicar sonidos de letras leyendo palabras sin sentido
Lectura de la historia/Lectura Oral	Habilidad de Lectura Oral	Leyendo un texto de con precisión y con una velocidad razonable durante un minuto.
Comprensión Lectora	Comprensión de lectura	Respondiendo correctamente a preguntas de recuperación de hechos sobre la historia leída.

(World Vision International, 2017; School the World, 2019)

a. Alfabeto.

En esta Subprueba el aplicador le muestra al estudiante la plantilla con 27 letras y le solicita que las lea en español y en voz alta, según el orden (de izquierda a derecha) y que señale con su dedo la letra mientras la va leyendo. El estudiante puede decir el nombre o el sonido de cada una de las letras (mayúsculas y minúsculas) las cuales no se encuentran en orden alfabético. Si el estudiante duda de las letras por más de 5 segundos se le debe de preguntar ¿si sabe su nombre, el sonido que tiene o si conoce cualquier palabra que empiece con esa letra? y se le solicita que nombre las que identifique, aunque no las diga en orden. En esta Subprueba no hay límite de tiempo. La regla de suspensión indica que si el estudiante duda por más de 10 segundos más, se le agradece por su esfuerzo y se marcan como incorrectas las letras que no reconoció o no intentó. Y se debe de continuar con el siguiente estudiante, ya que las siguientes Subpruebas se suspenden debido a que son de mayor dificultad.

b. Palabras simples.

En esta Subprueba el aplicador le muestra al estudiante la plantilla con 20 palabras simples y le solicita que las lea en español y en voz alta, según el orden (de izquierda a derecha). Se califican como correctas las palabras sin importar el orden en que las lea. Adicionalmente si el estudiante la lee mal y luego se autocorrije se califica como correcta.

Si el estudiante duda por 5 segundos se le pregunta si ¿Hay alguna palabra en la lista que pueda leer? Y se solicita que la señale y las lea. Se debe repetir la pregunta y animar al estudiante a continuar. En esta Subprueba no hay límite de tiempo. La regla de suspensión indica que si el estudiante duda por más de 10 segundos se debe agradecer el esfuerzo que hizo. Y se deben de marcar como incorrectas las palabras que el estudiante no identificó o no intentó leer. Y se debe de continuar con el siguiente estudiante, ya que las siguientes Subpruebas se suspenden debido a que son de mayor dificultad.

c. Palabras inventadas.

En esta Subprueba el aplicador le indica al estudiante que leerá nuevamente otras palabras. Le indica que son palabras que no tienen ningún significado. Antes de comenzar, se practica la lectura de dos palabras inventadas. Durante la práctica se puede corregir al estudiante y se le anima a dar lo mejor. Se le entrega la plantilla con 20 palabras inventadas y se le solicita que las lea en español y en voz alta, según el orden (de izquierda a derecha). Se califican como correctas las palabras sin importar el orden en que las lea. Adicionalmente si el estudiante la lee mal y luego se autocorrige se califica como correcta. Si el estudiante duda por 5 segundos se le pregunta si ¿hay alguna palabra en la lista que pueda leer? Y se solicita que la señale y las lea. Se debe repetir la pregunta y animar al estudiante a continuar. En esta Subprueba no hay límite de tiempo. La regla de suspensión indica que si el estudiante duda por más de 10 segundos se debe agradecer el esfuerzo que hizo. Y se deben de marcar como incorrectas las palabras que el estudiante no identificó o no intentó leer. Y se debe de continuar con la siguiente Subprueba.

d. Lectura oral.

En esta Subprueba el aplicador le indica al estudiante que leerá una historia, durante un minuto y al finalizar se le realizarán algunas preguntas sobre lo que leerá. Se le entrega la plantilla con la historia que contiene 129 palabras y se le solicita que la lea en español y en voz alta, según el orden (de izquierda a derecha). Se le anima a que lea cada palabra sin saltarse ninguna y si llega a alguna palabra que no sepa leer se le indica que el aplicador se

la puede decir. Se le anima a que lea lo mejor que pueda y se le pregunta si tiene alguna pregunta sobre lo que debe realizar.

Se comienza a contar el tiempo cuando el estudiante empiece a leer la primera palabra del título. Si el estudiante no puede leer la primera palabra después de unos cuantos segundos se debe marcar como incorrecta y se lee la palabra. Debe dejarse que el estudiante continúe leyendo. Si el estudiante no puede leer correctamente por lo menos 5 palabras en los primeros 30 segundos se le debe de pedir que se detenga y se le agradece por su esfuerzo. Si lee mal alguna palabra y luego se autocorrigió se califica como correcta. Si el estudiante lee al menos 5 palabras correctas en 30 segundos, se deja que continúe leyendo y se marca hasta la palabra que llegó a los 60 segundos y se deja que termine de leer el texto. Si el estudiante termina de leer la historia antes de un minuto se debe marcar el tiempo exacto. Al finalizar se le agradece al estudiante.

e. Comprensión lectora.

En esta Subprueba el aplicador le indica al estudiante que le hará unas preguntas sobre la historia que acaba de leer en voz alta. Y le solicita que responda correctamente con la información que recuerda. Le indica que puede mirar la historia de nuevo si lo necesita. Esta es la única Subprueba en la que el aplicador puede propiciar un diálogo con el estudiante. Se califican como correctas las respuestas si corresponden a las alternativas de respuesta para cada pregunta. En esta Subprueba no hay límite de tiempo (Visión Mundial, 2018).

5.Hojas impresas.

Documento de Word que contiene las plantillas de las hojas impresas con las letras, palabras e historia que los estudiantes leyeron durante la aplicación de la Prueba de lectura inicial.

6. Caracterización de las escuelas.

Matriz en documento de Word que contiene la información actualizada sobre cada una de las nueve escuelas que participaron en la evaluación.

F. Procedimiento

1. Limpieza de la base de datos en Excel

Este proceso se realizó durante las primeras semanas del mes de mayo. Se hizo una doble verificación en cada uno de los siguientes procesos para tener la base completamente limpia:

a. Acondicionamiento previo.

Para tener la base limpia se eliminó el color del relleno de algunas celdas en la base de datos, también se eliminaron las filas o columnas que incluían información repetida o que no aportaría al análisis de la prueba. (Grado 1_2_4_5_6) A su vez en cada una de las palabras de la historia, se eliminaron los signos de puntuación (, . ;).

b. Verificación.

Para asegurar que la base se encontraba limpia y lista para el análisis, se verificó si existían filtros en las columnas y se mostraron las columnas ocultas. Respecto a la información acerca del tiempo y algunos datos generales de los estudiantes se consultó con el encargado de STW para completar la información de la base de datos y así identificar los cambios que se habían realizado en la adaptación de la prueba STAR. Se verificó que el nombre de las escuelas y las hojas del archivo en Excel estuvieran escritas correctamente.

c. Actualización.

Se identificaron las palabras que estaban escritas con inicial mayúscula y les correspondía la letra inicial minúscula. Se unificaron las palabras que se encontraban escritas de diferente forma, para que fueran iguales como en la hoja que el estudiante leyó. Cada una de las columnas se recodificó por medio de un código alfanumérico, se puede verificar el código en el Anexo 4.

d. Recodificación de variables.

Para identificar la información de cada una de las escuelas y los estudiantes evaluados, se agregaron columnas en las cuales se codificaron cada uno de los datos. Se agregó una columna para registrar el total de respuestas en cada una de las Subpruebas. En la Subprueba de lectura oral se recodificaron los 129 ítems agregando (_r) para identificar las respuestas que se calificaron durante el minuto disponible y fue con este dato con el que se realizaron todos los análisis.

2. Análisis en SPSS

Al tener la base de datos limpia, en la segunda quincena de mayo se cargó en SPSS la base de datos en donde se generó la sintaxis y los resultados de los siguientes procedimientos:

a. Tablas de frecuencia.

Para identificar que se encontraban los datos correspondientes a cada una de las escuelas, los 60 estudiantes, la variable del idioma, sexo y los ítems de la lectura oral que se leyeron en el límite de tiempo. Ver la codificación en el Anexo 5 y 6. Con este proceso se verificó la presencia de datos perdidos previo a iniciar los análisis. Así mismo se obtuvo una muestra representativa y aleatoria de la población que participó en la aplicación de la

prueba, para identificar el comportamiento de los ítems en la población y en otra subpoblación la cual fue seleccionada aleatoriamente (30 estudiantes).

b. Análisis de confiabilidad de la prueba y en cada Subprueba.

En total la prueba está compuesta por 206 ítems de los cuales los ítems que tenían una varianza 0 fueron eliminados en la escala. Estos fueron tres en la Subprueba de alfabeto y dos ítems en la Subprueba de lectura oral. Por medio del análisis de confiabilidad se obtuvo en SPSS el Alfa de Cronbach para toda la prueba incluyendo el total global de respuestas en la prueba y sin el total global de la prueba. Así también se obtuvo la confiabilidad para cada una de las Subpruebas incluyendo el total de respuestas de la Subprueba y sin el total global de respuestas en la Subprueba en el software de jMetrik. Se decidió omitir los ítems que presentaron un índice de correlación de Pearson no significativo. Se realizó en diferente software debido a que no fue posible adquirir un software gratuito que permitiera trabajar con los 206 ítems y 60 respuestas recopiladas.

c. Análisis de confiabilidad según mitad.

Se dividió la prueba en dos mitades, el criterio para separarlas fue dividir cada Subprueba en dos, los primeros ítems de cada Subprueba conformaron la mitad 1 y los últimos ítems de cada Subprueba conformaron la mitad 2. Los ítems que conformaron la mitad 1 fueron: en la Subprueba de alfabeto del ítem A1 al ítem A14, en la Subprueba de palabras simples del ítem S1 al ítem S10, en la Subprueba de palabras inventadas del ítem I1 al ítem I10, en la Subprueba de lectura oral del ítem L1 al ítem L65 y en la Subprueba de comprensión lectora del ítem C1 al ítem C5. Los ítems que conformaron la mitad 2 fueron: en la Subprueba de alfabeto del ítem A15 al ítem A27, en la Subprueba de palabras simples del ítem S11 al ítem S20, en la Subprueba de palabras inventadas del ítem I11 al ítem I20, en la Subprueba de lectura oral del ítem L66 al ítem L129 y en la Subprueba de comprensión lectora del ítem C6 al ítem C10. Se realizó el análisis de confiabilidad para cada mitad, incluyendo el total de cada Subprueba y sin el total de cada Subprueba.

d. Correlaciones entre cada Subprueba.

Se realizó el análisis de correlaciones por cada Subprueba y se incluyó el total de cada Subprueba y el total global de la prueba. Se obtuvo el índice de correlación de Pearson.

e. Análisis factorial exploratorio³.

Al realizar el AFE no se obtuvo un indicador válido de KMO para toda la prueba y cada una de las Subpruebas. Posterior a varios análisis se obtuvo el KMO únicamente para la Subprueba de palabras simples con rotación Varimax y extracción con el método de Componentes principales y de la Subprueba de comprensión lectora al seleccionar las dos escuelas (Chulumal Jornada Matutina y Panimaché) que habían tenido mayor variabilidad en el total global de la prueba (primera corrida). Se obtuvo el KMO para toda la prueba considerando únicamente con 38 ítems, los cuales fueron: (10 ítems) en la Subprueba de alfabeto 6, 8, 10, 11, 13, 16, 20, 21, 22 y 24; en la Subprueba de palabras simples (2 ítems) 8 y 18; en la Subprueba de palabras inventadas (8 ítems) 1, 2, 7, 8, 10, 11, 13 y 17; en la Subprueba de lectura oral (10 ítems) 8, 9, 10, 15, 17, 20, 22, 25, 37 y 59; y en la Subprueba de comprensión lectora (8 ítems) 1, 2, 3, 4, 6, 7, 8 y 9. Y con 16 ítems los cuales fueron: en la Subprueba de alfabeto (2 ítems) 10 y 24; en la Subprueba de palabras simples (1 ítem) 8; en la Subprueba de palabras inventadas (4 ítems) 1, 2, 8 y 11; en la Subprueba de lectura oral (6 ítems) 17, 20, 22, 25, 37 y 59; y en la Subprueba de comprensión lectora (2 ítems) 1 y 3 para obtener el KMO global con rotación Varimax y extracción con el método de componentes principales. Y para todas las Subpruebas a excepción de la Subprueba de lectura oral se obtuvo el KMO con Promax y Máxima verosimilitud. (Segunda corrida) Para todos los análisis se excluyeron todos los ítems que tenían varianza 0.

f. Preparación de la base para cargarla en jMetrik.

Para calificar la prueba en jMetrik fue necesario recodificar cada una de las respuestas en SPSS. Las respuestas correctas codificadas con el número 1 se codificó con Z y las

³ AFE, Análisis Factorial Exploratorio

respuestas incorrectas o no contestadas codificadas con el número 0 se recodificaron con la letra Y.

g. Tablas personalizadas.

Se generaron tablas con las diferentes variables que se recopilaron durante la aplicación de la Prueba de lectura inicial en las nueve escuelas en el grado de 3ro. primaria.

3. Análisis en jMetrik

A finales de mayo se cargó en jMetrik, la base de datos en Excel en donde se generaron los siguientes procedimientos:

a. Calificación de la prueba.

Se realizó la calificación de los ítems por cada Subprueba y de toda la prueba, en la cual se obtuvo el bparam que representa la dificultad del ítem en TRI, la desviación estándar (se), Infit o cuadrado medio ponderado (wms), cuadrado medio ponderado estandarizado (Std. wms), Outfit o cuadrado medio no ponderado (ums) y cuadrado medio no ponderado estandarizado (Std. ums). Linacre (2012), Meyer sugiere que el valor de UMS y WMS pueden tener un rango de 0.5 a 1.5 con valores mayores que dos indicadores del ítem actual que degrada la medida. Los valores estandarizados de UMS y WMS son útiles para muestras pequeñas (menos de 300) y los valores absolutos de esta estadística no pueden ser mayores a 3. Al calificar a los estudiantes se obtuvieron los residuales los cuales se compartieron al equipo de STW en un documento de Excel.

b. Análisis de ítems.

Para cada uno de los ítems en cada Subprueba se obtuvo el indicador de dificultad, la desviación estándar y el índice de discriminación. Al obtener dichos índices se ordenaron

para identificar los primeros 10 ítems más difíciles y los que mejor discriminan, así como el promedio para cada Subprueba respecto al índice de dificultad y de discriminación.

c. Confiabilidad.

A través de jMetrik se obtuvo el análisis de confiabilidad para toda la prueba por medio de los siguientes métodos: Guttman's L2, Coeficiente de Alpha, Feldt-Gilmer, Feldt-Brennan y Raju's Beta.

d. Confiabilidad si el ítem es eliminado.

Para cada uno de los ítems en cada Subprueba se obtuvo la confiabilidad si el ítem es eliminado por medio de los métodos de Guttman's L2, Coeficiente de Alpha, Feldt-Gilmer, Feldt-Brennan y Raju's Beta.

e. Análisis diferencial de los ítems.

Se realizó el análisis diferencial de los ítems (DIF) para las cinco Subpruebas con las siguientes variables: escuela (10 códigos), edad (desde 8 a 13 años), sexo (mujer y hombre) e idioma (K'iché, Castellano y otro).

f. Calibración de los ítems.

En la calibración de los ítems se obtuvieron los estimados de los parámetros de los ítems con MMNL Apar (dificultad) y su error estándar, Bpar (discriminación) y su error estándar y Cpar (adivinación) con su error estándar. Así mismo se obtuvo las estadísticas de ajuste de los ítems para cada una de las Subpruebas.

g. Puntuaciones.

Para cada Subprueba se obtuvo la tabla de puntuaciones en la escala de Rasch, incluyendo la habilidad y el error estándar.

h. Calificación.

Para cada estudiante e ítem se obtuvo la tabla de calificaciones en Rasch, así como las estadísticas de calidad de la escala.

G. Diseño del método

El enfoque de la investigación fue cuantitativo, por medio de él se buscó la objetividad, bajo la perspectiva de que la realidad es una y que se la puede observar sin afectarla. El tipo de investigación cuantitativa, no experimental permite que se observen situaciones ya existentes, no provocadas intencionalmente en la investigación por quien la realiza. En la investigación no experimental las variables independientes ocurren y no es posible manipularlas, no se tiene control directo sobre dichas variables, ni se puede influir en ellas, porque ya sucedieron, al igual que sus efectos (Hernández, R, 2014).

H. Análisis de los datos o información

Para el análisis de los datos se utilizaron los siguientes softwares estadísticos:

1. Excel

Por medio de la herramienta análisis de datos, en Excel se obtuvieron las estadísticas descriptivas de los diferentes resultados obtenidos en el análisis de los ítems.

2. SPSS

Por medio de la plataforma de software IBM SPSS que ofrece el análisis estadístico avanzado, con una amplia biblioteca de algoritmos de aprendizaje automático, análisis de texto, extensibilidad de código abierto, integración con big data e implementación sin problemas en aplicaciones. Su facilidad de uso, flexibilidad y escalabilidad hace que SPSS sea accesible para usuarios con todos los niveles de habilidad y proyectos de todos los tamaños y complejidad. Ya que permite mejorar la eficiencia y minimiza el riesgo (IBM, 2019).

3. jMetrik

Es un sistema integrado para realizar análisis psicométricos con fines de investigación y operación sin el costo asociado como algún otro software psicométrico. (Gotzmann & Bahry, 2012) jMetrik es un software psicométrico gratuito y de código abierto, sin un límite para el tamaño de la muestra o el número de tablas en las que se pueden almacenar la base de datos. Los métodos psicométricos incluyen el análisis de ítems clásicos, estimación de confiabilidad, escalado de prueba, funcionamiento diferencial de los ítems, teoría de respuesta de ítems no paramétrica, modelos de medición Rasch, modelos de respuestas de ítems y la vinculación y ecuación de la teoría de respuesta de los ítems (Psychomeasurement Systems, 2018).

Durante el proceso de la limpieza de la base se realizó el método de doble verificación con dos fuentes diferentes, para realizar el proceso simultáneamente y corroborar la integridad de los datos a partir de los cuales se realizaron los siguientes análisis en los softwares mencionados anteriormente.

I. Consideraciones éticas

Se solicitó a cada una de las autoridades de las escuelas, la autorización para realizar la aplicación de la prueba de lectura. Una escuela que se había considerado en la muestra, no aceptó participar por temor a que la prueba significara una supervisión. Debido a ello,

STW eligió las dos jornadas de una de las escuelas contempladas en la muestra. La participación de los estudiantes fue voluntaria y al iniciar la aplicación se les compartió el propósito de la aplicación y cada uno de ellos expresó si deseaba participar. A los estudiantes se les indicó que tenía el derecho de retirarse de la prueba durante la aplicación y que no se tendrían beneficios ni riesgos específicos por participar.

Se resguardó la información de la base de datos, así como de cada uno de los documentos que compartió el equipo de STW en donde se pudo conocer la Prueba de Lectura Inicial aplicada por STW. En este informe se dan a conocer las propiedades psicométricas de la prueba guardando la confidencialidad del contenido de cada uno de los ítems en las cinco Subpruebas. A su vez las calificaciones individuales se entregaron en un archivo de Excel al equipo de STW. En esta prueba no había datos sensibles aparte de su nombre, sin embargo, se decidió utilizar un código para cada estudiante y así guardar la confidencialidad de su identidad.

IV. RESULTADOS DE LA INVESTIGACIÓN

A continuación, se presentan los resultados que se obtuvieron en los estadísticos descriptivos por Subprueba; la correlación de Pearson para cada Subprueba en la población y la muestra con y sin el total de la prueba. Se presenta el análisis de ítems para cada Subprueba acorde a su índice de dificultad, de discriminación y desviación estándar. También se identifican los primeros 10 ítems de mayor dificultad y de mejor discriminación, así como el promedio de dificultad y discriminación por cada Subprueba. Adicionalmente se presenta en esta sección los indicadores de ajuste del modelo para cada Subprueba: RMSE y error estándar. Se incluye el análisis cuantitativo de la habilidad requerida para responder la Subprueba, los descriptivos de la Subprueba incluyendo la dificultad, el WMS estándar y el UMS estándar. Se presentan los indicadores según dificultad, desviación estándar y estadísticos descriptivos.

Para determinar la confiabilidad se realizó para el total de la prueba, confiabilidad si el ítem es eliminado, confiabilidad el ítem-total, análisis de la prueba de dos mitades sin el total de cada Subprueba y con el total de cada Subprueba. La validación incluyó un Análisis Factorial Exploratorio analizando el KMO en cada análisis realizado. Se calibraron los ítems con TRI y se obtuvieron las puntuaciones para cada Subprueba.

A. Descriptivos por cada Subprueba

El promedio de respuestas correctas brindadas por los estudiantes evaluados fue de 113 y la mediana de 112 sobre un total de 206 ítems, como se puede observar en la Tabla 12. Se observó que la menor cantidad de respuestas correctas fue de 1 por cada 10 ítems (26/206) y el máximo de respuestas correctas por un estudiante fue de 9 por cada 10 ítems (200/206). La media es de 113,2 que equivale al 56,61% y la mediana es de 112,5, con moda de 1112, una desviación estándar de 30,66 que equivale a 3,43 y un rango intercuartil de 31, lo cual distribuye a la población evaluada en 6 rangos. La curtosis es de 1,68 lo cual indica que la distribución de las frecuencias es platicúrtica porque es menos apuntada y con colas menos gruesas que la normal y la oblicuidad es de 0,26. Para esta Prueba de lectura inicial el KR21 es de 0,9504. En la Tabla 13 y gráfica 1 se pueden corroborar los datos.

Tabla 13. *Estadísticos de respuesta correcta de la Prueba de lectura inicial*

Número de ítems=	206
Número de examinados=	60
Mínimo=	26
Máximo=	200
Media=	113,2167
Error típico de la media=	3,9882
Mediana=	112,5
Moda=	112
Desviación estándar =	30,6606
Varianza=	954,360
Asimetría=	.26
Error típico de asimetría=	.309
Rango intercuartil =	31
Oblicuidad=	0,2654
Curtosis =	1,6843
Error típico de la curtosis=	,608
Suma=	6795,00
KR21=	0,9504

Gráfica 1. *Histograma del total de puntuaciones obtenidas por los estudiantes*

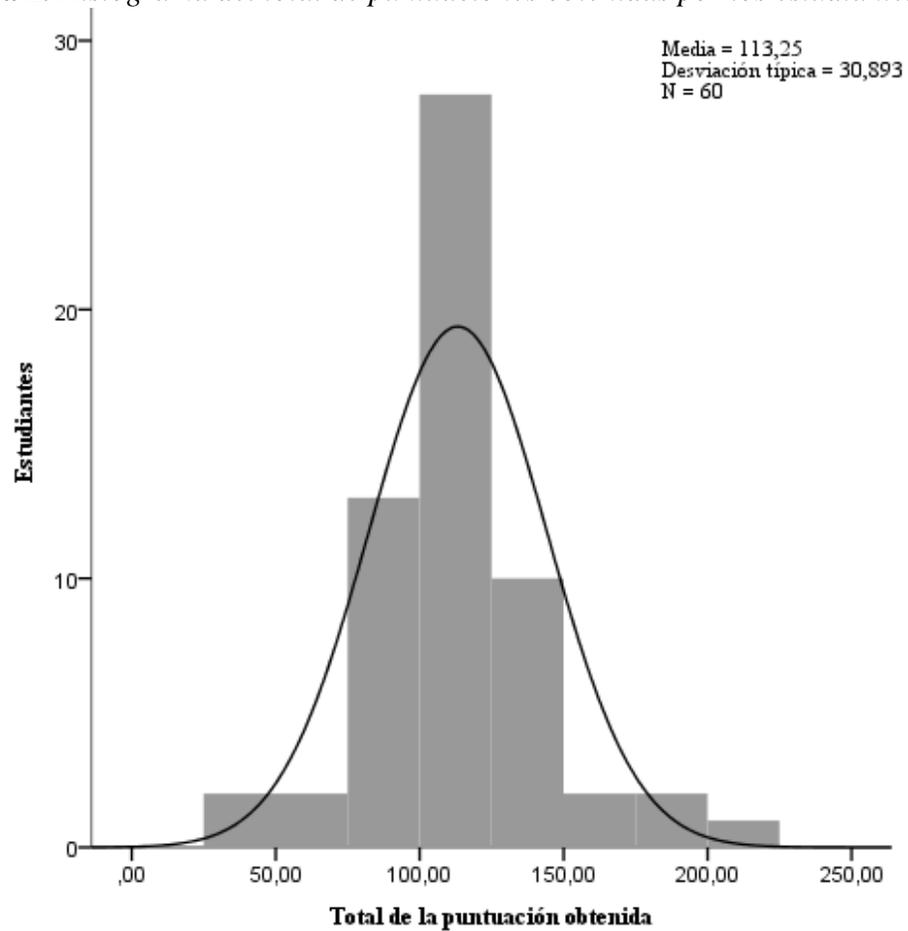


Tabla 14. Descriptivos para la Subprueba de alfabeto

	Bases de datos con el total de la Subprueba de alfabeto		Bases de datos con el total de la Prueba de lectura inicial	
	Población	Muestra	Población	Muestra
Media	0,50266667	0,57689474	0,33733333	0,42883333
Error típico	0,02979648	0,02539695	0,01130522	0,0177712
Mediana	0,482	0,576	0,326	0,417
Moda		0,612	0,316	0,439
Desviación estándar	0,13654462	0,11070275	0,03916244	0,06156125
Varianza de la muestra	0,01864443	0,0122551	0,0015337	0,00378979
Curtosis	-0,79171311	-0,63553111	-0,24134173	1,07773565
Coefficiente de asimetría	0,39313902	0,04087322	0,80235485	1,38553494
Rango	0,46	0,379	0,125	0,186
Mínimo	0,303	0,383	0,285	0,368
Máximo	0,763	0,762	0,41	0,554
Suma	10,556	10,961	4,048	5,146
Cuenta	21	19	12	12
Nivel de confianza (95,0%)	0,06215437	0,05335702	0,02488263	0,03911415

Se observó una media menor cuando se analizaron los ítems con el total de la Prueba que con el de la Subprueba de alfabeto. La mediana de los indicadores de correlación es de 0.32 como mínimo y de 0.57 como máximo. Los menores coeficientes de correlación se encontraron cuando se relacionó cada ítem de la Subprueba alfabeto y el total de la prueba en la población.

Tabla 15. Descriptivos para la Subprueba de palabras simples

	Bases de datos con el total de la Subprueba de Palabras Simples		Bases de datos con el total de la Prueba de Lectura Inicial	
	Población	Muestra	Población	Muestra
Media	0,58665285	0,6718487	0,35084438	0,44917404
Error típico	0,03023111	0,02858805	0,01664096	0,01770331
Mediana	0,563	0,640	0,347	0,439
Moda	0,782	0,853	0,347	0,439
Desviación estándar	0,13177434	0,12461241	0,06226477	0,06856461
Varianza de la muestra	0,01736448	0,01552825	0,0038769	0,00470111
Curtosis	-0,12706001	-1,19299946	-1,18149395	-0,98465696
Coefficiente de asimetría	0,01094954	0,23263406	0,16578471	0,6150205
Rango	0,47202539	0,39165806	0,18208072	0,19408985
Mínimo	0,31204482	0,46197316	0,2663372	0,36553302
Máximo	0,7840702	0,85363123	0,44841792	0,55962287
Suma	11,1464042	12,7651254	4,91182134	6,73761056
Cuenta	19	19	14	15
Nivel de confianza (95,0%)	0,0635132	0,06006126	0,03595061	0,03796982

La media es menor cuando se analizaron los ítems con el total de la prueba que con el de la Subprueba de palabras simples. La mediana de los indicadores de correlación es de 0.34 como mínimo y de 0.64 como máximo. Los menores coeficientes de correlación se encontraron cuando se relacionó cada ítem de la Subprueba palabras simples y el total de la prueba en la población.

Tabla 16. Descriptivos para la Subprueba de palabras inventadas

	Bases de datos con el total de la Subprueba de palabras inventadas		Bases de datos con el total de la Prueba de lectura inicial	
	Población	Muestra	Población	Muestra
Media	0,48188357	0,62478319	0,36848412	0,47806879
Error típico	0,0257398	0,02051718	0,02321101	0,02207762
Mediana	0,476	0,626	0,376	0,457
Moda				
Desviación estándar	0,11511188	0,08459449	0,06963304	0,06623287
Varianza de la muestra	0,01325074	0,00715623	0,00484876	0,00438679
Curtosis	-1,48054726	0,71725394	-1,60765568	-1,28628617
Coefficiente de asimetría	-0,19490313	0,43267036	0,07727263	0,55764141
Rango	0,3469996	0,35645381	0,19024006	0,1780077
Mínimo	0,29725731	0,46447011	0,27085529	0,40748707
Máximo	0,64425691	0,82092392	0,46109535	0,58549477
Suma	9,63767146	10,6213141	3,31635711	4,3026191
Cuenta	20	17	9	9
Nivel de confianza (95,0%)	0,05387402	0,04349447	0,05352469	0,05091109

La media es menor cuando se analizaron los ítems con el total de la prueba que con el de la Subprueba de palabras inventadas. La mediana de los indicadores de correlación es de 0.37 como mínimo y de 0.62 como máximo. Los menores coeficientes de correlación se encontraron cuando se relacionó cada ítem de la Subprueba palabras inventadas y el total de la prueba en la población.

Tabla 17. Descriptivos para la Subprueba de lectura oral

	Bases de datos con el total de la Subprueba de lectura oral		Bases de datos con el total de la Prueba de lectura inicial	
	Población	Muestra	Población	Muestra
Media	0,66090385	0,66598417	0,61353379	0,60264001
Error típico	0,01047032	0,00773339	0,00824912	0,00664129
Mediana	0,691	0,657	0,627	0,587
Moda	0,691	0,657	0,607	0,563
Desviación estándar	0,11179247	0,08220709	0,09036457	0,07305422
Varianza de la muestra	0,01249756	0,00675801	0,00816576	0,00533692
Curtosis	2,66873872	4,62619429	1,69917558	0,56539833
Coefficiente de asimetría	-1,71985685	-2,0246944	-1,33797478	-0,34303788
Rango	0,50654144	0,3873914	0,43507359	0,36082476
Mínimo	0,25659004	0,36462139	0,29013641	0,37266223
Máximo	0,76313147	0,75201278	0,72521	0,73348699
Suma	75,3430395	75,2562115	73,6240549	72,9194407
Cuenta	114	113	120	121
Nivel de confianza (95,0%)	0,0207436	0,01532273	0,01633408	0,0131493

La media es menor cuando se analizaron los ítems con el total de la prueba que con el de la Subprueba de lectura oral. La mediana de los indicadores de correlación es de 0.58 como mínimo y de 0.69 como máximo. Los coeficientes de correlación fueron muy similares cuando se relacionó cada ítem de la Subprueba lectura oral y el total de la prueba en la población y la muestra.

Tabla 18. Descriptivos para la Subprueba comprensión lectora

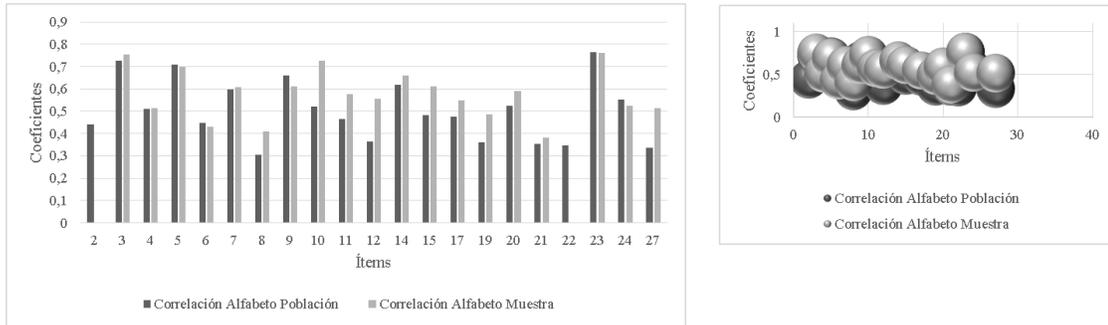
	Bases de datos con el total de la Subprueba de comprensión lectora		Bases de datos con el total de la Prueba de lectura inicial	
	Población	Muestra	Población	Muestra
Media	0,55978496	0,54894165	0,35959753	
Error típico	0,03040357	0,03710114	0	
Mediana	0,568	0,528	0,359	
Moda				
Desviación estándar	0,09614453	0,10493787		
Varianza de la muestra	0,00924377	0,01101196		
Curtosis	-0,31781981	-0,6532708		
Coefficiente de asimetría	0,02311953	0,36852999		
Rango	0,31545836	0,30885036	0	
Mínimo	0,41264465	0,39964564	0,35959753	
Máximo	0,72810301	0,708496	0,35959753	
Suma	5,59784965	4,39153318	0,35959753	
Cuenta	10	8	1	
Nivel de confianza (95,0%)	0,06877765	0,08773025		

La media es menor cuando se analizaron los ítems con la Subprueba de comprensión lectora. La mediana de los indicadores de correlación es de 0.35 como mínimo y de 0.56 como máximo. Al correlacionar cada ítem de la Subprueba comprensión lectora y el total de la prueba, solo se logró obtener en la población la correlación del ítem 5.

B. Correlación de Pearson en cada Subprueba

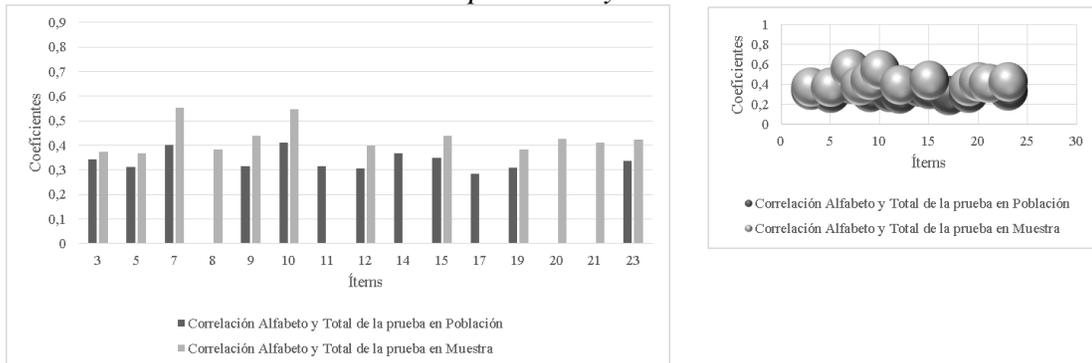
A continuación, se presenta la correlación de Pearson del total en cada Subprueba con el total de la Prueba para la población y la muestra aleatoria. El propósito fue identificar el comportamiento de los ítems en la muestra y en otra submuestra la cual fue seleccionada aleatoriamente.

Gráfica 2. *Correlación de los ítems-total de la Subprueba de alfabeto en la población y la muestra*



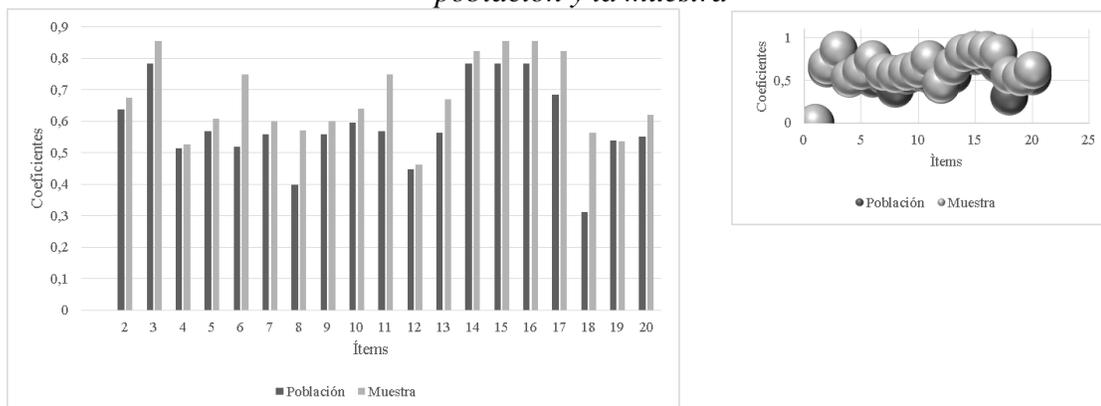
El mayor indicador de correlación obtenido en los ítems de la Subprueba de alfabeto fue de 0.70 y el menor fue de 0.30 para la muestra y la población. Los ítems 2 y 22 tuvieron una correlación no significativa. El criterio de identificación 0.10 evidenció que los ítems: 8, 10, 11, 12, 15, 19 y 27 tuvieron una correlación mayor en la muestra, que en la población. La correlación ítem-total de la Subprueba de alfabeto fue de 0.10 mayor en $\frac{1}{4}$ de los ítems de la muestra, con relación al indicador obtenido entre la correlación de cada ítem con el total de la Subprueba, en la población evaluada.

Gráfica 3. *Correlación ítems-total de la Subprueba de alfabeto con el total de la prueba en la población y la muestra*



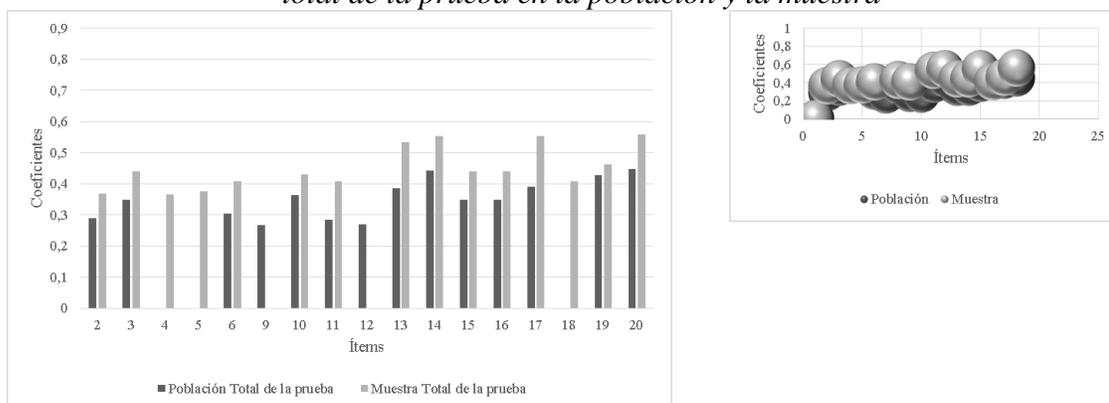
El mayor indicador de correlación obtenido en los ítems de la Subprueba de alfabeto fue de 0.50 y el menor fue de 0.20 para la muestra y la población. Los ítems: 8, 11, 14, 17, 20 y 21 tuvieron una correlación no significativa. El criterio de identificación 0.10 evidenció que los ítems: 7, 9, 10 y 12 tuvieron una correlación mayor en la muestra, que en la población. La correlación ítem-total de la Subprueba de alfabeto con el total de la prueba fue de 0.10 mayor en $\frac{1}{4}$ de los ítems de la muestra, en la población evaluada.

Gráfica 4. *Correlación de los ítems-total de la Subprueba de palabras simples en la población y la muestra*



El mayor indicador de correlación obtenido en los ítems de la Subprueba de palabras simples fue de 0.80 y el menor fue de 0.30 para la muestra y la población. El criterio de identificación 0.10 evidenció que los ítems: 6, 8, 11, 13, 17 y 18 tuvieron una correlación mayor en la muestra, que en la población. La correlación ítem-total de la Subprueba de palabras simples fue de 0.10 mayor en $\frac{1}{4}$ de los ítems de la muestra, con relación al indicador obtenido entre la correlación de cada ítem con el total de la Subprueba, en la población evaluada.

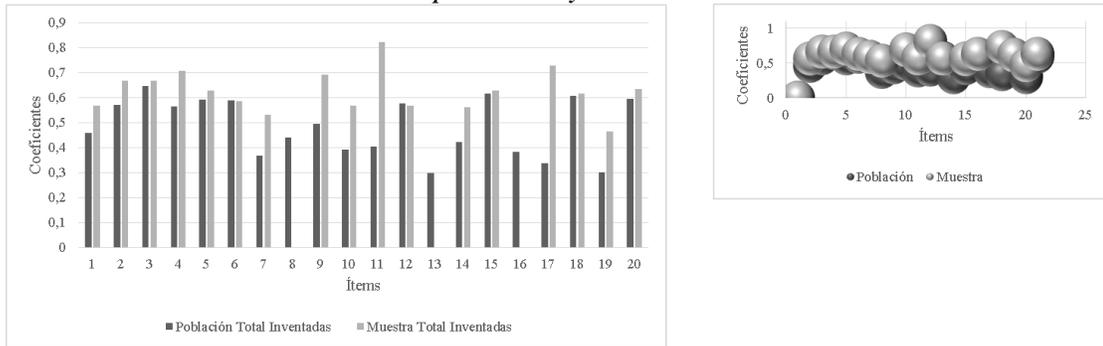
Gráfica 5. *Correlación de los ítems-total de la Subprueba de palabras simples con el total de la prueba en la población y la muestra*



El mayor indicador de correlación obtenido en los ítems de la Subprueba de palabras simples fue de 0.50 y el menor fue de 0.20 para la muestra y la población. Los ítems: 4, 5, 9, 12 y 18 tuvieron una correlación no significativa. El criterio de identificación 0.10 evidenció que los ítems: 3, 6, 11, 13 14, 17 y 20 tuvieron una correlación mayor en la

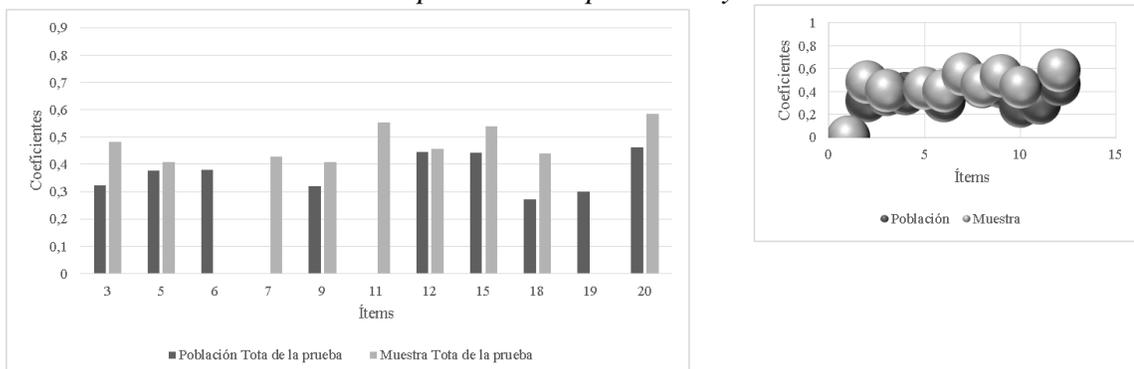
muestra, que en la población. La correlación ítem-total de la Subprueba de Palabras Simples con el total de la Prueba fue de 0.10 mayor en casi 2/4 de los ítems de la muestra, en la población evaluada.

Gráfica 6. *Correlación de los ítems-total de la Subprueba de palabras inventadas en la población y la muestra*



El mayor indicador de correlación obtenido en los ítems de la Subprueba de palabras inventadas fue de 0.80 y el menor fue de 0.30 para la muestra y la población. Los ítems 8, 13 y 16 tuvieron una correlación no significativa. El criterio de identificación 0.10 evidenció que los ítems: 1, 4, 7, 9, 10, 11, 14, 17 y 19 tuvieron una correlación mayor en la muestra, que en la población. La correlación ítem-total de la Subprueba de palabras inventadas fue de 0.10 mayor en 1/3 de los ítems de la muestra, con relación al indicador obtenido entre la correlación de cada ítem con el total de la prueba, en la población evaluada.

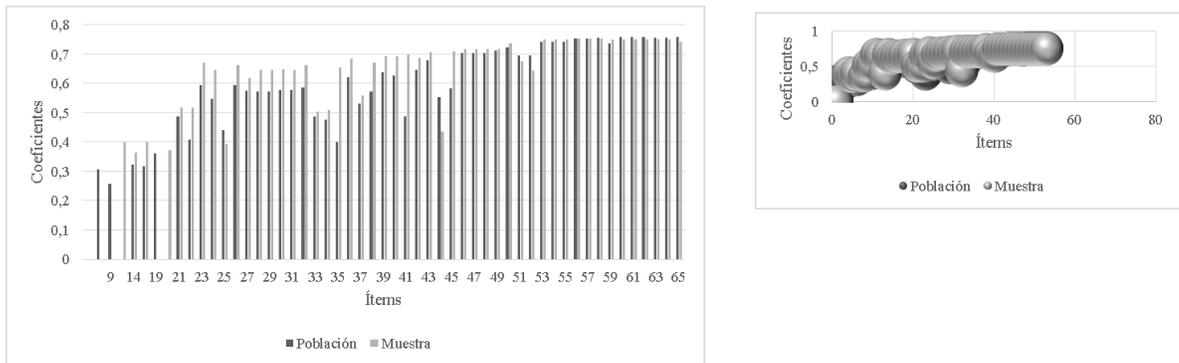
Gráfica 7. *Correlación de los ítems-total de la Subprueba de palabras inventadas con el total de la prueba en la población y la muestra*



El mayor indicador de correlación obtenido en los ítems de la Subprueba de palabras inventadas fue de 0.50 y el menor fue de 0.20 para la muestra y la población. Los ítems: 6,

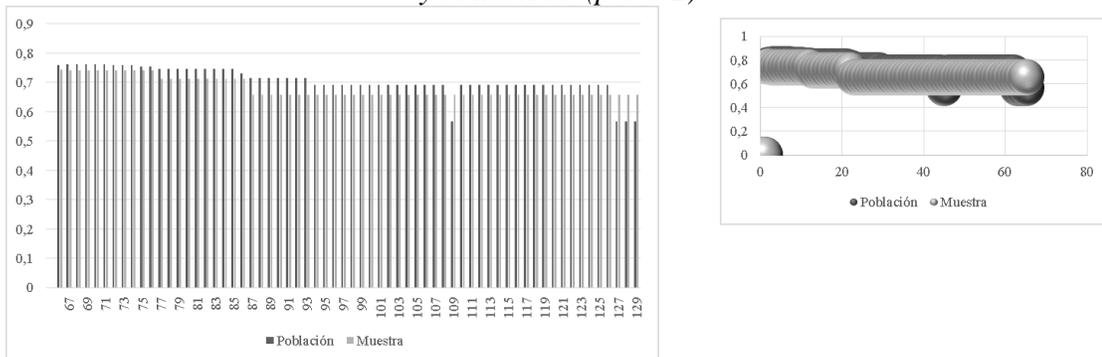
7, 11 y 19 tuvieron una correlación no significativa. El criterio de identificación 0.10 evidenció que los ítems: 3, 18 y 20 tuvieron una correlación mayor en la muestra, que en la población. La correlación ítem-total de la Subprueba de palabras inventadas con el total de la prueba fue de 0.10 mayor en menos de 1/4 de los ítems de la muestra, en la población evaluada.

Gráfica 8. *Correlación de los ítems-total de la Subprueba de lectura oral en la población y la muestra (parte 1)*



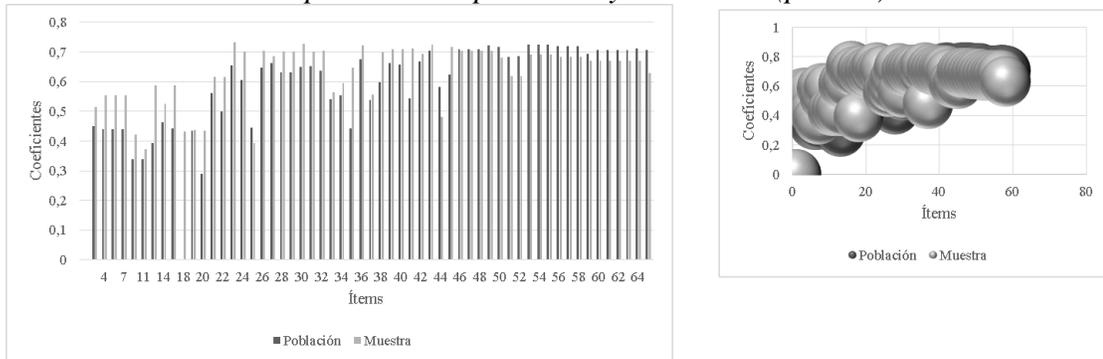
Para presentar la gráfica de correlación de los ítems de la Subprueba de lectura oral, se dividieron en dos partes debido a que son 129 en total. El mayor indicador de correlación obtenido en la primera parte de los ítems (65 ítems) de la Subprueba de lectura oral fue de 0.70 y el menor fue de 0.20 para la muestra y la población. Los ítems 2, 9, 13, 19 y 20 tuvieron una correlación no significativa. El criterio de identificación 0.10 evidenció que los ítems: 22, 24, 35, 37, 41, y 45 tuvieron una correlación mayor en la muestra y el ítem 44 tuvo una correlación mayor en la población.

Gráfica 9. *Correlación de los ítems-total de la Subprueba de lectura oral en la población y la muestra (parte 2)*



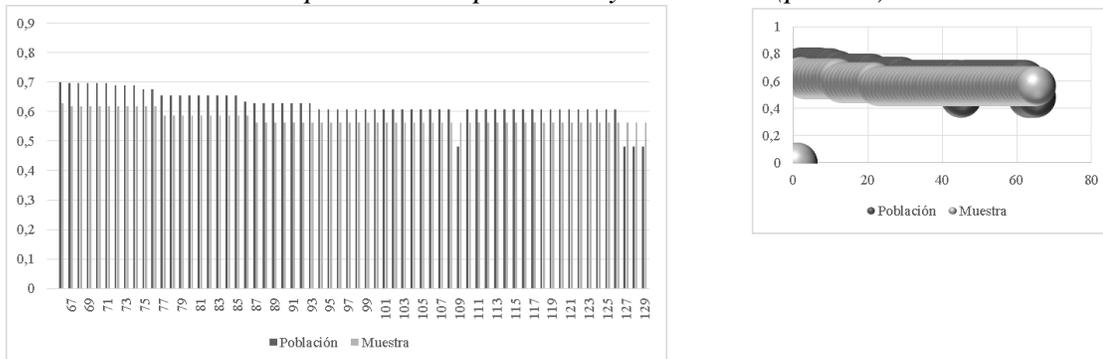
El mayor indicador de correlación obtenido en la segunda parte de los ítems (64 ítems) de la Subprueba de lectura oral con el total de la Prueba fue de 0.70 y el menor fue de 0.50 para la muestra y la población. La mayoría de los ítems presentaron una correlación levemente mayor en la población y no en la muestra, a diferencia de las Subpruebas anteriores. Los ítems 109, 127, 128 y 129 tuvieron una correlación mayor en la muestra.

Gráfica 10. *Correlación de los ítems-total de la Subprueba de lectura oral con el total de la prueba en la población y la muestra (parte 1)*



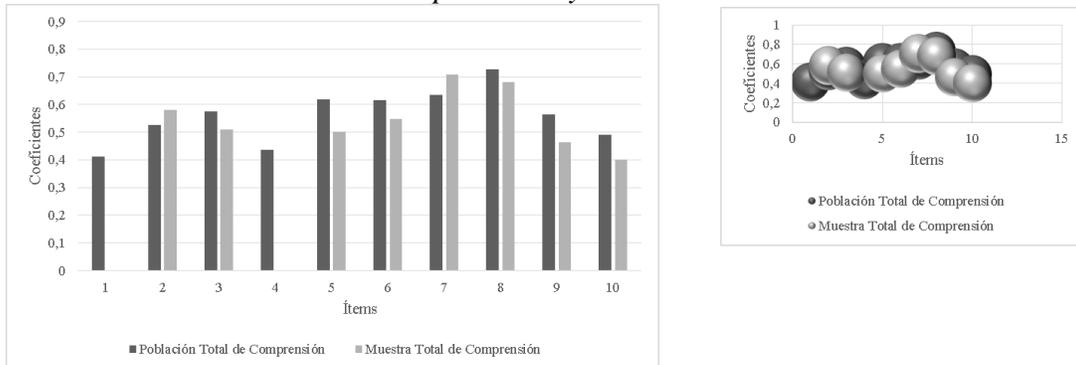
El mayor indicador de correlación obtenido en la primera parte de los ítems (65 ítems) de la Subprueba de lectura oral con el total de la prueba fue de 0.70 y el menor fue de 0.20 para la muestra y la población. El ítem 18 tuvo una correlación no significativa. El criterio de identificación 0.10 evidenció que los ítems: 4, 5, 7, 13, 17, 20, 22, 24, 35, 37, 41, y 45 tuvieron una correlación mayor en la muestra y el ítem 44 tuvo una correlación mayor en la población. A partir del ítem 46 los demás ítems muestran mayor correlación en la población.

Gráfica 11. *Correlación de los ítems-total de la Subprueba de lectura oral con el total de la prueba en la población y la muestra (parte 2)*



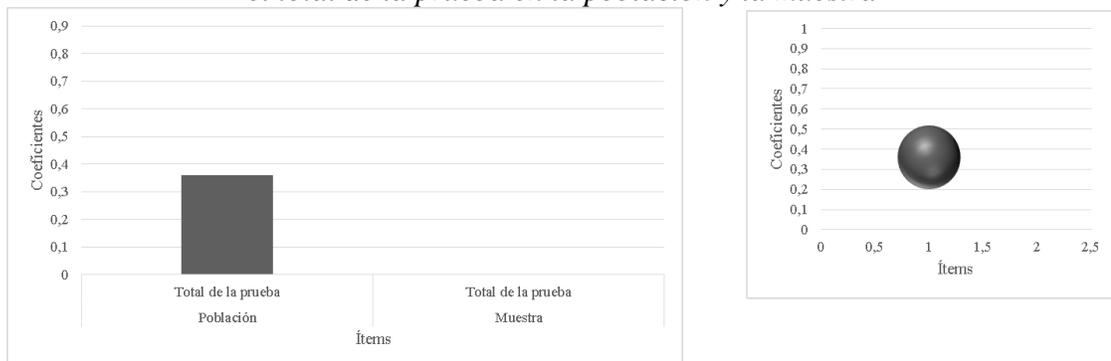
El mayor indicador de correlación obtenido en la segunda parte de los ítems (64 ítems) de la Subprueba de lectura oral con el total de la prueba fue de 0.70 y el menor fue de 0.40 para la muestra y la población. La mayoría de los ítems tuvieron una correlación mayor en la población. Los ítems 109, 127, 128 y 129 tuvieron una correlación mayor en la muestra.

Gráfica 12. *Correlación de los ítems-total de la Subprueba de comprensión lectora en la población y la muestra*



El mayor indicador de correlación obtenido en los ítems de la Subprueba de comprensión lectora fue de 0.70 y el menor fue de 0.40 para la muestra y la población. Los ítems 1 y 4 tuvieron una correlación no significativa. El criterio de identificación 0.10 evidenció que el ítem 5 tuvo una correlación mayor en la población, que en la muestra. La correlación ítem-total de la Subprueba de comprensión lectora fue mayor en la población. El ítem 2 y el 7 tienen un indicador mayor de correlación en la muestra.

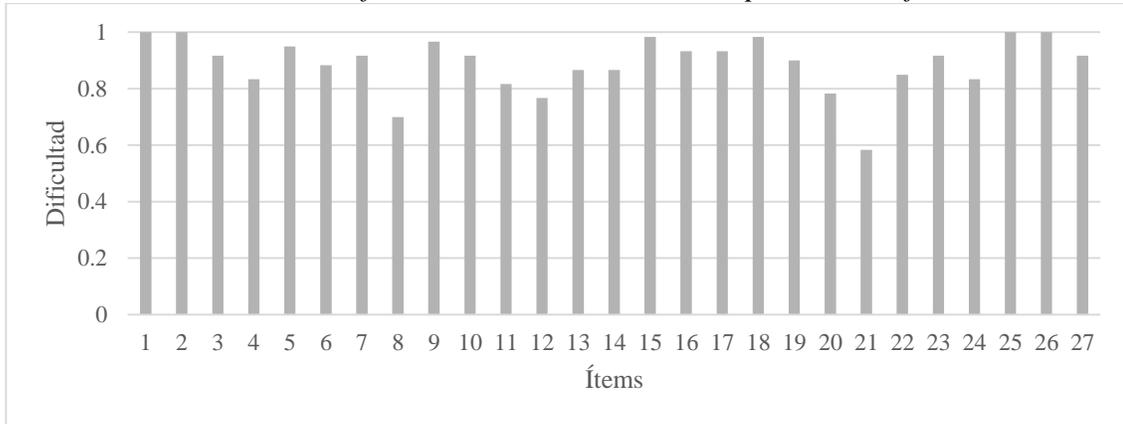
Gráfica 13. *Correlación de los ítems-total de la Subprueba de comprensión lectora con el total de la prueba en la población y la muestra*



En la correlación de la Subprueba de comprensión lectora con el total de la prueba, el ítem 5 fue el único que mostró un indicador de correlación de 0.35 en la población.

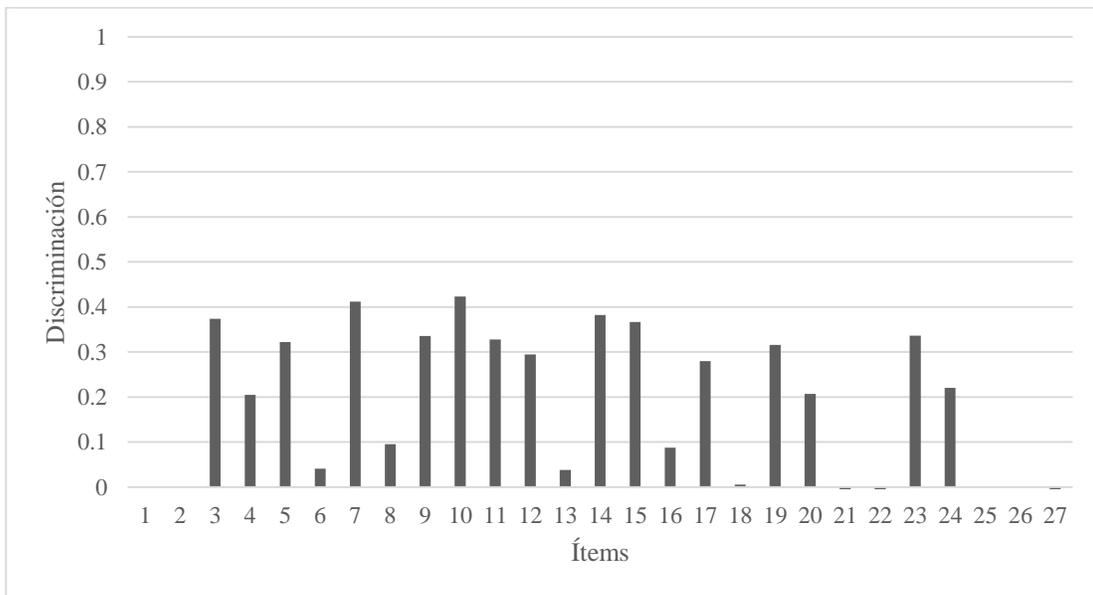
C. Análisis de los ítems

Gráfica 14. *Dificultad de los ítems en la Subprueba de alfabeto*



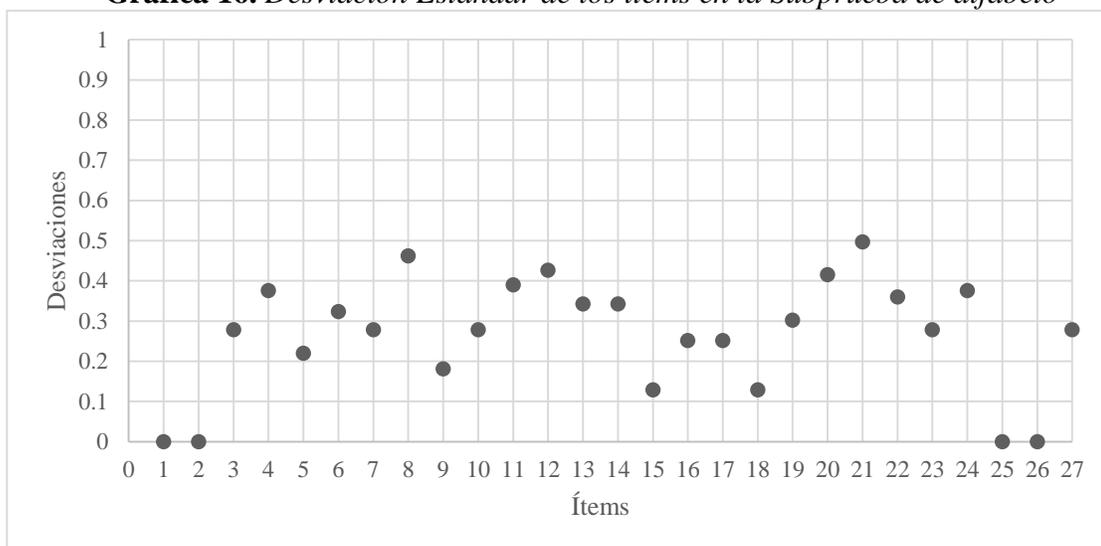
En la mayoría de los ítems de la Subprueba de alfabeto, la dificultad es mayor a 0.70 y únicamente dos ítems tuvieron menor indicador, lo cual significa que la mayoría de los ítems fueron fáciles de contestar, para los estudiantes de 3° primaria.

Gráfica 15. *Discriminación de los ítems en la Subprueba de alfabeto*



Más de la mitad de los ítems tienen una discriminación superior a 0.20 y el 18% de los ítems tienen una discriminación menor a 0.10. Existen tres ítems con discriminación negativa, menor a 0.10 lo cual representa el 11% de los ítems en la Subprueba de alfabeto.

Gráfica 16. *Desviación Estándar de los ítems en la Subprueba de alfabeto*



El 85% de los ítems tienen una desviación estándar superior a 0.10 y cuatro ítems tienen una desviación estándar de 0.

Tabla 19. *Resultado promedio de la Subprueba alfabeto*

Promedio	Dificultad	Discriminación
ALFABETO	0,88950741	0,18128148

En promedio la Subprueba de alfabeto es muy fácil para los estudiantes. La discriminación es justa, ya que se encuentra entre 0.10 y 0.30.

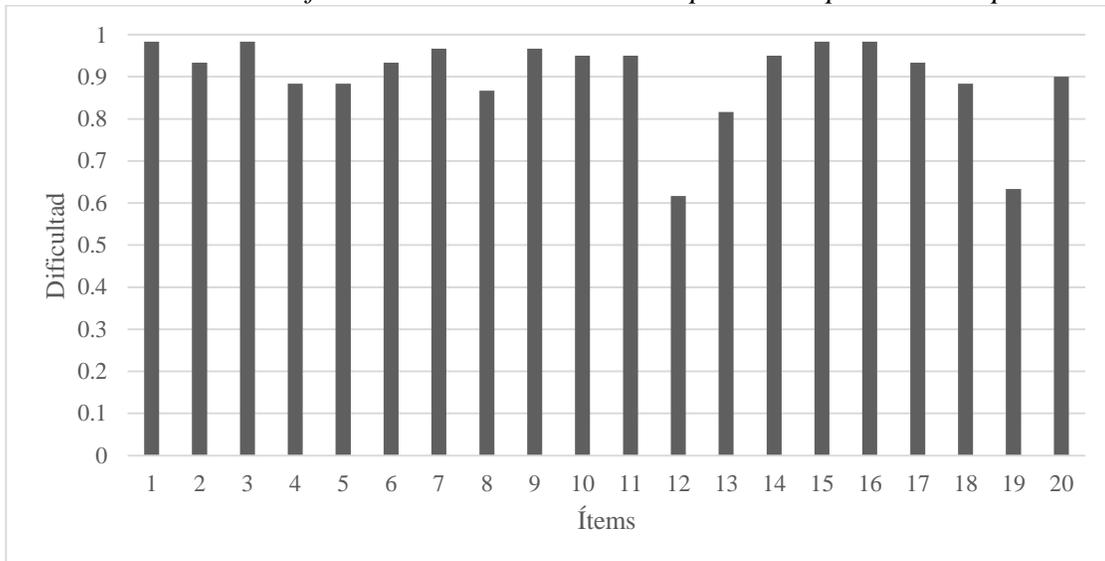
Tabla 20. *Ítems más difíciles y que mejor discriminan de la Subprueba de alfabeto*

Ítem	Dificultad	Ítem	Discriminación
14	0,8667	10	0,4235
13	0,8667	7	0,4116
22	0,85	14	0,3819
24	0,8333	3	0,374
4	0,8333	15	0,3667
11	0,8167	23	0,3364
20	0,7833	9	0,3353
12	0,7667	11	0,3279
8	0,7	5	0,322
21	0,5833	19	0,3161

En la Tabla 20 se muestran los ítems de la Subprueba de alfabeto, según el indicador de dificultad están ordenados de menor a mayor dificultad, se evidencia que para más de la

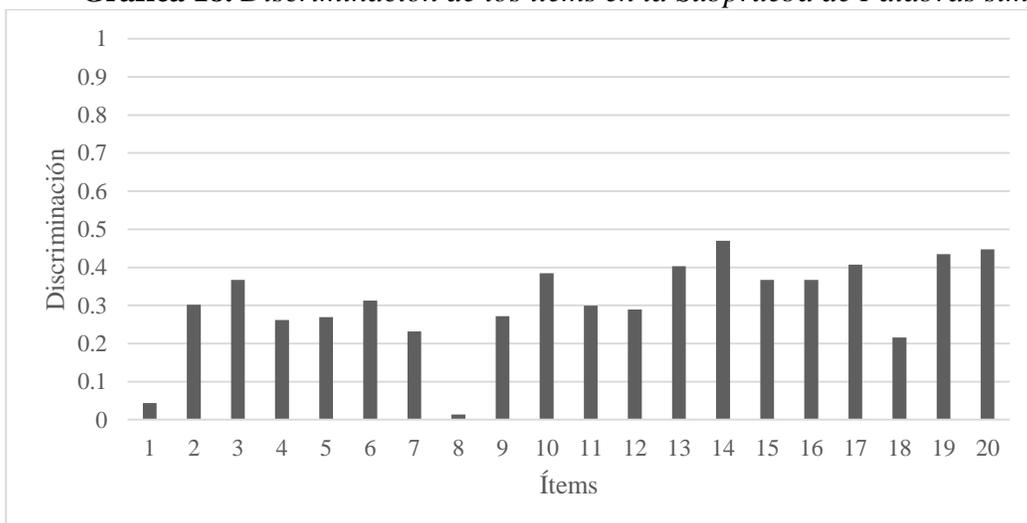
mitad de la población fue fácil responderlos correctamente. Los ítems según el indicador de discriminación están ordenados de mayor a menor, estos son superiores a 0.30.

Gráfica 17. *Dificultad de los ítems en la Subprueba de palabras simples*



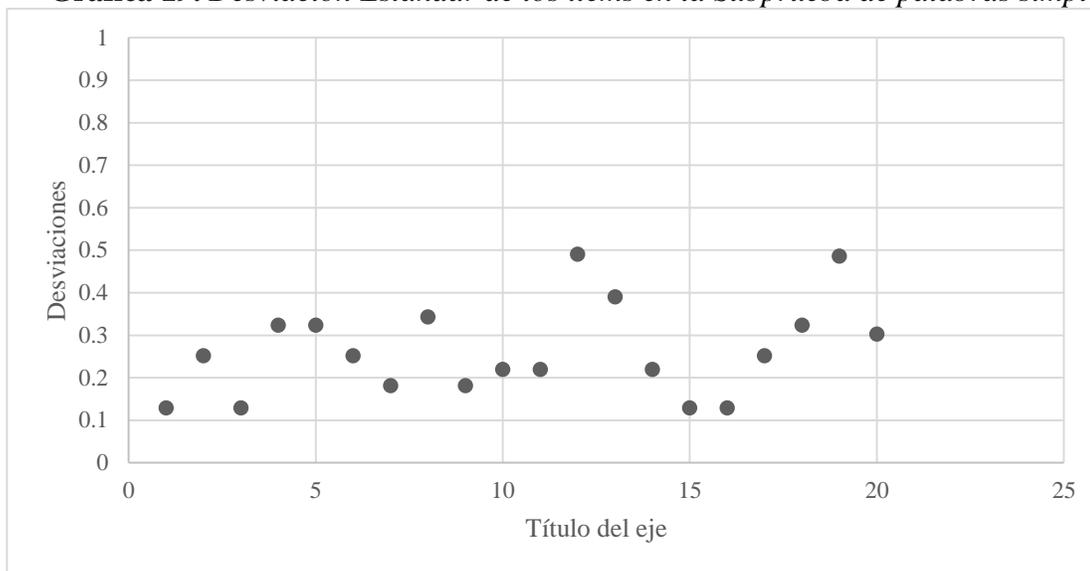
En la mayoría de los ítems de la Subprueba de Palabras simples, la dificultad es mayor a 0.80 y únicamente dos ítems tuvieron menor indicador, lo cual significa que la mayoría de los ítems fueron fáciles de contestar, para los estudiantes de 3° primaria.

Gráfica 18. *Discriminación de los ítems en la Subprueba de Palabras simples*



Más de la mitad de los ítems tienen una discriminación superior a 0.20 y el 10% de los ítems tienen una discriminación menor a 0.10 en la Subprueba de palabras simples.

Gráfica 19. *Desviación Estándar de los ítems en la Subprueba de palabras simples*



El 100% de los ítems tienen una desviación estándar superior a 0.10.

Tabla 21. *Resultado promedio de la Subprueba palabras simples*

Promedio SIMPLES	Dificultad	Discriminación
	0,89999	0,307945

En promedio la Subprueba de palabras simples es muy fácil para los estudiantes. La discriminación es justa, ya que se encuentra entre 0.10 y 0.30.

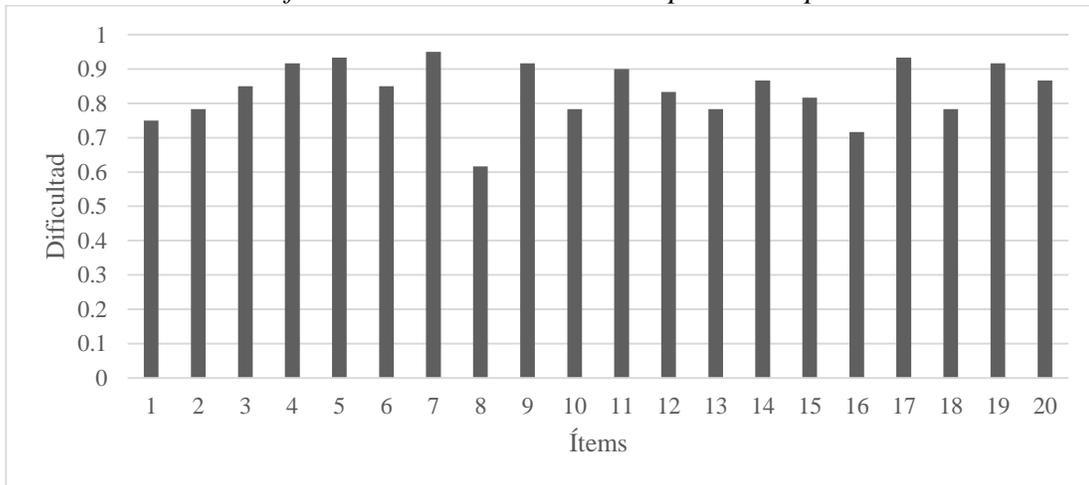
Tabla 22. *Ítems más difíciles y que mejor discriminan de la Subprueba de palabras simples*

Ítem	Dificultad	Ítem	Discriminación
6	0,9333	20	0,4658
17	0,9333	15	0,4556
20	0,9	12	0,4545
4	0,8833	6	0,3858
5	0,8833	5	0,3741
18	0,8833	3	0,3459
8	0,8667	9	0,3225
13	0,8167	18	0,2976
19	0,6333	19	0,2869
12	0,6167	1	0,2659

En la Tabla 22 se muestran los ítems de la Subprueba de palabras simples, según el indicador de dificultad están ordenados de menor a mayor, se evidencia que para más de la

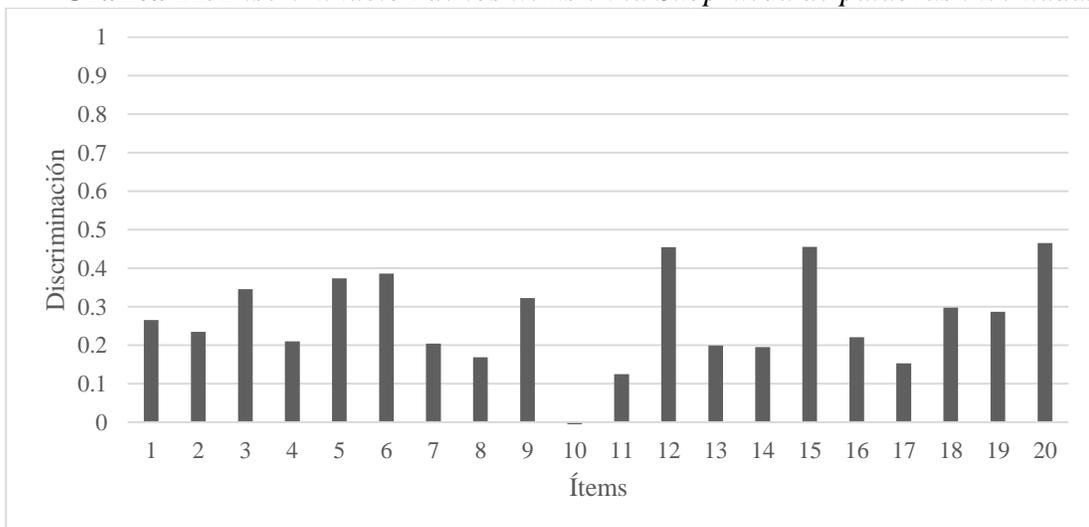
mitad de la población fue fácil responderlos correctamente. Los ítems según el indicador de discriminación están ordenados de mayor a menor, estos son superiores a 0.25.

Gráfica 20. *Dificultad de los ítems en la Subprueba de palabras inventadas*



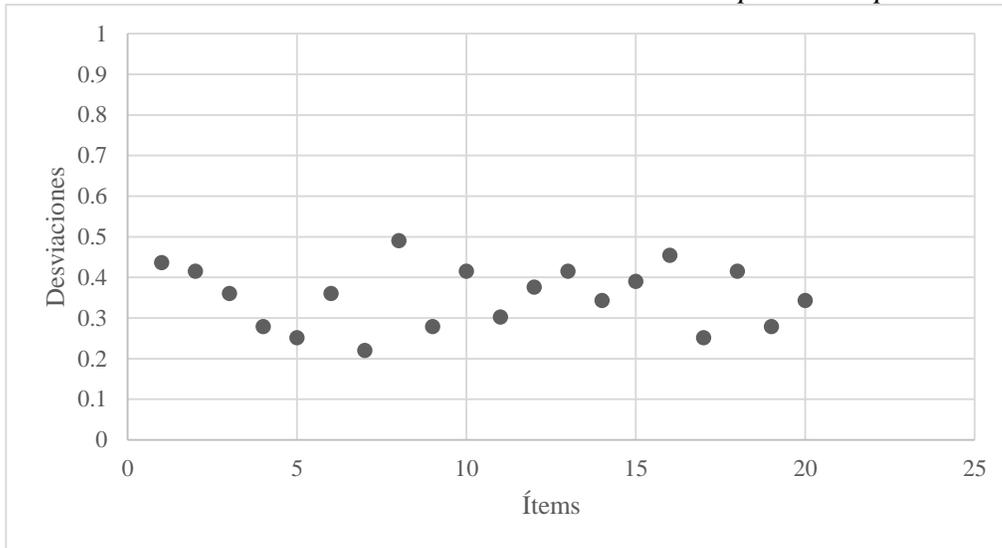
En la mayoría de los ítems de la Subprueba de palabras inventadas, la dificultad es mayor a 0.70 y únicamente un ítem tuvo un menor indicador, lo cual significa que la mayoría de los ítems fueron fáciles de contestar, para los estudiantes de 3° primaria.

Gráfica 21. *Discriminación de los ítems en la Subprueba de palabras inventadas*



Más de la mitad de los ítems tienen una discriminación superior a 0.20. Existe un ítem con discriminación negativa, menor a 0.10 lo cual representa el 5% de los ítems en la Subprueba de palabras inventadas.

Gráfica 22. Desviación Estándar de los ítems en la Subprueba de palabras inventadas



El 100% de los ítems tienen una desviación estándar superior a 0.10.

Tabla 23. Resultado promedio de la Subprueba palabras inventadas

Promedio INVENTADAS	Dificultad	Discriminación
	0,838335	0,265265

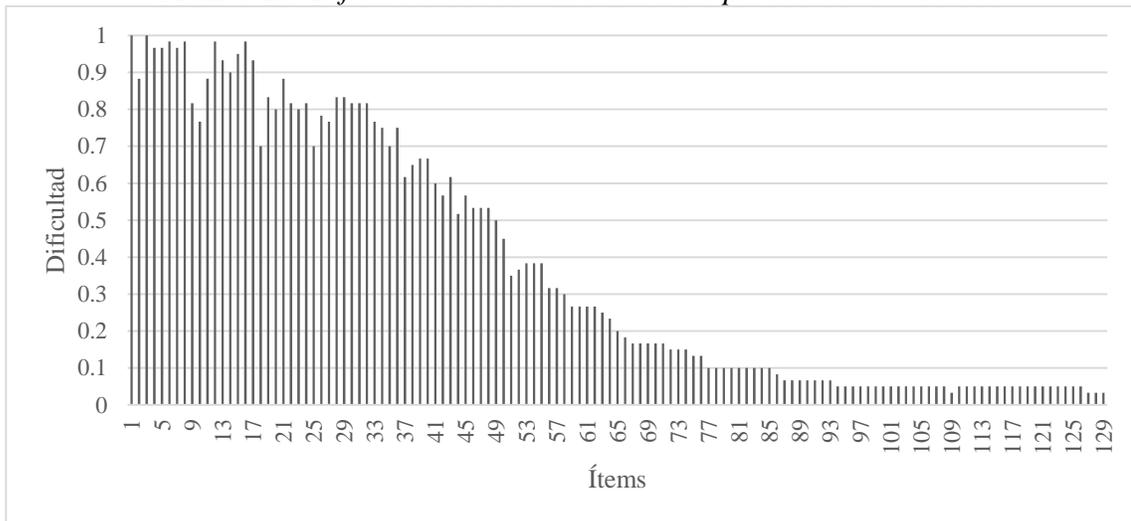
En promedio la Subprueba de palabras inventadas es muy fácil para los estudiantes. La discriminación es justa, ya que se encuentra entre 0.10 y 0.30.

Tabla 24. Ítems más difíciles y que mejor discriminan de la Subprueba de palabras inventadas

Ítem	Dificultad	Ítem	Discriminación
8	0,6167	20	0,4658
3	0,85	15	0,4556
12	0,8333	12	0,4545
15	0,8167	6	0,3858
2	0,7833	5	0,3741
10	0,7833	3	0,3459
13	0,7833	9	0,3225
18	0,7833	18	0,2976
1	0,75	19	0,2869
16	0,7167	1	0,2659

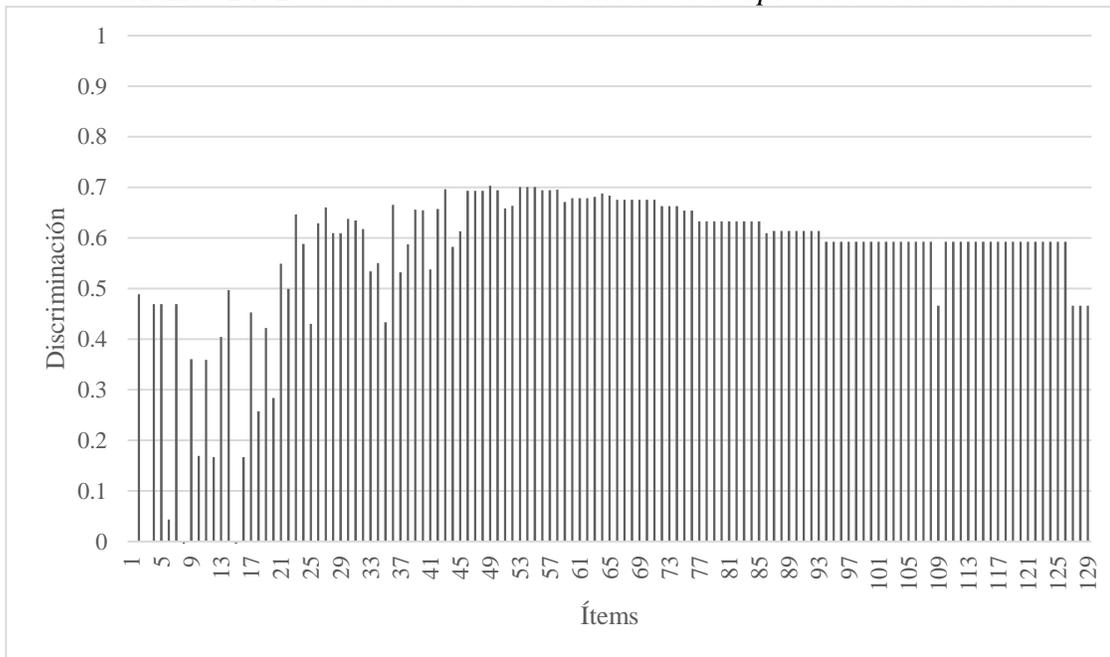
En la Tabla 24 se muestran los ítems de la Subprueba de palabras inventadas, según el indicador de dificultad están ordenados de menor a mayor, se evidencia que para más de la mitad de la población fue fácil responderlos correctamente. Los ítems según el indicador de discriminación están ordenados de mayor a menor, estos son superiores a 0.25.

Gráfica 23. *Dificultad de los ítems en la Subprueba de lectura oral*



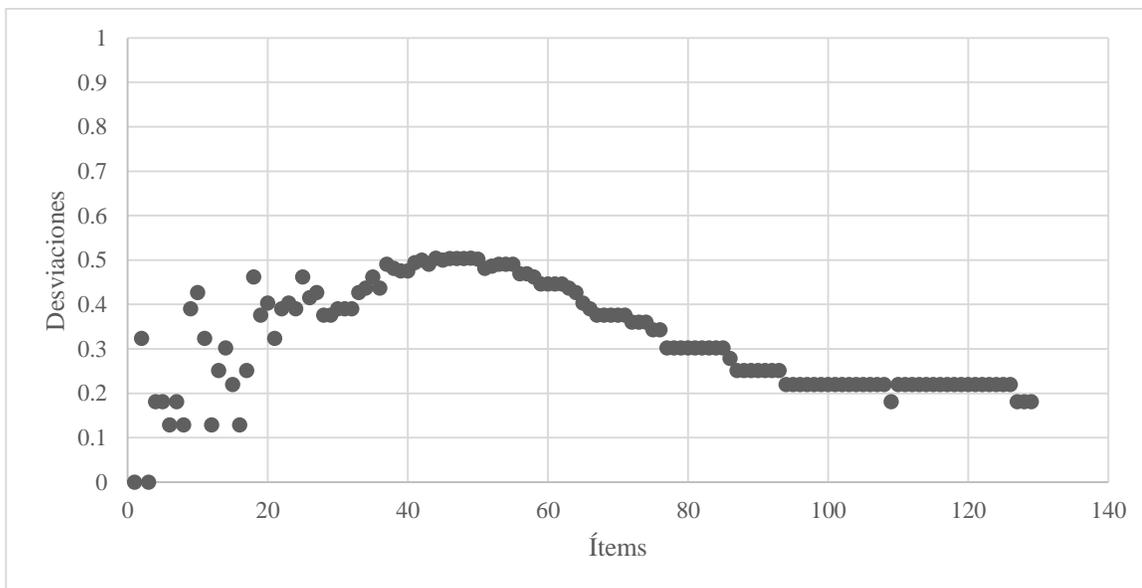
El 39% de los ítems en la Subprueba de lectura oral, tuvieron una dificultad mayor a 0.40 lo cual es un grado de dificultad normal. El 10% tuvo una dificultad entre 0.20 y 0.39 lo cual significa que fueron difíciles para los estudiantes y el 51% de los ítems obtuvieron una dificultad menor a 0.20 lo que indica que la mitad de los ítems fueron muy difíciles de contestar, para los estudiantes de 3° primaria.

Gráfica 24. *Discriminación de los ítems en la Subprueba de lectura oral*



Más de la mitad de los ítems tienen una discriminación superior a 0.20 y uno de los ítems tiene una discriminación menor a 0.10. Existen dos ítems con discriminación negativa, menor a 0.10 lo cual representa el 1% de los ítems en la Subprueba de lectura oral.

Gráfica 25. *Desviación Estándar de los ítems en la Subprueba de lectura oral*



El 98% de los ítems tienen una desviación estándar superior a 0.10 y dos ítems tienen una desviación estándar de 0.

Tabla 25. *Promedio de la Subprueba lectura oral*

Promedio LECTURA	Dificultad	Discriminación
ORAL	0,378426357	0,5408189

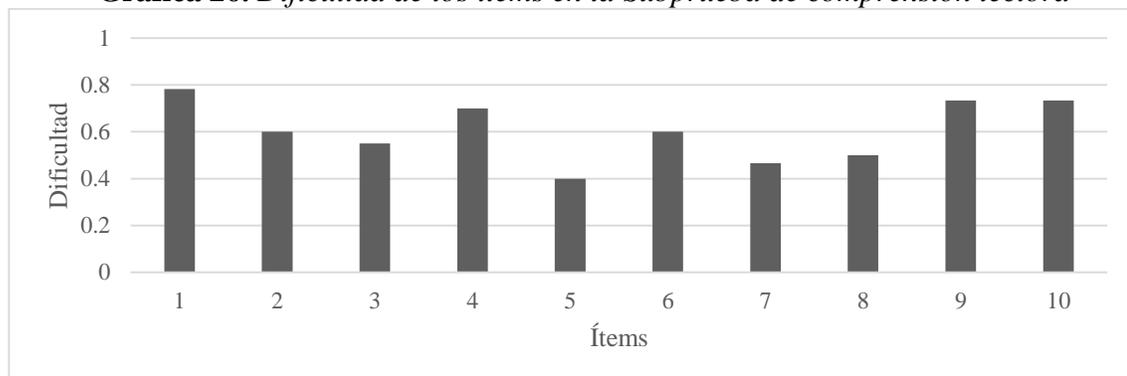
En promedio la Subprueba de lectura oral es difícil para los estudiantes. La discriminación es buena, ya que es mayor a 0.30.

Tabla 26. *Ítems más difíciles y que mejor discriminan de la Subprueba de lectura oral*

Ítem	Dificultad	Ítem	Discriminación
94	0,05	49	0,7034
95	0,05	53	0,7008
96	0,05	54	0,7008
97	0,05	55	0,7008
98	0,05	43	0,6962
99	0,05	58	0,6959
109	0,0333	50	0,6944
127	0,0333	56	0,6944
128	0,0333	57	0,6944
129	0,0333	46	0,6934

En la Tabla 26 se muestran los ítems de la Subprueba de lectura oral, según el indicador de dificultad están ordenados de menor a mayor, se evidencia que para la mitad de la población fue difícil responderlos correctamente. Los ítems según el indicador de discriminación están ordenados de mayor a menor, estos son superiores a 0.69.

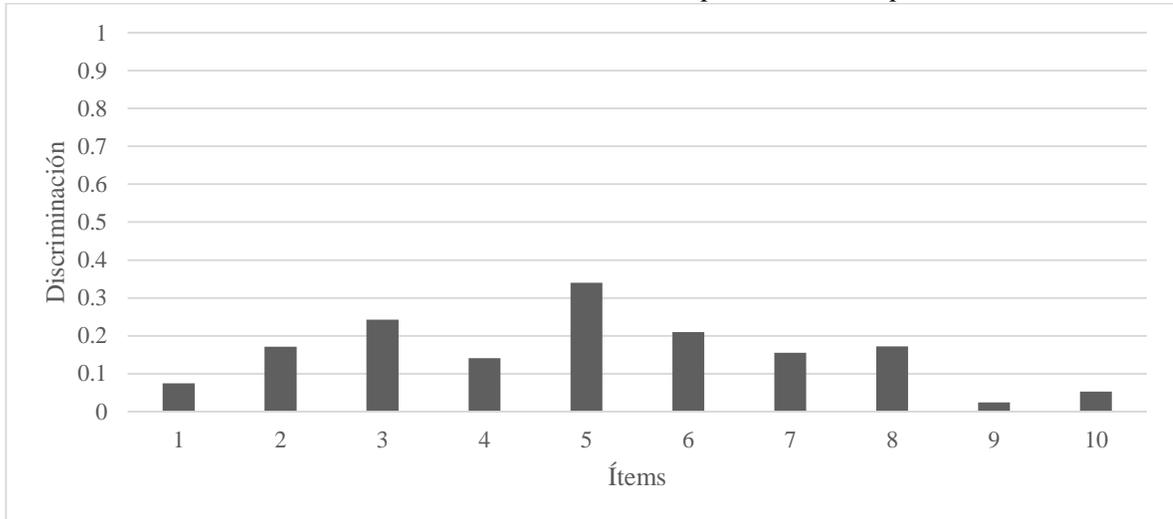
Gráfica 26. *Dificultad de los ítems en la Subprueba de comprensión lectora*



En la mayoría de los ítems de la Subprueba de comprensión lectora, la dificultad es mayor a 0.50 y dos ítems tuvieron un indicador menor a 0.45 lo cual significa que fueron

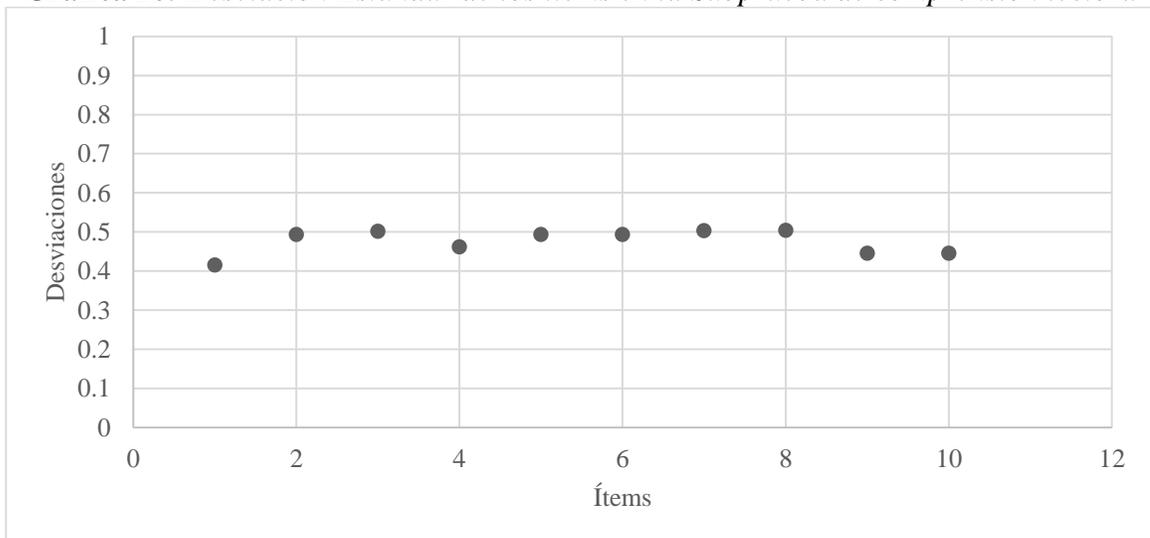
difíciles de contestar, La mayoría de los ítems fueron fáciles de contestar, para los estudiantes de 3° primaria.

Gráfica 27. *Discriminación de los ítems en la Subprueba de comprensión lectora*



El 30% de los ítems tienen una discriminación superior a 0.20 y el 30% de los ítems tienen una discriminación menor a 0.10 en la Subprueba de comprensión lectora.

Gráfica 28. *Desviación Estándar de los ítems en la Subprueba de comprensión lectora*



El 100% de los ítems tienen una desviación estándar superior a 0.10.

Tabla 27. *Promedio de la Subprueba comprensión lectora*

Promedio COMPRENSIÓN	Dificultad	Discriminación
	0,60666	0,15825

En promedio la Subprueba de comprensión lectora es fácil para los estudiantes. La discriminación es justa, ya que se encuentra entre 0.10 y 0.30.

Tabla 28. *Ítems más difíciles y que mejor discriminan de la Subprueba de comprensión lectora*

Ítem	Dificultad	Ítem	Discriminación
1	0,7833	5	0,3396
9	0,7333	3	0,2425
10	0,7333	6	0,2102
4	0,7	8	0,1717
2	0,6	2	0,1711
6	0,6	7	0,1551
3	0,55	4	0,1406
8	0,5	1	0,0748
7	0,4667	10	0,0526
5	0,4	9	0,0243

En la Tabla se muestran los ítems de la Subprueba de comprensión lectora, según el indicador de dificultad están ordenados de menor a mayor, se evidencia que la mitad de los ítems fueron fáciles o normales para responder. Los ítems según el indicador de discriminación están ordenados de mayor a menor, estos son superiores a 0.25.

D. Resultados con Rasch

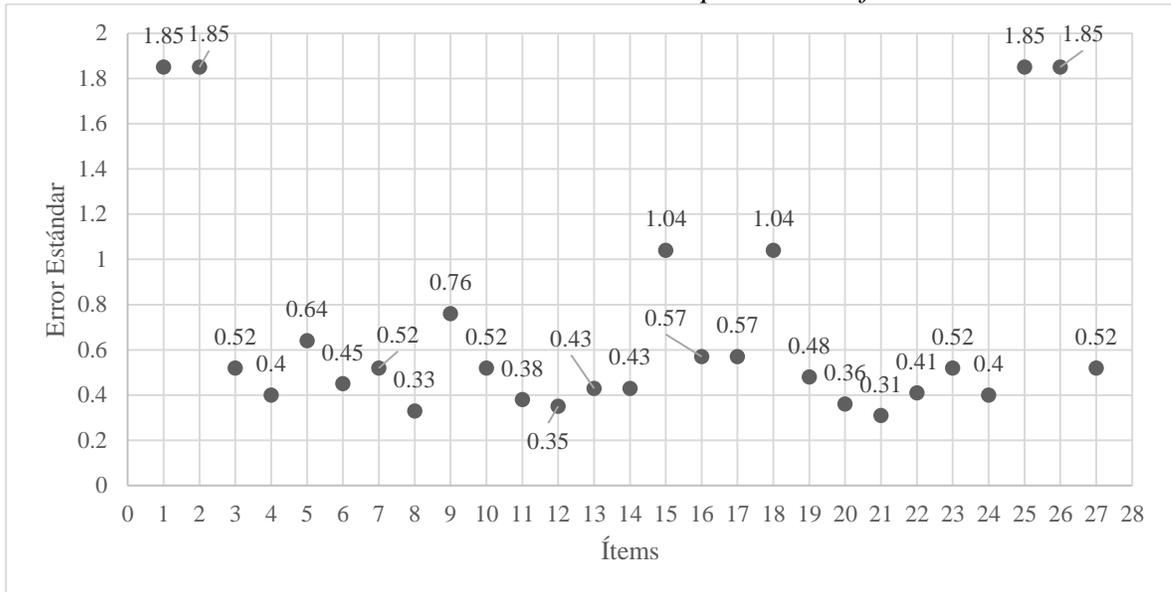
Tabla 29. Estadística de calidad de la escala

Estadísticas	Toda la prueba		Alfabeto		Palabras Simples		Palabras Inventadas		Lectura Oral		Comprensión Lectora	
	Ítems	Personas	Ítems	Personas	Ítems	Personas	Ítems	Personas	Ítems	Personas	Ítems	Personas
Varianza observada	11,649	4,1354	1,2129	1,2863	3,1056	2,0197	0,7284	0,872	36,7762	16,338	0,6878	1,4079
Desviación Estándar observada	3,4131	2,0336	1,1013	1,1342	1,7623	1,4211	0,8535	0,9338	6,0643	4,042	0,8293	1,1866
Mean Square Error	0,3818	0,0662	0,3059	0,6703	0,8866	0,8283	0,2163	0,5522	1,1582	0,1853	0,1182	0,6108
Raíz MSE	0,6179	0,2574	0,5531	0,8187	0,9416	0,9101	0,4651	0,7431	1,0762	0,4305	0,3438	0,7816
Varianza ajustada	11,2671	4,0691	0,9071	0,6161	2,219	1,1914	0,5121	0,3197	35,6179	16,1526	0,5695	0,7971
Desviación Estándar ajustada	3,3567	2,0172	0,9524	0,7849	1,4896	1,0915	0,7156	0,5655	5,9681	4,019	0,7547	0,8928
Índice de separación	5,432	7,8383	1,722	0,9587	1,582	1,1993	1,5386	0,7609	5,5454	9,3353	2,1948	1,1424
Número de estrato	7,5761	10,7844	2,6293	1,6116	2,4427	1,9324	2,3848	1,3479	7,7272	12,7804	3,2598	1,8565
Confiabilidad	0,9672	0,984	0,7478	0,4789	0,7145	0,5899	0,703	0,3667	0,9685	0,9887	0,8281	0,5662

El error cuadrático medio (RMSE) de la prueba total es de 0.38 en todos los ítems de la prueba y de 0.06 en los estudiantes, lo cual es menor al parámetro establecido por Linacre. En la Subprueba de alfabeto el RMSE es de 0.30 en todos los ítems de la Subprueba y de 0.67 en los estudiantes. Para la Subprueba de palabras simples el RMSE es de 0.88 en todos los ítems de la Subprueba y de 0.82 en los estudiantes. En la Subprueba de palabras inventadas el RMSE es de 0.21 en todos los ítems de la Subprueba y de 0.55 en los estudiantes. En la Subprueba de lectura oral el RMSE es de 1.15 en todos los ítems de la Subprueba y de 0.18 en los estudiantes. Y para la Subprueba de comprensión lectora el RMSE es de 0.11 en todos los ítems de la Subprueba y de 0.61 en los estudiantes. Al analizar los RMSE se interpretan con relación a que, mientras más bajo es, el error disminuye.

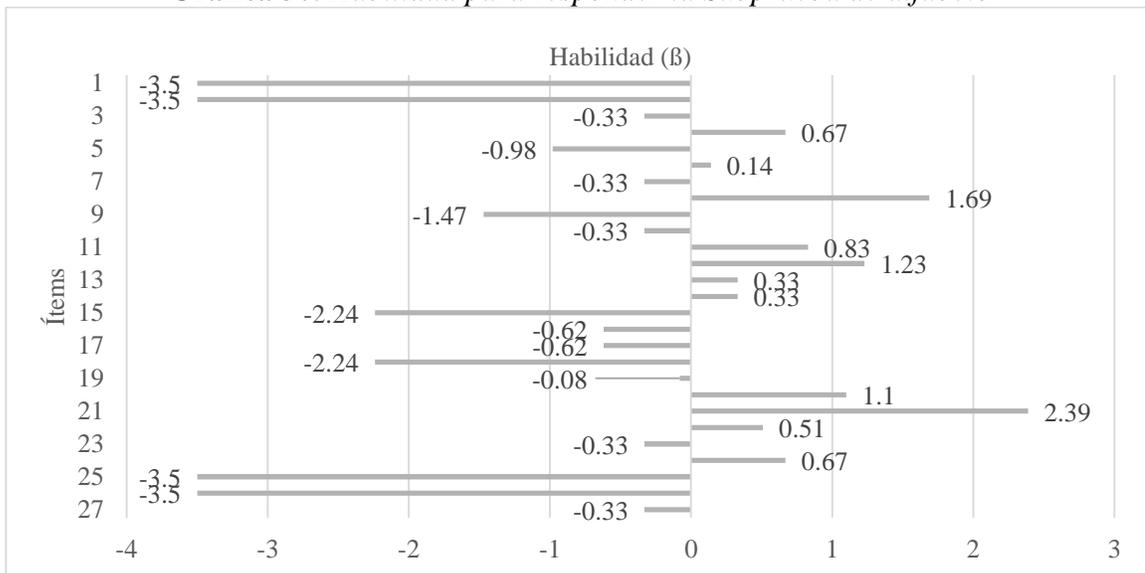
E. Indicadores de ajuste del modelo

Gráfica 29. Error Estándar de la Subprueba de alfabeto



El 41% de los ítems (11 ítems) en la Subprueba de alfabeto, se encuentran por debajo del criterio de MRSE que es 0.48 y más de la mitad de los ítems tiene un error mayor a 0.48.

Gráfica 30. Habilidad para responder la Subprueba de alfabeto



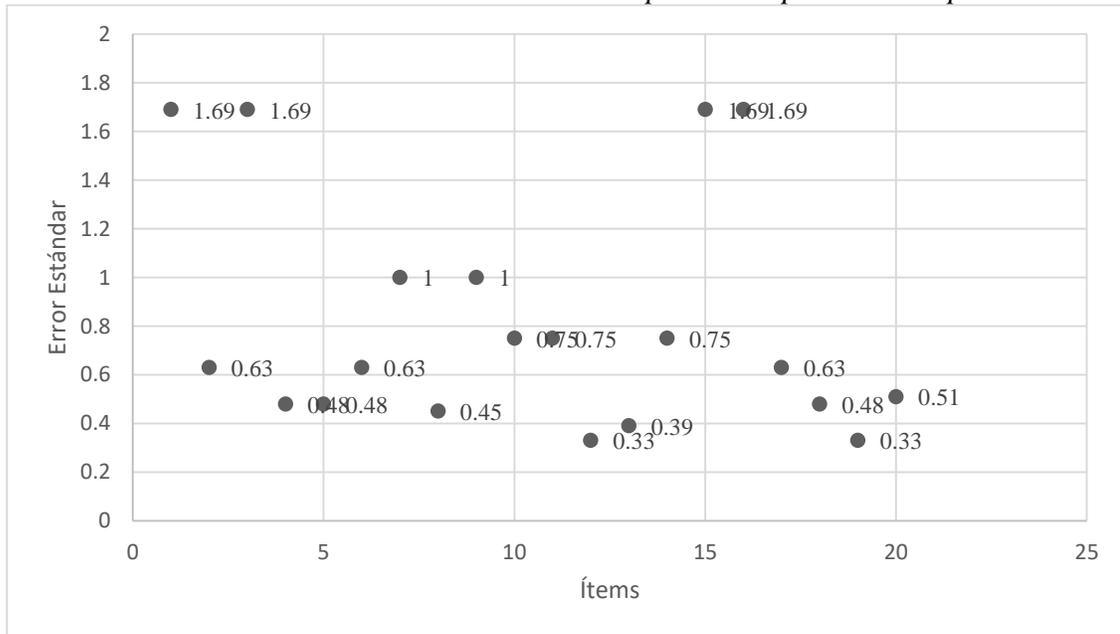
El 60% de los ítems de la Subprueba de alfabeto requirieron de una habilidad negativa, es decir fueron respondidos fácilmente por los estudiantes de 3° Primaria. El 15% de los ítems requirieron de una habilidad mayor a uno para ser respondidos correctamente.

Tabla 30. Descriptivos del WMS y UMS estándar de la Subprueba de alfabeto

	Dificultad de alfabeto	Std. WMS	Std. UMS
Media	-0,51888889	-0,02518519	-0,23703704
Error típico	0,314832445	0,16086161	0,16650575
Mediana	-0,33	-0,08	-0,22
Moda	-0,33	-1,03	-1,27
Desviación estándar	1,63591737	0,83586146	0,86518924
Varianza de la muestra	2,676225641	0,69866439	0,74855242
Curtosis	-0,296884839	-0,50140762	-0,9273464
Coefficiente de asimetría	-0,603561189	0,54211171	0,35751393
Rango	5,89	2,84	2,89
Mínimo	-3,5	-1,12	-1,47
Máximo	2,39	1,72	1,42
Suma	-14,01	-0,68	-6,4
Cuenta	27	27	27
Nivel de confianza (95,0%)	0,647147358	0,33065578	0,34225746

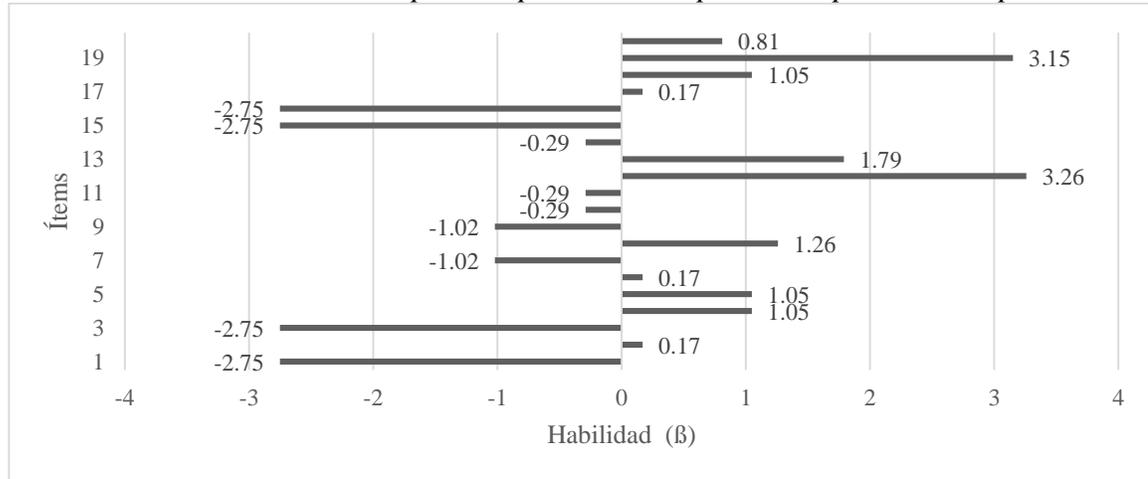
La dificultad promedio observada en la Subprueba de alfabeto fue de -0.518 la cual es baja. La mediana de la dificultad es -0.33 la cual es negativa. En el WMS estandarizado el valor es de -0.08 y el UMS estandarizado es de -0.22. Se obtuvo una desviación estándar de 1.63. En los indicadores de ajuste, el promedio de WMS fue de -0.02 y en el UMS de -0.23 lo cual demuestra que el modelo está razonablemente ajustado.

Gráfica 31. Error Estándar de la Subprueba de palabras simples



El 20% de los ítems (4 ítems) en la Subprueba de palabras simples, se encuentran por debajo del criterio de MRSE que es 0.48 y más de la mitad de los ítems tiene un error mayor a 0.48.

Gráfica 32. *Habilidad para responder la Subprueba de palabras simples*



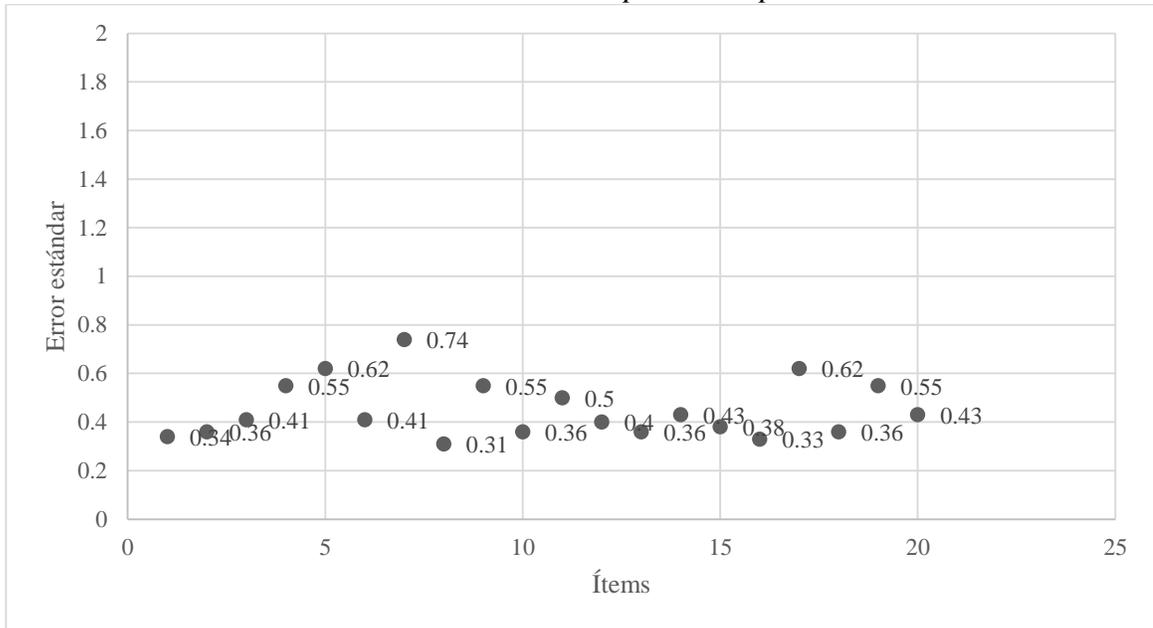
El 45% de los ítems de la Subprueba de palabras simples requirieron de una habilidad negativa, es decir fueron respondidos fácilmente por los estudiantes de 3° Primaria. El 35% de los ítems requirieron de una habilidad mayor a uno para ser respondidos correctamente.

Tabla 31. *Descriptivos del WMS y UMS estándar de la Subprueba palabras simples*

	Dificultad de Palabras Simples	Std. WMS	Std. UMS
Media	0,001	0,011	0,0985
Error típico	0,403847222	0,18379322	0,22959607
Mediana	0,17	0,07	0,01
Moda	-2,75	-0,63	-1,19
Desviación estándar	1,806059683	0,82194826	1,02678484
Varianza de la muestra	3,261851579	0,67559895	1,05428711
Curtosis	-0,42107237	2,12449261	0,66967899
Coefficiente de asimetría	-0,078544386	-0,09072254	0,76632885
Rango	6,01	3,99	3,92
Mínimo	-2,75	-2,08	-1,19
Máximo	3,26	1,91	2,73
Suma	0,02	0,22	1,97
Cuenta	20	20	20
Nivel de confianza (95,0%)	0,845261951	0,38468363	0,4805501

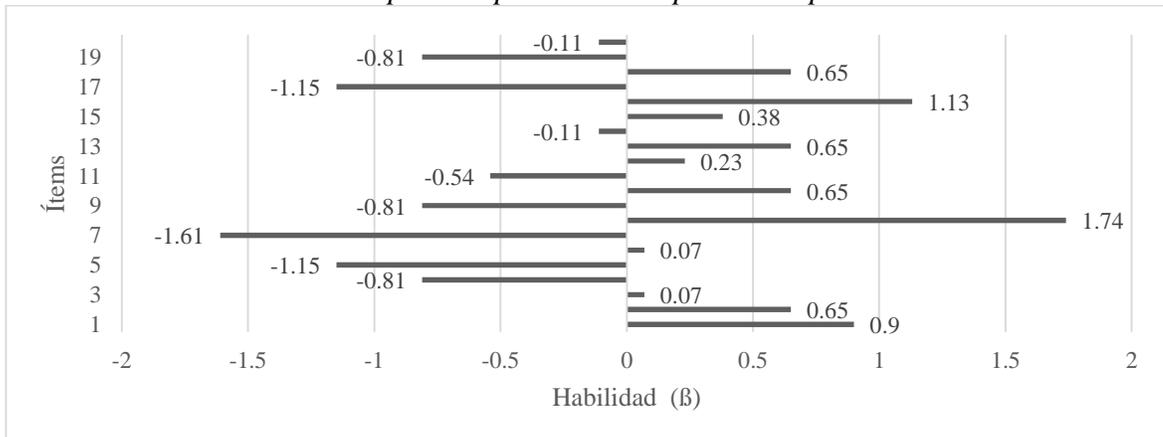
La dificultad promedio observada en la Subprueba de palabras simples fue de 0.001 la cual es baja. La mediana de la dificultad es 0.17 la cual es positiva. En el WMS estandarizado el valor es de 0.07 y el UMS estandarizado es de 0.01. Se obtuvo una desviación estándar de 1.80. En los indicadores de ajuste, el promedio de WMS fue de 0.011 y en el UMS de 0.09 lo cual demuestra que el modelo está razonablemente ajustado.

Gráfica 33. Error Estándar de la Subprueba de palabras inventadas



El 70% de los ítems (14 ítems) en la Subprueba de palabras inventadas, se encuentran por debajo del criterio de MRSE que es 0.48 y menos de la mitad de los ítems tiene un error mayor a 0.48.

Gráfica 34. Habilidad para responder la Subprueba de palabras inventadas



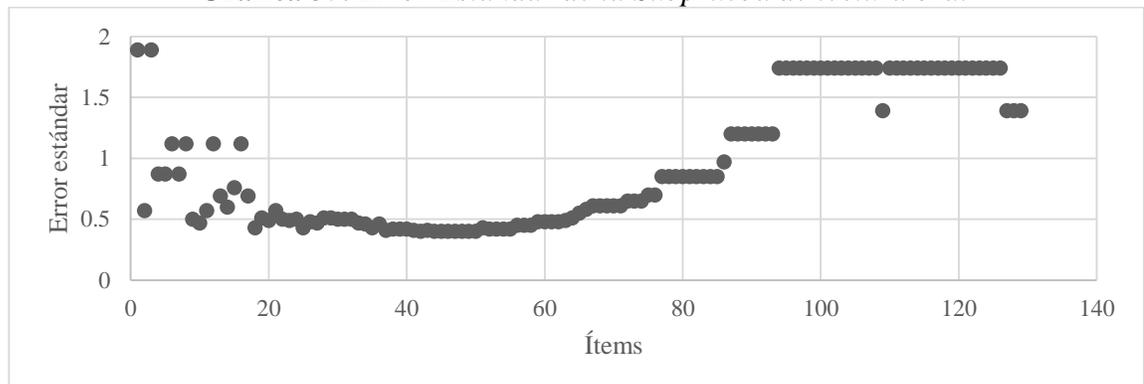
El 45% de los ítems de la Subprueba de palabras inventadas requirieron de una habilidad negativa, es decir fueron respondidos fácilmente por los estudiantes de 3° Primaria. El 10% de los ítems requirieron de una habilidad mayor a uno para ser respondidos correctamente.

Tabla 32. *Descriptivos del WMS y UMS estándar de la Subprueba palabras inventadas*

	Dificultad de palabras inventadas	Std. WMS	Std. UMS
Media	0,001	0,0655	0,071
Error típico	0,195561136	0,13972619	0,21767902
Mediana	0,07	0,22	0,01
Moda	0,65		
Desviación estándar	0,874575987	0,62487451	0,97349019
Varianza de la muestra	0,764883158	0,39046816	0,94768316
Curtosis	-0,602215358	-0,85894716	-1,42051387
Coefficiente de asimetría	-0,035574063	-0,04717363	0,10944042
Rango	3,35	2,24	3,05
Mínimo	-1,61	-0,96	-1,3
Máximo	1,74	1,28	1,75
Suma	0,02	1,31	1,42
Cuenta	20	20	20
Nivel de confianza (95,0%)	0,409314162	0,29245027	0,45560744

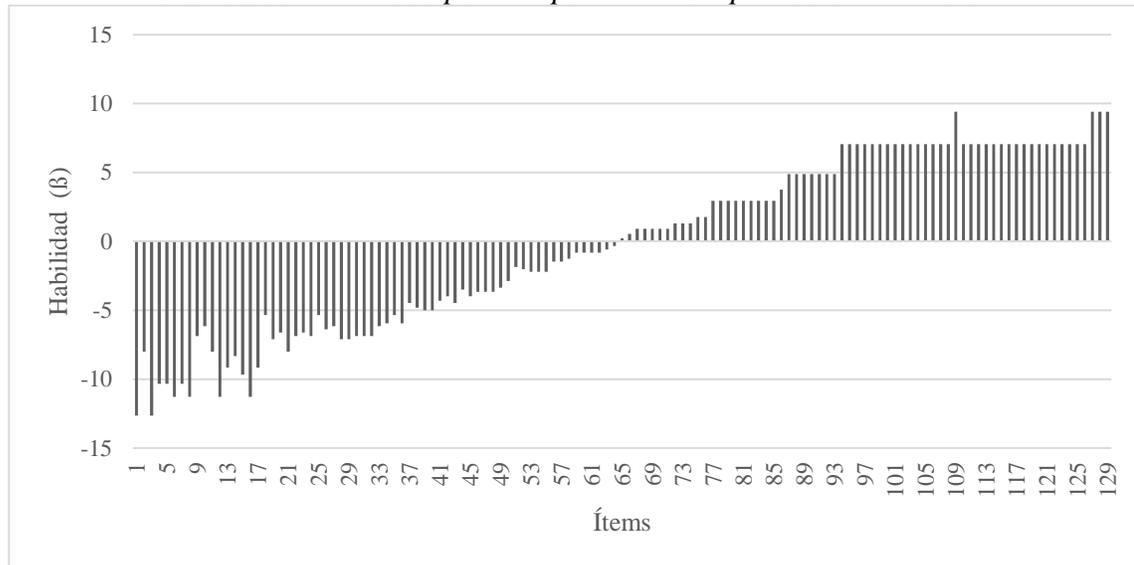
La dificultad promedio observada en la Subprueba de palabras inventadas fue de 0.001 la cual es baja. La mediana de la dificultad es 0.07 la cual es positiva. En el WMS estandarizado el valor es de 0.22 y el UMS estandarizado es de 0.01. Se obtuvo una desviación estándar de 0.87. En los indicadores de ajuste, el promedio de WMS fue de 0.06 y en el UMS de 0.07 lo cual demuestra que el modelo está razonablemente ajustado.

Gráfica 35. *Error Estándar de la Subprueba de lectura oral*



El 23% de los ítems (30 ítems) en la Subprueba de lectura oral, se encuentran por debajo del criterio de MRSE que es 0.48 y más de la mitad de los ítems tiene un error mayor a 0.48.

Gráfica 36. *Habilidad para responder la Subprueba de lectura oral*



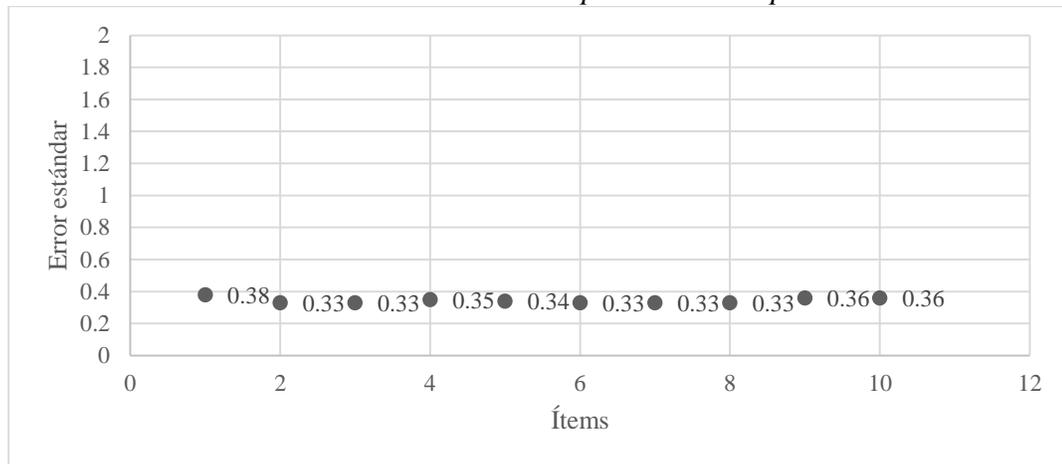
El 50% de los ítems de la Subprueba de lectura oral requirieron una habilidad negativa, es decir fueron respondidos fácilmente por los estudiantes de 3° Primaria. El 45% de los ítems requirieron de una habilidad mayor a uno para ser respondidos correctamente.

Tabla 33. *Descriptivos del WMS y UMS estándar de la Subprueba lectura oral*

	Dificultad de Lectura Oral	Std. WMS	Std. UMS
Media	-0,195271318	-0,92170543	5,28178295
Error típico	0,549598008	0,13009212	2,70642261
Mediana	0,22	-0,78	-0,71
Moda	7,06	-0,69	-1,25
Desviación estándar	6,242233434	1,47756246	30,7390519
Varianza de la muestra	38,96547825	2,18319082	944,889315
Curtosis	-1,237277205	3,0142867	77,456572
Coefficiente de asimetría	-0,178708886	1,19504974	8,24484285
Rango	22,06	8,26	311,18
Mínimo	-12,65	-3,26	-1,28
Máximo	9,41	5	309,9
Suma	-25,19	-118,9	681,35
Cuenta	129	129	129
Nivel de confianza (95,0%)	1,087473558	0,25740949	5,35511954

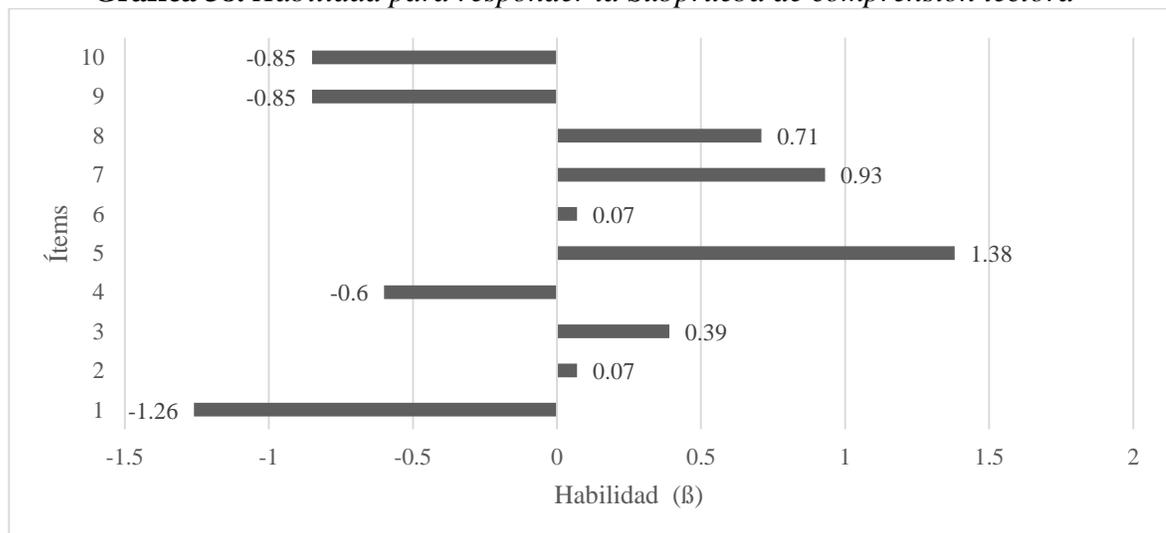
La dificultad promedio observada en la Subprueba de lectura oral fue de -0.19 la cual es baja. La mediana de la dificultad es 0.22 la cual es positiva. En el WMS estandarizado el valor es de -0.78 y el UMS estandarizado es de -0.71. Se obtuvo una desviación estándar de 6.24. En los indicadores de ajuste, el promedio de WMS fue de -0.9 y en el UMS de 5.28 lo cual demuestra que el modelo está razonablemente ajustado.

Gráfica 37. *Error Estándar de la Subprueba de comprensión lectora*



El total de los ítems (4 ítems) en la Subprueba de comprensión lectora, se encuentran por debajo del criterio de MRSE que es 0.48.

Gráfica 38. *Habilidad para responder la Subprueba de comprensión lectora*



El 40% de los ítems de la Subprueba de comprensión lectora requirieron de una habilidad negativa, es decir fueron respondidos fácilmente por los estudiantes de 3°

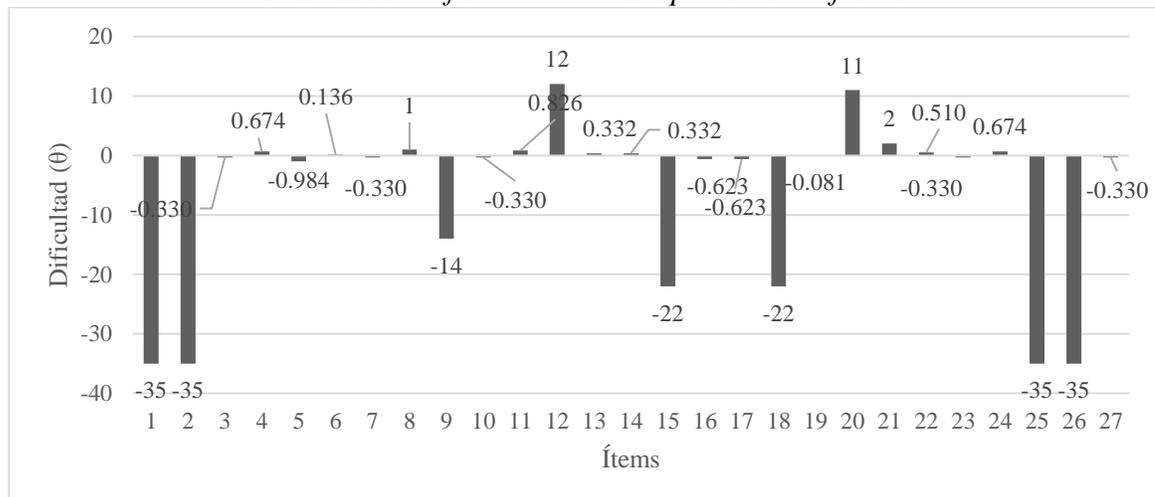
Primaria. El 10% de los ítems requirieron de una habilidad mayor a uno para ser respondidos correctamente.

Tabla 34. *Descriptivos del WMS y UMS estándar de la Subprueba comprensión lectora*

	Dificultad de Comprensión Lectora	Std. WMS	Std. UMS
Media	-0,001	0,037	-0,09
Error típico	0,275436825	0,28262677	0,24716167
Mediana	0,07	0,2	0,03
Moda	0,07		
Desviación estándar	0,871007718	0,89374431	0,78159381
Varianza de la muestra	0,758654444	0,79877889	0,61088889
Curtosis	-1,173742187	0,79562101	0,83422301
Coefficiente de asimetría	0,089581364	-0,65829798	-1,07106011
Rango	2,64	3,13	2,51
Mínimo	-1,26	-1,78	-1,71
Máximo	1,38	1,35	0,8
Suma	-0,01	0,37	-0,9
Cuenta	10	10	10
Nivel de confianza (95,0%)	0,623081386	0,63934616	0,55911853

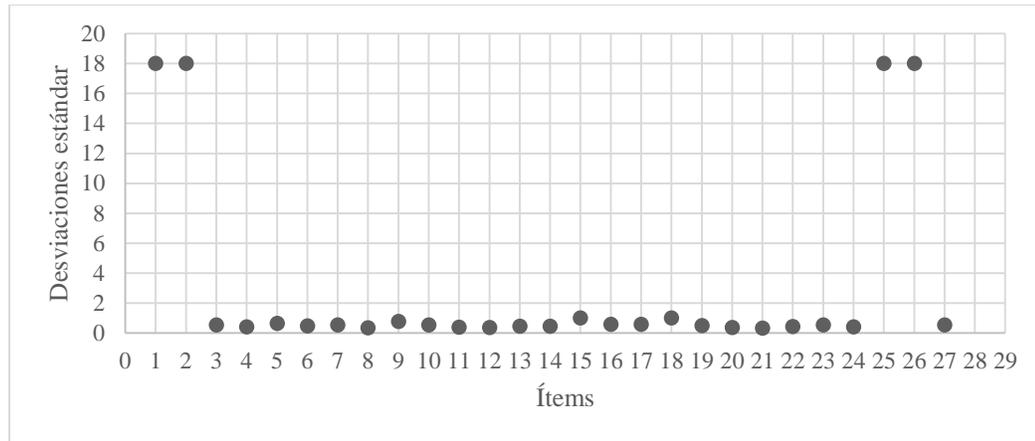
La dificultad promedio observada en la Subprueba de comprensión lectora fue de -0.001 la cual es baja. La mediana de la dificultad es 0.07 la cual es positiva. En el WMS estandarizado el valor es de 0.20 y el UMS estandarizado es de 0.03. Se obtuvo una desviación estándar de 0.87. En los indicadores de ajuste, el promedio de WMS fue de 0.04 y en el UMS de -0.09 lo cual demuestra que el modelo está razonablemente ajustado.

Gráfica 39. *Dificultad en la Subprueba de alfabeto*



El 59% (16 ítems) de los ítems tuvieron un indicador de habilidad menor a 0 y de ellos el 22% de los ítems (6 ítems) se encuentran a más de dos logits, lo cual indica que se requirió una menor habilidad para responder los ítems de la Subprueba de alfabeto.

Gráfica 40. *Desviación Estándar en la Subprueba alfabeto*



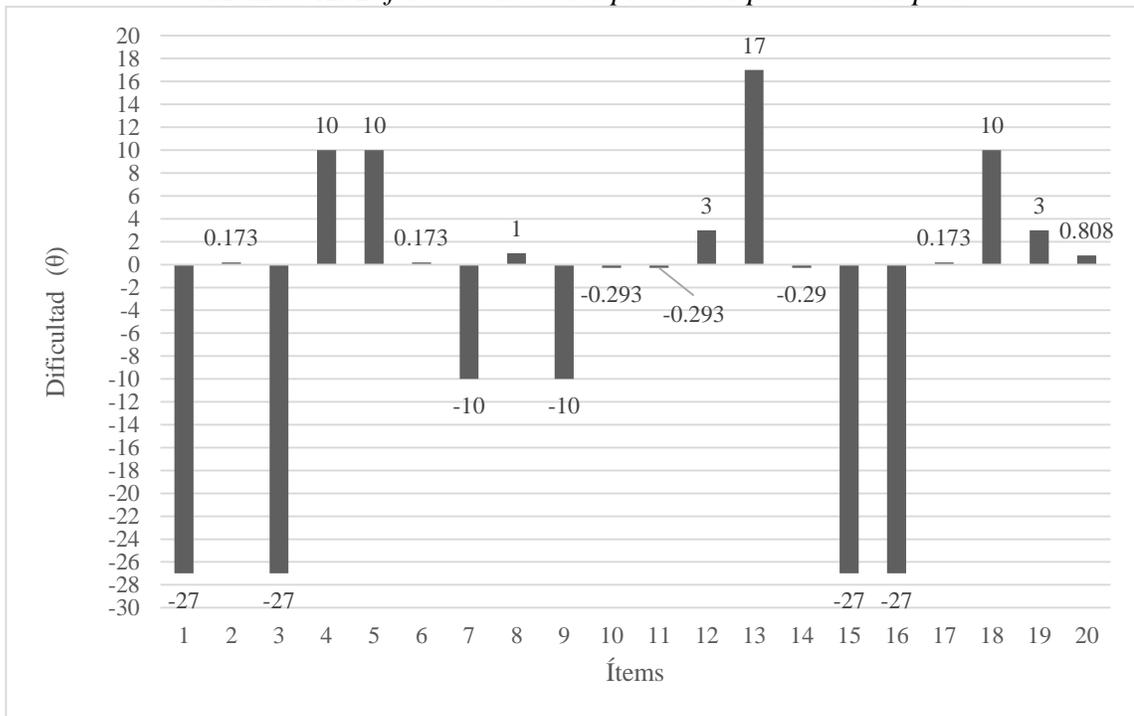
El parámetro de habilidad de los ítems en la Subprueba de alfabeto se encuentra a menos de dos desviaciones estándar.

Tabla 35. *Estadísticos descriptivos del modelo de habilidad de la Subprueba de alfabeto*

	Theta θ	Desviación Estándar
Media	-6.388037037	3.105518519
Error típico	2.73121556	1.218602134
Mediana	-0.33	0.517
Moda	-0.33	0.517
Desviación estándar	14.19181235	6.33204243
Varianza de la muestra	201.4075378	40.09476134
Curtosis	0.256278299	2.586559548
Coefficiente de asimetría	-1.236404269	2.096164764
Rango	47	17.694
Mínimo	-35	0.306
Máximo	12	18
Suma	-172.477	83.849
Cuenta	27	27
Nivel de confianza (95.0%)	5.614093988	2.50487256

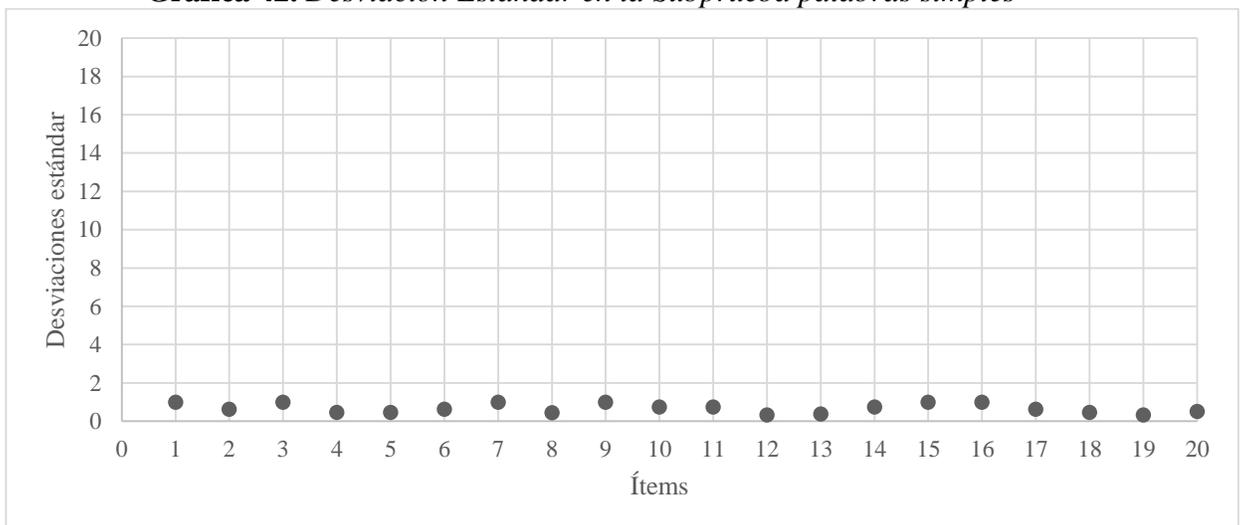
La mediana fue de -0.33 y el promedio de Theta θ fue de -6.38 lo cual indica que se requirió menor habilidad para responder la Subprueba de alfabeto.

Gráfica 41. Dificultad en la Subprueba de palabras simples



El 45% (9 ítems) de los ítems tuvieron un indicador de habilidad menor a 0 y de ellos el 20% de los ítems (6 ítems) se encuentran a más de dos lógit, lo cual indica que se requirió una menor habilidad para responder los ítems de la Subprueba de palabras simples.

Gráfica 42. Desviación Estándar en la Subprueba palabras simples



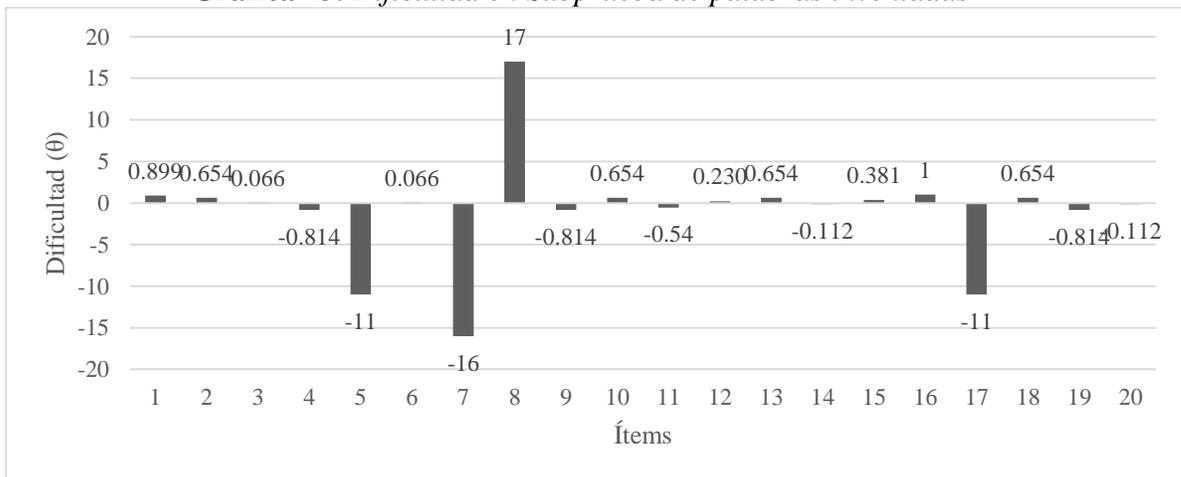
El parámetro de habilidad de los ítems en la Subprueba de palabras simples se encuentra a menos de dos desviaciones estándar.

Tabla 36. Estadísticos descriptivos del modelo de habilidad de la Subprueba de palabras simples

	Theta θ	Desviación Estándar
Media	-3.67745	0.67745
Error típico	3.01510482	0.0553963
Mediana	0.173	0.629
Moda	-27	1
Desviación estándar	13.4839587	0.24773978
Varianza de la muestra	181.817141	0.061375
Curtosis	-0.30416459	-1.46375033
Coefficiente de asimetría	-0.80284464	0.19154677
Rango	44	0.671
Mínimo	-27	0.329
Máximo	17	1
Suma	-73.549	13.549
Cuenta	20	20
Nivel de confianza (95,0%)	6.31068691	0.11594578

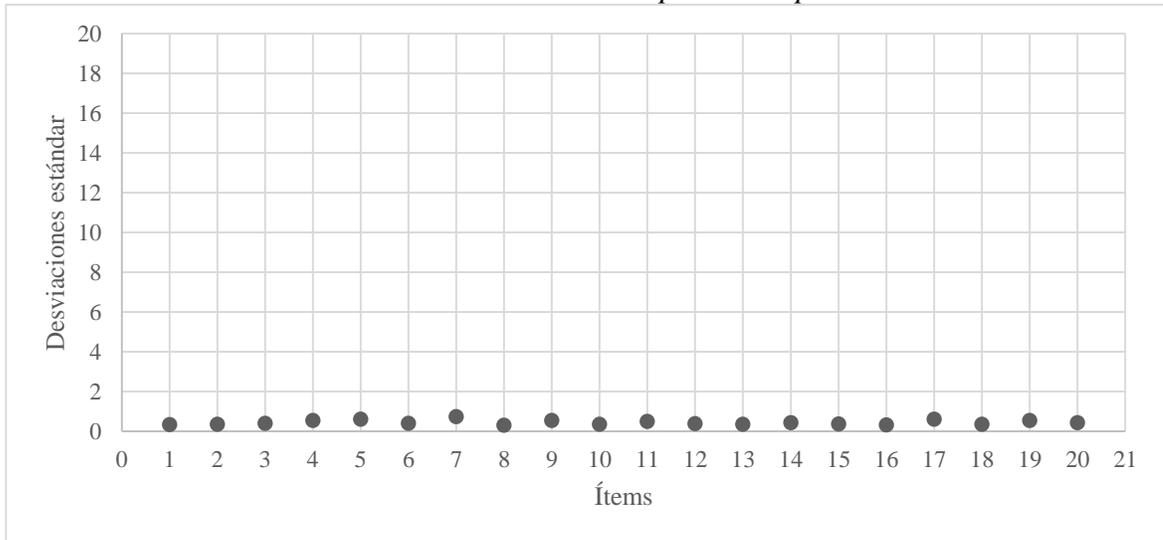
La mediana fue de 0.17 y el promedio de Theta θ fue de -3.67 lo cual indica que se requirió menor habilidad para responder la Subprueba de palabras simples.

Gráfica 43. Dificultad en Subprueba de palabras inventadas



El 45% (9 ítems) de los ítems tuvieron un indicador de habilidad menor a 0 y de ellos el 15% de los ítems (3 ítems) se encuentran a más de dos lógit, lo cual indica que se requirió una menor habilidad para responder los ítems de la Subprueba de palabras inventadas.

Gráfica 44. *Desviación Estándar en la Subprueba de palabras inventadas*



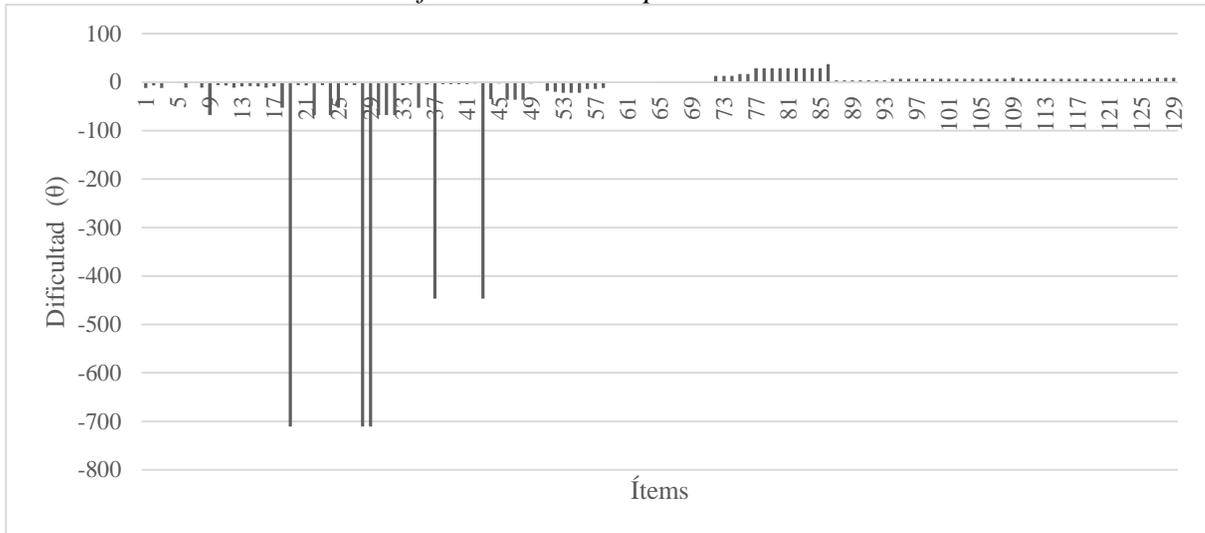
El parámetro de habilidad de los ítems en la Subprueba de palabras inventadas se encuentra a menos de dos desviaciones estándar.

Tabla 37. *Estadísticos descriptivos del modelo de habilidad de la Subprueba de palabras inventadas*

	Theta θ	Desviación Estándar
Media	-0.9474	0.45005
Error típico	1.42817419	0.02631314
Mediana	0.066	0.413
Moda	0.654	0.358
Desviación estándar	6.38698915	0.11767595
Varianza de la muestra	40.7936305	0.01384763
Curtosis	4.08261025	0.33898188
Coefficiente de asimetría	0.08440848	1.00963441
Rango	33	0.432
Mínimo	-16	0.311
Máximo	17	0.743
Suma	-18.948	9.001
Cuenta	20	20
Nivel de confianza (95,0%)	2.98920294	0.05507404

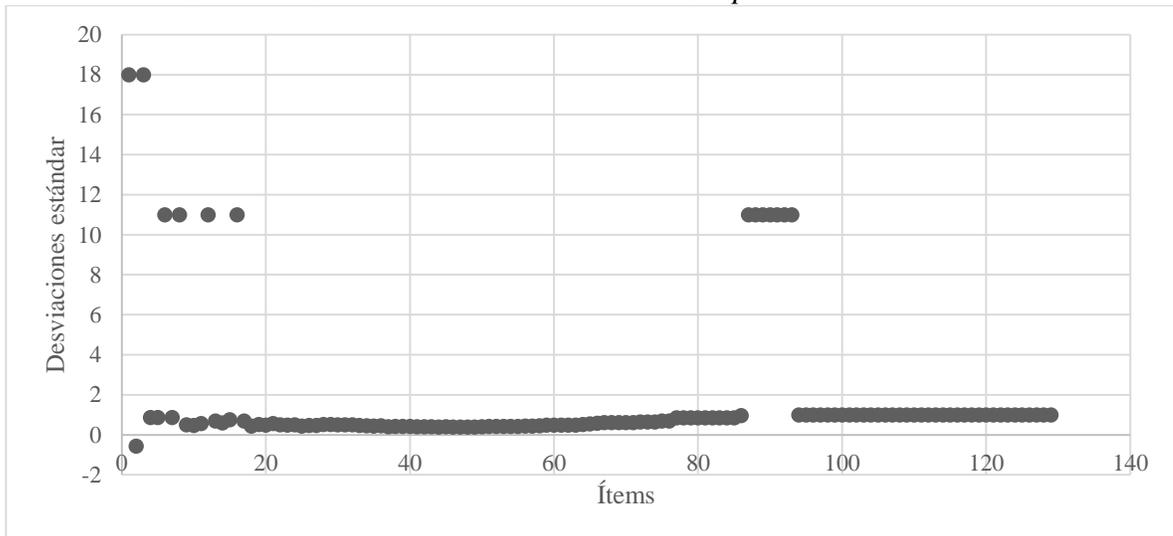
La mediana fue de 0.06 y el promedio de Theta θ fue de -0.94 lo cual indica que se requirió menor habilidad para responder la Subprueba de palabras inventadas.

Gráfica 45. Dificultad en la Subprueba de lectura oral



El 42% (54 ítems) de los ítems tuvieron un indicador de habilidad menor a 0 y de ellos el 17% de los ítems (22 ítems) se encuentran a más de dos lógit, lo cual indica que se requirió una menor habilidad para responder los ítems de la Subprueba de lectura oral.

Gráfica 46. Desviación Estándar en la Subprueba lectura oral



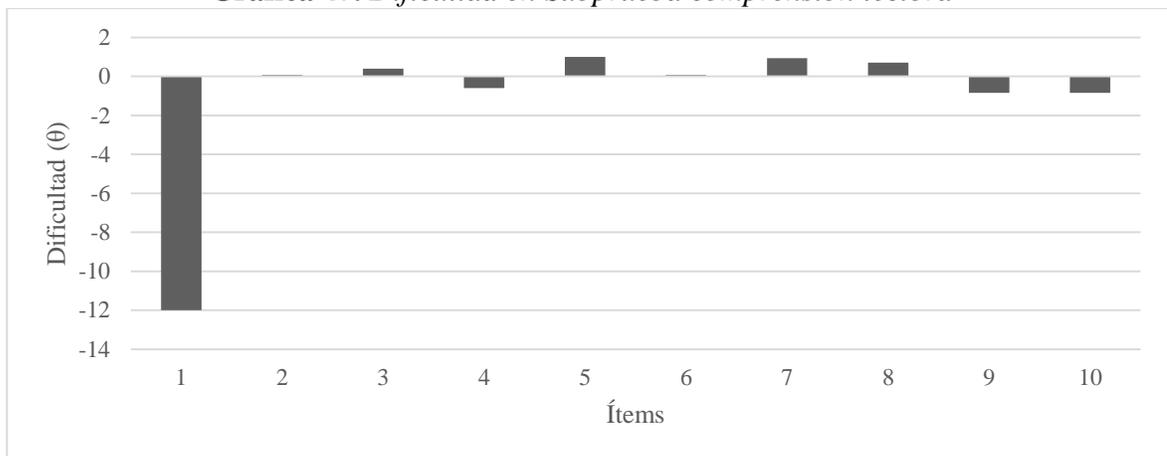
El parámetro de habilidad de la mayoría de los ítems en la Subprueba de lectura oral, se encuentra a menos de dos desviaciones estándar. El 6% (8 ítems) se encuentran entre 12 y 11 desviaciones estándar y el 1% se encuentran a 18 desviaciones estándar.

Tabla 38. Estadísticos descriptivos del modelo de habilidad de la Subprueba de lectura oral

	Theta θ	Desviación Estándar
Media	-26.519155	1.83455814
Error típico	10.6790984	0.31205477
Mediana	0.223	0.699
Moda	7	1
Desviación estándar	121.291241	3.5442609
Varianza de la muestra	14711.5653	12.5617853
Curtosis	24.2124419	8.05684621
Coefficiente de asimetría	-4.92292487	2.97877784
Rango	748	18.569
Mínimo	-711	-0.569
Máximo	37	18
Suma	-3420.971	236.658
Cuenta	129	129
Nivel de confianza (95,0%)	21.1304206	0.61745368

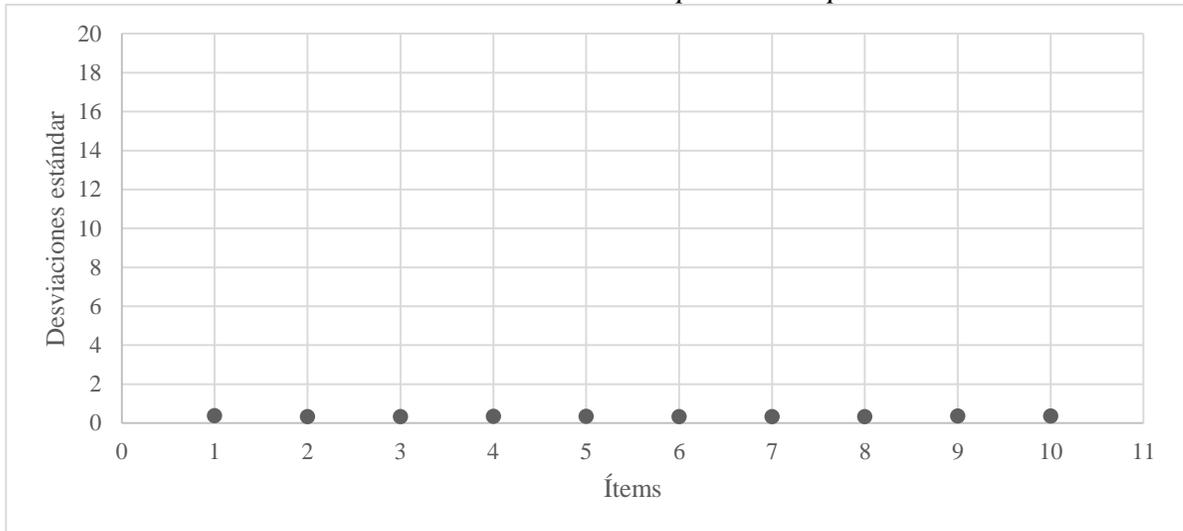
La mediana fue de 0.22 y el promedio de Theta θ fue de -26.51 lo cual indica que se requirió menor habilidad para responder la Subprueba de comprensión lectora.

Gráfica 47. Dificultad en Subprueba comprensión lectora



El 30% (3 ítems) de los ítems tuvieron un indicador de habilidad menor a 0 y de ellos el 1% de los ítems (1 ítem) se encuentran a más de dos lógit, lo cual indica que se requirió una menor habilidad para responder los ítems de la Subprueba de comprensión lectora.

Gráfica 48. *Desviación Estándar en la Subprueba comprensión lectora*



El parámetro de habilidad de la mayoría de los ítems en la Subprueba de comprensión lectora se encuentra a menos de dos desviaciones estándar.

Tabla 39. *Estadísticos descriptivos del modelo de la habilidad de la Subprueba de comprensión lectora*

	Theta θ	Desviación Estándar
Media	-1.1122	0.3428
Error típico	1.22921664	0.00602366
Mediana	0.074	0.337
Moda	0.074	0.328
Desviación estándar	3.88712432	0.01904848
Varianza de la muestra	15.1097355	0.00036284
Curtosis	9.14989281	0.64089485
Coefficiente de asimetría	-2.97893259	1.08310619
Rango	13	0.058
Mínimo	-12	0.325
Máximo	1	0.383
Suma	-11.122	3.428
Cuenta	10	10
Nivel de confianza (95,0%)	2.78068123	0.01362646

La mediana fue de 0.07 y el promedio de Theta θ fue de -1.11 lo cual indica que se requirió menor habilidad para responder la Subprueba de comprensión lectora.

F. Análisis diferencial de los ítems

Al realizar el análisis diferencial de los ítems con cada Subprueba, no se encontró una diferencia que beneficie a un grupo de acuerdo con la característica que posee. Ver las Tablas en el anexo 7.

G. Confiabilidad

1. Confiabilidad total

Tabla 40. *Confiabilidad de la prueba en la población basado en los elementos tipificados*

Todas las Subpruebas	Alfabeto	Palabras simples	Palabras inventadas	Lectura oral	Comprensión lectora
,982	,839	,890	,829	,988	,764

Los intervalos de confianza se encuentran en 0,9714 y 0,9862 lo cual indica que son cercanos a la estimación que corresponde a 0,9794 por lo tanto se concluye que la confiabilidad es alta y estable.

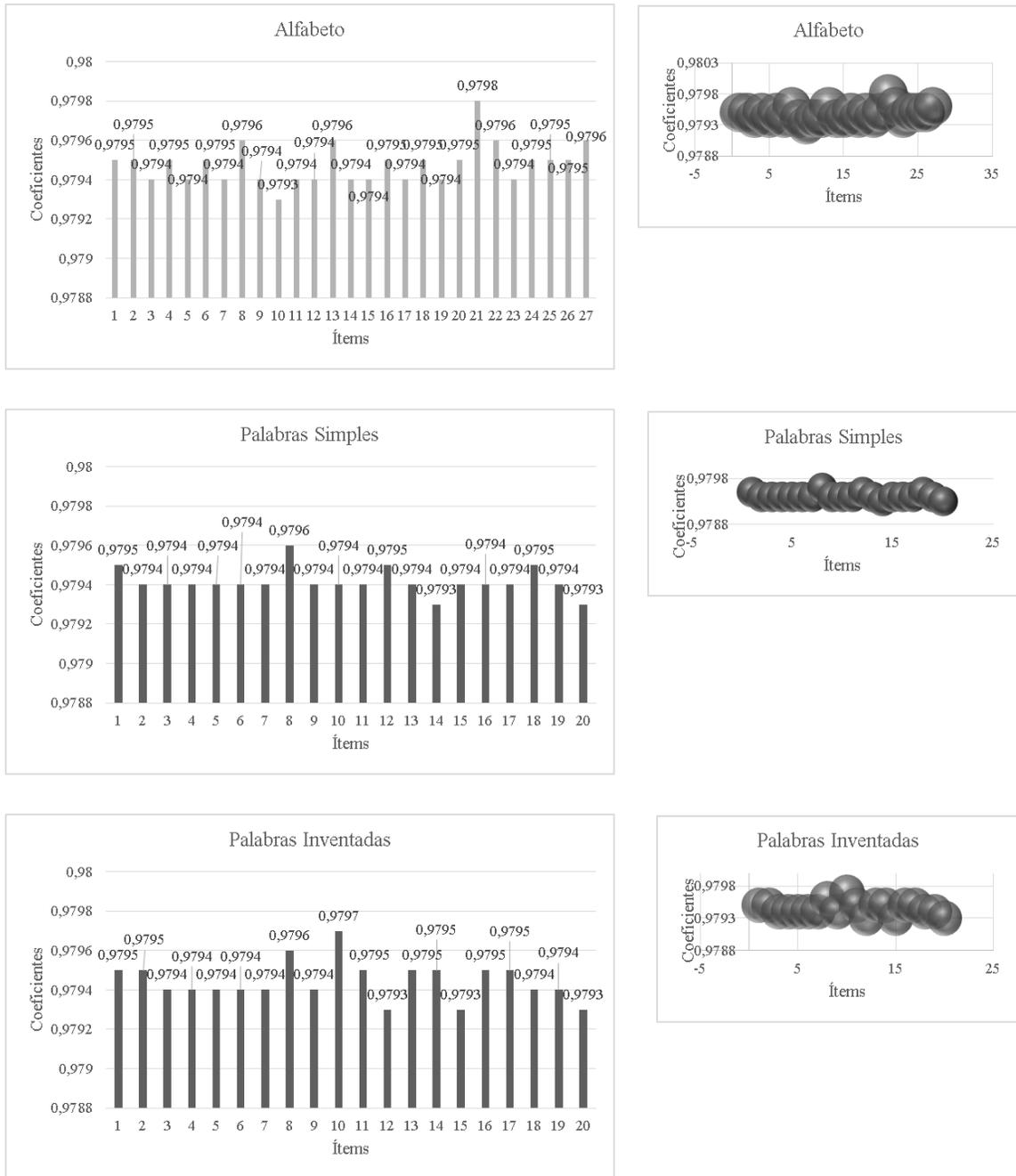
Tabla 41. *Indicadores de confiabilidad basados en diferentes métodos*

Método	Estimado	95% Intervalos de confianza	
Guttman's L2	0,983	(0,9764,	0,9886)
Coficiente de Alpha	0,9794	(0,9714,	0,9862)
Feldt-Gilmer	0,9815	(0,9742,	0,9876)
Feldt-Brennan	0,9814	(0,9741,	0,9875)
Raju's Beta	0,9794	(0,9714,	0,9862)
KR21	0,9504	Estimación generada por jMetrik	

La confiabilidad total es alta debido a que todos los indicadores estimados son superiores a 0,95. El rango entre el límite inferior (0,97) y superior (0,98) es de 0.01.

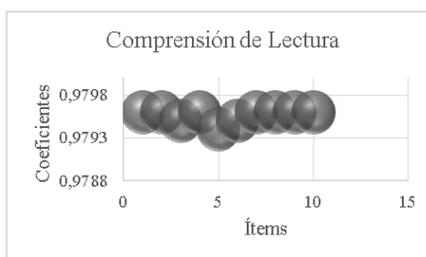
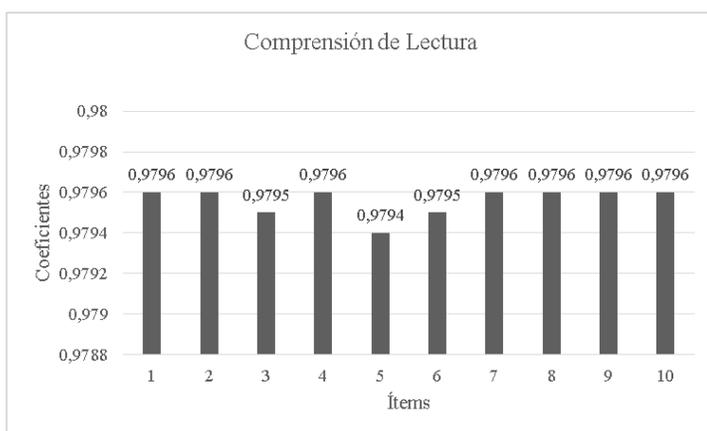
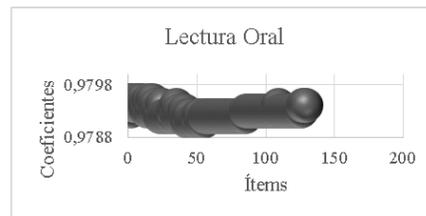
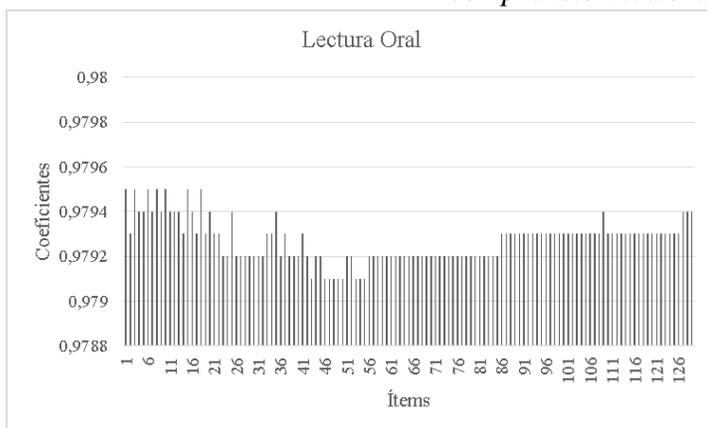
2. Confiabilidad si el ítem es eliminado.

Gráfica 49. Confiabilidad si el ítem es eliminado en la Subprueba de alfabeto, palabras simples e inventadas



Todos los ítems en la Subprueba de alfabeto, palabras simples y palabras inventadas miden con una confiabilidad superior a 0.97 aunque el ítem sea eliminado.

Gráfica 50. Confiabilidad si el ítem es eliminado en la Subprueba de lectura oral y comprensión lectora



Todos los ítems en la Subprueba de lectura oral y comprensión lectora miden con una confiabilidad superior a 0.97 aunque el ítem sea eliminado.

3. Confiabilidad de los ítems

Tabla 42. Confiabilidad de la prueba según Subpruebas de lectura inicial

	Todas las Subpruebas	Alfabeto	Simples	Inventadas	Lectura	Comprensión
Población	,982	,839	,890	,829	,988	,764
Muestra	,985	,859	,926	,898	,989	,658

Al analizar la confiabilidad en cada una de las Subpruebas para la población y la muestra se observó que en ambos grupos la confiabilidad es superior a 0.98. Al tomar como referencia la dificultad teórica de las pruebas, en donde la Subprueba de alfabeto requiere

la menor habilidad y la Subprueba de comprensión lectora requiere la mayor habilidad de lectura. La confiabilidad es mayor en la Subprueba de alfabeto que, en la Subprueba de palabras simples, tanto en la población como en la muestra. La confiabilidad en la Subprueba de palabras inventadas es menor en comparación de las Subpruebas anteriores. La Subprueba de lectura oral registró un indicador de 0.98 en ambos grupos. La menor confiabilidad se observó en la Subprueba de comprensión lectora, particularmente de la muestra.

4. Confiabilidad ítem-total

Tabla 43. *Confiabilidad de la prueba incluyendo los ítems y el total global*

	Todas las Subpruebas	Alfabeto	Simple	Inventadas	Lectura	Comprensión
Población	,982	,862	,906	,857	,988	,826
Muestra	,985	,880	,935	,913	,989	,759

Cuando se analizó la confiabilidad incluyendo el total de la prueba, se observó un leve incremento del indicador en todas las Subpruebas, con excepción en la Subprueba de lectura oral que es igual.

5. Confiabilidad en la prueba de dos mitades sin el total

Tabla 44. *Confiabilidad de la prueba de dos mitades sin el total*

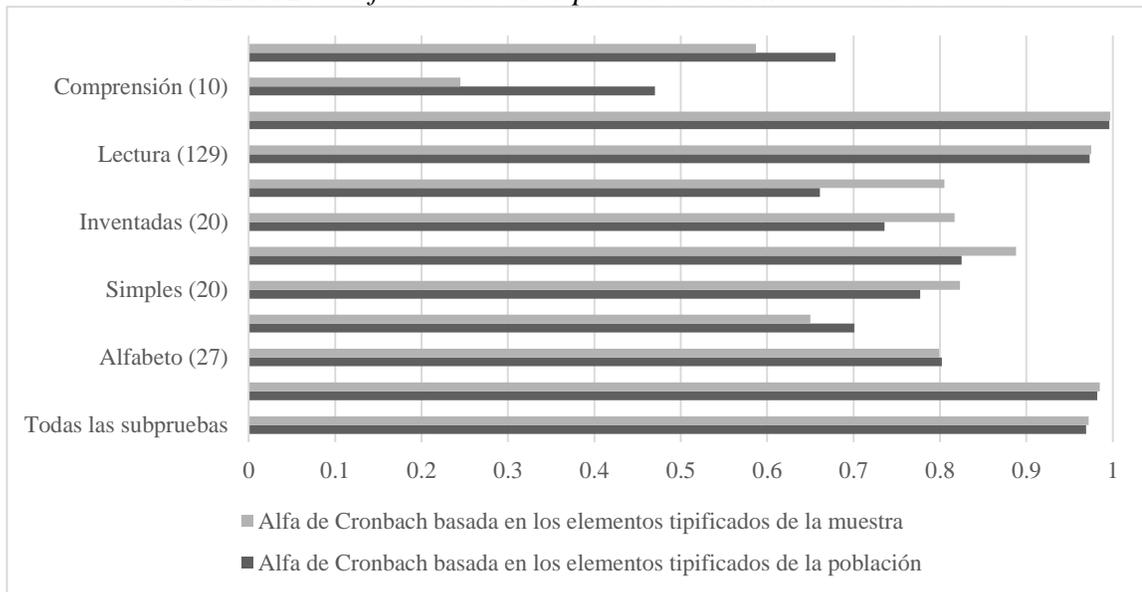
	Confiabilidad sin total	Alfa de Cronbach basada en los elementos tipificados de la población	Alfa de Cronbach basada en los elementos tipificados de la muestra
Todas las Subpruebas	Mitad 1	,969	,972
	Mitad 2	,982	,985
Alfabeto (27)	Mitad 1	,802	,799
	Mitad 2	,701	,650
Simple (20)	Mitad 1	,777	,823
	Mitad 2	,825	,888
Inventadas (20)	Mitad 1	,736	,817
	Mitad 2	,661	,805
Lectura (129)	Mitad 1	,973	,975
	Mitad 2	,996	,997
Comprensión (10)	Mitad 1	,470	,245
	Mitad 2	,679	,587

Al partir la Prueba en dos mitades, se identificó en la primera mitad, que la confiabilidad fue de 0.97 y en la segunda mitad de 0.98 tanto para la población como para la muestra. En la Subprueba de alfabeto la primera mitad de los ítems obtuvo el mayor indicador (0.80) en comparación con la segunda mitad de los ítems, que obtuvieron una confiabilidad menor de (0.70) tanto en la población como en la muestra.

En la Subprueba de palabras simples la confiabilidad fue menor a 0.88 en donde la primera mitad obtuvo los coeficientes más bajos, tanto en la población como en la muestra. En la Subprueba de palabras inventadas en la primera mitad se obtuvo un indicador más alto que en la segunda mitad, particularmente en la población. Tanto en la población como en la muestra, se registró una confiabilidad sensiblemente menor que el indicador de confiabilidad obtenido en toda la prueba (se obtuvo 0.15 como resultado de restar la confiabilidad obtenida en la primera mitad de toda la Prueba en la muestra, menos la confiabilidad obtenida en la primera mitad de la Subprueba de palabras inventadas en la muestra).

La Subprueba de lectura oral reflejó un indicador de confiabilidad de 0.97 en la primera mitad y de 0.99 en la segunda mitad, tanto en la población como en la muestra. Los menores indicadores de confiabilidad se identificaron en la Subprueba de comprensión lectora, particularmente en la primera mitad, tanto en la población como en la muestra. A diferencia de la mayoría de las Subpruebas anteriores, los indicadores de la población obtuvieron una confiabilidad mayor que en la muestra; sin embargo, el mayor indicador de confiabilidad en la segunda mitad de ítems de la población representó el 30% menor confiabilidad que el indicador más bajo de toda la prueba.

Gráfica 51. Confiabilidad de la prueba de dos mitades sin el total



Cuando se dividió la prueba en dos mitades, la confiabilidad más alta se observó en la prueba total y en la Subprueba de lectura oral. Esta fue ligeramente mayor en la muestra que en la población. En la Subprueba de alfabeto se observó una diferencia ligeramente mayor en la población, en ambas mitades de la población con relación a la muestra. Por el contrario, en la Subprueba de comprensión de lectura, la confiabilidad fue mucho mayor en la población en comparación a la muestra en ambas mitades. En las Subpruebas de palabras simples y palabras inventadas se identificó la confiabilidad más alta en la muestra, en ambas mitades.

6. Confiabilidad en la prueba de dos mitades con el total

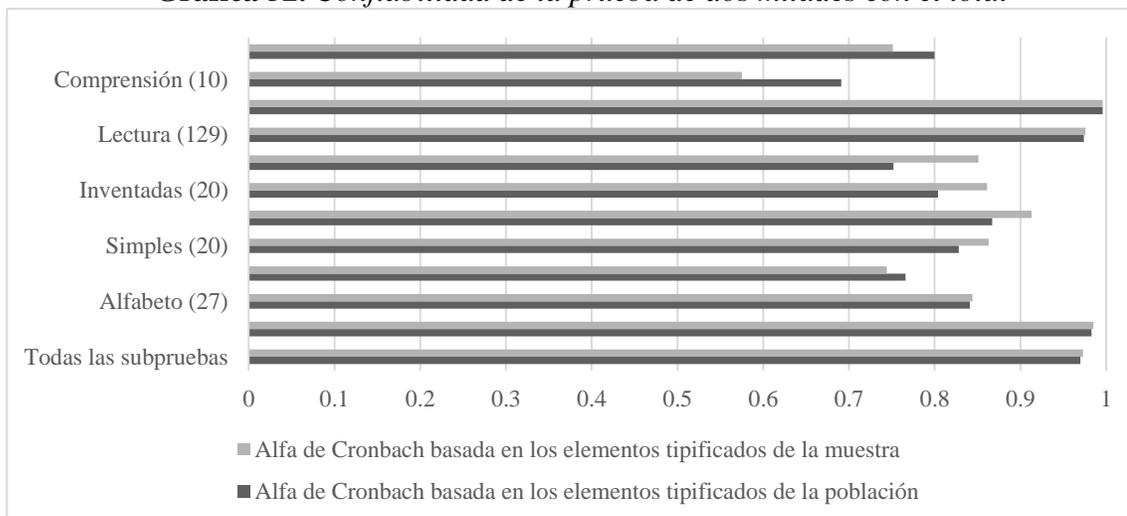
Tabla 45. *Confiabilidad de la prueba de dos mitades con el total*

	Confiabilidad con total	Alfa de Cronbach basada en los elementos tipificados de la población	Alfa de Cronbach basada en los elementos tipificados de la muestra
Todas las Subpruebas	Mitad 1	,970	,973
	Mitad 2	,983	,985
Alfabeto (27)	Mitad 1	,841	,844
	Mitad 2	,766	,744
Simples (20)	Mitad 1	,828	,863
	Mitad 2	,867	,913
Inventadas (20)	Mitad 1	,804	,861
	Mitad 2	,752	,851
Lectura (129)	Mitad 1	,974	,976
	Mitad 2	,996	,996
Comprensión (10)	Mitad 1	,691	,575
	Mitad 2	,800	,751

Al partir la prueba en dos mitades incluyendo el total, los indicadores de confiabilidad se mantuvieron en 0.97 en la primera mitad de la prueba y 0.98 en la segunda mitad de la prueba, tanto para la población como para la muestra. En la Subprueba de alfabeto la confiabilidad se mantuvo en 0.8 en la primera mitad y en 0.7 en la segunda mitad, en ambos grupos. Al incluir el total global en la Subprueba de palabras simples, la confiabilidad se incrementó a 0.91 en la segunda mitad de la muestra y se registró un coeficiente de 0.83 como el menor en la primera mitad de la población.

En la Subprueba de palabras inventadas, en la primera mitad se obtuvo un indicador más alto que en la segunda mitad, particularmente en la muestra. Los indicadores de confiabilidad se mantuvieron igual al incluir el total global de la prueba con la Subprueba de lectura oral. Los menores indicadores de confiabilidad se identificaron en la Subprueba de comprensión lectora, particularmente en la primera mitad tanto en la población como en la muestra. Al incluir el total global, se identificó un incremento considerable en ambas mitades de los dos grupos.

Gráfica 52. Confiabilidad de la prueba de dos mitades con el total



En general los indicadores de confiabilidad mejoraron al incluir el total global. Particularmente en la Subprueba de comprensión lectora, tanto en la primera mitad y la segunda, en los dos grupos de estudiantes. En la Subprueba de alfabeto incrementó ligeramente el indicador de confiabilidad en la población en la primera mitad, a diferencia de la confiabilidad sin el total.

H. Análisis Factorial Exploratorio

Tabla 46. KMO y prueba de Bartlett de la Prueba con 38 ítems

		Máxima verosimilitud y Promax en población
Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,367
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	938,900
	Grados de libertad	703
	Sig.	,000

El análisis factorial exploratorio en la Prueba de Lectura Inicial obtuvo un KMO de 0,367 con máxima verosimilitud. Lo anterior indica que la Prueba posee un indicador bajo para establecer si los resultados de la habilidad de la prueba son adecuados.

Tabla 47. Varianza total explicada de la Prueba con 38 ítems

Factor	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total
1	,730	12,128	12,128	,730	12,128	12,128	,360
2	,610	10,136	22,264	,610	10,136	22,264	,466
3	,477	7,928	30,192	,477	7,928	30,192	,366
4	,387	6,429	36,621	,387	6,429	36,621	,353
5	,345	5,737	42,358	,345	5,737	42,358	,320
6	,323	5,361	47,720	,323	5,361	47,720	,449
7	,306	5,075	52,794	,306	5,075	52,794	,299
8	,276	4,579	57,373	,276	4,579	57,373	,300
9	,251	4,161	61,534	,251	4,161	61,534	,417
10	,231	3,838	65,372	,231	3,838	65,372	,257
11	,206	3,421	68,793	,206	3,421	68,793	,281
12	,197	3,274	72,067	,197	3,274	72,067	,306
13	,181	2,999	75,066	,181	2,999	75,066	,274
14	,167	2,765	77,831	,167	2,765	77,831	,239
15	,155	2,572	80,403				
16	,141	2,338	82,741				
17	,124	2,063	84,804				
18	,117	1,948	86,752				
19	,107	1,773	88,525				
20	,091	1,515	90,039				
21	,081	1,346	91,385				
22	,066	1,089	92,474				
23	,061	1,016	93,491				
24	,054	,904	94,394				
25	,049	,816	95,210				
26	,046	,757	95,967				
27	,040	,660	96,627				
28	,035	,576	97,203				
29	,034	,558	97,760				
30	,033	,545	98,306				
31	,025	,410	98,715				
32	,022	,361	99,076				
33	,016	,272	99,348				

Factor							Total
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	
34	,013	,211	99,559				
35	,011	,178	99,737				
36	,007	,123	99,861				
37	,004	,074	99,935				
38	,004	,065	100,000				

El 77% de la varianza de la habilidad para responder la prueba, se explica con 14 factores.

Tabla 48. *KMO y prueba de Bartlett de la Prueba con 16 ítems*

		Máxima verosimilitud y Promax en población
Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,619
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	224,648
	Grados de libertad	120
	Sig.	,000

El análisis factorial exploratorio en la Prueba de lectura inicial con 16 ítems obtuvo un KMO de 0,619 con máxima verosimilitud. Lo anterior indica que la prueba posee un indicador bajo para establecer si los resultados de la habilidad de la prueba son adecuados.

Tabla 49. *Varianza total explicada de la prueba con 16 ítems*

Factor	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total
1	,562	21,411	21,411	,562	21,411	21,411	,289
2	,292	11,146	32,557	,292	11,146	32,557	,396
3	,279	10,638	43,195	,279	10,638	43,195	,249
4	,237	9,020	52,215	,237	9,020	52,215	,221
5	,207	7,902	60,117	,207	7,902	60,117	,286
6	,188	7,173	67,290	,188	7,173	67,290	,235
7	,179	6,830	74,120	,179	6,830	74,120	,268
8	,146	5,571	79,691				
9	,119	4,531	84,223				
10	,102	3,886	88,109				
11	,091	3,465	91,574				
12	,070	2,664	94,238				
13	,057	2,174	96,412				
14	,045	1,702	98,114				
15	,032	1,222	99,337				

16 ,017 ,663 100,000

El 74% de la varianza de la habilidad para responder la prueba, se explica con 7 factores.

Tabla 50. Estadísticos de adecuación (KMO y prueba de Bartlett) de la Subprueba de alfabeto

		Máxima verosimilitud y Promax en población
Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,587
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	786,729
	Grados de libertad	253
	Sig.	,000
	Iteraciones	Fijo

El análisis factorial exploratorio en la Subprueba de alfabeto obtuvo un KMO de 0,587 con máxima verosimilitud. Lo anterior indica que la Subprueba posee un indicador bajo para establecer si los resultados de la habilidad de la prueba son adecuados.

Tabla 51. Varianza total explicada de la Subprueba de alfabeto

Factor	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total
1	5,221	32,633	32,633	3,720	23,251	23,251	3,219
2	1,895	11,843	44,476	1,152	7,199	30,450	3,740
3	1,683	10,519	54,995	1,210	7,560	38,010	2,859
4	1,153	7,204	62,199	1,730	10,811	48,821	2,612
5	,957	5,982	68,181	1,475	9,221	58,042	2,889
6	,882	5,511	73,692	,834	5,213	63,255	1,960
7	,836	5,223	78,915	,717	4,479	67,734	2,111
8	,703	4,396	83,311	,434	2,715	70,449	1,157
9	,649	4,056	87,367	,408	2,548	72,998	1,403
10	,564	3,524	90,891	,275	1,718	74,716	1,644
11	,410	2,561	93,452				
12	,345	2,158	95,610				
13	,247	1,541	97,151				
14	,213	1,334	98,485				
15	,203	1,270	99,754				
16	,039	,246	100,000				

El 74% de la varianza de la habilidad para reconocer las letras del alfabeto, se explica con diez factores.

Tabla 52. Estadísticos de adecuación (KMO y prueba de Bartlett) de la Subprueba palabras simples

		Componentes principales y Varimax en muestra	Máxima verosimilitud y Promax en población
Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,740	,676
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	50,624	448,152
	Grados de libertad	21	120
	Sig.	,000	,000
	Iteraciones	predefinido	fijo

El análisis factorial exploratorio en la Subprueba de palabras simples obtuvo un KMO de 0,740 con componentes principales y de ,676 con máxima verosimilitud. Lo anterior indica que la subprueba posee un indicador suficiente, es decir que los resultados de la habilidad de la prueba son adecuados.

Tabla 53. Varianza total explicada de la Subprueba palabras simples en la muestra

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	4,185	59,780	59,780	4,185	59,780	59,780	3,716	53,083	53,083
2	1,469	20,980	80,760	1,469	20,980	80,760	1,937	27,677	80,760
3	,759	10,837	91,597						
4	,244	3,490	95,087						
5	,144	2,063	97,151						
6	,121	1,730	98,881						
7	,078	1,119	100,000						

El 27% de la varianza de la habilidad para leer palabras simples se explica con dos factores.

Tabla 54. Varianza total explicada de la Subprueba palabras simples en la población

Factor	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total
1	5,221	32,633	32,633	3,720	23,251	23,251	3,219
2	1,895	11,843	44,476	1,152	7,199	30,450	3,740

3	1,683	10,519	54,995	1,210	7,560	38,010	2,859
4	1,153	7,204	62,199	1,730	10,811	48,821	2,612
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total
5	,957	5,982	68,181	1,475	9,221	58,042	2,889
6	,882	5,511	73,692	,834	5,213	63,255	1,960
7	,836	5,223	78,915	,717	4,479	67,734	2,111
8	,703	4,396	83,311	,434	2,715	70,449	1,157
9	,649	4,056	87,367	,408	2,548	72,998	1,403
10	,564	3,524	90,891	,275	1,718	74,716	1,644
11	,410	2,561	93,452				
12	,345	2,158	95,610				
13	,247	1,541	97,151				
14	,213	1,334	98,485				
15	,203	1,270	99,754				
16	,039	,246	100,000				

El 74% de la varianza de la habilidad para leer palabras simples se explica con diez factores.

Tabla 55. Estadísticos de adecuación (KMO y prueba de Bartlett) de la Subprueba de palabras inventadas

		Máxima verosimilitud y Promax en población
Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,561
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	320,490
	Grados de libertad	136
	Sig.	,000
	Iteraciones	Fijo

El análisis factorial exploratorio en la Subprueba de palabras inventadas obtuvo un KMO de 0,561 con máxima verosimilitud. Lo anterior indica que la Subprueba posee un indicador bajo para establecer si los resultados de la habilidad de la prueba son adecuados.

Tabla 56. Varianza total explicada de la Subprueba de palabras inventadas

Factor	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total
	1	4,450	26,179	26,179	4,123	24,253	24,253
2	1,632	9,599	35,779	1,369	8,051	32,304	2,179
3	1,471	8,655	44,434	1,262	7,426	39,730	2,348
4	1,435	8,444	52,878	,953	5,607	45,336	1,284

5	1,156	6,801	59,679	1,145	6,733	52,069	2,247
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total
6	1,108	6,515	66,194	,800	4,704	56,773	2,406
7	,917	5,394	71,589	,562	3,304	60,077	1,947
8	,886	5,213	76,802	,960	5,646	65,723	1,282
9	,825	4,854	81,656	,517	3,043	68,766	1,715
10	,737	4,336	85,992	,843	4,961	73,727	2,058
11	,556	3,272	89,264				
12	,478	2,812	92,076				
13	,423	2,487	94,563				
14	,331	1,947	96,510				
15	,274	1,609	98,119				
16	,214	1,261	99,380				
17	,105	,620	100,000				

El 73% de la varianza de la habilidad para leer palabras inventadas se explica con diez factores.

Tabla 57. Estadísticos de adecuación (KMO y prueba de Bartlett) de la Subprueba comprensión lectora

		Componentes principales y Varimax en muestra	Máxima verosimilitud y Promax en población
Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,607	,781
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	44,904	100,425
	Grados de libertad	15	36
	Sig.	,000	,000
	Iteraciones	predefinido	fijo

El análisis factorial exploratorio en la Subprueba de comprensión lectora obtuvo un KMO de 0,607 con componentes principales y de ,781 con máxima verosimilitud. Lo anterior indica que la Subprueba posee un indicador suficiente, es decir que los resultados de la habilidad de la prueba son adecuados.

Tabla 58. Varianza total explicada de la Subprueba comprensión lectora en la muestra

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
	1	3,742	62,360	62,360	3,742	62,360	62,360	3,085	51,409
2	1,122	18,696	81,056	1,122	18,696	81,056	1,779	29,648	81,056
3	,608	10,134	91,191						
4	,389	6,491	97,682						

5	,096	1,606	99,288						
		% de la	%	% de la	%	% de la	%		
	Total	varianza	acumulado	Total	varianza	acumulado	Total	varianza	acumulado
6	,043	,712	100,000						

Como resultado del análisis de varianza se identificaron dos componentes. Debido a que el KMO en la Subprueba de comprensión lectora posee un valor menor, al criterio de suficiencia de la Subprueba se omite la interpretación de la explicación de la varianza.

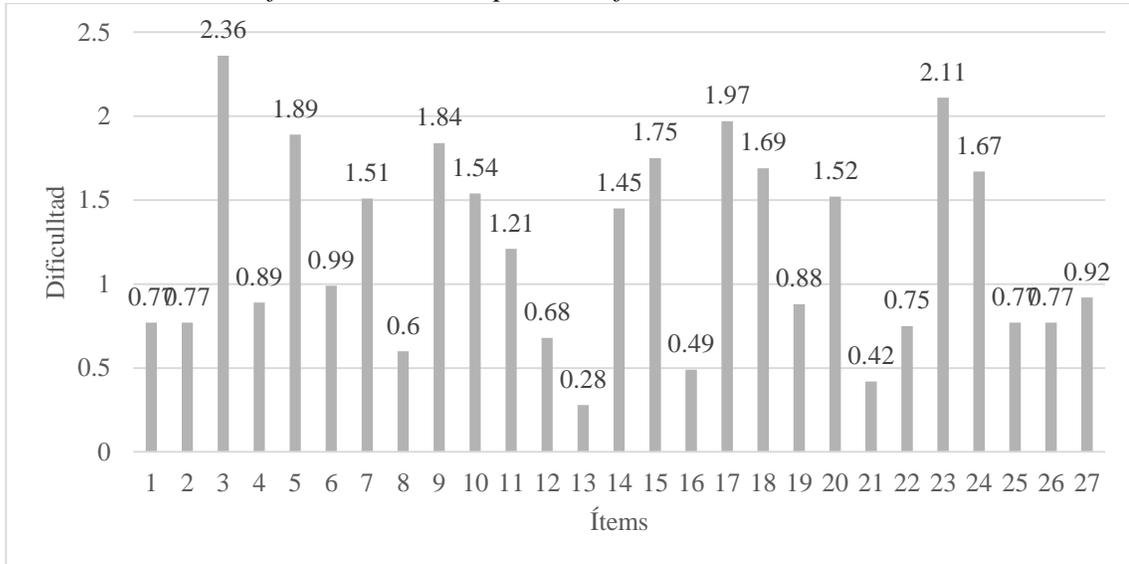
Tabla 59. Varianza total explicada de la Subprueba comprensión lectora en la población

Factor	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total
1	4,450	26,179	26,179	4,123	24,253	24,253	2,989
2	1,632	9,599	35,779	1,369	8,051	32,304	2,179
3	1,471	8,655	44,434	1,262	7,426	39,730	2,348
4	1,435	8,444	52,878	,953	5,607	45,336	1,284
5	1,156	6,801	59,679	1,145	6,733	52,069	2,247
6	1,108	6,515	66,194	,800	4,704	56,773	2,406
7	,917	5,394	71,589	,562	3,304	60,077	1,947
8	,886	5,213	76,802	,960	5,646	65,723	1,282
9	,825	4,854	81,656	,517	3,043	68,766	1,715
10	,737	4,336	85,992	,843	4,961	73,727	2,058
11	,556	3,272	89,264				
12	,478	2,812	92,076				
13	,423	2,487	94,563				
14	,331	1,947	96,510				
15	,274	1,609	98,119				
16	,214	1,261	99,380				
17	,105	,620	100,000				

El 73% de la varianza de la habilidad para Comprender la Lectura se explica con diez factores.

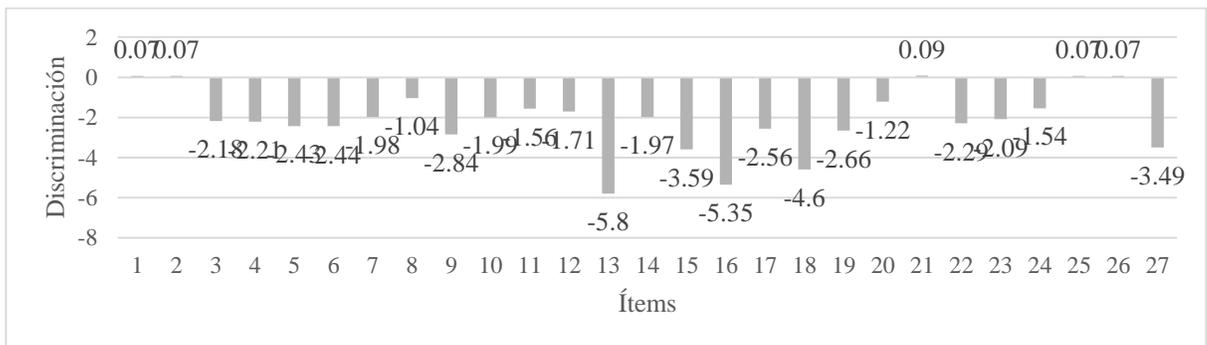
I. Calibración de los ítems con TRI

Gráfica 53. Dificultad de la Subprueba alfabeto en la calibración de los ítems



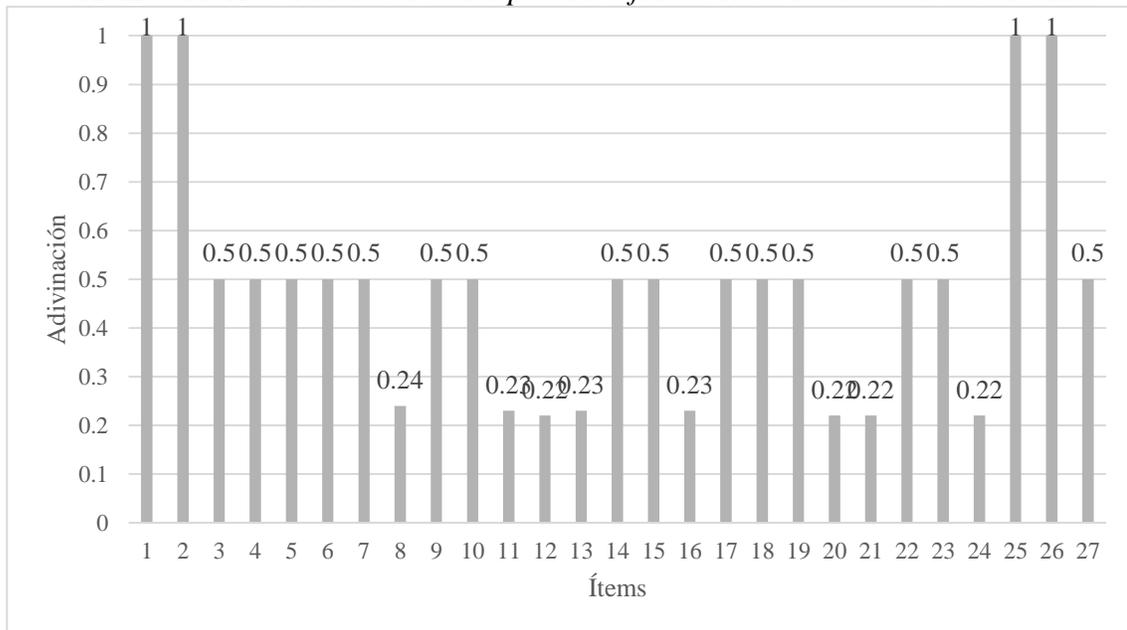
El 48% de los ítems en la Subprueba de alfabeto, registraron una dificultad superior a 1 y el 52% se observó con indicadores positivos menores a 1.

Gráfica 54. Discriminación de la Subprueba alfabeto en la calibración de los ítems



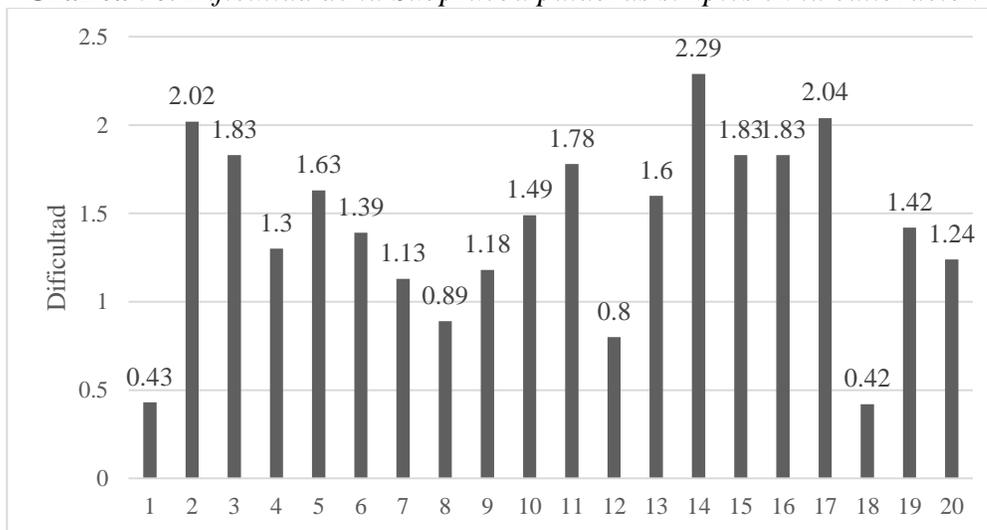
Los ítems de la Subprueba de alfabeto no discriminan la habilidad de lectura en los estudiantes evaluados, debido a que se encuentran debajo de 0.1.

Gráfica 55. Adivinación de la Subprueba alfabeto en la calibración de los ítems



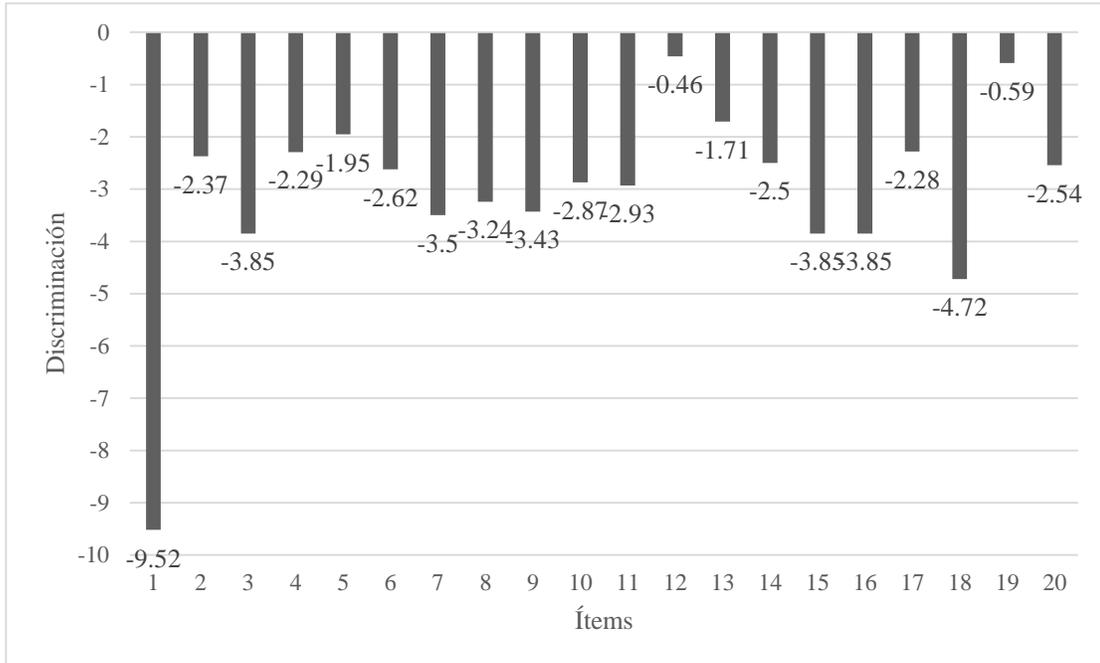
La adivinación en los ítems de la Subprueba de alfabeto no se evidenció, debido a que superaron el criterio de adivinación que es igual a 0.

Gráfica 56. Dificultad de la Subprueba palabras simples en la calibración de los ítems



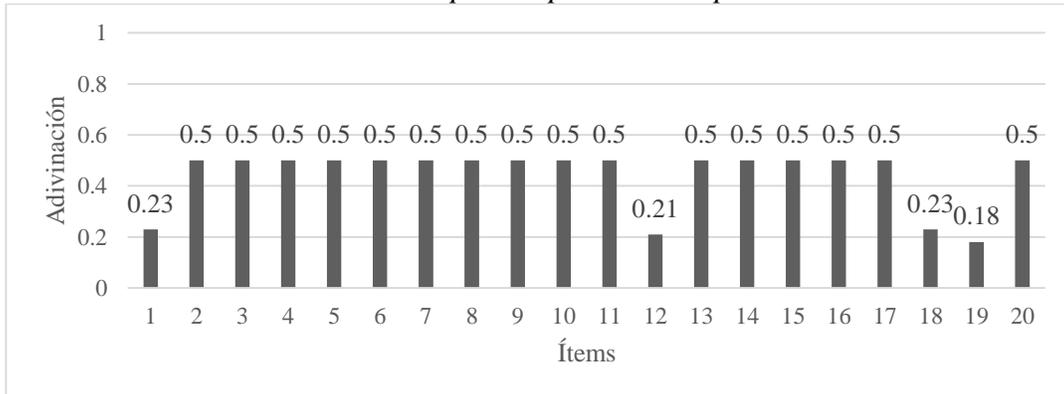
El 80% de los ítems en la Subprueba de palabras simples, registraron una dificultad superior a 1 y el 20% se observó con indicadores positivos menores a 1.

Gráfica 57. Discriminación de la Subprueba palabras simples en la calibración de los ítems



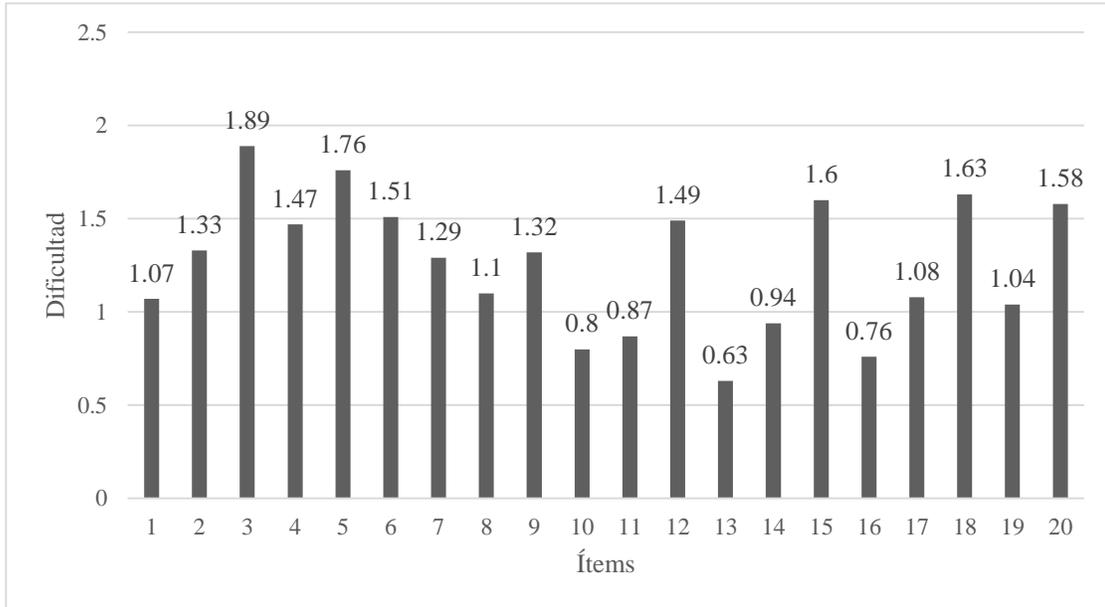
Los ítems de la Subprueba de palabras simples no discriminan la habilidad de lectura en los estudiantes evaluados, debido a que se encuentran debajo de 0.1.

Gráfica 58. Adivinación de la Subprueba palabras simples en la calibración de los ítems



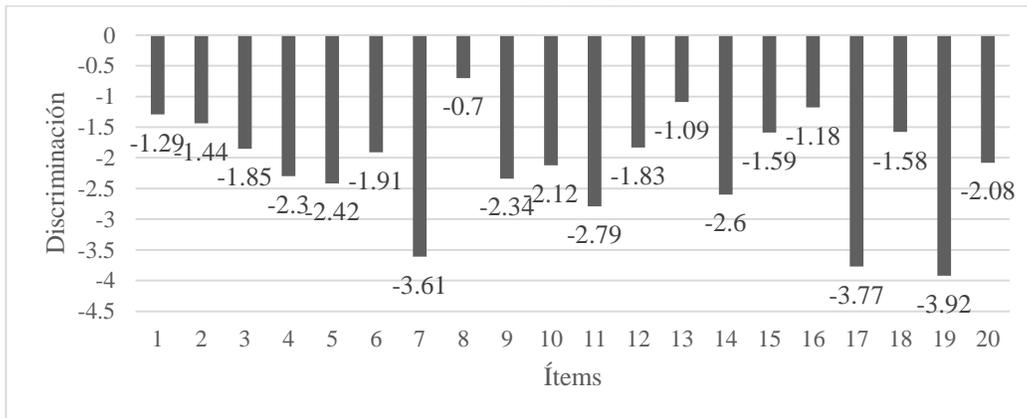
La adivinación en los ítems de la Subprueba de palabras simples no se evidenció, debido a que superaron el criterio de adivinación que es igual a 0.

Gráfica 59. Dificultad de la Subprueba palabras inventadas en la calibración de los ítems



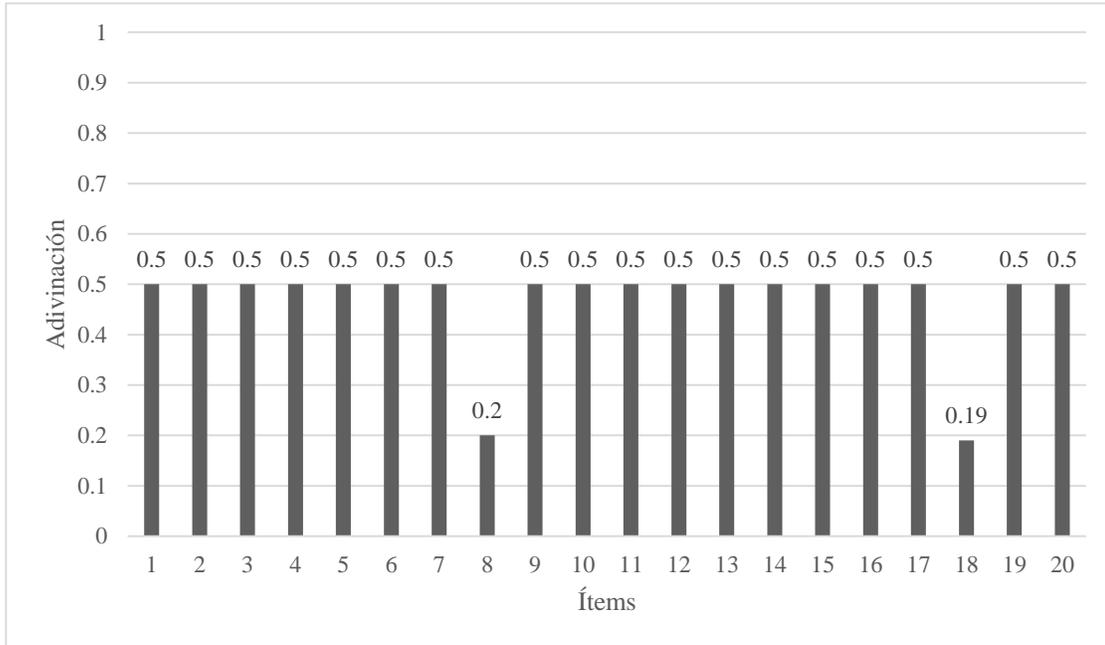
El 75% de los ítems en la Subprueba de palabras inventadas, registraron una dificultad superior a 1 y el 25% se observó con indicadores positivos menores a 1.

Gráfica 60. Discriminación de la Subprueba de palabras inventadas en la calibración de los ítems



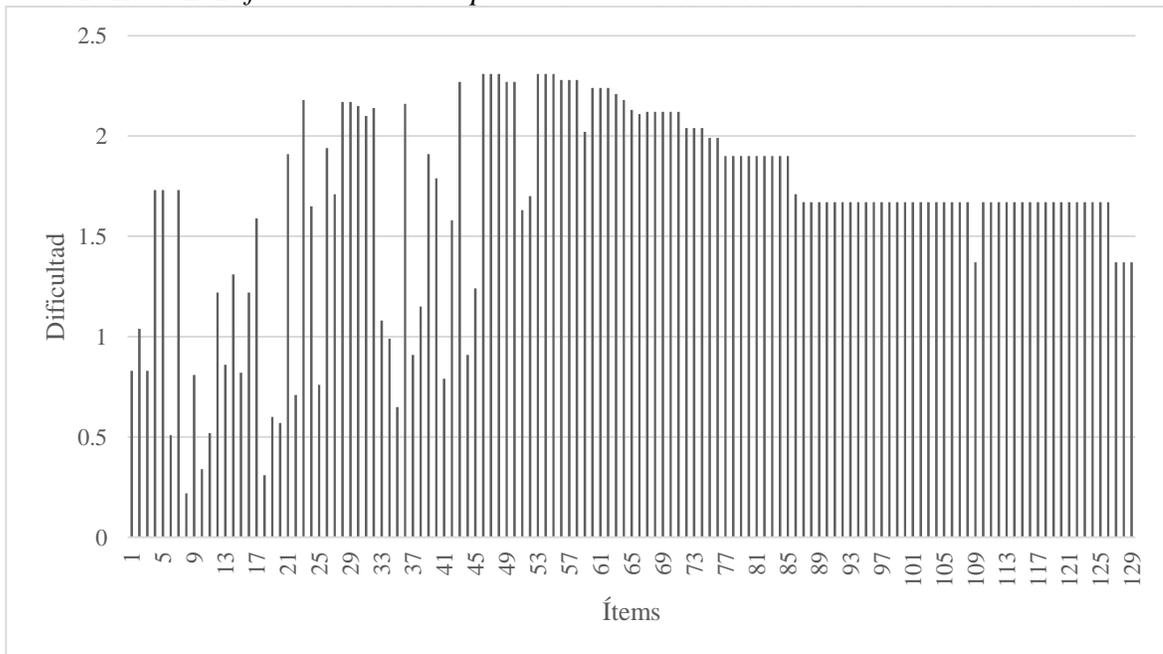
Los ítems de la Subprueba de palabras inventadas no discriminan la habilidad de lectura en los estudiantes evaluados, debido a que se encuentran debajo de 0.1.

Gráfica 61. Adivinación de la Subprueba palabras inventadas en la calibración de los ítems



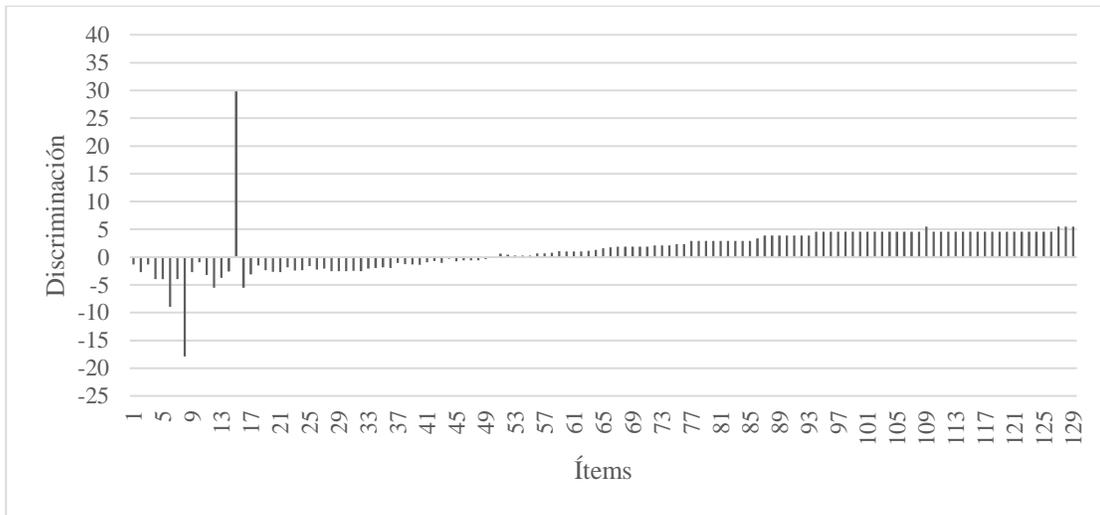
La adivinación en los ítems de la Subprueba de palabras inventadas no se evidenció, debido a que superaron el criterio de adivinación que es igual a 0.

Gráfica 62. Dificultad de la Subprueba lectura oral en la calibración de los ítems



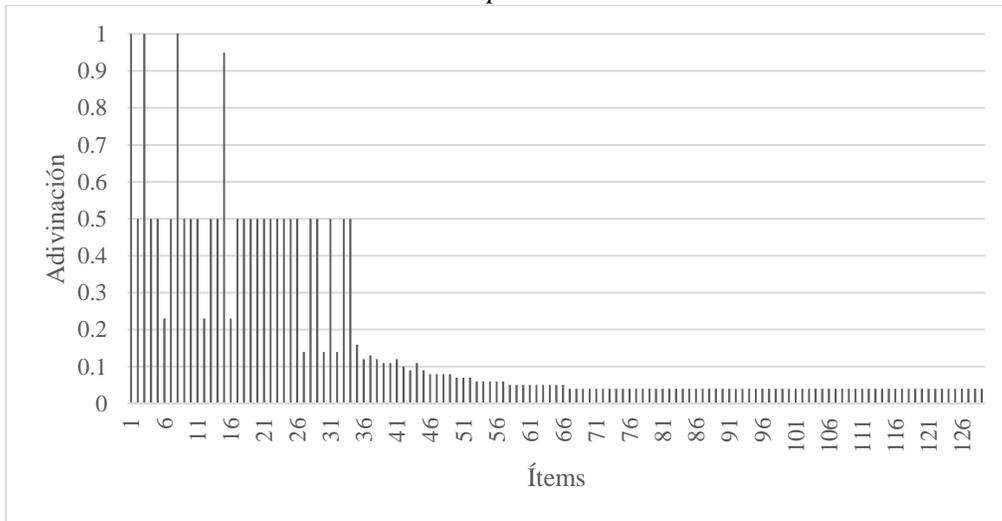
El 15% de los ítems en la Subprueba de lectura oral, registraron una dificultad superior a 1 y el 85% se observó con indicadores positivos menores a 1.

Gráfica 63. *Discriminación de la Subprueba lectura oral en la calibración de los ítems*



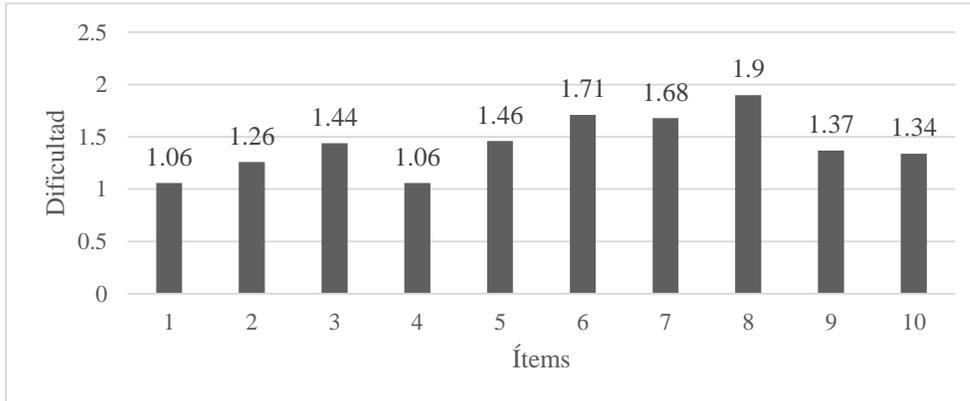
El 38% de los ítems de la Subprueba de lectura oral, no discriminan la habilidad de lectura en los estudiantes evaluados. El ítem 15 mostró una discriminación más alta, a diferencia del 62% de los ítems que discriminan con un indicador menor a 0.60 la habilidad de lectura oral de los estudiantes de 3ro. Primaria.

Gráfica 64. *Adivinación de la Subprueba lectura oral en la calibración de los ítems*



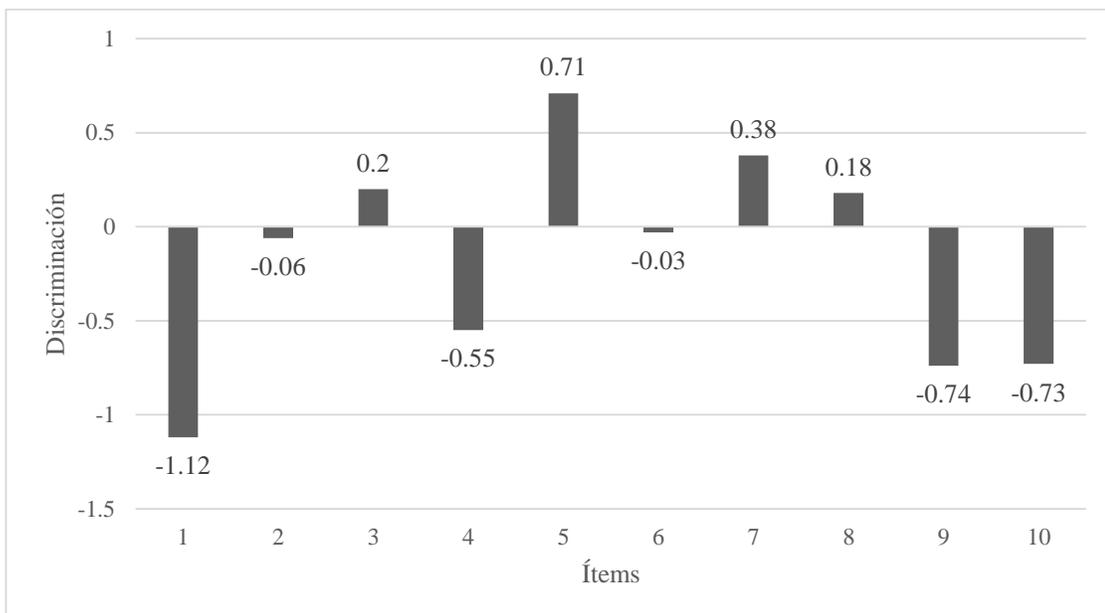
La adivinación en los ítems de la Subprueba de comprensión lectora no se evidenció, debido a que superaron el criterio de adivinación que es igual a 0.

Gráfica 65. *Dificultad de la Subprueba comprensión lectora en la calibración de los ítems*



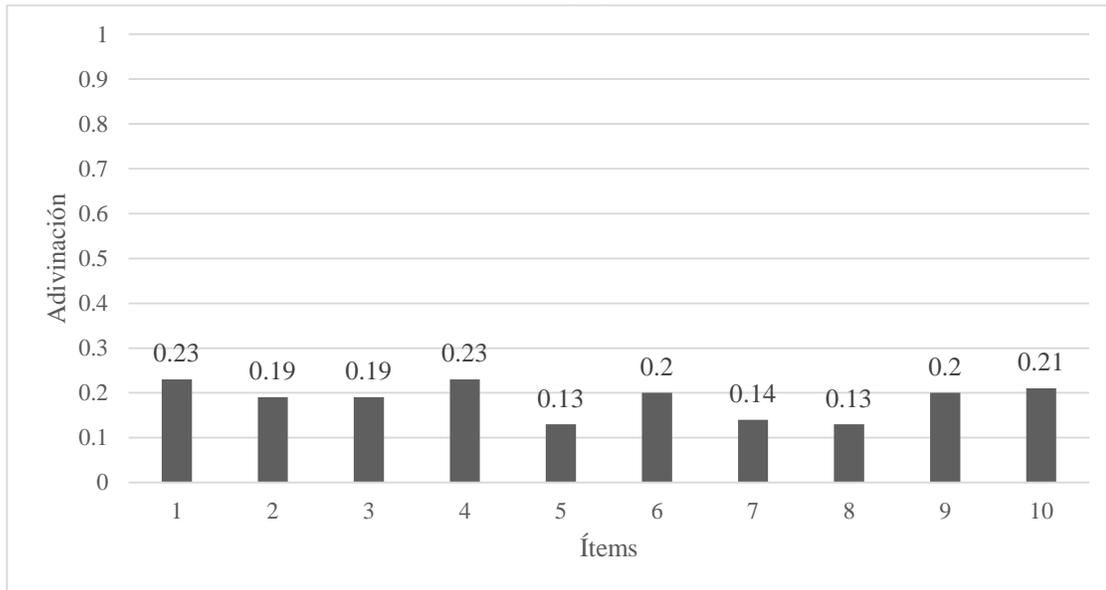
El 100% de los ítems en la Subprueba de comprensión lectora, registraron una dificultad superior a 1.

Gráfica 66. *Discriminación de la Subprueba comprensión lectora en la calibración de los ítems*



El 40% de los ítems con un indicador menor a 0.71 discriminan la habilidad que tienen los estudiantes para comprender lo que leen. El 60% de los ítems de la Subprueba de comprensión lectora, no discriminan la habilidad de los estudiantes evaluados.

Gráfica 67. Adivinación de la Subprueba comprensión lectora en la calibración de los ítems



La adivinación en los ítems de la Subprueba de comprensión lectora no se evidenció, debido a que superaron el criterio de adivinación que es igual a 0.

J. Calificación de los ítems

Para la calificación por estudiantes se utilizó la Teoría de Respuesta al Ítem (TRI) la cual se basa en la métrica de -6 a 6. Los archivos de calificación por estudiantes son de uso exclusivo para *School the World*. Se encuentran en formato de Excel adjunto con este informe preliminar.

Se encuentran ordenados de acuerdo a los ítems con menor dificultad a los que tienen mayor dificultad. En el proceso de análisis, el software omitió el cálculo de las puntuaciones theta para los ítems sin varianza o que habían sido respondidos por todos los estudiantes.

Tabla 60. *Tabla de puntuaciones en la Subprueba alfabeto*

Punteo	Theta	Error estándar
0	-4.9245	1.86
1	-3.6361	1.06
2	-2.8244	0.79
3	-2.3035	0.67
4	-1.903	0.6
5	-1.569	0.56
6	-1.2769	0.53
7	-1.0128	0.5
8	-0.7681	0.49
9	-0.5368	0.48
10	-0.3142	0.47
11	-0.0968	0.46
12	0.1187	0.46
13	0.3353	0.47
14	0.5561	0.47
15	0.7849	0.48
16	1.0259	0.5
17	1.285	0.52
18	1.5708	0.55
19	1.8969	0.59
20	2.2883	0.66
21	2.7991	0.78
22	3.6011	1.05
23	4.8838	1.86

El 85% de los ítems de la Subprueba de alfabeto tienen un valor theta negativo y el 15% fue omitido por el software por no contar con la varianza.

Tabla 61. *Tabla de puntuaciones en la Subprueba palabras imples*

Punteo	Theta	Error estándar
0	-5.4954	1.87
1	-4.1632	1.09
2	-3.2708	0.84
3	-2.6504	0.74
4	-2.1388	0.69
5	-1.6913	0.65
6	-1.2901	0.62
7	-0.9242	0.59
8	-0.5847	0.57
9	-0.2634	0.56
10	0.0468	0.55
11	0.3524	0.55

Punteo	Theta	Error estándar
12	0.6601	0.56
13	0.9769	0.57
14	1.3114	0.59
15	1.6753	0.62
16	2.0862	0.67
17	2.5727	0.73
18	3.1916	0.85
19	4.1105	1.11
20	5.4743	1.89

El 45% de los ítems de la Subprueba de palabras simples tienen un valor theta negativo.

Tabla 62. *Tabla de puntuaciones en la Subprueba palabras inventadas*

Punteo	Theta	Error Estándar
0	-4.5241	1.85
1	-3.2583	1.04
2	-2.4747	0.77
3	-1.976	0.65
4	-1.5918	0.59
5	-1.2686	0.55
6	-0.9821	0.52
7	-0.7189	0.5
8	-0.4708	0.49
9	-0.2317	0.49
10	0.0028	0.48
11	0.2371	0.49
12	0.4753	0.49
13	0.722	0.5
14	0.9836	0.52
15	1.2682	0.55
16	1.5896	0.59
17	1.9722	0.65
18	2.47	0.77
19	3.2537	1.04
20	4.5203	1.85

El 45% de los ítems en la Subprueba de palabras inventadas tienen un valor theta negativo.

Tabla 63. *Tabla de puntuaciones en la Subprueba lectura oral*

Score	Theta	Std. Err
0	-14,2257	1,86
1	-12,9284	1,07
2	-12,094	0,8
3	-11,5382	0,7
4	-11,0919	0,64
5	-10,7031	0,61
6	-10,3499	0,58
7	-10,022	0,56
8	-9,7146	0,55
9	-9,4252	0,53
10	-9,153	0,51
11	-8,8971	0,5
12	-8,6567	0,48
13	-8,4307	0,47
14	-8,2177	0,45
15	-8,0163	0,44
16	-7,8251	0,43
17	-7,6427	0,42
18	-7,4678	0,41
19	-7,2992	0,41
20	-7,136	0,4
21	-6,977	0,4
22	-6,8216	0,39
23	-6,669	0,39
24	-6,5184	0,39
25	-6,3693	0,39
26	-6,2212	0,38
27	-6,0736	0,38
28	-5,9261	0,38
29	-5,7783	0,38
30	-5,6299	0,39
31	-5,4806	0,39
32	-5,3302	0,39
33	-5,1785	0,39
34	-5,0252	0,39
35	-4,8703	0,39
36	-4,7135	0,4
37	-4,5549	0,4
38	-4,3941	0,4
39	-4,2312	0,41
40	-4,0659	0,41

Score	Theta	Std. Err
41	-3,8983	0,41
42	-3,7283	0,41
43	-3,5558	0,42
44	-3,3807	0,42
45	-3,2034	0,42
46	-3,0238	0,42
47	-2,8422	0,43
48	-2,659	0,43
49	-2,4745	0,43
50	-2,289	0,43
51	-2,1028	0,43
52	-1,9163	0,43
53	-1,7297	0,43
54	-1,543	0,43
55	-1,3564	0,43
56	-1,1699	0,43
57	-0,9835	0,43
58	-0,7971	0,43
59	-0,6109	0,43
60	-0,4248	0,43
61	-0,2389	0,43
62	-0,0534	0,43
63	0,1317	0,43
64	0,3163	0,43
65	0,5004	0,43
66	0,6839	0,43
67	0,867	0,43
68	1,0498	0,43
69	1,2323	0,43
70	1,4149	0,43
71	1,5975	0,43
72	1,7803	0,43
73	1,9635	0,43
74	2,147	0,43
75	2,3309	0,43
76	2,5153	0,43
77	2,7001	0,43
78	2,8853	0,43
79	3,0711	0,43
80	3,2571	0,43
81	3,4432	0,43
82	3,629	0,43

Score	Theta	Std. Err
83	3,8139	0,43
84	3,9972	0,43
85	4,1782	0,42
86	4,356	0,42
87	4,53	0,41
88	4,6995	0,41
89	4,864	0,4
90	5,0232	0,4
91	5,1771	0,39
92	5,3255	0,38
93	5,4686	0,37
94	5,6067	0,37
95	5,7401	0,36
96	5,869	0,36
97	5,9938	0,35
98	6,1155	0,35
99	6,2338	0,34
100	6,3494	0,34
101	6,4627	0,34
102	6,5743	0,33
103	6,6845	0,33
104	6,7937	0,33
105	6,9024	0,33
106	7,011	0,33
107	7,1201	0,33
108	7,23	0,33
109	7,3413	0,33
110	7,4546	0,34
111	7,5705	0,34
112	7,6897	0,35
113	7,813	0,35
114	7,9414	0,36
115	8,0759	0,37
116	8,2181	0,38
117	8,3696	0,4
118	8,5325	0,41
119	8,7097	0,43
120	8,9048	0,45
121	9,1238	0,48
122	9,3739	0,52
123	9,6684	0,57
124	10,0311	0,64

Score	Theta	Std. Err
125	10,515	0,76
126	11,2895	1,04
127	12,5517	1,85

El 48% de los ítems en la Subprueba de lectura oral tiene un valor theta negativo y el 1% fue omitido por el software por no contar con la varianza.

Tabla 64. *Tabla de puntuaciones en la Subprueba comprensión lectora*

Punteo	Theta	Error Estándar
0	-3.7756	1,87
1	-2.4557	1,08
2	-1.5838	0,83
3	-0.9814	0,74
4	-0.4744	0,69
5	-0.0026	0,68
6	0.4697	0,7
7	0.9783	0,74
8	1.5839	0,83
9	2.4608	1,08
10	3.7854	1,87

El 50% de los ítems de la Subprueba de comprensión lectora tienen un valor theta negativo. En promedio los 60 estudiantes de 3ro. primaria leyeron en la historia 51 palabras por minuto.

V. DISCUSIÓN DE LOS RESULTADOS

La confiabilidad y validez corresponden a propiedades de las interpretaciones, inferencias o usos específicos de las medidas que las pruebas proporcionan (Prieto & Delgado, 2003). A partir de los resultados de los análisis estadísticos realizados con la base de datos de *School the World*, a continuación, se presentan la discusión de los resultados. Es importante considerar que si la medida tiene valores consistentes y precisos es posible confiar en los resultados obtenidos, los cuales se integran con el modelo con el cual se producen medidas de ajuste precisas y estables.

A. Correlación en cada Subprueba

La media fue menor cuando se correlacionaron los ítems con el total de la prueba que con el de las Subpruebas de alfabeto, palabras simples, palabras inventadas, lectura oral y comprensión lectora. La Subprueba de comprensión lectora tuvo un comportamiento diferente a las demás Subpruebas, ya que no se obtuvo la media de la muestra, cuando se correlacionaron los ítems con el total de la prueba de comprensión lectora. Los indicadores mínimos de la mediana (0.32 a 0.56) se encontraron dentro de los parámetros considerados altos, al igual que los indicadores máximos de la mediana (0.57 a 0.69). Existió mayor correlación entre los ítems de la Subprueba de lectura oral, ya que el intervalo fue de 11 en los ítems con el total de la prueba. La Subprueba de palabras simples fue la que tuvo el mayor intervalo (30); sin embargo, su correlación es considerable alta entre 0.34 a 0.64.

Los menores coeficientes de correlación se encontraron cuando se relacionó cada ítem de las Subpruebas de alfabeto, palabras simples y palabras inventadas con el total de la prueba en la población. En la Subprueba de lectura oral la mayoría de los ítems tuvieron una correlación mayor en la población a diferencia de las anteriores Subpruebas. Al correlacionar cada ítem de la Subprueba comprensión lectora y el total de la Prueba, únicamente se obtuvo una correlación significativa en la población con el ítem 5.

En la Subprueba de palabras simples no se obtuvo una correlación significativa, en tanto que en la Subprueba de lectura oral tuvo 0.03 y la Subprueba de alfabeto el 0.07. En

la Subprueba de palabras inventadas y comprensión lectora se observó una correlación no significativa. Se observó que la invarianza de los datos limitó la obtención de una medida de correlación de los ítems.

La correlación más fuerte se observó en la Subprueba de alfabeto con 0.26% de ítems que tuvieron correlación mayor en la muestra y la menor correlación de los ítems, fue en la Subprueba de lectura oral (0.08%) en los ítems con mayor correlación en la muestra y el ítem 44 que tuvo mayor correlación en la población. En la Subprueba de comprensión lectora se observaron las mayores correlaciones en la población a diferencia de las otras Subpruebas en donde la mayor correlación se observó en la muestra.

B. Dificultad (TCT/TRI)

En la Subprueba de alfabeto el 70% de los 27 ítems fueron fáciles según la TCT y desde la TRI el 60% de los ítems fueron fáciles de responder y el 15% de los ítems fueron difíciles para los estudiantes de 3ro. primaria. Lo cual refleja que la mayoría de los estudiantes de 3ro. primaria identificaron las 27 letras que conforman el abecedario, el cual es un microproceso que está más relacionado directamente con el acceso al léxico y al procesamiento sintáctico. Solo si los estudiantes tienen el prerrequisito del principio alfabético de forma integrada y eficaz pueden decodificar y comprender lo que leen. Ya que si desconocen algún grafema o fonema será difícil que puedan leer la palabra y comprender su significado. De acuerdo con los resultados es necesario abordar estrategias didácticas en el aula para apoyar a los estudiantes que aún desconocen el nombre o sonido de algunas letras, ya que se esperaría que en 3ro. primaria todos tengan esta habilidad desarrollada. Tal y como ha indicado RTI (2016) si las habilidades básicas fundamentales no se adquieren desde el principio, las brechas en los resultados de aprendizaje (entre los estudiantes que han dominado las habilidades fundamentales de lectura y los que no lo hacen) crecen con el tiempo.

En la Subprueba de palabras simples el 80% de los 20 ítems fueron fáciles de contestar según la TCT, lo cual significa que leyeron la palabra correctamente, pero no necesariamente implica que fueron los estudiantes que tienen una habilidad alta. Desde la TRI el 45% de los ítems se consideraron fáciles. Lo cual evidencia que la mayoría de los

estudiantes pueden leer palabras simples, sin embargo, existen palabras que debido a su longitud o composición son difíciles de leer en español, para los estudiantes de 3ro. primaria quienes en su mayoría el idioma materno es el K'iché.

En la Subprueba de palabras inventadas el 70% de los 20 ítems fueron fáciles desde la TCT y según la TRI el 45% de los ítems fue respondido correctamente. Lo cual evidenció que los estudiantes pueden decodificar palabras que no son conocidas y las cuales son difíciles de leer al reconocerlas visualmente.

En la Subprueba de lectura oral en TCT y TRI se confirmó que el 51% de los 129 ítems fueron difíciles para los estudiantes. La mitad de la Subprueba de lectura oral tenía ítems fáciles de responder para los estudiantes y la otra mitad de ítems fueron difíciles según TCT. Se infiere que lo anterior pudo deberse a la fluidez y velocidad, la cual se puede identificar debido a que esta fue la única Subprueba en la que los estudiantes tenían límite de tiempo (60 segundos) y la mayoría logró leer hasta la mitad de la historia con el límite de tiempo.

En la Subprueba de comprensión lectora el 50% de los ítems fueron fáciles desde la TCT y según la TRI el 40% de los ítems fueron difíciles. Los ítems con mayor índice de dificultad correspondieron a los que el estudiante debía de responder a partir de un detalle específico o una inferencia a partir de la comprensión de la historia. Los macroprocesos son las operaciones de alto nivel asociadas a la comprensión del texto o procesamiento semántico. La finalidad básica sobre el procesamiento semántico es determinar el grado de comprensión del mensaje de acuerdo con la Junta de Andalucía (s.f). Los ítems con menor índice de dificultad fueron los que el estudiante debía argumentar su respuesta por lo que había varias posibilidades para ser calificada como una respuesta correcta y se considera que influyó el criterio del aplicador para calificarla como correcta o incorrecta.

La dificultad de las Subpruebas de prueba de lectura inicial aplicada por STW, según TCT se identificó en la Subprueba de lectura oral con 0.38 como promedio de dificultad configurándose como la más difícil por tener menos respuestas correctas. La Subprueba de palabras simples fue la más fácil con 0.90 de promedio de dificultad, ya que fue la que concentró mayor cantidad de respuestas correctas. En las Subpruebas de alfabeto, palabras inventadas y comprensión lectora más de la mitad de los estudiantes evaluados acertaron los ítems. Al analizar la dificultad de las Subpruebas desde la TRI en la que se

determina a partir de la comparación del ítem y la habilidad del estudiante durante toda la prueba, se identificó que la Subprueba de comprensión lectora fue la más difícil, con 40% lo cual afirma que se requirió de mayor habilidad para responder correctamente a las 10 preguntas sobre la comprensión de la historia, que previamente el estudiante leyó en voz alta. La Subprueba de alfabeto fue la más fácil ya que obtuvo el 60% de ítems con dificultad negativa, lo que quiere decir que los ítems se ubicaron como son fáciles en la escala logística y se requirió menor habilidad para responderlos correctamente, tal y como lo establece la Junta de Andalucía (s.f) los microprocesos son los que están más directamente relacionados con el acceso al léxico y el procesamiento sintáctico. La finalidad básica de los microprocesos está en el procesamiento léxico y sintáctico, para determinar si el sujeto realiza la decodificación de forma integrada y eficaz, la cual se espera que en 3ro. Primaria sea un proceso de afianzamiento debido a que en los grados anteriores se tuvieron que adquirir estos microprocesos.

C. Discriminación y habilidad

Considerando que el índice de discriminación del reactivo es una medida para determinar si las competencias o habilidades que mide la prueba también las mide el ítem, se identificó que en la Subprueba de alfabeto el promedio de discriminación fue 0.18. Esto significa que la mayoría de los estudiantes dominaron la habilidad de reconocimiento de las letras. La habilidad de los estudiantes ya está instalada, sin embargo, la prueba logró discriminar. Más de la mitad de los ítems tienen un índice de discriminación correspondiente a 0.20 e indica que los estudiantes requirieron de una habilidad menor a 0 para responderlos correctamente, lo que indica que los ítems son muy fáciles. Hay 6 ítems a dos lógit de distancia, los cuales hacen referencia a las vocales y a una de las primeras consonantes que los estudiantes aprenden, los cuales requirieron una menor habilidad para responderlos.

En la Subprueba de palabras simples el promedio de discriminación fue de 0.31 lo cual significa que identifica bien la habilidad que se requiere para responder cada uno de los ítems. El 45% de los ítems de la Subprueba de palabras simples requirieron de una habilidad negativa, es decir fueron respondidos fácilmente por los estudiantes de 3°.

Primaria. Hay 4 ítems a dos lógit de distancia, los cuales hacen referencia a las palabras cortas que regularmente se aprenden a leer al combinar las vocales y las primeras consonantes, los cuales requirieron una menor habilidad para responderlos, debido a que son reconocidos por la mayoría de los estudiantes.

En la Subprueba de palabras inventadas el promedio de discriminación fue 0.26 lo cual se consideró que discrimina poco. Para responder al 40% de los ítems se requirió de una habilidad menor a 0 porque los estudiantes ya lo sabían. Hay 3 ítems a más de un lógit de distancia, dos de ellos están formados por cuatro letras las cuales podrían considerarse comunes y uno de los ítems tiene discriminación negativa, en estos ítems se requirió una menor habilidad para responderlos, porque fueron reconocidos por la mayoría de los estudiantes.

La Subprueba de lectura oral obtuvo el mayor promedio de discriminación en la prueba, con 0.54 lo cual confirmó que discriminó muy bien. Más de la mitad de los ítems discriminaron superior a 0.20 aunque 6 por cada 10 tiene una habilidad menor a 0.20. Se observó que en total 38 ítems fueron los que discriminaron la habilidad de los estudiantes. Los primeros ítems aparentemente fueron fáciles, porque más estudiantes acertaron en función del tiempo límite (60 segundos). Según la TCT la Subprueba parece difícil, sin embargo, en TRI se evidenció que se requirió de menor habilidad para responderla correctamente, lo cual expone que los estudiantes de 3ro. primaria en promedio están leyendo menos de las palabras que se esperaría que leyeran en un minuto. En todas las Subpruebas se observó que las respuestas están a menos de una desviación estándar a excepción de la Subprueba de lectura oral, el 7% de los ítems están a más de 11 desviaciones estándar.

La Subprueba de comprensión lectora obtuvo un promedio de 0.16, lo cual es considerado como una discriminación que puede mejorar. Esta fue la Subprueba que menos discriminó entre la habilidad del estudiante que se requería para responder los ítems. Se considera que pudo haber influido el criterio del evaluador al calificar al estudiante, ya que la mayoría de las respuestas tenían varias alternativas para determinar que la respuesta era correcta. Por lo que se pudo calificar una respuesta correcta, aunque esta estuviera incompleta. Al igual que la Subprueba de alfabeto que obtuvo un promedio de 0.18 considerado como una discriminación que puede mejorar. Sin embargo, en esta Subprueba

se esperaría que todos los estudiantes de 3ro. primaria pudieran reconocer cada una de las letras del abecedario por su nombre y sonido, como un prerrequisito para leer y comprender lo que leen.

D. RMSE

Se obtuvo un indicador de RMSE para todos los ítems de cada Subprueba y para los estudiantes. En el total de la prueba de lectura inicial el error cuadrático medio (RMSE) fue de 0.38 en todos los ítems de la prueba y de 0.06 en los estudiantes, lo cual es menor al parámetro establecido por Linacre. Se interpretan en relación con que mientras más bajo es el RMSE, disminuye el error (Linacre, J., 2002).

Respecto a todos los ítems en cada Subprueba, el RMSE mayor fue en la Subprueba de lectura oral (1.15), seguido de la Subprueba de palabras simples (0.88), la Subprueba de alfabeto (0.30), la Subprueba de palabras inventadas (0.21) y la Subprueba de comprensión lectora obtuvo el RMSE más bajo (0.11) en todos los ítems de la Subprueba.

El RMSE para los estudiantes fue mayor en la Subprueba de palabras simples (0.82), seguido de la Subprueba de alfabeto (0.67), la Subprueba de comprensión lectora (0.61), la Subprueba de palabras inventadas (0.55) y el más bajo RMSE fue en la Subprueba de lectura oral (0.18).

E. Infit (WMS) y Outfit (UMS)

El promedio de WMS y UMS osciló entre -0.9 a 0.06 el menor es en la Subprueba de lectura oral (Std. WMS 0.2 – Std. UMS 0.03) y el mayor se obtuvo en la Subprueba de palabras inventadas (Std. WMS 0.22 – Std. UMS 0.01). Lo que indicó que, en promedio el modelo se ajustó y los datos son razonablemente predecibles en el infit, que es la medida más difícil.

Para el WMS (Infit) y el UMS (Outfit) de todas las Subpruebas, la de alfabeto se ajusta el 100% debido a que los ítems son razonablemente predecibles, debido a que la mayoría respondió correctamente al ítem, aunque no necesariamente acertó en los otros

ítems. La varianza pareció estar afectada por la homogeneidad de la muestra y por la cantidad de datos.

Se observó que la habilidad que se requiere para responder está cerca de 0 en todas las Subpruebas. La única mediana negativa fue observada en la Subprueba de alfabeto, debido a que se alejó -0.33 de la media, es decir que requirió menor habilidad para responder. Para los estudiantes fue relativamente fácil, debido a que en 3ro. Primaria para poder leer y escribir se esperaría que ya conozcan todas las letras del abecedario.

F. Análisis diferencial

La Prueba de lectura inicial es una prueba referida al criterio debido a que identifica el dominio (o la falta del dominio) en relación con las competencias específicas y predeterminadas. Este tipo de prueba es utilizado cada vez más en los sistemas educativos, para evaluar el dominio que han logrado los estudiantes de las habilidades académicas esperadas en cada grado escolar. La información que brindan proporciona la base para determinar la intervención que se dará a los estudiantes rezagados, evaluar el currículo y determinar el éxito que tiene cada escuela para implementar el currículo (Gregory, 2012). Al realizar el análisis diferencial se confirmó que ninguna de las Subpruebas funciona diferencialmente con las características de la escuela a la que pertenecen los estudiantes; ni con la diferencia de edad; ni si es hombre o mujer; y tampoco con el idioma que hablan (K'iché, Castellano u otro).

G. Confiabilidad

La confiabilidad total fue alta en virtud que todos los indicadores estimados fueron superiores a 0,95. El rango entre el límite inferior (0,97) y superior (0,98) fue de 0.01. Al analizar la confiabilidad de las Subpruebas por separado, se confirmó que al eliminar cualquiera de los ítems dentro de ellas mismas, la confiabilidad fue superior a 0.97 lo cual significa que la prueba de lectura inicial posee una confiabilidad muy alta y cada uno de los ítems aporta consistentemente a cada Subprueba. Kaplan y Saccuzo (2009) sugieren que coeficientes de fiabilidad alrededor de .90 es un nivel elevado de confiabilidad. A pesar

de que en los análisis que se realizaron continuamente se identificó que la población evaluada es homogénea, la confiabilidad no se vio afectada por la poca varianza de los resultados.

Cuando se introdujeron todos los ítems, todas las Subpruebas presentan una confiabilidad mayor en la muestra (0.86 – 0.99) a excepción de la Subprueba de comprensión lectora (0.66) que obtuvo una confiabilidad mayor en la población (0.76). Al tomar la dificultad teórica de la prueba en la que se espera que las primeras Subpruebas son más fáciles, ya que requieren de menor habilidad para responderlas correctamente y la Subprueba de comprensión lectora requiere de mayor habilidad para responderla correctamente.

De igual manera cuando en el análisis de la confiabilidad se incluyó el total de la prueba de lectura inicial, se observó un leve incremento del indicador en todas las Subpruebas, con excepción de la Subprueba de lectura oral la cual se mantuvo igual.

1. Confiabilidad en la prueba de dos mitades sin el total

Toda la Prueba de lectura inicial en la mitad 1 y mitad 2 obtuvo un coeficiente de confiabilidad entre 0.97 y 0.98 en la población y la muestra, sin incluir el total de la prueba. En la Subprueba de alfabeto y comprensión lectora se observó una diferencia ligeramente mayor en la población, en ambas mitades. La mitad 1 tuvo un coeficiente de confiabilidad mayor que la mitad 2 en las Subpruebas de alfabeto (0.802 en la población y 0.799 en la muestra) y palabras inventadas (0.736 en la población y 0.817 en la muestra), tanto en la población como en la muestra, sin incluir el total de la Subprueba. La Subprueba de lectura oral fue la que mayor índice de confiabilidad obtuvo en comparación de las otras Subpruebas.

2. Confiabilidad en la prueba de dos mitades con el total

Toda la prueba de lectura inicial obtuvo en la mitad 1 y mitad 2 un coeficiente de confiabilidad entre 0.97 y 0.98 en la población y la muestra al incluir el total de la prueba.

En la Subprueba de comprensión lectora se observó una diferencia ligeramente mayor en la población, en ambas mitades y en la Subprueba de alfabeto en la mitad 2. La mitad 1 obtuvo un coeficiente de confiabilidad mayor que la mitad 2 en las Subpruebas de alfabeto (0.841 en la población y 0.844 en la muestra) y palabras inventadas (0.804 en la población y 0.861 en la muestra), tanto en la población como en la muestra al incluir el total de la Subprueba. La Subprueba de lectura oral fue la que mayor índice de confiabilidad obtuvo en comparación de las otras Subpruebas. Se confirmó que, al incluir el total de las Subpruebas, la confiabilidad se incrementó.

H. Análisis factorial exploratorio

Primeramente, se excluyeron los ítems que tenían varianza 0 y se obtuvo el KMO global a partir de 36 ítems los cuales se agruparon en un componente conformado por los 16 ítems, los cuales representan el 7.76% del total de los ítems de la prueba. Con un 10% de la varianza de la habilidad de esos ítems. Lo cual se comprobó que es necesario analizar el KMO por Subpruebas ya que cada una representa una variabilidad distinta y gradual.

En la Subprueba de alfabeto el KMO fue de 0.587 y en la Subprueba de palabras inventadas el KMO es de 0.561 ambos con máxima verosimilitud, lo cual indicó que poseen un indicador bajo para establecer si los resultados de la habilidad de la prueba son adecuados. La Subprueba de palabras simples obtuvo un KMO aceptable 0.74 de componentes principales y 0.67 con máxima verosimilitud, los clasifica en los ítems que tienen mayor correlación. Esta Subprueba explica menos de 1/3 de la habilidad de los estudiantes lo cual es bajo. En esta investigación solo se incluyeron los ítems de la Subprueba y se agruparon los ítems en dos factores. En 0.6 y 0.7 están los mayores puntajes del KMO en las Subpruebas, lo cual significa que no necesariamente los ítems deben agruparse en un solo factor porque no tienen relación directa entre ellas, ya que la habilidad de lectura es progresiva, tal y como lo indican Armbruster et al. (2003); Hudson, Lane, Pullen Torgesen (2005) la decodificación es el primer paso para el reconocimiento de las palabras, los lectores deben avanzar en su capacidad de decodificación hasta el punto en que se convierta en automática; entonces su atención será libre de pasar de las letras y palabras individuales a las ideas contenidas en el texto.

El modelo válido de KMO fue con componentes principales en la Subprueba de comprensión lectora con 0.607 y de 0.781 con máxima verosimilitud. Lo cual indica que la Subprueba posee un indicador suficiente y los resultados de la habilidad de la prueba son adecuados. Se considera que los ítems no tienen que estar relacionados entre sí, para formar un solo componente. Ya que se pueden formar varios factores lo cual responde al contexto, el aplicador, la concentración, el cansancio para responder, entre otros. La explicación de la varianza fue comprometida en la Subprueba de alfabeto en la primera corrida y en la Subprueba de palabras simples en la segunda corrida.

I. Calibración (Dificultad-Discriminación-Adivinación)

Al analizar las Subpruebas que conforman la prueba de lectura inicial se evidencia el aporte de los ítems a la prueba y se reconoce que no todos aportaron en el mismo nivel, ya que cada ítem responde a un grado de dificultad, discriminación y adivinación; sin embargo, hay una representación de todos los niveles en los que está distribuida la habilidad Lectora de los estudiantes evaluados.

Se identificó que la mitad de los ítems que conforman la Subprueba de alfabeto fueron fáciles según la calibración. En la Subprueba de palabras simples 2/10 ítems fueron fáciles. A pesar de que la población es homogénea se pudo identificar el índice de dificultad. En la Subprueba de palabras inventadas $\frac{1}{4}$ fue difícil para los estudiantes. En la Subprueba de lectura oral 8/10 ítems fueron difíciles. Y en la Subprueba de comprensión lectora fue la más difícil ya que el 40% de los ítems tuvieron un indicador menor a 0.71 que discriminan la habilidad que tienen los estudiantes para comprender lo que leen.

J. Total de puntuaciones

Se obtuvieron las puntuaciones de cada una de las cinco Subpruebas, en las que se identificaron que los primeros y últimos punteos tiene un error estándar mayor a uno y la dificultad está distribuida en negativos y positivos.

K. Validez

Debido a que la validez es un concepto unitario en el que se considera toda la evidencia acumulada con respecto a la interpretación de las puntuaciones de un instrumento para el propósito establecido (Mertens, 2005; Bostwick y Kyte, 2005) se puede establecer que la prueba de lectura inicial es válida debido a los diferentes análisis anteriormente discutidos, en los cuales se evidencia que cada Subprueba aporta a la medición de las habilidades lectoras de los estudiantes de 3ro. Primaria.

VI. CONCLUSIONES

La prueba de lectura inicial, aplicada por *School the World* en 9 escuelas ubicadas en Quiché, cuenta con las propiedades psicométricas de confiabilidad alta (límite inferior 0.97 y límite superior 0.98) y validez; a pesar de que la población evaluada se caracterizó por ser homogénea. La media de la prueba fue mayor cuando se correlacionaron los ítems con de cada una de las Subpruebas de alfabeto, palabras simples, palabras inventadas, lectura oral y comprensión lectora. El promedio de respuestas correctas fue de 113, lo cual corresponde a la mitad del total de ítems de la prueba. Los intervalos de confianza se encuentran en 0,9714 y 0,9862 lo cual indica que son cercanos a la estimación que corresponde a 0,9794 por lo tanto se concluye que la confiabilidad es alta y estable.

La mayoría de los ítems en la prueba tuvieron un indicador de dificultad bajo, lo cual significa que fue fácil de responder para la mayoría de los estudiantes de 3ro. Primaria según la TCT. Lo anterior también aporta al análisis que hará el equipo técnico de STW, para identificar si es más pertinente evaluar a los estudiantes de 2do. primaria considerando que sea en dos momentos, al inicio del ciclo escolar y al finalizar, a fin de que se pueda brindar un acompañamiento a los docentes durante el ciclo escolar.

En la Subprueba de alfabeto se observó que la mayoría de los estudiantes de 3ro. Primaria identificaron las 27 letras que conforman el abecedario, lo cual se confirmó con el promedio de discriminación de 0.18. Lo anterior confirmó que la mayoría de los estudiantes dominan la habilidad de reconocimiento de las letras y con el WMS (infit) y UMS (outfit) en el que se ajusta al 100% debido a que los ítems son razonablemente predecibles.

La Subprueba de palabras simples fue la que obtuvo el mayor intervalo (30); sin embargo, su correlación fue considerablemente alta entre 0.34 a 0.64 y sin correlaciones no significativas. La Subprueba fue la más fácil con 0.90 de promedio de dificultad, debido a que los estudiantes tuvieron más respuestas correctas. Obtuvo el mayor RMSE (0.82) para los estudiantes y un KMO aceptable 0.74 de componentes principales y 0.67 con máxima verosimilitud, que los clasificó en los ítems con mayor correlación.

Para la Subprueba de palabras inventadas el promedio de discriminación es 0.26 lo cual se considera que discrimina poco la habilidad de los estudiantes y sus respuestas. Para responder al 40% de los ítems se requirió de una habilidad menor a 0. Lo anterior debido a

que los contenidos son reconocidos por la mayoría de los estudiantes quienes en 3ro. Primaria se espera que puedan decodificar palabras nuevas, con lo cual aportan al incremento de su vocabulario. Esta Subprueba obtuvo 0.21 de RMSE el cual es uno de los más bajos y obtuvo el mayor promedio de Std. WMS 0.22 – Std. UMS 0.01.

La Subprueba de lectura oral es la que tuvo más ítems y en la que existe mayor correlación entre los ítems, ya que tiene un intervalo de 11 ítems con el total de la prueba. La mayoría de los ítems tuvieron una correlación mayor en la población a diferencia de las anteriores Subpruebas. La mitad de la Subprueba estuvo formada por ítems fáciles y la otra mitad contuvo ítems difíciles para los estudiantes; según la dificultad por TCT esta es la Subprueba más difícil, ya que es en la que menos respuestas correctas obtuvieron los estudiantes, debido a que la mayoría leyó menos de 129 palabras por minuto. Esta Subprueba obtuvo el mayor promedio de discriminación en la prueba, con 0.54 lo cual indica que discrimina muy bien. Y también obtuvo el mayor RMSE (1.15). Obtuvo el menor promedio de Std. WMS 0.2 – Std. UMS 0.03. En el análisis de la confiabilidad con y sin el total de la prueba de lectura inicial, fue la única Subprueba en que se mantuvo igual la confiabilidad.

En la Subprueba de comprensión lectora se obtuvieron mayores correlaciones en la población a diferencia de las otras Subpruebas en donde se observó la mayor correlación en la muestra. Según TRI esta Subprueba fue la más difícil con 40% lo cual afirma que se requiere de mayor habilidad del estudiante para responder las diez preguntas correctamente. A su vez esta Subprueba fue la que menos discriminó entre la habilidad del estudiante que se requería para responder los ítems, se considera que el criterio del aplicador influyó al momento de calificar esta Subprueba. En esta Subprueba el modelo válido de KMO es con componentes principales de 0.607 y de 0.781 con máxima verosimilitud. Lo anterior indicó que la Subprueba posee un indicador suficiente y los resultados de la habilidad de la prueba son adecuados.

Por medio de los datos numéricos se puede dar sentido a la información, la cual está vinculada con el fundamento teórico y la experiencia en la temática. Es importante destacar que a pesar de que los participantes tuvieron un comportamiento muy homogéneo, se logró obtener la confiabilidad, lo cual refleja la calidad y solidez de la prueba de lectura inicial aplicada por *School the World* con lo cual se rechaza la hipótesis nula.

VII. RECOMENDACIONES

Al equipo de School the World se le propone:

- Redactar un manual técnico para la aplicación de la prueba y el manual para los aplicadores con directrices que permitan unificar los criterios en un protocolo a seguir antes, durante y después de la aplicación de la prueba de lectura inicial. En el mismo se puede describir la base, los usos recomendados de la prueba y datos esenciales sobre la confiabilidad y validez.
- Realizar un muestreo estratificado (por asignación proporcional u óptima) para seleccionar la muestra probabilística en la que se puedan realizar las generalizaciones de los análisis.
- Contemplar para futuras evaluaciones la elección de muestras utilizando otro diseño en el que se incluya un grupo control, para que se puedan dar los resultados a los docentes de las escuelas e identificar las fortalezas y áreas de mejora para desarrollar las habilidades lectoras en los estudiantes por medio de estrategias para desarrollar los diferentes niveles de comprensión lectora. (literal, inferencial y crítico).
- Incluir en la prueba de lectura inicial, las preguntas relacionadas a los factores asociados al aprendizaje de la lectoescritura, con el cual se podrán hacer nuevos análisis estadísticos a partir de la información recopilada de los estudiantes, los docentes y el contexto.
- Considerar la aplicación de la prueba de lectura en 2do. primaria por medio de un pre-test (a inicio del ciclo escolar) y post-test (al finalizar el ciclo escolar) a fin de que se pueda trabajar un plan de intervención en el grado y darle continuidad en los siguientes grados en los cuales se afianzan las habilidades lectoras de los estudiantes.
- Aprovechar la aplicación de pruebas en la plataforma tecnológica en la cual se reduce el número de errores de aplicación y desde el cual se facilita el procesamiento de los datos.
- Explorar los recursos que hay en la página del Digecade⁴ con relación al proceso lector para cada uno de los grados de la Primaria, ya que hay material mediado que puede ser

⁴ <https://www.mineduc.gob.gt/DIGECADE/>

de apoyo en los planes de mejora o para la integración curricular, principalmente en las aulas que son multigrado.

- Crear una versión de la prueba en el lenguaje materno de los estudiantes del Quiché, quienes se comunican en el idioma K'iché. Se debe considerar que debido a la cantidad total de ítems (206) se requiere una muestra más grande. Se sugiere elegir muestras aleatorias y que puedan tener un idioma materno diferente a fin de realizar el análisis diferencial.

VIII. REFERENCIAS

- AERA, APA, & NCME. (1999). *Standards for educational and psychological testing*.
- Aliaga, J. (2007). *Psicometria: Tests Psicométricos, confiabilidad y validez*. Psicología: Tópicos de actualidad.
- American Psychological Association, American Educational Research Association, and National Council on Measurement in Education. (1985). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Anastasi, A., & Urbina, S. (1998). *Tests psicológicos*. México: Prentice Hall. ISBN -970-17-0186-0.
- Angoff, W. (1993). Perspectives on differential item functioning methodology. P.W. Holland y H. Wainer (eds.): *Differential Item Functioning*. Hillsdale, NJ: Lawrence Erlbaum.
- Armbruster, B. B. (2003). *Put reading first: The research building blocks for teaching children to read, kindergarten through grade 3*. Jessup, MD: National Institute for Literacy.
- Badenas, N. (2009). *El Modelo de Rasch aplicado a la medición multidimensional de la pobreza en España: bienaventurados los que tienen tele, aunque no coman carne*. Granada, España: Dialnet ISBN-978-84-691-8950-4.
- Barraza, A. (2007). *Apuntes sobre metodología de la Investigación. Confiabilidad?* México: Universidad Pedagógica de Durango.
- Barrios, M., & Cosculluela, A. (2013). *Fiabilidad*. GNUFDL PID_00198628: Universitat Oberta de Catalunya.
- Bogliaccini, J., Cardoso, M., & Rodríguez, F. (2005). *Construcción de índices. Confiabilidad: Alpha de Cronbach*. Uruguay: Laboratorio Metodológico, Facultad de Ciencias Humanas.
- Bogoya, D., Barragán, S., Contento, M., & Ocaña, A. (2014). *Calibración de instrumentos de evaluación - clasificación en matemáticas en la Universidad Jorge Tadeo Lozano*. Bogotá, Colombia: Revista Complutense de Educación. Vol. 25 Núm. 4 ISSN: 1130-2496.
- Bond, T., & Fox, C. (2015). *Applying the Rasch Model. Fundamental Measurement in the Human Sciences. Third Edition*. New York and London.: Routledge. Taylor & Francis Group.

- Bostwick, G. y Kyte, N. . (2005). Measurement. En Grinell, R.M. y Unrau, Y.A. (Eds.). *Social work: Research and evaluation. Quantitative and qualitative approaches (7ª ed.)* Nueva York: Oxford University Press., (pp.97-111).
- Chall, J. (1996). *Learning to Read, the Great Debate*. Harcourt Brace.
- Consejo Nacional de Educación. (2010). Consejo Nacional de Educación de Políticas Educativas. Guatemala.
- Cortina, J. (1993). *What is coefficient alpha? An examination of theory and applications*. J Appl Psychol. ;78:98-104.
- Cotto, E., & Del Valle, M. (2017). *Modelo para explicar y predecir el aprendizaje de la lectura para Guatemala*. Guatemala: Dirección General de Evaluación e Investigación Educativa, Ministerio de Educación.
- Del Valle, M., Cotto, E., & Mirón, R. (2017). *Evaluación de lectura inicial en estudiantes de segundo primaria*. Guatemala: Dirección General de Evaluación e Investigación Educativa, Ministerio de Educación. Disponible en red: <http://www.mineduc.gob.gt/digeduca>.
- Digeduca. (2018). Resultados generales de la evaluación educativa. Guatemala.
- Drasgow, F. (2016). *Technology and testing. Nueva*. Nueva York: Routledge.
- Embretson, S. y Reise, S. (2000). *Item response theory for psychologists*. Mahwah, NJ.: LEA.
- Facultad de Educación. (2018). *Informe para STW*. Guatemala: UVG.
- Family Search. (2019). *Family Search*. Retrieved from Plantilla:Quiché Departamento Mapa: https://www.familysearch.org/wiki/es/Plantilla:Quich%C3%A9_Departamento_Mapa
- Feldt, L. y Brennan, R. (1989). *Reliability*. In Linn LR. *Educational measurement*. 3rd ed. New York: Macmillan. p. 105-46.
- Fortín, Á. (2013). *Evaluación Educativa Estandarizada en Guatemala: Un camino recorrido*. Guatemala: Ministerio de Educación.
- Frías, D. (2013). *Alfa de Cronbach y consistencia interna de los ítems de un instrumento de medida*. . Valencia, España.: Facultad de Psicología. Universidad de Valencia. Recuperado de <http://www.uv.es/~friasnav/AlfaCronbach.pdf>.
- García, J., & González, M. y. (2001). *Introducción a la investigación en educación. Primera Edición*. Madrid, España: Universidad Nacional de Educación a Distancia. Editorial IMPRESA. .

- Gómez, J., & Hidalgo, M. (2003). *Desarrollos recientes en Psicometría*. Avances en Medición, 1 (1), 17-36 .
- González M. (2008). *El análisis de reactivos en el modelo de Rasch. Manual Técnico A.* . Sonora, México.: Serie Medición y Metodología. México: Universidad de Sonora, Instituto Nacional para la Evaluación de la Educación. [
- Gotzmann, A., & Bahry, L. (2012). *Software Review*. Research & Practice in Assessment.
- Gregory, R. (2012). *Pruebas Psicológicas. Historia, principios y aplicaciones. Sexta edición*. México: Pearson.
- Hasbrouck, J., & Tindal, G. (2006). *Oral reading fluency norms: A valuable assessment tool for reading teachers*. International Reading Association (pp. 636–644) doi:10.1598/RT.59.7.3.
- Hernández, R. (2014). *Metodología de la Investigación*. México D.F.: McGRAW-HILL, 6ta Edición.
- Hudson, R., Lane, H., Pullen, P., & Torgesen, J. (2009). *The Complex Nature of Reading Fluency: A Multidimensional View*. Taylor & Francis Group, LLC.
- Ibáñez, N. (2016). *Validación del Cuestionario de Evaluación de las Relaciones Familiares Básicas (CERFB)*. Barcelona, España: Departament de Psicologia.
- IBM. (2019). *Analytics*. Retrieved from IBM SPSS software: <https://www.ibm.com/analytics/spss-statistics-software>
- Instituto Nacional de Estadística. (2010). Boletín informativo departamento de Quiché. *Volumen 4, No. 4*.
- Instituto Nacional de Estadística. (2011). *Mapas de pobreza rural en Guatemala. Resumen Ejecutivo*. Guatemala.
- Instituto Nacional de Estadística. (2014). *Caracterización departamental Quiché*. Guatemala.
- Jiménez, K., & Montero, E. (2013). *Aplicación del modelo de Rasch, en el análisis psicométrico de una prueba de diagnóstico en matemática*. Costa Rica: Revista digital Matemática, Educación e Internet. Vol 13. No. 1.
- Joya, A., Us, K., & López, K. (2019, marzo 30). Comunicación por correo electrónico. Guatemala - Honduras, Guatemala - Honduras.
- Junta de Andalucía. (s.f). *Guía de evaluación de destrezas lectoras. Educación Primaria*. Sevilla, España: Consejería de Educación, Cultura y Deporte.
- Kaplan, R. y. (2009). *Psychological testing : principles, applications, and issues*. Belmont, CA : Wadsworth Cengage Learning,.

- Karabatsos, G. (2000). *A critique of Rasch residual fit statistics*. . Journal of Applied Measurement, 1, 152-176.
- Kristof, W. (1974). *Estimation of reliability and true score variance from a split of a test into three arbitrary parts*. Psychometrika. 39:491-99.
- Learning Renaissance. (2015). *STAR Early Literacy™ Technical Manual*. United Kingdom. Australia.
- Ley de Educación Nacional. (1991, enero 9). Guatemala: Congreso de la República de Guatemala.
- Linacre, J. (2002). *What do Infit and Outfit, Mean-square and Standardized mean?* Retrieved from Rasch Measurement Transactions: www.rasch.org/rmt/rmt162f.htm
- López, K. (2019). *Caracterización de las escuelas de School the World*. Guatemala.
- Martín, N., Díaz, C., Córdoba, G., & Picquart, M. (2011). *Calibración de una prueba de química por el modelo de Rasch*. México: Revista electrónica de investigación educativa. REDIE vol.13 no.2. versión On-line ISSN 1607-4041.
- Martínez, R. (2005). *Psicometría: Teoría de los test psicológicos y educativos*. Madrid, España.: Síntesis Psicología.
- Mertens, D. (2005). *Research and evaluation in Education and Psychology: Integrating diversity with quantitative, qualitative, and mixed methods (2ª ed.)*. Thousand Oaks: Sage. .
- Messick, S. (1988). The once and future issues of validity: Assessing the meaning and consequences of measurement. *En H. Wainer y H.l. Braun (Eds.), Test validity*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Messick, S. (1989). Validity. The specification and development of tests of achievement and ability. *En R. L. Lino (Ed.), Educational Measurement (3th edition)* Washington, DC: American Council on Education.
- Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. . *Educational Measurement: Issues and Practice*, 15, 5-12.
- Meyer, J. (2014). *Applied Measurement with jMetrik*. New York.: Routledge. Taylor & Francis.
- Mineduc. (2008, agosto 12). Acuerdo Gubernativo Número 225-2008. Guatemala: Diario de Centro América.
- Mineduc. (2010, julio 15). Acuerdo Ministerial No. 1171-2010. Guatemala: Diario de Centro América.

- Mineduc. (2012, septiembre 14). *Currículo Nacional Base Guatemala*. Retrieved from Competencias Tercer Grado: http://cnbguatemala.org/index.php?title=Competencias_Tercer_Grado&redirect=no
- Mineduc. (2018). *Digeduca*. Retrieved from Resultados generales de la evaluación educativa: www.mineduc.gob.gt/digeduca
- Mineduc. (2019). Anuario Estadístico de la Educación. *Sistema de Registro Educativo - SIRE-*, Guatemala.
- Mineduc. (2019). *Base legal de Digeduca*. Retrieved from Digeduca: <https://www.mineduc.gob.gt/digeduca/>
- Mineduc. (s.f). *Mineduc*. Retrieved from DIGEBI: <https://www.mineduc.gob.gt/DIGEBI/mapaLinguistico.html>
- Mineduc. (2018). *Guatemala en PISA-D. Programa Internacional de Evaluación de Estudiantes*. . Guatemala: Dirección General de Evaluación e Investigación Educativa.
- Morales, P. (2008). *Estadística aplicada a las Ciencias Sociales*. Madrid, España: Universidad Pontificia Comillas.
- Muñiz, J. (1996). *Psicometría*. Madrid, España: Editorial Universitaria.
- Muñiz, J. (2003). *Toría Clásica de los Test*. Madrid, España: Ediciones Pirámide.
- Muñiz, J. (2010). *Las teorías de los test: Teoría Clásica y Teoría de Respuesta a los ítems*. Papeles del psicólogo, 3 (1), 57-66.
- Muñiz, J. (2018). *Introducción a la Psicometría*. Madrid, España: Editoriales Pirámide.
- Muñiz, J., Fidalgo, A., & García-Cueto, E. y. (2005). *Análisis de los Ítems. Primera Edición*. . Madrid: España: Editorial La Muralla S.A.
- Murphy, K. y. (2005). *Psychological Testing: Principles and Applications*, 6th Edition. *Pearson*.
- OCDE. (2015). PISA, Resultados clave.
- OCDE/ Instituto Nacional de Calidad y Evaluación. (2000). *Proyecto PISA. La medida de los conocimientos y destrezas de los alumnos: Un nuevo marco para la evaluación*. Madrid.
- OREALC. (2005). *Competencias para la vida en las evaluaciones de lectura y escritura (SERCE – LLECE)*. Buenos Aires, Argentina.
- Osburn, h. (2000). *Coefficient alpha and related internal consistency reliability coefficients*. . Psychol Methods. 5:343-55.

- Pasquali, L. (2009). *Psychometrics*. Brasilia, Brasil.: Rev Esc Enferm USP.
- Prieto, G., & Delgado, A. (2003). *Análisis de un test mediante el modelo de Rasch*. Universidad de Salamanca, Oviedo, España: Psicothema 2003. Vol. 15, nº 1, pp. 94-100.
- Psychomeasurement Systems. (2018). *Psychomeasurement Systems, LLC*. Retrieved from Software and Consulting Services: <https://itemanalysis.com/jmetrik-download/>
- Quim, M. y Santos, J. (2015). *Análisis de ítems con jMetrik*. Guatemala. Recuperado de: http://www.mineduc.gob.gt/digeduca/documents/cuadernillosTecnicos/Analisis_ite ms_Jmetrik.pdf: Guatemala: Dirección General de Evaluación e Investigación Educativa, Ministerio de Educación.
- Quim, M.; Cruz, A. y Santos, J. (2014). *Informe departamental y municipal de primaria 2013*. Guatemala: Dirección General de Evaluación e Investigación Educativa, Ministerio de Educación.
- RTI, International. (2012). *Tangerine™ - Electronic Data Collection Tool for Early Reading and Math Assessments*. Kenya Field Trial Report: SUMMARY.
- RTI, International. (2016). *Early Grade Reading Assessment (EGRA) Toolkit, Second Edition*. Washington, DC.
- s.n. (2019). *Creating and Editing Instruments in Tangerine: User Manual*. <http://www.tangerinecentral.org/project-1>.
- School the World. (2019). Retrieved from <https://schooltheworld.org/founders-story/>
- School the World. (2019). *Prueba de lectura inicial en Tangerine*. Guatemala.
- School the World. (2019, 05 02). *School the World*. Retrieved from <http://schooltheworld.org/index.php>
- SEGEPLAN. (2018). Retrieved from Prioridades Nacionales de Desarrollo: <http://pnd.gt/>
- Silva, A. (1987). *Algunas consideraciones sobre la utilización del Coeficiente r de Pearson como índice de acuerdo entre observadores*. México: Anuario de Psicología. Escuela Nacional de Estudios Profesionales de la Universidad Nacional Autónoma de México .
- Sireci, S. y Faulkner-Bond, M. (2016). *The times they are A-changing, but the song remains the same. Future issues and practices in test validation*. . Nueva York: Guilford Press.: En C. S. Wells y M. Faulkner-Bond (eds.), Educational measurement. From foundations to future (pp. 435-448).
- Soriano, M., Miranda, A., Soriano, E., Nieves, F., & Félix, V. (2011). *Examining the efficacy of an intervention to improve fluency and reading comprehension in Spanish children with Reading Disabilities*. . International Journal of Disability, Development and Education, 58(1), 47-59 .

- SPSS. (2010). *IBM SPSS Statistics. Editor de datos.*
- Tabachnick, B. y. (2011). *Using multivariate statistics.* . New York.: Harper & Row.
- Tennant, J. (2015). *A placement Test Revisited: An ongoing validation study of one university`s test.*
- Torgesen, J. K. (2004). Avoiding the devastating downward spiral: The evidence that early intervention prevents reading failure. . *American Educator.*, 8.
- Tristán, A. (2001). *Análisis de Rasch para todos: una guía simplificada para evaluadores educativos.* México: Centro Nacional de Evaluación.
- UNESCO. (2015). Informe de resultados Tercer Estudio Regional Comparativo y Explicativo. *Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación,* Santiago, Chile.
- University of Washington. (2019, junio 23). *University of Washington.* Retrieved from Office of Educational Assessment. Understanding Item Analyses.: <http://www.washington.edu/assessment/scanning-scoring/scoring/reports/item-analysis/>
- USAID. (2006). *Increasing education access quality and equity in Guatemala. Latin American and Caribbean Education Profiles 1999–2004.* Guatemala.
- USAID. (2009). *Manual para la evaluación inicial de la lectura en niños de educación primaria .* EdData II.
- USAID. (2009). *Using Opportunity to Learn and Early Grade Reading Fluency to Measure School Effectiveness in Guatemala.* Washington.
- USAID. (2013). *Aprendizaje de la Lectoescritura.* Guatemala.
- USAID-Mineduc. (2017). *La enseñanza de la comprensión lectora. Xnaq'tzb'il tu'n tel tniky' ti'j u'jb'il (mam). Libro para docentes del Nivel Primario.* Guatemala.
- Vaughn Gross Center for Reading & Language Arts. (2005). *Introduction to the 3-tier reading model (3rd ed.).* Austin, TX: The University of Texas.
- Vaughn, S., & Linan-Thompson, S. (2004). *Researchbased methods of reading instruction grades K-3.* Alexandria, VA: Association for Supervision and Curriculum Development.
- Visión Mundial. (2018). *Literacy Boost instrument Guatemala_ESP.* Guatemala.
- World Vision. (2017). *STAR Guidance.*
- World Vision International. (2017). *STAR Guidance (School-based Test About Reading).*

IX. ANEXOS

Anexo 1. Mapa del departamento de Quiché

Municipios del Departamento Quiché



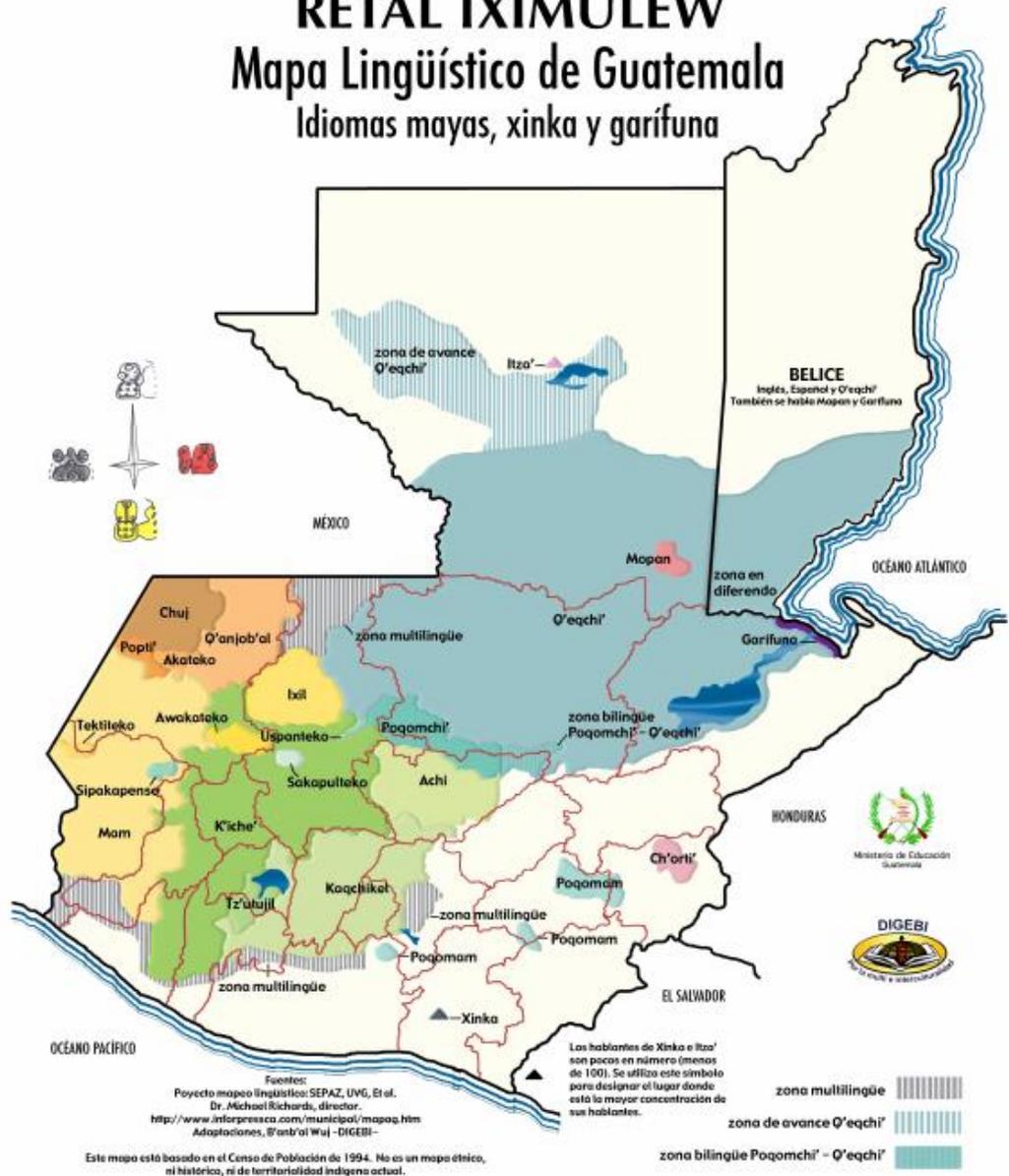
(Family Search, 2019)

Anexo 2. Mapa lingüístico de Guatemala

RETAL IXIMULEW

Mapa Lingüístico de Guatemala

Idiomas mayas, xinka y garífuna



(Mineduc, s.f)

Anexo 3. Caracterización de las escuelas de STW

Escuela	Distancia de la cabecera departamental	Tiempo de intervención	Proyecto de infraestructura	Intervenciones	Estado actual
EORM ⁵ Agua Viva	A 61 Km desde Santa Cruz del Quiché.	---	Inició en el año 2012	Talleres a padres y madres de familia, compra de libros y actividades de lectura.	---
EORM Chocruz	A 15.5 Km desde Santa Cruz del Quiché.	*1er año.	Se inauguró la construcción del edificio escolar en julio 2018	Talleres con padres y madres de familia y la promoción de la lectura con los niños y niñas de la escuela.	---
EORM Chuguexa	A 63 Km en carretera asfaltada sobre la ruta interamericana.	6 años	Inició en el año 2010 a 2018	No más capacitaciones por decisión del Consejo de Padres.	Monitoreo
EORM Chulumal JM	A 18.5 Km desde Santa Cruz del Quiché, y menos de 5 minutos de Chichicastenango.	---	Inició en 2012*	Capacitación a padres y madres de familia, compra de libros y lecturas con los niños y niñas de la escuela.	Monitoreo
EORM Chulumal JV	Ida y vuelta 37 Km.	7 años	Inició en enero 2012 a 2018.	---	---
EORM Mactzul VII	A 10 Km, pertenece a la municipalidad de Chichicastenango, en las montañas y de difícil acceso, esta entre Santa Cruz y Chichicastenango.	6 años	Inició en agosto del año 2010 a 2018	---	Monitoreo 3 o 4 veces al año.
EORM Panimaché Quinto	Ida y vuelta 60 km.	6 años	Inaugurado en marzo del año 2012.	Construcción de un área de juego.	Etapas de monitoreo.

Escuela	Distancia de la cabecera departamental	Tiempo de intervención	Proyecto de infraestructura	Intervenciones	Estado actual
---------	--	------------------------	-----------------------------	----------------	---------------

⁵ Escuela Oficial Rural Mixta

EORM Río Mactzul	A 15.5 Km desde Santa Cruz del Quiché a la comunidad Río Mactzul. Es parte Chichicastenango se encuentran en las montañas y no es de fácil acceso.	6 años	Se construyó el edificio escolar en el 2010.	Talleres a padres, madres de familia, compra de libros, acompañamiento en lectura con niños y niñas en el año 2012.	Monitoreo
EORM Xalbaquej	A 18 Km desde Santa Cruz del Quiché. Parte de Chichicastenango, pero a larga distancia y profundo en un valle, 40 minutos desde Chichicastenango.	---	Inició actividades en el 2010.	Talleres a padres y madres de familia y lectura con los niños y niñas.	---
EORM Xesic	A 4 km en carretera de terracería.	5 meses *1er año.	Inició en julio 2018		

(López, 2019)

Anexo 4. Códigos de las variables y los ítems

Código	Etiqueta de la variable
D1	NÚMERO DE ESCUELA
D2	LETRA DE LA ESCUELA
D3	CÓDIGO DE LA ESCUELA
D4	SI ESTÁ EN LA ESCUELA
D5	GRADO 3RO.
D11	NOMBRE DE LA ESCUELA
D12	NOMBRE ESTUDIANTE
D13	SEXO HOMBRE
D14	SEXO MUJER
D14_1	SEXO UNIDA (MUJER Y HOMBRE)
D15	EDAD
D16	IDIOMA K'ICHÉ
D17	IDIOMA CASTELLANO
D18	IDIOMA OTRO
D19_T	IDIOMA TOTAL
D20	IDIOMA K'ICHÉ CASTELLANO
D21	IDIOMA K'ICHÉ OTRO
D22	IDIOMA CASTELLANO OTRO
TC	TIEMPO CALCULADO
TD	TIEMPO DESTINADO
AT2	ALFABETO TOTAL
ST2	PALABRAS SIMPLES TOTAL
IT2	PALABRAS INVENTADAS TOTAL
CT	COMPRESIÓN LECTORA TOTAL
T_G	TOTAL GLOBAL DE LA PRUEBA

Anexo 5. Código de las Subpruebas de la prueba de lectura inicial

Subprueba	Código de la Subprueba
Alfabeto	Alfabeto (a) (A)
Palabras Simples	Simple (s) (S)
Palabras Inventadas	Inventadas (i) (I)
Lectura Oral	Lectura (l) (L)
Comprensión de Lectura	Comprensión (c) (C)

Anexo 6. Código de las escuelas en las que se aplicó la prueba de lectura inicial

Nombre de la Escuela	Código de la Escuela
Agua Viva	A
Chocruz	B
Chuguexa	C
Chulumal Jornada Matutina	D
Chulumal Jornada Vespertina	E
Mactzul VII	F
Panimaché	G
Río Mactzul	H
Xalbaquej	I
Xesic	J

Anexo 7. Tablas del análisis diferencial

Tabla 65. Análisis diferencial entre Escuelas con la Subprueba alfabeto

	A	B	C	D	E	F	G	H	I	J
A	0	0	0	0	0	0	0	0	0	0
B	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0	0
D	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	0	0
J	0	0	0	0	0	0	0	0	0	0

Tabla 66. *Análisis diferencial entre edades con la Subprueba alfabeto*

	8	9	10	11	12	13	14
8	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0

Tabla 67. *Análisis diferencial entre idioma con la Subprueba alfabeto*

	Castellano	K'iché	Otro
Castellano	0	0	0
K'iché	0	0	0
Otro	0	0	0

Tabla 68. *Análisis diferencial entre sexo con la Subprueba alfabeto*

	Mujer	Hombre
Mujer	0	0
Hombre	0	0

Tabla 69. *Análisis diferencial entre Escuelas con la Subprueba palabras simples*

	A	B	C	D	E	F	G	H	I	J
A	0	0	0	0	0	0	0	0	0	0
B	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0	0
D	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	0	0
J	0	0	0	0	0	0	0	0	0	0

Tabla 70. *Análisis diferencial entre edades con la Subprueba palabras simples*

	8	9	10	11	12	13	14
8	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0

Tabla 71. *Análisis diferencial entre idiomas con la Subprueba palabras simples*

	Castellano	K'iché	Otro
Castellano	0	0	0
K'iché	0	0	0
Otro	0	0	0

Tabla 72. *Análisis diferencial entre sexo con la Subprueba palabras simples*

	Mujer	Hombre
Mujer	0	0
Hombre	0	0

Tabla 73. *Análisis diferencial entre Escuelas con la Subprueba palabras inventadas*

	A	B	C	D	E	F	G	H	I	J
A	0	0	0	0	0	0	0	0	0	0
B	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0	0
D	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	0	0
J	0	0	0	0	0	0	0	0	0	0

Tabla 74. *Análisis diferencial entre edades con la Subprueba palabras inventadas*

	8	9	10	11	12	13	14
8	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0

Tabla 75. *Análisis diferencial entre idiomas con la Subprueba palabras inventadas*

	Castellano	K'iché	Otro
Castellano	0	0	0
K'iché	0	0	0
Otro	0	0	0

Tabla 76. *Análisis diferencial entre sexo con la Subprueba palabras inventadas*

	Mujer	Hombre
Mujer	0	0
Hombre	0	0

Tabla 77. *Análisis diferencial entre Escuelas con la Subprueba lectura oral*

	A	B	C	D	E	F	G	H	I	J
A	0	0	0	0	0	0	0	0	0	0
B	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0	0
D	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	0	0
J	0	0	0	0	0	0	0	0	0	0

Tabla 78. Análisis diferencial entre edades con la Subprueba lectura oral

	8	9	10	11	12	13	14
8	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0

Tabla 79. Análisis diferencial entre idiomas con la Subprueba lectura oral

	Castellano	K'iché	Otro
Castellano	0	0	0
K'iché	0	0	0
Otro	0	0	0

Tabla 80. Análisis diferencial entre sexo con la Subprueba lectura oral

	Mujer	Hombre
Mujer	0	0
Hombre	0	0

Tabla 81. Análisis diferencial entre Escuelas con la Subprueba comprensión lectora

	A	B	C	D	E	F	G	H	I	J
A	0	0	0	0	0	0	0	0	0	0
B	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0	0
D	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	0	0
J	0	0	0	0	0	0	0	0	0	0

Tabla 82. *Análisis diferencial entre edades con la Subprueba comprensión lectora*

	8	9	10	11	12	13	14
8	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0

Tabla 83. *Análisis diferencial entre idiomas con la Subprueba comprensión lectora*

	Castellano	K'iché	Otro
Castellano	0	0	0
K'iché	0	0	0
Otro	0	0	0

Tabla 84. *Análisis diferencial entre sexo con la Subprueba comprensión lectora*

	Mujer	Hombre
Mujer	0	0
Hombre	0	0