

UNIVERSIDAD DEL VALLE DE GUATEMALA

Facultad de Educación



**Análisis de reactivos y validación de la prueba diagnóstica de
lectura y escritura emergente**

Trabajo de graduación presentado por Rossen Jenniffer Mirón López, para optar
al grado académico de Magister en Medición, Evaluación e Investigación
Educativas

Guatemala

2016

UNIVERSIDAD DEL VALLE DE GUATEMALA

Facultad de Educación



**Análisis de reactivos y validación de la prueba diagnóstica de
lectura y escritura emergente**

Trabajo de graduación presentado por Rossen Jenniffer Mirón López, para optar
al grado académico de Magister en Medición, Evaluación e Investigación
Educativas

Guatemala

2016

Vo.Bo. (f) 
M.A. Bianca Lissette Argueta Pensamiento

Tribunal Examinador:

(f) 
M.A. Bianca Lissette Argueta Pensamiento

(f) 
M.A. Leslie Vanessa Rosales Flores

(f) 
M.A. María Raquel Montenegro Mañoz

Guatemala, 14 de enero de 2016

PREFACIO

Este triunfo no es solo mío, por ello agradezco a todas las personas que contribuyeron a alcanzar esta meta profesional. Especialmente doy las gracias...

A Dios, por ser siempre el sostén en mi vida y en mi camino.

A mi madre Vilma López, porque ella me ha dado luz, guía, amor, dedicación, valores y todo lo que soy.

A mi esposo Mario Soto y mi hija Isabella, por brindarme de manera incondicional su apoyo, por los desvelos que el trabajo de tesis nos implicó a todos y por las largas conversaciones sostenidas en beneficio de mi aprendizaje... en realidad gracias por todo.

A mi hermana Nineth de Scott y su familia por estar siempre a mi lado apoyándome y creyendo en mí.

A mis abuelos Rosenda Barrios y Francisco López⁺, porque mi vida no sería la misma sin ellos.

A mi madrina Silvia de O'Connell por creer en mí, y porque esa confianza me ha permitido alcanzar y disfrutar muchas alegrías y triunfos.

A mi asesora de tesis y tribunal examinador porque provocaron el deseo de superarme, el anhelo de aprender, pero aprender bien y en definitiva porque son profesionales que admiro por su trayectoria en el campo de la evaluación y en la disciplina en que se desenvuelven.

A la Doctora Mariela Zelada por su paciencia y enseñanza durante el tiempo que compartimos durante la carrera.

A la Universidad del Valle de Guatemala, por ser mi casa de estudios.

A las autoridades de la Dgeduca por brindarme la oportunidad de continuar con mis estudios.

CONTENIDO

Prefacio	v
Índice de gráficas	x
Índice de tablas.....	xi
Resumen.....	xii
I. INTRODUCCIÓN	1
II. JUSTIFICACIÓN	4
III. ANTECEDENTES.....	6
1. Currículum Nacional Base del Nivel de Educación Preprimaria.....	13
IV. MARCO CONCEPTUAL.....	16
A. La lectura y escritura en los primeros grados de escolaridad.....	17
1. Elementos para el desarrollo de la lectura y escritura emergente	19
a. Desarrollo del lenguaje oral	20
c. Desarrollo de la expresión oral	21
d. Desarrollo de la comprensión oral	21
e. Desarrollo del vocabulario	22
f. Animación a la lectura	23
g. Desarrollo de la conciencia fonológica	24
h. Desarrollo de la conciencia del lenguaje escrito	25
i. Desarrollo de la grafomotricidad	25
2. La lectura y escritura en el Currículum Nacional Base (CNB).....	28
a. Estándares Educativos del Nivel de Educación Preprimaria relacionados al aprendizaje de la lectura y escritura emergente.....	29

B.	Conceptos básicos del test.....	31
1.	Tipos de test	32
2.	Clasificación del test según su función	33
3.	Construcción del test	35
a.	Finalidad del test.....	35
b.	Marco del test.....	36
4.	Construcción del ítem	37
a.	El formato	37
b.	Revisión de los ítems	38
c.	Revisión por expertos	39
d.	Estudio piloto.....	39
e.	Estudio de campo.....	40
5.	Propiedades de los ítems	41
a.	Validez	41
1)	Fuentes internas de validez	42
b.	Dimensionalidad	43
1)	Análisis factorial exploratorio.....	45
2)	Análisis factorial confirmatorio	45
c.	Independencia local	45
d.	Los modelos teóricos de la TRI	46
1)	Modelo de un parámetro	47
2)	Modelo de dos parámetros	48
6.	Tamaño de la muestra	51
7.	Uso adecuado de los test	51
a.	Ámbito de aplicación	52
b.	Destinatarios	52
c.	Conocimientos de aplicación	53

VI.	METODOLOGÍA	54
A.	Definición del problema	54
B.	Objetivos	55
C.	Diseño.....	55
D.	Descripción de la población y de la muestra	56
E.	Obtención y análisis de datos	56
1.	Obtención de los datos	57
a.	<i>Software</i> utilizado	58
2.	Análisis de los datos.....	59
a.	Análisis factorial exploratorio.....	60
b.	Análisis Rash	60
c.	Parámetro de dificultad del ítem	61
d.	Parámetro de discriminación del ítem.....	62
e.	Coefficiente Kappa.....	63
3.	Codificación de los datos	63
F.	Instrumento para la recolección de información	64
1.	Contextualización del manual de aplicación.....	68
G.	Procedimientos	69
VII.	PRESENTACIÓN DE RESULTADOS	72
A.	Tabla de alcance y secuencia.....	72
B.	Juicio de expertos	73
C.	Comprobación de los supuestos de la TRI	74
1.	Dimensionalidad de la prueba e independencia local	74
2.	Parámetros de los ítems.....	76
a.	Curvas características del ítem.....	78
b.	Ajuste de las respuestas a los ítems	82

c.	Función de información de los ítems y del test.....	84
d.	Ajuste del test.....	85
VIII.	DISCUSIÓN DE LOS RESULTADOS / HALLAZGOS.....	86
A.	Tabla de alcance y secuencia.....	86
B.	Juicio de expertos	88
C.	Comprobación de los supuestos de la TRI	89
D.	Parámetros de los ítems	90
IX.	CONCLUSIONES	93
X.	RECOMENDACIONES	95
XI.	REFERENCIAS BIBLIOGRÁFICAS.....	96
XII.	ANEXOS	103
	Anexo 1. Carta de solicitud para realizar aplicación en el establecimiento seleccionado para la obtención de datos.	103
	Anexo 2. Instrumento para validación de expertos	105
	Anexo 3. Matriz de correlación.....	108
	Anexo 4. Sintaxis de r	109

ÍNDICE DE GRÁFICAS

Gráfica	Página
Gráfica 1. <i>CCI de 5 ítems de conciencia fonológica</i>	79
Gráfica 2. <i>CCI de 6 ítems de escritura emergente e inicial</i>	80
Gráfica 3. <i>CCI de 7 ítems de decodificación</i>	80
Gráfica 4. <i>CCI de 8 ítems de desarrollo de la conciencia del lenguaje escrito</i>	81
Gráfica 5. <i>CCI de 6 ítems de comprensión oral</i>	82
Gráfica 6. <i>Mapa de ítems (Infit y Outfi)</i>	83
Gráfica 7. <i>Función de información del Ítem y del test</i>	84
Gráfica 8. <i>Ajuste del test</i>	85

ÍNDICE DE TABLAS

Tabla	Página
Tabla 1. <i>Estándares educativos del área de Comunicación y Lenguaje</i>	30
Tabla 2. <i>Estándares educativos del área Destrezas de Aprendizaje</i>	30
Tabla 3. <i>Distribución de la muestra evaluada para la aplicación piloto de la Prueba diagnóstica para primero primaria</i>	57
Tabla 4. <i>Determinación del nombre que corresponde a las subpruebas que contiene la Prueba diagnóstica para primer grado de primaria</i>	65
Tabla 5. <i>Modificaciones al Manual de aplicador</i>	69
Tabla 6. <i>Tabla de alcance y secuencia para equiparlo al contexto guatemalteco</i>	73
Tabla 7. <i>Fuerza de concordancia entre juicio de expertos y validación con docentes</i> ...	74
Tabla 8. <i>Índices de ajuste para un factor</i>	75
Tabla 9. <i>Cuatro factores agrupados para la Prueba diagnóstica</i>	76
Tabla 10. <i>Resultados 2pl</i>	77
Tabla 11. <i>Frecuencia de respuestas según escala utilizada</i>	88
Tabla 12. <i>Resultados 2pl para la sección Decodificación</i>	90

RESUMEN

El propósito de este estudio es analizar la consistencia interna y validez de la Prueba diagnóstica para primer grado de primaria que mide lectura y escritura emergente, aplicada a 300 estudiantes en Guatemala, a partir de la aplicación de la Teoría de Respuesta al Ítem (TRI) desde modelo de Rasch y 2pl (discriminación y dificultad).

La metodología delimitó llevar a cabo un estudio piloto para la validación del instrumento de evaluación que es propiedad del Centro de Investigación Educativa (CIE). Se realizó un muestreo no probabilístico por conveniencia, conformado por 300 sujetos con 6 a 7 años, ubicados principalmente en la capital y en dos municipios de Guatemala con distribución en el área urbana y rural, además de pertenecer a centros educativos que según los resultados publicados por la Dgeduca con la prueba para primer grado del 2013 obtuvieron resultados por debajo del logro y en el logro. La observación inicial se realizó con el análisis factorial exploratorio como parte de la comprobación de los supuestos de la TRI, las propiedades psicométricas de la escala se analizaron según el modelo de Rasch, entre ellas el ajuste, con el modelo de dos parámetros (2pl) se calcularon las curvas características del ítem, la Función de información y el ajuste del test.

Los resultados indican que la prueba diagnóstica para primer grado de primaria tiene un mejor ajuste cuando se elimina la subprueba de decodificación de los datos al modelo de Rasch. Se observó consistencia interna de los ítems. Las curvas características de los ítems presentaron consistencia hacia la izquierda de la distribución.

Entre las conclusiones se observó que el instrumento de evaluación medido identifica el rasgo en función de la habilidad que requiere el ítem, que por ser una prueba diagnóstica se encuentra apropiado que la tendencia de la habilidad sea hacia la izquierda, asimismo que los ítems presentaron una consistencia adecuada y que es un instrumento que mide el constructo de lectura y escritura en las etapas emergente (67%) e inicial (33%).

I. INTRODUCCIÓN

Las funciones cognitivas que tienen relación con el lenguaje oral, la memoria, la atención, las nociones de espacio y tiempo, y el lenguaje escrito son herramientas fundamentales para conocer el mundo. Por ello, en los primeros años escolares se debe desarrollar y estimular estas funciones para que los aprendizajes sean significativos.

El proceso de aprendizaje en la primera infancia consiste en una transición de lo natural (el lenguaje oral-auditivo) a lo cultural (el lenguaje lecto-escrito) donde el lenguaje oral es el fundamento de la lectura (Fundar, 2011). Por tanto, las habilidades de comunicación y de aprendizaje que son la base para desarrollar competencias se forman por medio del dominio del lenguaje oral. Bravo (2003) asegura que el éxito en el aprendizaje de la lectura depende del desarrollo cognitivo y psicolingüístico que se adquiere en el nivel de educación preprimaria. En cuanto a las aptitudes de decodificación escrita es importante aproximar estas aptitudes lo más posible a la comunicación oral y auditiva (Hirsh, 1997).

Pero ¿cómo se corrobora que los estudiantes han adquirido o no las aptitudes de lectura y escritura del nivel de educación preprimaria? Una evaluación diagnóstica estandarizada, es una opción para ello. Con la evaluación diagnóstica, el tomador de decisiones (puede ser el docente, el administrador educativo, el funcionario público) tiene en sus manos la información pertinente para abordar las habilidades o conocimientos que no se alcanzado en beneficio del logro de competencias. Este proceso es el punto de partida para tomar decisiones pertinentes. Entonces, ¿qué tipo de evaluación es el instrumento adecuado? Aquel que permita saber si el estudiante está bien o no en relación con un dominio definido. Permitiendo diseñar o seleccionar estrategias de intervención que favorezcan el aprendizaje, en este caso, de la lectura y la escritura.

En Guatemala, la Universidad del Valle de Guatemala por medio del Centro de Investigaciones Educativas (CIE) cuenta con instrumentos de evaluación para estudiantes y para personal de instituciones que lo requieran. Entre las áreas que evalúa están: aptitud académica, orientación vocacional, lectura, habilidad general, intereses vocacionales, personalidad y conocimiento (UVG, 2015). Por tal razón el CIE posee un instrumento de lectura y escritura para ser administrado a los estudiantes inscritos en primer grado durante los primeros dos meses del ciclo escolar (CETT-CA-RD, S.f.), a razón de lo anterior y por ser de carácter diagnóstico, según lo indica el programa Centros de Excelencia para la Capacitación de Maestros (CETT), se determina que el constructo medido con el instrumento es de lectoescritura emergente. La prueba para primer grado surge del Programa Aulas Creativas que se inicia con el Programa Centros de Excelencia para la Capacitación de Maestros en Centroamérica y República Dominicana (CETT-CA/RD), por sus siglas en inglés; dicha evaluación requiere de parámetros de medida que arrojen información sobre la consistencia interna y validez del instrumento con la finalidad de administrar de forma adecuada la prueba diagnóstica para primer grado cuando se requiera en el contexto guatemalteco.

La evaluación diagnóstica para primer grado de primaria es una prueba de dominio que permitirá conocer el nivel de adquisición de los aprendizajes de la lectura y escritura emergente, necesarios para la adquisición de lectura y escritura inicial que están por desarrollar en el Nivel Primario. Para asegurar que los datos obtenidos con el instrumento de evaluación son válidos y confiables, se ha procedido a analizar los reactivos del instrumento, así como a obtener la validez del instrumento de evaluación. Es decir, se aplicó la prueba diagnóstica para primer grado en el contexto guatemalteco, situación que se consideró inherente porque el instrumento fue creado en República Dominicana. País que generó un manual para guiar la aplicación y la interpretación del instrumento.

Razón por la cual, se consideró importante analizar la pertinencia del lenguaje utilizado, situación que dio lugar a modificación del manual de aplicación de la prueba. Asimismo, se llevó a cabo un juicio de expertos para obtener validez de contenido, dando sustento al constructo que mide la prueba, desde la perspectiva técnica del diseño y

objetivo de esta. Los datos cuantitativos y cualitativos obtenidos representan el insumo indispensable de la unidad de análisis para redactar las conclusiones y recomendaciones hacia la prueba diagnóstica para primer grado de primaria.

II. JUSTIFICACIÓN

La prueba diagnóstica se desarrolla dentro del marco del Programa Aulas Creativas que surge del Programa de Centros de Excelencia para la Capacitación de Maestros en Centroamérica y República Dominicana (CETT-CA/RD). El propósito del programa es lograr la excelencia en el aprendizaje de la lectura y escritura de los niños de primero a tercer grado de educación primaria, mediante un programa intensivo de capacitación a maestros. La iniciativa apoya a los países con mayores dificultades en la enseñanza y a los maestros que trabajan en áreas urbano-marginales y rurales. (UVG, 2015).

El programa Centros de Excelencia para la Capacitación de Maestros (CETT) ha promovido, entre la comunidad educativa, las aptitudes en la enseñanza de lectura y escritura necesarias para alcanzar el éxito en dichos aprendizajes, además ha permitido la adquisición de las competencias de las distintas dimensiones de la enseñanza de la lectura y escritura (Aguirre International, 2005) correspondientes a los grados iniciales con el desarrollo del programa. El CETT, creó y distribuyó al consorcio de instituciones colaboradoras, en el caso de Guatemala al Centro de Evaluación Educativa (CIE), de la Universidad del Valle de Guatemala (UVG), un conjunto de herramientas de diagnóstico y evaluación para primero, segundo y tercero primaria que tienen la finalidad de contribuir en la tomar de decisiones basadas el historial educativo de los estudiantes que se recogen como evidencia con dichas herramientas.

Por tanto, el CIE observó una oportunidad y una necesidad al contar con las herramientas de diagnóstico y evaluación que el CETT le proveyó. Es decir, los instrumentos de evaluación requieren ser validados en el contexto guatemalteco para asegurar que son pertinentes y que poseen características psicométricas adecuadas para utilizarse con los estudiantes del país. Es preciso aclarar que el CIE realizó una priorización entre los tres instrumentos de evaluación, reconociéndose que la demanda de evaluación que el Centro se atiende se encuentra en primer grado, por tanto, en el presente trabajo de tesis se realizó el proceso de validación de la prueba diagnóstica para estudiantes de primer grado de primaria.

Por lo anterior, es necesario analizar los reactivos y validar la prueba diagnóstica para primer grado elaborada por el CETT, que evalúa aspectos de la lectura y escritura emergente con la finalidad de calcular los estadísticos psicométricos de los ítems y de la prueba para que el instrumento sea utilizado con confianza.

III. ANTECEDENTES

En Guatemala existen instituciones gubernamentales y no gubernamentales comprometidas con la evaluación externa para asegurar el uso correcto de los datos y la mejora de los procesos basados en la evidencia. El Centro de Investigación Educativa (CIE), (2016) es una institución no gubernamental que entre sus servicios “incluyen estudios de evaluación del rendimiento escolar” (párr. 1), por tanto, es inherente recoger información por medio de instrumentos de evaluación estandarizados que en su construcción cumplen con alta calidad técnica. De acuerdo con Tiana, citado por Backhoff, Bouzas, González, Andrade, & Hernández (2008), la evaluación es el medio para mejorar procesos y tomar decisiones oportunas, Tiana explícitamente expone:

“La primera de las funciones es, precisamente, proporcionar información confiable e interpretaciones válidas que permitan conocer de manera precisa el estado cómo se encuentra un sistema educativo y sus componentes en un momento determinado. Es una función diagnóstica de la evaluación educativa, con la cual se puede alcanzar un doble objetivo: 1) servir de base para que las autoridades educativas tomen decisiones documentadas y 2) satisfacer la demanda social de información de la sociedad. Esta función es especialmente relevante en los momentos en que un país experimenta cambios o reformas educativas.” (Pág. 17).

La construcción de la evaluación, como lo indica Tiana, debe cumplir con un proceso riguroso que implica cumplir con una serie de pasos concatenados que tienen como fin último servir de referente para la toma de decisiones. Las evaluaciones estandarizadas son cuestionarios desarrollados para medir diversas variables que tienen sus propios procedimientos de aplicación, codificación e interpretación (Martínez, Hernández, & Hernández, 2006) para asegurar que la interpretación de los resultados dé información generalizable para la población meta, el administrador de la evaluación estandarizada realiza el análisis de los reactivos que conforman dicha prueba para asegurar que el instrumento cumple con los parámetros de confiabilidad y validez requeridos.

Por tanto, a nivel internacional, regional y nacional existe diversidad de instrumentos de evaluación válidos y confiables que miden lectura y escritura en los grados iniciales de educación primaria.

A nivel internacional, por ejemplo, se ha utilizado la prueba EGRA, por su sigla en inglés (*Early Grade Reading Assessment*) para la evaluación temprana de la lectura. El instrumento incluye los principales componentes que según el *National Reading Panel* (NRP) predicen las dificultades de aprendizaje en lectura. Los datos normativos de la adaptación al español de la prueba en mención, afirma que es un instrumento que puede ser utilizado para la detección temprana del riesgo de presentar dificultades de aprendizaje en la lectura, y también para la planificación de la enseñanza de la lectura en los primeros grados de escolaridad (Jiménez, Gove, & Rodríguez, 2014).

Cuetos, Rodríguez & Arribas (2007), señala que España cuenta con una serie de pruebas para evaluar la capacidad cognitiva tanto de la lectura como de escritura. El Test de Análisis de Lectura y Escritura (TALE) está destinado a estudiantes de 6 a 10 años de edad que cursan los primeros cuatro grados de educación primaria. TALE es un test que determina los niveles generales y las características específicas de lectura y escritura que los estudiantes han alcanzado y poseen. Además, posee la batería de evaluación de los procesos lectores (PROLEC-R) que valora la lectura en español. Se centra en los procesos que intervienen en la comprensión del material escrito: identificación de letras, reconcomiendo de palabras, procesos sintácticos y procesos semánticos. Destinada para estudiantes de 6 a 12 años.

Entre los instrumentos que miden escritura, se encuentra el Test de Rendimiento en la Escritura (Proesc). Puede ser aplicado individualmente o en grupo. La batería de pruebas contempla dominios de palabras con ortografía arbitraria, dictado de pseudopalabras (palabras desconocidas), dictado de frases y redacción.

La evaluación externa del proyecto (PEC), aplicada en México y Chile, es un instrumento utilizado a nivel regional para ser administrado en los grados de primero y

segundo primaria con la finalidad de obtener al inicio del trabajo académico evidencias válidas y confiables del nivel de conocimiento sobre habilidades de escritura. La prueba de escritura evalúa los niveles presilábico, silábico, silábico/alfabético, alfabético inicial y escritura del alfabético convencional (Ramos, 2007).

Un instrumento utilizado por su consistencia interna es el Test de Habilidades prelectoras (THP). La prueba está constituida por 117 reactivos, distribuidos en cuatro áreas: conciencia fonológica, identificación de letras, memoria verbal y procesos sintácticos y semánticos del lenguaje oral (Consoli, Canales, & Lingán, 2011). El TPH fue aplicado en Perú a estudiantes de entre 4 y 6 años.

Las etapas del aprendizaje de la lectura propuestas por Frith en 1989 son tres. González Álvarez (2003) y Jiménez y Ortiz (2007) definen la etapa Logográfica, en la que reconoce de forma global algunas palabras que le son familiares. La segunda etapa es la Alfabética, en donde se toma conciencia de que las palabras se componen de elementos o unidades fónicas que tienen representaciones gráficas. La etapa Ortográfica, desarrolla la capacidad de reconocer un buen número de palabras sin necesidad de decodificar a partir de sus grafemas. Lo anterior, destaca que a nivel internacional son muchos los instrumentos que se encuentran disponibles para evaluar lectura y escritura en los grados iniciales de escolaridad.

En el ámbito guatemalteco se cuenta, entre otros, con instrumentos de evaluación de lectura que miden específicamente comprensión lectora para grados iniciales como habilidades que intervienen en el proceso de aprender a leer, es decir, el constructo de lectura inicial. Específicamente, las pruebas de rendimiento académico aplicadas por el Ministerio de Educación desde el 2006 hasta el 2010, para el primer grado de primaria sobre la comprensión lectora, evidencia de manera reiterativa tendencias de rezagos respecto a otros países de América Latina. Es decir, en el 2010 el 47.5% de los estudiantes de primer grado de primaria alcanzaron el nivel de logro a nivel nacional. Con aras comparativas, al respecto, se resalta que la Dgeduca reporta para el 2008 el 55.87% de

logro y la medida más baja se registra en el 2006 con el 39.84% de logro en comprensión lectora a nivel nacional.

Respecto al instrumento de evaluación que mide lectura inicial, en 2015, el Centro de investigaciones Educativas (CIE) realizó una investigación con la Evaluación de Lectura para Grados Iniciales (ELGI) la cual está adaptada al contexto nacional y ha sido administrada por la Digeduca con la finalidad de establecer la norma o los baremos para primer grado en Guatemala. En 2014, la Digeduca realiza la evaluación de Primaria Nacional que tiene representatividad muestral a nivel municipal y departamental, con estudiantes de primer grado de primaria, aplicando ELGI. Producto de dichos esfuerzos institucionales, la Digeduca publica que el 13 % de estudiantes alcanza al finalizar el primer grado las habilidades de lectura requeridas, según el CNB (Del Valle & Cotto, 2015); la publicación se sustenta en el *Modelo para explicar y predecir el aprendizaje de lectura para Guatemala*.

La prueba LEE (*Emergent Literacy profile*) mide lecto-escritura emergente. Se administra a estudiantes de primer grado del nivel primario. Consta de 4 subpruebas; la primer subprueba se refiere a la identificación de letras. La segunda, es de escritura emergente. La tercera, mide la lectura inicial y la cuarta, contempla conceptos impresos y comprensión oral. En Guatemala, en el 2011 fue adaptada al español por USAID para evaluar el alfabetismo emergente (Cotto E. , 2012).

Otro instrumento disponible en Guatemala para medir habilidades, en este caso de escritura, es la evaluación de Escritura para el nivel primario que administra el Ministerio de Educación. El instrumento es aplicado a estudiantes de primero, tercero y sexto grado de primaria. El Mineduc por medio de la Digeduca expone que el objetivo es analizar las destrezas de escritura adquiridas por los estudiantes de primaria para reconocer cómo ha sido el proceso de enseñanza por parte del docente. Con base en el objetivo mencionado, se observa que la pretensión es realizar recomendaciones sobre cómo mejorar los procesos de escritura en el aula (Mineduc, 2014).

Al respecto de las evaluaciones, en Guatemala existen instituciones gubernamentales como no gubernamentales que se interesan por contar con instrumentos de evaluación válidos y confiables para la medición del constructo del dominio lector y escrito, pero que han sido mayores los esfuerzos por medir la lectura y escritura en el proceso inicial (primero, segundo y tercero primaria), en comparación con la medición en grados del Nivel de Educación Preprimaria, en donde el constructo por medir es la lectura y escritura emergente.

También es importante es reconocer que, en el país, los procesos de evaluación están normados en la legislación existente. En las líneas siguientes se observa cuál es la normativa que avala el sistema de evaluación en Guatemala:

Reglamento de evaluación de los aprendizajes. Acuerdo Ministerial 1171-2010.

- a. Artículo 46, capítulo XV. Define la evaluación externa: *«es la evaluación realizada 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 y 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, capítulo XV. Define el propósito como: *«la evaluación externa se realiza con el propósito de obtener información directa que... d. Brinde elementos que conforman un marco de referencia para la mejora continua del Sistema Educativo Nacional.»*

Hace más de una década se inició en Guatemala el proceso de transición educativa con la Reforma Educativa (Mineduc, 2012), con esto se adquiere el término de calidad, el cual Unesco indica que la determinan tres dimensiones fundamentales: el respeto de los derechos, equidad y pertinencia. Pudiendo añadir la relevancia, y dos dimensiones más de carácter operativo: eficacia y eficiencia.

La calidad educativa se manifiesta en dos niveles, es decir a nivel del estudiante, la educación necesita buscar y reconocer el conocimiento previo de los estudiantes, reconocer los modos formales e informales, practicar la no discriminación y proveer un entorno de aprendizaje seguro y apoyado. A nivel del sistema de aprendizaje, se necesita una estructura de apoyo para implementar políticas, establecer normas, distribuir recursos y medir los resultados de aprendizaje de modo que se logre el mejor impacto posible sobre un aprendizaje para todos (Unesco, 2007).

En la línea de trabajo de la calidad educativa, en el 2005 se crea el acuerdo ministerial No. 35 que en su artículo sobre Autorización, se lleva a cabo una acción que sustenta el modelo de Reforma Educativa. El cual dice literalmente:

Artículo 1. Autorización: “Se autoriza el Currículum Nacional Base, para el nivel de educación primaria, diseñado en el marco del proceso de transformación curricular, el cual se orienta con una nueva visión de país acorde al diseño de la reforma educativa y las aspiraciones contenidas en los Acuerdos de Paz, con características de flexibilidad, integralidad, perfectibilidad y participación”.

A. Políticas educativas

Con fundamento en lo anterior, el Ministerio de Educación junto a catorce instituciones más: Universidad de San Carlos de Guatemala -USAC-, Universidades Privadas, Academia de las Lenguas Mayas de Guatemala -ALMG-, Comisión Nacional Permanente de Reforma Educativa -CNPRE-, Consejo Nacional de Educación Maya -CNEM-, Consejo Nacional para la Atención de las Personas con Discapacidad -CONADI, Alianza Evangélica de Guatemala -AEG-, Conferencia Episcopal de Guatemala -CEG-, Asamblea Nacional del Magisterio -ANM-, Comité Coordinador de Asociaciones Agrícolas, Comerciales, Industriales y Financieras -CACIF-, Sector de Mujeres Organizadas, Organizaciones del Pueblo Xinka, Organizaciones del Pueblo Garífuna y Asociación de Colegios Privados, integran el consejo nacional de educación para que entre otras políticas educativas se elabore la política de Calidad, estando consientes que la

calidad no es una política aislada, sino que, cuando se habla de calidad se refiere a todo el Sistema Educativo. Siendo entonces que la calidad abarca desde el desarrollo de las competencias básicas de la lectoescritura y Matemática hasta la infraestructura.

El Consejo Nacional de Educación, publica la Política 2 sobre Calidad, que literalmente manifiesta perseguir el «mejoramiento de la calidad del proceso educativo para asegurar que todas las personas sean sujetos de una educación pertinente y relevante» (p. 6). Además, tiene como tercer objetivo estratégico el fortalecimiento del sistema de evaluación para garantizar la calidad educativa (Consejo Nacional de Educación, 2010).

B. Currículum Nacional Base

Como respuesta a los avances definitorios en el tema de calidad y con base en el primer objetivo estratégico que el Consejo Nacional de Educación (2010) por medio del documento Políticas educativas expone que «Contar con diseños e instrumentos curriculares que respondan a las características y necesidades de la población y a los avances de la ciencia y la tecnología» (p. 6) se vuelve una prioridad en beneficio de la educación. Con el objetivo de relacionar los esfuerzos de tantas entidades guatemaltecas comprometidas con el tema expuesto, es posible resaltar que el Ministerio de Educación para el Nivel de Educación Primaria inició la implementación del Currículo Nacional Base -CNB- en el año 2005. Su centro es la persona como ente promotor del desarrollo individual, social, de las características culturales y de los procesos participativos que favorecen la convivencia armónica. Hace énfasis en la valoración de la identidad cultural, en la interculturalidad y en las estructuras organizativas para el intercambio social en los centros y ámbitos educativos, de manera que las interacciones entre los sujetos no solamente constituyen un ejercicio de democracia participativa, sino fortalecen la interculturalidad (Digezur, 2008).

El Currículum Nacional Base está organizado en competencias, ejes y áreas para el desarrollo de los aprendizajes, en donde cada área tiene sus competencias, indicadores de logro y contenidos organizados en declarativos, procedimentales y actitudinales,

descansando todo ello en los estándares educativos. Convirtiéndose por sus características en la herramienta pedagógica que orienta el trabajo del docente, trabajo, que la política educativa de Calidad sugiere monitorear y evaluar.

El currículum, entonces, refleja linealidad con lo establecido en la política de Calidad, reconociéndose tal linealidad, pertinente y relevante porque responde a las expectativas de los involucrados en el ámbito educativo. Por otro lado, pero con relación al tema abordado, la Ley de Educación Nacional 12-91, manifiesta que la educación formal está conformada por niveles educativos, los cuales han sido el ente rector de la clasificación de los niveles como se presentan en el Currículo Nacional Base que rige el quehacer del proceso de enseñanza-aprendizaje-evaluación en Guatemala.

Con dicho sustento, en el presente trabajo de graduación en modalidad de tesis, se observa adecuado relacionar la malla curricular que el Mineduc de Guatemala expone en los documentos curriculares que avalan el Nivel de Educación Preprimaria porque el constructo, de la Prueba diagnóstica para primer grado, objeto de estudio de la presente tesis, es lectura emergente y tiene como propósito identificar los aprendizajes que los estudiantes poseen al inicio del primer grado de primaria.

1. Currículum Nacional Base del Nivel de Educación Preprimaria

Además de la legislación y las políticas educativas, el Sistema Educativo, hasta la fecha, posee para la ejecución de la educación en la población entre 4 y 6 años el Currículum Nacional Base (CNB) del Nivel de Educación Preprimaria, el cual está organizado en áreas que responden a la estructura del conocimiento en las diferentes etapas del desarrollo humano. Como se observa en el CNB de dicho nivel educativo, en la preprimaria se establecen las bases y los fundamentos esenciales para el posterior desarrollo del comportamiento humano, así como la existencia de grandes reservas y posibilidades que en ella existen para la formación de diversas capacidades, cualidades personales y el establecimiento inicial de rasgos del carácter. También se forma la personalidad tomando como base la plasticidad que tiene el cerebro infantil. De esa

manera, la socialización y la estimulación desarrolladas en forma simultánea permiten preparar los niños para la vida y para la adquisición de aprendizajes permanentes (Digezur, 2008).

C. Entidades en Guatemala que hacen evaluación educativa externa

En el país se encuentra diversidad de instituciones que tienen la capacidad instalada para realizar evaluación educativa externa. Entre ellos, se resaltan dos instituciones por su trayectoria histórica en la evaluación educativa del país.

1. Dirección General de Evaluación e Investigación Educativa (Digeduca)

En Guatemala, desde 1992 se han realizado esfuerzos para evaluar el rendimiento escolar, alcanzando actualmente, avances significativos en el ámbito de la evaluación educativa, porque de acuerdo con Fortín, A. (2013), el país ha logrado crear un sistema de evaluación sostenible, basado en principios técnicos y útil para desarrollar políticas educativas. Desde el año 2008 ha sido el Ministerio de Educación por medio de la Digeduca que ha estado a cargo del proceso de evaluación del Sistema Educativo Nacional por medio de la Dirección General de Evaluación e Investigación Educativa (Digeduca).

La Digeduca es la dirección encargada de evaluar y divulgar el desempeño tanto del estudiante como del docente para proveer información de calidad de los aprendizajes que está basada en criterios sistemáticos que aseguran la objetividad de la información para la toma de decisiones y planificar acciones correspondientes (Ministerio de Educación, 2010).

2. El Centro de Investigaciones Educativas (CIE)

El CIE pertenece a la Universidad del Valle de Guatemala. La brújula de trabajo del Centro de Investigaciones Educativas es contribuir al desarrollo educativo del país por medio de la investigación, realización de proyectos, prestación de servicios de evaluación,

asesoría y capacitación en material psicométrico. Porque su visión es ser reconocidos como un centro de investigación educativa de excelencia, por su rigor científico, calidad técnica y ética en el manejo de la información al proveer resultados válidos y confiables, que contribuyen al mejoramiento de la calidad educativa nacional (UVG, 2015).

Aunado a lo anterior, se resalta que el CIE posee el objetivo de «realizar estudios de evaluación del rendimiento escolar, capacitaciones a docentes y directores, capacitación en el uso y manejo de pruebas psicométricas a profesionales del área de recursos humanos». Asimismo, se encarga de la “elaboración de materiales didácticos de apoyo. Además de brindar el servicio de la aplicación de pruebas, con estándares internacionales fundamentados en los conceptos pedagógicos y psicométricos más actuales, construidas por equipos multidisciplinarios internacionales” (UVG, 2015).

Por ello, cuenta con instrumentos de evaluación para estudiantes y personal de instituciones que lo requieran. Entre las áreas que evalúa están: aptitud académica, orientación vocacional, lectura, habilidad general, intereses vocacionales, personalidad y conocimiento (UVG, 2015). Siendo la razón por la que el CIE cuenta con el instrumento que evalúa lectoescritura emergente, como diagnóstico de los aprendizajes que los estudiantes de primer grado de primaria poseen al momento de comenzar los aprendizajes en dicho grado. El instrumento de evaluación mencionado responde a la participación que tuvo Guatemala en el Programa Aulas Creativas que se inició con el Programa Centros de Excelencia para la Capacitación de Maestros (CETT-CA/RD).

Por ser un instrumento de evaluación construido para el programa mencionado, el CIE observa que requiere conocer los parámetros de validez y confiabilidad del instrumento para tomar la decisión de administrar o no el instrumento cuando se requiera con la población objetivo en Guatemala.

IV. MARCO CONCEPTUAL

En español se usa el sistema de escritura alfabético, es decir, se basa en un alfabeto. Este sistema de escritura tiene ortografía altamente transparente porque cada sonido se corresponde con una letra del alfabeto, aunque cabe mencionar que las letras c, s, j, g, x son excepciones de dicha regla (Montenegro, 2020).

La lectura y la escritura integran procesos inversos de utilización de un mismo código. Es decir, al leer, se decodifica un mensaje que procede de la lengua hablada y al escribir se codifica un mensaje desde esa misma lengua (Junta de Andalucía, 2015). Es innegable la relación funcional entre ambas habilidades y que, por lo general, en los grados iniciales de escolarización se plantea de forma conjunta la lectura y la escritura.

Las etapas emergente e inicial de la lectura y escritura promueven el desarrollo de componentes comunes para aprender a leer y escribir. La conciencia fonológica y el desarrollo del principio alfabético son base para dichos aprendizajes. Estos componentes, menciona Montenegro, R., (2020), son el fundamento para desarrollar la decodificación y la codificación. Ambos aprendizajes están cimentados en el lenguaje oral y en el conocimiento del vocabulario propio del idioma del niño.

La adquisición de la lectura y la escritura, como competencia para la vida desarrolla, según Atorresi, (s.f.), 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.

Leer es necesario para el aprendizaje, pues la mayoría de las actividades escolares requieren de la lectura. El aprendizaje de la habilidad lectora es una de las metas principales en la educación y, a la vez, es una herramienta usada para aprender (Cotto & Arriaga, 2014). Por lo tanto, la lectura es la puerta de acceso a la cultura escrita y a la socialización, conocimientos, información, entre otros (Atorresi, Unesco_ LLECE, s.f.).

En Guatemala, enseñar a leer y a escribir es la tarea principal de los grados iniciales del Nivel de Educación Primaria, en donde cabe aclarar, que el Nivel de Educación Preprimaria prepara al estudiante para que la adquisición de los aprendizajes (Digezur, 2008) se dé sin contratiempos. Sustento de ello, se encuentra en la caracterización del nivel que el Currículo Nacional Base expone.

Así que, la lectura y la escritura constituyen dos de los requisitos indispensables para la comunicación de la persona. Es por medio de estas que el ser humano despliega un mundo de imaginación, comprensión y criterio por lo que las letras le transmiten. (Atorresi, Unesco_ LLECE, s.f.).

A. La lectura y escritura en los primeros grados de escolaridad

El aprendizaje de la lectura y escritura se inicia con el lenguaje oral y se va incrementando de acuerdo con la exposición a diferentes experiencias de aprendizaje dentro del contexto letrado en que se desenvuelve el ser humano durante sus años escolares. Específicamente la adquisición de la habilidad tanto de leer como de escribir se adquiere por medio del dominio de los sonidos y letras del idioma propio del niño, y del vocabulario. Como se ha mencionado anteriormente, dichos aprendizajes son indispensables para la decodificación que con su automatización da paso al logro de leer con fluidez, para posteriormente desarrollar intencionalmente la comprensión lectora. (Camargo, Montenegro, Linan-Thompson, & Loeza, 2013). Por tanto, es indisoluble la unión entre la lectura y la escritura cuando el proceso de aprendizaje comienza. Además, es indispensable mencionar que el aprendizaje de la lectura y escritura es gradual.

El aprendizaje formal para leer y escribir se da en tres etapas. La emergente, inicial y finalmente, la establecida. De acuerdo con Camargo, Montenegro y otros (2013) se definen:

- Emergente o prealfabética

Inicia en el nacimiento y culmina en el Nivel de Educación Preprimaria. Se le conoce como la etapa preparatoria, en la cual se busca fortalecer en el niño el desarrollo del

lenguaje oral y ampliación del vocabulario. Asimismo, la familiarización con materiales impresos, el conocimiento de las convenciones de la escritura, tales como la direccionalidad de izquierda a derecha y de arriba abajo son aprendizajes preparatorios. Además, que las palabras sirven para nombrar objetos y que los sonidos se relacionan con las letras impresas.

También que las historias tienen un inicio y un final, que se usan estrategias para entender, recordar y comunicar lo que se lee o lo que se escucha y desarrollar las destrezas visuales, auditivas y motrices necesarias para la lectura y escritura.

Específicamente, la escritura en la etapa emergente se caracteriza con que el estudiante usa el dibujo y el trazo de algunas letras para expresarse por escrito. Pero esta tiene componentes como el vocabulario, conciencia del lenguaje escrito, expresión escrita y la grafomotricidad que deben enseñarse formalmente en el Nivel de Educación Preprimario.

- Inicial

Establece que los estudiantes deben aprender a leer en primer grado de primaria y afianzar el aprendizaje durante los dos siguientes años de educación escolarizada. Durante la etapa inicial del aprendizaje de la lectura se continúa con el fortalecimiento de la conciencia fonológica, del principio alfabético y del vocabulario para asegurar el aprendizaje de la decodificación que posteriormente dará lugar a la fluidez. Ahora bien, para la escritura, el aprendizaje se centra en consolidar el proceso de codificación o transcripción, así como de la producción de palabras, oraciones, párrafos y textos, esperándose que los estudiantes progresen en el aprendizaje sintáctico y léxico.

En cuanto a la escritura en la etapa inicial requiere atención en los componentes de codificación, legibilidad de escritura a mano, fluidez de escritura, escritura de oraciones (cada vez más complejas), producción de diversos tipos de texto con coherencia y cohesión y de la ortografía en la producción de textos (Montenegro, 2020). Es necesario

aclarar que en el presente informe no se abordará la etapa inicial porque sobrepasa los objetivos de la investigación.

- Establecida

Etapa de la adquisición de lectura y escritura que se da desde cuarto grado de primaria en lo sucesivo hasta llegar a ser lectores y escritores independientes o autónomos. Es decir, el aprendizaje ahora se centra en la adquisición de habilidades y destrezas que lleven a la comprensión global del texto.

En la etapa inicial se enlistaron, en los tres últimos componentes, lo referido a la producción o redacción de textos con coherencia y cohesión, además del uso de oraciones complejas y con ortografía correcta, respondiendo a que estos aprendizajes deben recibir énfasis justo en la etapa establecida de la primaria (Montenegro, 2020), de ahí el nombre de la etapa, porque son justo dichos componentes los que deben afianzarse en el segundo ciclo del nivel primario, pero sin dejar de mencionarse que este aprendizaje está influido por el papel del docente.

1. Elementos para el desarrollo de la lectura y escritura emergente

Como se ha afirmado anteriormente, en el Nivel de Educación Preprimaria, para desarrollar la lectura y escritura emergentes, es importante realizar un plan de estimulación del lenguaje. Camargo, Montenegro y otros (2013) observaron varios componentes para desarrollar la lectura y escritura en la etapa emergente, resaltándose que, en esta etapa, componentes como el desarrollo de la conciencia del lenguaje escrito y el desarrollo del lenguaje se enriquecen con el aprendizaje del vocabulario, y que aunado a ello, el desarrollo de la conciencia fonológica aportan tanto a la lectura como a la escritura (Jiménez, Gove, & Rodríguez, 2014).

a. Desarrollo del lenguaje oral

El lenguaje oral implica saber escuchar y hablar. El proceso de escuchar es aprendido, para desarrollarlo es importante relacionarlo con todas las actividades que requieran expresión oral y percepción auditiva. Es decir, motivar al niño para que participe en conversaciones formales para que practique la organización de sus ideas y la selección de las palabras.

Las destrezas transferibles e indispensables para el aprendizaje de la lectura Olilla (1981) son cuatro: a) entender y recordar hechos b) localizar sucesos en una secuencia c) seguir instrucciones y d) interpretar y evaluar ideas en las historias. Además, estas destrezas se relacionan con el aprendizaje de la comprensión oral porque están relacionadas con la escucha.

En el caso específico del aporte para el aprendizaje de la escritura, el lenguaje oral desarrolla aspectos como la conciencia de usar el lenguaje oral “con diversos propósitos, es decir comunicarse de manera oral o por escrito” (Montenegro, 2020, p. 72). Así como con la sintaxis.

b. Crecimiento sintáctico

En los primeros años de escolaridad, es decir, en los años donde se desarrolla la etapa emergente de la escritura, la sintaxis definida como el estudio de la forma de combinación de las palabras en una oración (Real Academia Española, 2005) se da desde la oralidad. Por tanto, escribir con enunciados completos se aprende, en concordancia con Montenegro, R., (2020) al participar y practicar la expresión de enunciados completos en discusiones, conversaciones, exposiciones, entre otros. Es importante indicar que el aprendizaje sintáctico en la preprimaria se da a nivel de oración, principalmente.

c. Desarrollo de la expresión oral

Hablar y escuchar son dos capacidades indispensables de atender cuando se trata el tema de la comunicación oral, porque es un proceso que va madurando conforme el estudiante va desarrollándose en un proceso formal de instrucción. Montenegro, (2022) asegura que la comunicación se presenta en distintos contextos y situaciones, porque el estudiante no llega a la escuela sin haber tenido experiencias con la familia y en la comunidad en la expresión oral, es decir en poner en práctica el componente de comunicarse de manera oral.

En el aprendizaje de la lectura y escritura emergente, la expresión oral aporta conocimiento al propiciarse actividades de exposición, narración, descripción, conversación, argumentación, entre otros (Moreno, 2018) citado en (Montenegro, 2020) principalmente cuando el docente dibuja lo que se conversa, escribe en un área visible al estudiante, algunas ideas relevantes de lo que se trata en clase.

d. Desarrollo de la comprensión oral

Comprender es un proceso que permite encontrar significado a lo que se oye o se lee (Camargo, Montenegro, y otros, 2013). La comprensión oral conforma los cimientos de la comprensión de la lectura. En su conjunto, el lenguaje oral se ve formado por medio de la interacción diaria entre compañeros, docentes y familia al escuchar el lenguaje hablado, permitiendo desarrollar la comprensión oral.

El primer paso, para aprender a leer y escribir es tomar conciencia de que lo escrito es una manera de representar lo que se habla, de ahí que el componente en la lectura emergente que desarrolla esta habilidad se llama conciencia del lenguaje escrito, que para los primeros grados de escolaridad se le conoce también como conceptos de impresión, porque incluye el conocimiento de los procedimientos para la lectura y escritura, además del uso de un libro o texto. Dicho conocimiento que incluye el uso de un libro se da con mayor facilidad cuando el estudiante está inmerso en un ambiente letrado, es decir en el

espacio de aprendizaje hay variedad de libros y material escrito que tiene como finalidad promover la interacción de los estudiantes con el lenguaje escrito y hablado.

Por lo expuesto, la comprensión del lenguaje oral (lo hablado) favorece el desarrollo de destrezas que serán transferibles a la comprensión lectora, que primero, requiere del desarrollo de estrategias de comprensión (Olilla, 1976) a nivel oral, es decir, leer en voz alta los textos a los estudiantes o contarles historias permite adquirir poco a poco las destrezas de: a) entender y recordar; b) localizar sucesos en una secuencia; c) seguir instrucciones, y d) interpretar y evaluar ideas en las historias que forman parte de las habilidades de la escucha. Siendo importante mencionar también, que la comprensión de un texto oral depende en gran medida, del conocimiento y entendimiento de las palabras que lo componen, es decir del manejo del vocabulario.

Al escuchar historias leídas por alguien más, los niños: a) relacionan las palabras escuchadas con las escritas y descubren que cada letra o grafema tiene un sonido y en conjunto un significado. Así adquieren nuevo vocabulario (Thompson, 2004) y estructuras oracionales, propias del lenguaje escrito. b) se familiarizan con el manejo de los libros: direccionalidad de la lectura, diferentes formas, letras y palabras, entre otras.

En este sentido, se observó en el CNB del Nivel de Educación Preprimaria que establece el estándar 5 del área de Comunicación y Lenguaje L1, que los niños escuchen y comprendan por lo menos cinco libros recreativos apropiados a su nivel de lectura por año (Digezur, 2008).

e. Desarrollo del vocabulario

El desarrollo del vocabulario es un aprendizaje transversal para la adquisición de habilidades necesarias que llevarán al logro de la lectura y escritura. Como parte de la etapa emergente se observa necesario porque se requiere un vocabulario expresivo para emitir las ideas., asegura que la enseñanza del vocabulario es un proceso arduo y que implica el conocimiento del léxico del idioma. En este sentido, Esquerra, A. (2003) citada

por Siquiera Loureiro (s.f.) asegura que la enseñanza del léxico consiste en instruir a los estudiantes en el conocimiento, aprendizaje y uso adecuados de las palabras de la lengua.

En la etapa emergente, donde el niño debe estar expuesto al aprendizaje del vocabulario, este se da en el Nivel de Educación Preprimaria por medio del trazo de una imagen que representa la nueva palabra por adquirir, esto con la finalidad que el niño se apropie del significado. En el estándar 10 del Currículo Nacional Base del Nivel de Educación Preprimaria se establece que utilizar el lenguaje oral permitirá al estudiante expresar su opinión de lo aprendido en clase y adquirir más vocabulario (Digezur, 2008).

f. Animación a la lectura

Otra de las claves para el aprendizaje de la lectura y escritura, es el entorno donde se desarrolle un niño en edad escolar. Este debe ser letrado, para que estimule el aprendizaje de la lectura. El rol del docente en la animación a la lectura es fundamental, porque debe elegir los libros literarios y texto funcionales para ofrecer una variedad de oportunidades de lectura a los niños.

La oportunidad de manipular, hojear, descubrir cómo se utilizan los libros propicia el ambiente para que el niño desee aprender a leer. Para motivar la lectura se pueden usar libros de gran tamaño (Camargo, Montenegro, y otros autores, 2013) porque permiten la manipulación y la visión clara de los conceptos impresos que presenta el libro.

Parte de la motivación a la lectura es jugar a leer (Condemarin, 2001), la importancia de esta actividad es que el infante logra acumular gran cantidad de palabras en su memoria que pueda articular y modular con la entonación y la sintaxis adecuada, que además le permite captar el significado del texto completo. Con esta actividad se favorece la iniciación de la lectura porque permite repetir varias veces los cuentos, las letras de las canciones y los poemas.

g. Desarrollo de la conciencia fonológica

La conciencia fonológica es la habilidad de reconocer y usar los sonidos en el lenguaje hablado (Villagran & et.al, 2010). El conocimiento fonológico es una habilidad necesaria para el aprendizaje de la lectura y escritura, además de ser consistente con la posibilidad que el proceso de aprender a leer pueda facilitar el desarrollo de tal conocimiento (Castles & Coltheart, 2004). Es decir, existirían varias vías para el desarrollo del conocimiento fonológico: una sería aprender a leer y escribir en un sistema de escritura alfabético, y la otra la enseñanza de habilidades fonológicas (Hulme, Snowling, & Carrol, 2005). Según esta idea se puede decir que existe relación bidireccional entre el conocimiento fonológico y el aprendizaje de la lectura (Wise & Wolfe, 2008).

La conciencia fonológica supone desarrollar la habilidad de discriminación, conteo, combinación, segmentación y manipulación. El Currículum Nacional Base del Nivel de Educación Preprimaria propone el uso de onomatopeyas en las actividades de enseñanza-aprendizaje.

Para el aprendizaje de la conciencia fonológica Borrero (2008) propone que se debe enseñar por medio de: a) motivar al estudiante con actividades que le lleven a poner atención a las unidades mínimas de sonido, es decir al fonema. b) Practicar la segmentación y combinación de sílabas, esto permitirá que se le facilite la segmentación de fonemas. c) como las actividades de conciencia fonémica se inician con actividades de combinación de fonemas, luego con la segmentación y, por último, con la manipulación del fonema. d) Es importante modelar varias veces los ejemplos de lo que se requiere que ejecute el niño. Finalmente, e) Practicar la segmentación de oraciones de palabras simples, luego dividir las compuestas en las dos sílabas que las forman.

Importante recordar en este punto, que el lenguaje oral y el componente que se enlista en el numeral 6, son aprendizajes que aportan tanto a la escritura como a la lectura (Montenegro, 2020, p. 46).

h. Desarrollo de la conciencia del lenguaje escrito

Otra forma de denominarle al desarrollo de la conciencia del lenguaje escrito es el aprendizaje de los conceptos impresos. Se reconoce que este elemento de la lectura y escritura emergente es la capacidad de comprender el lenguaje escrito y que este aprendizaje tiene relación con el lenguaje oral y con expresar mensajes.

En esta etapa del aprendizaje, el niño toma conciencia que la lectura se realiza de izquierda a derecha, de arriba abajo, que las letras se pueden leer (aprendizajes de los conceptos impresos) y que la estructura del lenguaje escrito es distinta al hablado (Camargo, Montenegro, y otros autores, 2013) porque hay normas que hacen comprensible la expresión escrita.

En la medida que el niño tiene la oportunidad de acercarse a materiales impresos, puede practicar junto a su docente las convenciones de la lectura y de la escritura, es decir con esta última, reconoce que hay letras mayúsculas, minúsculas, signos de puntuación, diferentes tipos de letras como cursiva y molde.

En este sentido, en el CNB del Nivel de Educación Preprimaria se observa que está establecido en el estándar 4 que el estudiante al concluir ese nivel debe ser capaz de comprender el concepto de texto impreso (seguimiento de izquierda a derecha y otros), así como relacionar imágenes, dibujos y signos contenidos en los textos (Digezur, 2008).

i. Desarrollo de la grafomotricidad

La grafomotricidad está referida al movimiento realizado con la mano al momento de escribir (psicopedagogia.com, 2016). En el desarrollo de la grafomotricidad se pone en práctica la habilidad motora relacionada con la escritura.

Piaget (1969), menciona que el dinamismo motor y los concretos aprendizajes motrices están estrechamente ligados a la actividad mental infantil. Por tanto, Swearingen

y Calder (2009) exponen que la habilidad motriz se adquiere por medio de destrezas que permiten el desarrollo psicomotor, estimulando las funciones motrices, la tonicidad muscular, las funciones del equilibrio, el control y la disociación del movimiento, la rapidez y precisión del mismo, así como el control de las relaciones espaciales y el dominio de las relaciones temporales.

La grafomotricidad incluye la coordinación visomotora, ubicación espacial, direccionalidad, segmentación, legibilidad y rapidez de la escritura. El sistema de escritura se afianza al momento de alcanzar el nivel alfabético. Osuna, García y Torres (2004), indican que hasta en este nivel el niño puede diferenciar entre la lengua oral y la lengua escrita, y es cuando debe consolidarse el concepto de palabra.

El aprendizaje de la escritura se da en distintos niveles. Ferreiro y Teberosky (1995) conceptualizan los niveles de escritura en cuatro:

- Los primeros intentos de escritura se realizan por medio del dibujo infantil. El niño intenta comunicar una idea.
- Posteriormente diferenciará entre dibujos y letras, sin asociar los fonemas con los grafemas plasmados en el papel. Es decir, la diferenciación se da cuando trata de representar los fonemas con un símbolo gráfico.
- En el siguiente nivel, relaciona el sonido con una sílaba para representarlo con un símbolo gráfico. El niño comienza a utilizar tantas letras como sílabas tengan las palabras. Está en el proceso de la escritura silábica.
- Al posicionarse en el conocimiento silábico pasará al silábico-alfabético en donde une vocales y consonantes e incluye algunas sílabas completas. Finalmente, conseguirá la correspondencia letra-sonido o escritura alfabética. A partir de este momento tendrá preocupación por la corrección ortográfica.

j. Expresión escrita

En el Nivel de Educación Preprimaria se puede observar que los estudiantes se expresan por escrito al hacer uso de garabatos, dibujos y algunos otros por medio de letras

y palabras que ya han logrado identificar durante el proceso de aprendizaje. (Montenegro, 2020).

Vigotsky (1988), señala que «el niño garabatea en el papel rayas y puntos indiferenciados, (aparentemente) carentes de sentido, pero cuando reproduce las frases se tiene la impresión de que las está leyendo al tiempo que señala determinados trazos y sin ningún error, varias veces seguidas, muestra que rayas y puntos determinan dichas frases» (pág. 65) Por su parte, Padilla (2007), realiza una comparación entre la línea en el dibujo y en la escritura, indicando que la línea como primer elemento fundamental del dibujo evidencia que tiene dos valores descriptivos, es decir tiene un valor de forma o un signo del que emana un significado y el otro valor es propio del significado plástico expresivo de la propia manera de concretarse gráficamente. La palabra signos para Padilla hace referencia a signos pictográficos, ideogramáticos, silábicos o fonéticos, en todos los casos se observa que su valor atiende tanto a la calidad gráfica como al contenido del texto.

En este punto se entiende el carácter del dibujo como lenguaje. Al igual que la escritura se usa para comunicar y expresar pensamientos y sentimientos, y aunque son diferentes medios, tienen concurrencia de sentido, es decir, el lenguaje de las palabras y el lenguaje de las formas coinciden en que ambos recurren a equivalencias para traducir la realidad profunda del ser. Por tanto, en la etapa emergente de la escritura, se contempla una evolución desde la expresión escrita por medio de imágenes hasta la representación gráfica de los grafemas.

Finalmente, el aprendizaje de la lectura en la etapa emergente es importante, ya que forma las bases para el aprendizaje futuro del código escrito (Camargo, Montenegro, y otros autores, 2013). Por su parte Cotto y Arriaga (2014), indican que el proceso de lectura emergente culmina al terminar la educación preprimaria y permite que los niños estén listos para aprender a leer. Quiñonez (2014), en la serie de cuadernillos pedagógicos, Aprender del error, menciona que durante los años 2011 y 2012 la mayoría de los estudiantes evaluados que ingresaron a primer grado de primaria, reconocen que la direccionalidad de la escritura es de izquierda a derecha y de arriba hacia abajo; además

que, al momento de realizar un dictado, pudieron escribir una o dos letras de las palabras dictadas.

Quiñonez (2014) asegura que la enseñanza de la escritura debe enfocarse al desarrollo de competencias para comunicar, expresar y crear ideas. A razón que la escritura “permite construir y reconstruir el conocimiento, la escritura transforma al sujeto y las sociedades” (Atorresi, 2010, p. 16).

Después de la práctica oral del crecimiento sintáctico, donde se espera su fortalecimiento hasta primer grado de primaria, para luego afianzar el aprendizaje sintáctico en los procesos de escritura de oraciones, párrafos, entre otros.

En definitiva, la etapa emergente de la lectura y la escritura es un aprendizaje continuo, interrelacionado y sistemático que requiere de componentes para la lectura como el vocabulario, desarrollo del lenguaje oral, la expresión y comprensión oral, la conciencia fonológica, la conciencia del lenguaje escrito, animación a la lectura y de los conceptos impresos, mientras que la lectura requiere también del vocabulario, del desarrollo del lenguaje y expresión oral, la conciencia fonológica y de aprendizajes específicos para la escritura, como la grafomotricidad, la expresión escrita y el crecimiento sintáctico, este último relacionado con el desarrollo del componente del lenguaje oral.

Por ser el Nivel de Educación Preprimaria que abarca la etapa fundamental donde se conforman las funciones cerebrales que se relacionan con el desarrollo del lenguaje (Gutiérrez Duarte & Ruiz León, 2018) la intervención oportuna del docente y de los padres de familia (Montenegro, 2020) para apoyar los aprendizajes tanto de la lectura como de la escritura en la etapa emergente es indispensable.

2. La lectura y escritura en el Currículum Nacional Base (CNB)

Para Guatemala el Currículum Nacional Base es el producto de la necesidad del cambio en el sistema educativo que se vio evidenciado en el diseño de la Reforma

Educativa, en 1998. (USAID/G, 1998) La Reforma Educativa busca que los procesos sean llevados con calidad hasta el aula.

La calidad en aula, esta explicada en el CNB por medio del Modelo de calidad en el aula que refleja un ciclo que inicia con los aprendizajes esperados que tienen como referente el currículo el cual está regido por la modalidad de entrega que es monitoreada con la evaluación que se da tanto en el aula como por entes externos, información que permite la revisión a los aprendizajes, completándose en este punto el proceso e iniciando nuevamente el ciclo de la calidad en el aula.

Como dentro de este modelo se observa la importancia que tiene el currículo, es preciso exponer que este está organizado por estándares, competencias, indicadores de logro, contenidos y actividades de enseñanza-aprendizaje-evaluación sugeridos al docente del nivel para que refuerce y desarrolle, en este caso, los aprendizajes, que para el caso específico del tema que se trata en el presente informe de tesis, se desarrollen las destrezas relacionadas con la lectura y la escritura. Los aprendizajes esperados o estándares son enunciados que establecen criterios claros, sencillos y medibles, que los docentes deben considerar como meta del aprendizaje de sus estudiantes (Digezur, 2008).

a. Estándares Educativos del Nivel de Educación Preprimaria relacionados al aprendizaje de la lectura y escritura emergente

Los estándares que desarrollan las habilidades para la adquisición de la lectura emergente son los que se enlistan en la siguiente tabla:

Tabla 1. Estándares educativos del área de Comunicación y Lenguaje
Estándar Descripción del estándar

1	Escucha y actúa de acuerdo con los mensajes orales de su contexto familiar y escolar.
4	Relaciona imágenes, dibujos y signos contenidos en los textos de lectura infantil haciendo predicciones, identificando el tema, el personaje principal y comprendiendo el concepto de texto impreso (seguimiento de izquierda a derecha y otros).
5	Escucha comprensivamente cinco libros recreativos apropiados a su nivel de lectura por año.
7	Comprende que un sonido o fonema está representado por una letra o grafía, que, a su vez, se combina con otros para formar palabras.
8	Utiliza en su comunicación, palabras de su entorno escolar y familiar, y le suma significados nuevos valiéndose del contexto y de sus conocimientos previos.
10	Utiliza el lenguaje oral y gestual para expresar su opinión de lo aprendido en clase.
11	Sigue instrucciones orales con precisión al realizar actividades cotidianas de tres indicaciones simples.

Fuente: Currículum Nacional Base. Nivel Preprimario, 2008.

Los estándares que desarrollan las habilidades para la adquisición de la escritura emergente son los que se enlistan en la tabla:

Tabla 2. Estándares educativos del área Destrezas de Aprendizaje
Estándar Descripción del estándar

2	Dibuja líneas y formas, siguiendo trazos
8	Representa gráfica y verbalmente sus razonamientos

Fuente: Currículum Nacional Base. Nivel Preprimario, 2008.

B. Conceptos básicos del test

La palabra test tiene origen inglés, que proviene del vocablo latino *testa-testis*. Entre sus significados destacan, como nombre se refiere a prueba o reactivo, por mencionar algunos; como verbo es ensayar o comprobar. En 1999 la *American Psychological Association* -APA- definió al test como «procedimiento por medio del cual una muestra de comportamiento de un dominio especificado es obtenida y posteriormente puntuada, empleando un proceso estandarizado» (Tornimbeni, Pérez, & Olaz, 2008). La definición anterior comprende tanto al test de ejecución máxima, donde las respuestas son evaluadas por su corrección y calidad, como a los de comportamiento típico en donde ninguna respuesta puede ser marcada como correcta o incorrecta.

El término evaluación se refiere al proceso que permite integrar la información obtenida por medio de test con la proveniente de otras fuentes, como la que se obtiene de la historia clínica, familiar, ocupacional o educacional de la persona evaluada.

La meta de las ciencias es establecer los principios generales que permitan la descripción, predicción y explicación de los fenómenos de su área de interés. Los investigadores requieren de instrumentos que les permitan recoger y comparar datos que permitan fundamentar la teoría, siendo la medida un elemento fundamental para el proceso (Martínez, Hernández, & Hernández, 2006). Dicha medición de los fenómenos debe realizarse de forma adecuada, es decir la Psicometría, que es la ciencia encargada de reunir el conjunto de formalidades que establecen las bases de la medición, permite realizar dos clasificaciones entre sus fenómenos de estudio: el psicológico y el educativo. Siendo el ámbito educativo el que dirigirá la exposición precedente.

Martínez, Hernández y Hernández (2006) citan a Nunnally y Berstein, (1994) quienes indican que la Teoría clásica del Test es la asignación de números a atributos de los sujetos, de tal forma que los números reflejen los diferentes grados del atributo que es evaluado. Es importante resaltar que esta definición no aplica para el modelo de la Teoría de Respuesta al Ítem (TRI). La Teoría de Respuesta al ítem surgió en los años cincuenta

como una reacción a los problemas y limitaciones que presentaba la Teoría Clásica del Test (TCT). El objetivo principal de la TRI es conseguir medidas invariantes respecto de los sujetos evaluados y de los instrumentos utilizados (Muñiz, 1997).

En los años 60 el matemático danés Georg Rash propuso un modelo para solventar las deficiencias de la TCT y construir pruebas más adecuadas y eficientes (Prieto & Delgado, 2003).

1. Tipos de test

Referirse a una clasificación de los test lleva a pensar en la existencia de evaluación sumativa, formativa y diagnóstica, pero la evaluación también puede delimitarse de acuerdo a la interpretación que el evaluador hará con los resultados (Miller, 2012) (Miller, Linn, & Gronlund, 2012), es decir con los datos que recoge una evaluación se pueden dar resultados referidos a norma y o a criterio. Debiéndose construir el instrumento de evaluación, con las características necesarias para la interpretación de los resultados propuestos, por ello, se pueden reconocer los test referido a criterio y test referidos a norma.

A modo de referencia histórica sobre los test referidos a criterio, se resalta que estos se comenzaron a desarrollar a inicios del siglo XX. Esta línea de investigación se vio en desuso durante 1914 a 1945, fechas coincidentes con el tiempo que duró la primera y segunda guerra mundial. Siendo retomado a mediados del mismo siglo de surgimiento (Martínez, 1995). Se registra que en los años setenta se comienza a aplicar sistemáticamente este tipo de prueba en el ámbito educativo.

La evaluación educativa parte como medio para seleccionar estudiantes, siendo los test referidos a norma los aplicados para cumplir este objetivo y los resultados se interpretaban en función de un grupo normativo o baremo. Posteriormente, se construyen con la finalidad de evaluar la calidad de un sistema educativo o el rendimiento académico. Al respecto surgieron opiniones sobre la inadecuada forma de exponer los resultados que

arrojan los instrumentos contruidos con este objetivo. Es entonces, cuando emerge la evaluación referida a criterio, la cual supone interpretar los puntajes para medir los cambios de cada individuo a lo largo del aprendizaje (Ravela, 2006).

Según Bond (1999) el test referido a norma es seleccionado cuando el fin es obtener información que permita discriminar entre estudiantes, mientras que, en los test referidos a criterio, el contenido se selecciona sobre la base de su importancia para el currículo. Por consiguiente, estos test resultan adecuados para fines de diagnóstico y prescripción de las experiencias de aprendizaje requeridas para asegurar el logro de determinados objetivos (Tornimbeni, Pérez, & Olaz, 2008).

2. Clasificación del test según su función

La función de un test se determina desde que se planea, es decir elaborar un test requiere la consideración cuidadosa de sus propósitos específicos. Cuando se elabora un test o una prueba su propósito hará que su construcción varíe en cierto grado. De acuerdo con Aiken, L. (2003) “la elaboración de un instrumento inicia con la definición de los constructos que van a medirse y con el esbozo del contenido propuesto” (p. 18).

Martínez, Hernández y Hernández (2006), al respecto evidencian de manera sencilla la clasificación habitual de los test en el contexto educativo de acuerdo a la función que cumplen:

- Según las consecuencias para el sujeto: se clasifican dos, los test de altas consecuencias utilizadas en procesos de selección y de bajas consecuencias que son útiles cuando el objetivo es investigar. Resaltando que es importante dicha clasificación en la interpretación de los resultados.
- Según el planteamiento y tipo de respuesta, es decir, los test están contruidos a partir de una muestra de conductas que exige, por parte del sujeto, una determinada forma de planteamiento y resolución del problema presentado. De acuerdo con esta función se denomina el formato de los ítems.

- Según el área del comportamiento acotada: Se reconocen dos clasificaciones de test; el primero se refiere a los test cognitivos conocidos también como test de ejecución máxima que miden aptitudes, inteligencia, rendimiento académico, entre otros. Y los test no cognitivos llamados también de ejecución típica que miden personalidad, intereses, motivación, actitudes, etc.
- De acuerdo con la modalidad de aplicación: se refiere a los test de aplicación individual o colectiva, de papel y lápiz o de ordenador, adaptativo o no adaptativo, entre otros.
- En función de las demandas temporales: se clasifican de rapidez o velocidad a test de potencia.
- De acuerdo con el modelo estadístico en que se basan las puntuaciones del test. La construcción de test debe basarse en algún modelo probabilístico que permita dar un significado a las puntuaciones y hacer inferencias a partir de la muestra de conductas planteadas en el test. Se clasifican en este ámbito a la Teoría de los Test y a la Teoría de Respuesta al Ítem.
- Acorde al tipo de interpretación se distingue entre el nivel de ejecución sobre algún dominio o criterio definido de antemano, estos hacen alusión a los Test Referido a Criterio (TRC), Popham (1975) afirmó que la interpretación de los test referido a criterio tiene la finalidad de evaluar la posición absoluta de un individuo con respecto a algún dominio previamente definido. Los de aproximación basada en normas, que sitúa a los sujetos o estímulos en función de estadísticos calculados en algún grupo de referencia al que pertenece el sujeto, se les llama Test Referido a Norma (TRN).

Los aspectos sobre el planteamiento y tipo de respuestas, así como el modelo estadístico que se enlistaron anteriormente, serán tratados con mayor detenimiento en los subtítulos siguientes.

3. Construcción del test

En la práctica, la construcción del test requiere llevar el proceso bajo procesos rigurosos de creación, en el cual se interactúa con expertos en la materia o constructo a evaluar y en medición. Se hará un recorrido por los aspectos indispensables en la construcción del test.

a. Finalidad del test

El objetivo para construir los ítems es el aspecto que se refiere a definir qué evaluar. Al objetivo se le denomina la definición del constructo (Cronbach, 1985). Constructo es la representación abstracta de un conjunto de comportamientos, es decir, Ravela (2006) indica que es la construcción conceptual de lo que se quiere evaluar; la definición de la realidad que involucra conocimientos, visiones y valores. A este proceso se le conoce como la construcción del referente del test. Otro aspecto, que surge del objetivo del test es que se pueden considerar dos grupos al momento de establecer el objetivo. Uno de estos grupos puede surgir con la finalidad de estudiar el nivel óptimo o máximo del sujeto en determinadas competencias de tipo memorístico, razonamiento o rendimiento y el otro grupo trata sobre actitudes, personalidad o emociones típicas de cada persona (Martínez, Moreno, & Muñiz, 2005).

De acuerdo con la finalidad que tiene un test se identifican siete clasificaciones; entre ellas las de decisiones relativas al nivel de instrucción o habilidad, de selección, de asignación, de clasificación de consejo, de screening y el de decisiones de diagnóstico. Martínez, Hernández y Hernández (2006), indican que la finalidad de prueba diagnóstica tiene como objetivo la identificación de fallos o deficiencias en el comportamiento para emprender alguna intervención o tratamiento. Aunque las puntuaciones suelen compararse con datos normativos, este tipo de decisión también es propia de los Test Referidos a Criterio.

b. Marco del test

Al delimitar el propósito del test y el constructo a medir. Se procede a verificar que las cuestiones del test correspondan con las metas importantes de los programas, el currículo educativo y la instrucción. Como segundo paso, se especifican las conductas que revelan los constructos y como tercer paso especificar qué tareas o situaciones obtienen dichas conductas. Los pasos descritos anteriormente se refieren a la denominada variable latente. Dicha variable se caracteriza porque no es directamente observable, pero se plantea como causa de las respuestas a los ítems construidos que tienen el propósito de medir dicha variable (Martínez, Hernández, & Hernández, 2006).

Definir el dominio involucra también determinar aspectos complementarios como las instrucciones de administración y el modo de calificación e interpretación de las respuestas.

Además, se delimitan las especificaciones derivadas del marco. En esta, se plasma las conductas concretas a evaluar, la cantidad de ítems que evaluarán las conductas. Es una práctica usual construir mayor cantidad de ítems que los especificados para utilizarse en el test. Es decir, Tornimbeni, Pérez & Olaz (2008) para la elaboración de pruebas de rendimiento referidas a normas se parte de la construcción de una tabla de especificaciones, la cual consiste en una tabla de doble entrada en donde se relacionan los contenidos específicos y se determina la cantidad de ítems que conformarán la prueba. Cuando la prueba es con referencia a criterio, se elabora y se delimita el dominio a cada objetivo. Después de definidos los objetivos se debe proceder a elaborar los ítems (Urbina, 1998).

4. Construcción del ítem

El término ítem se ha revitalizado por influjo del inglés y su empleo es muy frecuente en diversas disciplinas científico-técnicas, especialmente en psicología. El significado denotativo refiere a esta palabra como el elemento o conjunto de elementos que constituyen una unidad de información dentro de un conjunto (Real Academia Española, 2005).

La construcción de los ítems debe responder a determinados criterios objetivos que, una vez conocidos, deben ser utilizados para realizar la construcción de manera reglada y sistemática, produciendo así ítems adecuados al diseño y constructo medido (Martínez, Moreno, & Muñiz, 2005). Para seguir sistemáticamente el proceso, es conveniente tener en cuenta:

a. El formato

El aspecto del formato de los ítems responde a cómo hacerlo. Este tema abarca varias posibilidades, cuando es de elaboración del producto o llamado también de respuesta abierta, se adquieren las opciones de preguntas abiertas, de completación y de sustitución.

Por otro lado, existen ítems que constan de un enunciado, sea pregunta o proposición (Diguca, 2014) que se responde o completa con varias opciones entre las que el sujeto debe elegir un determinado número de ellas, frecuentemente solo una que considere la correcta cuando se evalúa ejecución máxima o la más acorde a su opinión cuando se evalúa ejecución típica. Cuando se solicita al sujeto seleccionar alguna de las opciones de respuesta se le conoce como respuestas cerradas. Las opciones pueden adoptar el formato de falso-verdadero, acuerdo-desacuerdo con lo expresado en el enunciado, tipo Likert que ofrece la graduación entre las opciones dadas. Cuando las opciones ofrecidas poseen su propio contenido y no son solo calificaciones del enunciado, como lo requieren las opciones anteriores, el formato es llamado de elección múltiple y selección múltiple. En

las siguientes líneas se describe cada uno de los formatos (Martínez, Moreno, & Muñiz, 2005):

- Preguntas abiertas. Cuando el producto solicitado es verbal, y el sujeto debe expresar por escrito u oralmente un determinado contenido. Las preguntas abiertas se distinguen por su longitud, la cual dependerá de lo solicitado al sujeto.
- De completación. Son aquellos ítems que presentan un estímulo en el que se han suprimido algunos de sus elementos pidiendo al sujeto completarlo. Este formato puede considerar dar el estímulo de forma verbal o escrita solicitando al sujeto que añada uno o más elementos de una oración.
- De sustitución. El último formato del primer grupo de elaboración del producto, es el ítem de sustitución llamado también de corrección. Consiste en solicitar que se sustituya, cambie, o elija entre opciones dadas determinado elemento por una alternativa correcta.
- Verdadero – falso y acuerdo – desacuerdo. Llamado así, cuando las opciones ofrecidas dicen literalmente esas frases. La variante de acuerdo-desacuerdo, es utilizada en las evaluaciones que miden actitudes y personalidad.
- Tipo Likert. Se presentan escalas de acuerdo o de números de 1 a 5 para valorar el enunciado.
- Elección alternativa o selección múltiple. Se denominan de elección múltiple cuando se presentan dos opciones a elegir. Generalmente este tipo de formato solicita que el sujeto se posicione en la opción que identifica su opinión. En cuanto al de selección múltiple se le llama al de más de dos opciones presentadas en el ítem.

b. Revisión de los ítems

Llevar a cabo este paso en la construcción de los test previene errores altamente costosos. Por tanto, conviene revisar los reactivos desde varios puntos de vista. Se encuentra importante contar con revisiones de expertos en el contenido, sujetos inmersos en la población objetivo y expertos en psicometría. El conjunto de observaciones que se deriven de estas revisiones detectará posibles problemas de infrarrepresentación del

constructo o introducción de varianza irrelevante al constructo. El hecho que los revisen sujetos de la población contribuye a evitar sesgos de los test, uso de lenguaje inadecuado u ofensas a grupos particulares.

c. Revisión por expertos

La revisión por expertos es un proceso que permite evidenciar validez de contenido. Los expertos que revisan los ítems redactados deben evaluar principalmente tres aspectos: a) claridad semántica y corrección gramatical; b) adecuación al nivel de comprensión de la población meta; c) congruencia con el constructo o dominio medido. La congruencia con el constructo o dominio medido es el principal parámetro que hace referencia al grado de consistencia que debe existir entre un ítem particular y el constructo o constructos a medir por el test (Tornimbeni, Pérez, & Olaz, 2008). Este aspecto contribuye significativamente en la confiabilidad y validez de las puntuaciones del test a construir (Oesterlind, 1990). Se pueden utilizar estadísticos como el coeficiente Kappa para la interpretación de concordancia entre jueces.

El proceso de la revisión por expertos es un proceso empírico que constituye una fuente de validez del instrumento. Por lo general se emplean escalas numéricas creadas para obtener medidas de acuerdo o desacuerdo entre jueces. Se recomienda que los ítems seleccionados sean aquellos que tienen al menos el 60% de criterios meritorios entre jueces (Herrera, 1998).

Es importante llevar a campo los ítems construidos, además del juicio de expertos para comparar los resultados entre los procesos con juicio de expertos y estudio piloto.

d. Estudio piloto

Esta fase del proceso involucra probar el funcionamiento del test. Llevar a campo el doble de ítems que se utilizará en el instrumento final (Martínez, Hernández, &

Hernández, 2006), puesto que muchos serán descartados en el proceso de revisión y en el análisis estadístico ulterior (Tornimbeni, Pérez, & Olaz, 2008).

Es importante recoger información sobre la comprensión de las instrucciones, los ítems y los procedimientos de respuesta. En esta fase también se recogen datos sobre el tiempo de aplicación (Martínez, Hernández, & Hernández, 2006). Además, se puede proceder a realizar algunos análisis de los ítems para seleccionar aquellos que mejor se adapten al objetivo del test.

e. Estudio de campo

Proceso utilizado para obtener los estadísticos del test como índice de confiabilidad, error de medida, errores de clasificación, validez, estructura interna, entre otros. Los estadísticos servirán para la calibración del test. A partir de los datos recogidos en campo, se obtendrán las escalas de puntuaciones en las que están basadas las normas. Es recomendable proceder posteriormente a establecer los distintos puntos de corte o los baremos de la prueba.

La distinción principal cuando se realizan test referidos a norma o a criterio (Martínez, Hernández, & Hernández (2006) indican que:

La puntuación de cada evaluado se interpreta en relación con los otros sujetos de la población a la que pertenece, por lo que se necesitan datos procedentes de los grupos de referencia. Los grupos de referencia están constituidos por las poblaciones y subpoblaciones a las que va dirigido el test. Generalmente, en los test referidos a criterio no es necesario realizar la medición como se describe anteriormente, a razón que el rendimiento o conducta de una persona en términos absolutos es suficiente para describir su estado con relación al dominio establecido. (p. 85)

Por tanto, el investigador debe aplicar el test a un grupo normativo, el cual es una muestra representativa de la población objetivo.

5. Propiedades de los ítems

La Teoría de Respuesta al Ítem (TRI) resalta que se interesa más por las propiedades de los ítems individuales que en las propiedades globales, como es el caso de la Teoría clásica del Test (TCT) (Abad, Garrido, Olea y Ponsoda, 2006) citado en (Tornimbeni, Pérez, & Olaz, 2008). Sin embargo, debe recordarse que ambas teorías tienen el mismo objetivo básico, es decir, obtener la puntuación que corresponde a una persona en un rasgo latente o dominio de conocimiento. Que en la TRI el rasgo latente se representa como θ y se simboliza: θ .

Los ítems son los ladrillos con los que se hacen los test y cada uno de estos ha de contribuir a que el test constituya un buen instrumento de medida (Martínez, Moreno, & Muñiz, 2005), es decir, que cumpla con los requisitos de validez y confiabilidad para generalizar los resultados obtenidos con el mismo y poder asegurar que los datos recogidos pueden ser generalizables en la toma de decisiones basada en evidencia.

a. Validez

Nunnally (1991) manifiesta que «las puntuaciones evidencian propiedades de validez cuando se verifica que el test realmente mide el constructo que pretenden medir, justificando adecuadamente las inferencias realizadas en función de sus resultados». (p. 59)

La situación ideal de validez del test es que este represente adecuadamente y mida la varianza relevante del constructo, es decir que las interpretaciones de los resultados de una prueba estén libres de sesgo. El hecho que las puntuaciones sean confiables es una condición indispensable pero no es suficiente para que sean válidas (Muñiz, 2001).

1) Fuentes internas de validez

Se distinguen dos grupos de fuentes de evidencia de validez de un test. Tornimbeni, Pérez & Olaz (2008) proponen dos categorías de fuentes, las internas y las externas. Entre las fuentes internas se clasifica la evidencia:

- a) Basada en el contenido del test. Se obtiene demostrando que el contenido (ítems) del test es una muestra representativa del constructo o dominio respecto del cual se desea hacer una inferencia. Los test que evalúan rasgos latentes poseen menos representatividad en relación con el dominio de comportamiento que intentan medir, puesto que no se basan en un dominio de conocimiento específico. De acuerdo con Martínez (1995), en la validación de contenido se deben cubrir los siguientes aspectos:
 - Definición del dominio de conocimiento o comportamiento a medir.
 - Identificación de expertos en ese dominio.
 - Juicio de los expertos acerca del grado en que el contenido del test es relevante y representativo del dominio. Los jueces deben establecer el nivel de pertinencia, relevancia y suficiencia.
 - Procedimientos estadísticos para resumir los datos de la fase precedente. Por medio de porcentajes de acuerdo o coeficientes de asociación o correlación.
- b) Basada en la estructura interna del test. Indica si las relaciones entre los ítems y las dimensiones (factores, escalas) permiten confirmar la existencia de los constructos que el test pretende medir. El marco conceptual de un test puede proponer una dimensión unitaria de comportamientos o varios factores. Este proceso se puede realizar por medio del análisis factorial (Carretero-Dios & Pérez, 2005).
- c) Además, la validez interna basada en el proceso de respuesta, desde un enfoque cualitativo indica que el análisis empírico y teórico del proceso de respuesta del test puede suministrar evidencia relacionada con la congruencia entre el constructo medido y la naturaleza del rendimiento o respuesta emitida por los examinados. Se

obtiene por medio de entrevistas o protocolos de pensamiento que emitan datos cualitativos.

Volviendo al tema de la confiabilidad de los datos recogidos en campo es preciso puntualizar en los postulados y supuestos de la TRI. Estos indican que todo modelo de la Teoría de Respuesta al Ítem relaciona matemáticamente tres aspectos (Tornimbeni, Pérez, & Olaz, 2008, p. 22-223):

- La probabilidad de emitir una respuesta determina a un ítem.
- Las características del individuo.
- Las propiedades del ítem (dificultad y/o discriminación).

Entonces, para comprobar que el rendimiento en un ítem depende de un único rasgo latente, es decir en una característica inobservable, se utilizan modelos de dimensionalidad.

b. Dimensionalidad

En la jerga común de la psicometría es habitual referirse a la confiabilidad cuando se habla de la consistencia entre los puntajes de un test obtenidos por los mismos individuos en distintas ocasiones o entre diferentes conjuntos de ítems equivalentes (APA, 1999). Con respecto a los supuestos de la TRI se destaca que la unidimensionalidad del rasgo latente es uno de los supuestos que debe demostrarse para identificar la confiabilidad de los ítems, en otras palabras, que los ítems de un test midan solo un rasgo (Cortada de Kohan & Cupani, 1998).

El término también se encuentra en la literatura como fiabilidad del test, se refiere al error de medida. Es decir, el resultado de la medición es un valor observado que no coincide con el valor verdadero, y en consecuencia siempre se mide con un margen de error (Olaz, 2008). Muñiz (2001) declara que el valor verdadero es un valor teórico que se reduce en un concepto matemático. El indicador más frecuente, en la TCT, aunque no el único, para expresar el grado de variabilidad o de error de medida es el coeficiente de

fiabilidad. Para estimarlo se elige, dependiendo del diseño, el test-retest, consistencia interna o test paralelos.

La estimación más utilizada bajo los principios de consistencia interna se basa en el estadístico Alfa de Cronbach. Se fundamenta, en su carácter práctico, ya que puede calcularse con una sola aplicación del test. El coeficiente Alfa parte de la correlación entre las dos mitades del test. Otro coeficiente utilizado es el coeficiente Omega y Theta que están basados en los datos del análisis factorial (Zeller, 1979). Ambos coeficientes indican en qué grado los ítems de un test convergen, están intercorrelacionados, a eso se refiere la consistencia interna. (Martínez, Moreno, & Muñiz, 2005) Los ítems cuya correlación con el total del test (discriminación) sea baja tenderán a debilitar la fiabilidad del instrumento.

Un coeficiente de Alpha de Cronbach de 0.80 sugiere que un 64% de la varianza observada es asimilable a la variabilidad de la puntuación verdadera, con un margen de error del 36% (Martínez, Moreno, & Muñiz, 2005). En el caso del análisis factorial, el cual se comprende como la técnica de reducción de los datos que tiene como objetivo agrupar un conjunto de variables en pocas dimensiones que expliquen la mayor cantidad de variabilidad de respuesta, es decir con el análisis factorial, por ejemplo, un gran número de ítems puede reducirse a un número pequeño de factores que confieran un significado teórico a la medición. En donde, De acuerdo con Pérez, E. y Tornimbeni, S. (2008) cada factor agrupa los ítems altamente correlacionados entre sí y que son, sin duda alguna, relativamente independientes con los otros factores (Aiken, 2003).

Antes de comenzar un análisis factorial se determina con pruebas estadísticas su viabilidad, una de ellas es la prueba de Kaiser-Mayer-Olikin (KMO), que su interpretación indica que cuando se encuentra en un rango de 0 a 1, se considera adecuado el valor ≥ 0.70 , el cual sugiere correlación satisfactoria entre los ítems. Si se cumple este principio, entonces es recomendable proceder con el análisis factorial.

Se distinguen dos tipos de análisis factorial, el exploratorio y el confirmatorio. Los cuales distinguen su utilidad por el objetivo del análisis.

1) Análisis factorial exploratorio

Se extraen e interpretan posibles factores que explican la covarianza entre las variables, en este caso los ítems del test, sin una estructura teórica previa conjeturada de modo explícito. La interpretación para el análisis factorial exploratorio se utilizan los mismos valores que en el confirmatorio.

2) Análisis factorial confirmatorio

La estructura factorial se define a priori sobre la base de un modelo teórico para explicar la covarianza entre ítems. El análisis intenta verificar cuan bien se ajustan los resultados observados al modelo teórico preestablecido mediante métodos como el de máxima verosimilitud y estadísticos como Chi cuadrado (X^2), la diferencia en Chi cuadrado entre los modelos comparados, el índice de bondad del ajuste (GFI) y el índice de ajuste comparativo (CFI).

Los valores entre estos índices de bondad de ajuste (GFI y CFI) oscilan entre 0 y 1, en donde 1 indica un ajuste perfecto. Valores superiores a 0.90 se refieren a un ajuste satisfactorio entre las estructuras teóricas y los datos empíricos. Mientras que valores de 0.95 o superiores establecen un ajuste óptimo. El X^2 debe ser no significativo para indicar un buen ajuste de los datos observados al modelo teórico de la medición. Como es sensible a la muestra este último estadístico, es necesario tener precaución al momento de interpretarlo (Thorndike, 1989) y se considera una buena práctica cuando se tiene varios modelos para comparar identificar aquellos que tienen una reducción significativa en el X^2 para considerarlo el que mejor ajuste tiene.

c. Independencia local

La independencia local corresponde al segundo supuesto de la TRI. Este se verifica la respuesta de un individuo a un ítem no depende de la respuesta que haya dado a los

otros ítems del test. Por tanto, este segundo supuesto se deriva del supuesto de unidimensionalidad, porque la respuesta a un ítem solo es explicada por el nivel de theta y los parámetros de los ítems: discriminación y/o dificultad (Tornimbeni, Pérez, & Olaz, 2008), según sea el modelo que se utilice.

En principio, el supuesto de independencia local, matemáticamente, indica que la «probabilidad de responder correctamente a un conjunto de ítems es igual al producto de las probabilidades de responder correctamente a cada uno de esos ítems individualmente» (bis, p. 225). En palabras de Hambleton y Rogers (1991), la correlación de Pearson es un procedimiento que se debe realizar para contrastar si el supuesto de independencia se cumple. Es decir, obtener la matriz de correlaciones entre los ítems, pero no en la muestra completa sino en submuestras que sean homogéneas en el nivel de habilidad de sus miembros deberá ser 0, porque si el supuesto se cumple, los ítems no deberían correlacionarse entre sí.

Por tanto, en concordancia con Kline (2000) la independencia local es el supuesto fundamental y distintivo de la TRI porque significa que el nivel de theta (rasgo o aptitud del individuo) debería explicar totalmente la variabilidad de respuesta al ítem.

d. Los modelos teóricos de la TRI

La finalidad de analizar los ítems de un test es mejorar las propiedades psicométricas de este. Esto asegura que el instrumento está midiendo el constructo que se pretende medir (validez), sin error (confiabilidad).

El análisis de los ítems dependerá de las líneas con que se concibe la construcción del test (Bechger, Maris, & Verstralen, 2003). Se reconocen dos teorías que permiten el análisis de los ítems. Por tanto, la TCT es un modelo teórico que ha tenido auge desde hace 100 años (Martínez, Hernández, & Hernández, 2006), del cual surgen, en vías de completar lo propuesto por esta teoría, la teoría de la Generalizabilidad que viene a solventar, entre otros aspectos, la concepción unitaria e indiferenciada del error de medida.

Posterior a esta, se desarrolla la TRI que consiste, en términos generales, en establecer las relaciones entre las medidas observadas y el constructo, relacionando con éste las puntuaciones de cada uno de los ítems y por medio de los patrones de las respuestas obtener una estimación del valor del sujeto en el constructo.

En concordancia con lo dicho anteriormente los parámetros de los ítems (Tornimbeni, Pérez, & Olaz, 2008) son: dificultad, discriminación. Entonces, en línea con la TRI, para establecer el proceso de obtener dichos parámetros, es necesario tomar decisiones sobre cuál de los modelos que esta teoría propone representa el mejor rendimiento en los ítems. En otras palabras, cuál de los modelos se seguirá para obtener tales parámetros.

1) Modelo de un parámetro

El modelo de Rash se fundamenta en dos supuestos, a) el atributo a medir puede representarse en una única dimensión en la que se situarían las personas y los ítems. Y b) la aptitud de la persona en el atributo y la dificultad del ítem determinan la probabilidad de que la respuesta sea correcta.

El modelo de Rash también hace referencia al modelo de 1 parámetro, es decir al parámetro de dificultad (b) que se representa con β . La dificultad del ítem es una medida que se interpreta de igual forma que la medida theta de las personas, es decir describe en dónde se sitúa el ítem en la escala de aptitud.

Lo anterior hace referencia a identificar qué cantidad de aptitud (habilidad) requiere el ítem para ser resuelto con éxito, autores como Martínez, Hernández y Hernández aseguran que el parámetro de dificultad es el índice de la posición o localización del ítem en la escala de medida del rasgo o aptitud (theta). Aseguran que, es el nivel de aptitud en el punto de inflexión de la curva, claro está sino existen aciertos al azar. Así que, la probabilidad de acierto en dicho punto es igual a 0.50.

El modelo de un parámetro no refiere de ninguna manera la no medida de discriminación, más bien, asume que todos los ítems tienen discriminación 1.

2) Modelo de dos parámetros

El modelo de dos parámetros, además del cálculo de la dificultad, incluye el de discriminación (a), pero el cálculo del parámetro de dificultad está en función de la medida de discriminación, en donde en la curva característica del ítem (CCI) indica hasta qué punto la discriminación permite diferenciar entre sujetos que tienen una aptitud o rasgo inferior a la posición del ítem (dificultad) y los que la tienen superior.

Es posible inferir que la tasa de cambio, indican Martínez, Hernández y Hernández (2006), refleja la probabilidad de éxito según aumenta la aptitud.

Baker (1985), proporcionó la guía para la interpretación de valores de discriminación cuando el modelo logístico no se multiplica por 1.7, siendo esta:

Interpretación	Valores
Ninguna	0
Muy baja	0.01 – 0.34
Baja	0.35 – 0.64
Moderada	0.65 – 1.34
Alta	1.35 – 1.69
Muy alta	>1.70
Perfecta	∞

En general, se identifica que las medidas de discriminación oscilan entre 0.65 a infinito, considerándose las medias cercanas a 2 como discriminación alta. Ahora bien, cuando el algoritmo se multiplica por 1.7 (función log normal) la interpretación se refiere a buena discriminación cuando el valor es 2, discriminación media con el valor 1 y bajo discriminación con valor 0.5.

3) Otros modelos de la TRI

El modelo de tres parámetros, además de contemplar los parámetros indicados en los modelos anteriores, es decir de calcular discriminación a y dificultad b añade el parámetro de adivinación (c) o de azar.

En términos generales, este parámetro calcula la probabilidad de responder acertadamente al ítem por el azar y es preciso indicar que el parámetro c se añade para caracterizar la asíntota inferior en la curva característica del ítem (CCI).

4) La curva característica del ítem

En donde los tres modelos usan la curva característica del ítem (CCI) para especificar la relación entre el rendimiento observado en los ítems de un test y los rasgos latentes que explican ese desempeño (Tornimbeni, Pérez, & Olaz, 2008).

La ventaja del análisis de los parámetros de las propiedades psicométricas de los ítems que se estiman a partir de las curvas características de los ítems es que son más estables que los estimados con los métodos de la TCT, porque los parámetros de los ítems son menos dependientes de las muestras y de los instrumentos.

5) Función de información del ítem y del test

Luego de administrado el conjunto de ítems de un test y conocidos sus parámetros, así como el nivel de habilidad de cada persona de la muestra, es preciso estimar la precisión y el error típico de medición para diferentes niveles del rasgo. Siendo en la TRI la función de medición el indicador de la precisión de una medición en un punto de la escala de medición del rasgo latente (0, 1, -1, entre otros).

Cuando la CCI es grande, indica Cortada de Kohan & Cupani (1998), cambios mínimos en la aptitud se reflejan en cambios considerables en la probabilidad de acertar

el ítem. Entonces, el error de medición estándar no es uniforme para todos los niveles de aptitud y se expresa por la desviación estándar del ítem en un nivel de rasgo puntual. En resumen, a menor varianza de un ítem en determinado nivel de aptitud, mayor será la información aportada y la precisión de medición. En otras palabras, cuanto mayor sea el valor de la sumatoria de la función de información de cada uno de los ítems del test (FIT) menor será el valor del error típico de estimación.

La función de información de un test en su conjunto (FIT) y el error típico de medida para un determinado nivel de rasgo depende de: a) los parámetros de discriminación de los ítems (mientras más alto a mayor valor de información) b) la convergencia entre nivel de rasgo (θ) y los parámetros de dificultad de los ítems (cuanto más próximos sean, mayor será el monto de información aportado). En síntesis, los valores inferiores a 1 son indicadores de baja información (precisión), entre 1 y 1.69 valores moderados y mayores a 1.70 o superiores son valores elevados (Martínez Arias, 1995).

Entre los aspectos propuestos por Birnbaum (Chacón Moscoso & Pérez-Gil, 2008) es el relacionado con la cantidad de información proporcionada por los ítems y el test respecto del nivel de habilidad de los sujetos. Birnbaum (1968) citado en Chacón *et al*, (2008) introdujo la medida de información de Fisher en el contexto de la información proporcionada por un test. Esta información viene dada por la función de información del test, y presenta la siguiente expresión:

$$I(\theta) = \sum I_i(\theta)$$

Donde $I_i(\theta)$ es la información del ítem i condicionada a θ . Esta expresión indica la precisión de las puntuaciones que ofrece el test condicionada a cada uno de los valores que puede tomar la aptitud bajo estudio. «Esta nueva aproximación al concepto de fiabilidad de las puntuaciones vino a resolver otro de los grandes inconvenientes de la Teoría Clásica de los Tests: el de la homocedastidad del error de medida a lo largo de toda la distribución de valores de la aptitud. También fue Birnbaum quien sugirió la utilidad

de esta función de información en la construcción de tests» (Chacón Moscoso & Pérez-Gil, 2008, p. 14).

6. Tamaño de la muestra

La selección de la muestra es un proceso que el investigador lleva a cabo según las características de la población objetivo. Es decir, es importante considerar las características de la población para que el grupo de sujetos que serán parte del estudio compartan dichas características distintivas. Este proceso permite la generalización de los resultados.

Una muestra de 300 personas para análisis de ítems es un estándar deseable (Tornimbeni, Pérez, & Olaz, 2008). En consecuencia, para el análisis factorial se utilizan muestras grandes, de 300 sujetos para obtener resultados útiles y relativamente estables (Tabachnick & Fidel, 2001). Se debe contar idealmente con 10 participantes por ítem o como mínimo 5 (Nunnally, 1991).

En este sentido, para que los parámetros sean estables en la TRI se requieren muestras >500 porque son las que posibilitan el ajuste a cualquier modelo de los mencionados. Por ello, cuando las muestras son menores, es decir cuando son pequeñas el mejor modelo es el de Rash, de allí que es el más popular. (Martínez, Hernández, & Hernández, 2006).

7. Uso adecuado de los test

Disponer de un test técnicamente correcto para evaluar el constructo deseado es indispensable para asegurar que los datos pueden generalizarse; pero por si solo no se puede alcanzar el objetivo, es decir por muy bueno que sea una prueba es necesario asegurar que quien utiliza, administra, aplica, y maneja el instrumento de evaluación lo haga correctamente.

Por tanto, el objetivo que se persigue con la redacción del manual de aplicación es sistematizar las directrices del uso del test para asegurar la forma ética y profesional del desempeño de los administradores, así como del conocimiento de la aplicación (Martínez, Moreno, & Muñiz, 2005).

a. Ámbito de aplicación

Dentro de la sistematización del proceso se recomienda considerar (Martínez, Moreno, & Muñiz, 2005):

- Los test son procedimientos diseñados para ser aplicados bajo condiciones controladas, es decir estandarizadas y conllevan la utilización de protocolos de puntuación rigurosos.
- Incluir procedimientos que proporcionan ordenamientos de las personas involucradas.
- Es necesario aportar datos que justifiquen las inferencias que se hacen a partir de las puntuaciones de los test. Los datos deben estar disponibles para los usuarios de los test, así como para los profesionales e investigadores que deseen llevar a cabo una evaluación o revisión independiente.

b. Destinatarios

Las directrices se aplican al uso de los test en la práctica profesional, se responden cuestionamientos que pueden surgir a:

- Los responsables de elegir el test y el uso adecuado del mismo.
- Quienes aplican, puntúan e interpretan el test.
- Los evaluadores y todos los involucrados en la aplicación del test.
- Usuarios del test con fin investigativo.

c. Conocimientos de aplicación

El uso inadecuado del test, afirman Martínez, Moreno y Muñiz (2005) se deriva del deficiente conocimiento sobre todos los aspectos implicados en el proceso evaluativo que puede poseer el aplicador del test. Es recomendable evidenciar en el Manual de aplicación del test:

- La estandarización de los procesos. Es decir, asegurar que el trato y condiciones son los adecuados y los mismos para todos los grupos evaluados.
- El actuar ético y profesional esperado en campo.
- Proporcionar la información pertinente al momento de aplicar el instrumento.
- Resguardar los procesos técnicos para evitar la publicación de tal modo que la utilidad del test quede deteriorada.
- Asegurar el almacenaje y traslado de los instrumentos para mantener la confidencialidad de los datos recogidos.

VI. METODOLOGÍA

A. Definición del problema

La prueba diagnóstica para primer grado surge del Centro de Excelencia para la Capacitación de Maestros de Centroamérica y República Dominicana con el propósito que el docente recoja información sobre competencias que los estudiantes poseen al inicio del año. Las habilidades que evalúa la prueba diagnóstica son comprensión oral, Desarrollo de la conciencia fonológica, y escritura emergente; saberes que debiesen poseer de lectura y escritura emergente al ingresar a primer grado después de haber recibido instrucción formal e intencional en el Nivel de Educación Preprimaria.

Es importante aclarar que «esta prueba no evalúa todas las subcompetencias relacionadas con la lectura y la escritura...» (USAID/CETT, 2009), entonces, por ser una prueba diagnóstica, la comunidad educativa que utilice el instrumento de evaluación puede detectar deficiencias y fortalezas para planificar y desarrollar ambientes de enseñanza-aprendizaje-evaluación que respondan a las necesidades detectadas de la lectura y escritura emergente.

Se reconoce que en Guatemala existen limitados instrumentos validados y contextualizados con los que se pueda evaluar lectura y escritura emergente en un instrumento de diagnóstico de los aprendizajes que los estudiantes adquirieron previamente a primer grado de primaria. Por tanto, resulta importante determinar: ¿Cuál es la consistencia interna y validez de la Prueba diagnóstica de primer grado de primaria?

Al determinar la consistencia interna y validez del diseño de la Prueba diagnóstica para primer grado de primaria se podrá identificar si la prueba mide los constructos de forma confiable y válida, lo cual permitirá tomar decisiones del uso de la prueba en aplicaciones que requiera el CIE.

B. Objetivos

1. Objetivo general

Analizar la consistencia interna y validez de la prueba diagnóstica para primer grado de primaria que mide lectura y escritura emergente, aplicada a 300 estudiantes en Guatemala.

2. Objetivos específicos

- a. Describir lo que miden las cinco subpruebas que comprende la Prueba diagnóstica para primer grado de primaria.
- b. Obtener la medida de confiabilidad del instrumento correspondiente a la Prueba diagnóstica para primer grado de primaria
- c. Evidenciar la validez de la Prueba diagnóstica para primer grado de primaria para asegurar que las inferencias que se realizarán con ella son fiables.
- d. Analizar las propiedades psicométricas de la Prueba diagnóstica para primer grado de primaria.
- e. Explorar los factores que integran el constructo de la Prueba diagnóstica para primer grado de primaria.

C. Diseño

Para cumplir con los objetivos propuestos; el enfoque de investigación utilizado fue el cualitativo, que de acuerdo con Hernández Sampieri, Fernández Collado y Baptista Lucio (2014) este enfoque “representa un conjunto de procesos, es secuencial y probatorio” (p. 4).

El proceso realizado comprendió la validación de contenidos, el cual fue analizado con los criterios que establece el Juicio de expertos. Otro dato que se obtuvo fue la alineación de la Prueba diagnóstica para primer grado de primaria y el Currículo Nacional

Base del Nivel de Educación Preprimaria. Asimismo, se recogieron datos para el análisis de los ítems bajo lineamientos regidos por la teoría de la calificación de instrumentos estandarizados basados en criterios, es decir se calcularon valores de discriminación y dificultad de los ítems y confiabilidad del test. Por tanto, el diseño responde al enfoque cualitativo de la investigación.

D. Descripción de la población y de la muestra

La población a quien está dirigido el test son estudiantes que cursan el primer grado de primaria tanto para el sector privado como el público. Los estudiantes seleccionados para la aplicación piloto del instrumento fueron niños inscritos en la Etapa 6 años o Párvulos III del Nivel de Educación Primaria a razón del momento del año en que los datos se recogieron en campo (septiembre-octubre) donde se analizó que los estudiantes de primer grado ya habían alcanzado otras habilidades de lectura y escritura inicial y los estudiantes de la Etapa 6 años, estaban culminando el aprendizaje que la prueba diagnosticará con su población objetivo, es decir la lectura y escritura emergente. Esta decisión se consideró pertinente porque el diseño de la prueba establece que es un instrumento diagnóstico de lectura y escritura emergente para estudiantes de primer grado de primaria que debe ser aplicado, según su Manual de aplicación y calificación, durante el primer mes del ciclo escolar.

El aprendizaje de la lectura y escritura emergente finaliza en el Nivel de Educación Preprimaria. Por ello, para realizar el análisis del instrumento que midió dicho constructo se aplicó el instrumento en los meses de septiembre y octubre del año 2015 con estudiantes del sector público y privado que estaban por egresar del Nivel de Educación Preprimaria.

E. Obtención y análisis de datos

En los apartados siguientes se describe el proceso para la selección de la población de donde se recogieron los datos y los procedimientos que se realizaron para el análisis de estos.

1. Obtención de los datos

Los datos se obtuvieron de centros educativos que imparten educación para el Nivel de Educación Preprimaria. La selección fue de tipo no probabilístico específicamente por conveniencia, porque se pretende validar el instrumento de evaluación y no a la población (Vivanco, 2005). Dado que se trata de un estudio para la determinación de los estadísticos de la prueba, dicho proceso se organizó en el departamento de Guatemala. Es importante mencionar que esto obedece a que la evidencia observada por medio de los resultados de logro reportados por el Ministerio de Educación indica que en este departamento los estudiantes se encuentran entre los porcentajes más altos del nivel de logro (56 %); este hecho es significativo para determinar el techo de los ítems, es decir, identificar si estos resultan difíciles para la población que se evaluó.

La muestra se obtuvo de la ciudad capital y de los municipios de Guatemala, Santa Catarina Pínula y Fraijanes. El 67 % de los datos se obtuvieron de la ciudad capital, del sector oficial, jornada matutina y el 50 % de los centros educativos seleccionados se encuentran en el área urbana. Los estudiantes evaluados tienen entre 6 y 6.11 años. El 100 % de los evaluados se encontraban cursando el último año del Nivel de Educación Preprimaria.

La tabla siguiente evidencia la distribución de la muestra por municipio y zona.

Tabla 3. *Distribución de la muestra evaluada para la aplicación piloto de la Prueba diagnóstica para primero primaria*

<i>Municipio</i>	<i>Zona</i>	<i>Frecuencia</i>	<i>Porcentaje</i>	<i>Porcentaje acumulado</i>
<i>Guatemala</i>	11	66	22.00	22.00
<i>Guatemala</i>	1	60	20.00	42.00
<i>Guatemala</i>	15	32	10.67	52.67
<i>Guatemala</i>	16	69	23.00	75.67

<i>Municipio</i>	<i>Zona</i>	<i>Frecuencia</i>	<i>Porcentaje</i>	<i>Porcentaje acumulado</i>
<i>Fraijanes</i>	1	19	6.33	82.00
<i>Santa Catarina</i>	Km. 16 carretera al Salvador	54	18.00	100.00
<i>Pínula</i>				
<i>Total</i>		300	100.00	

La cantidad de estudiantes que conformó la muestra fue de 300 casos. Se tomó en cuenta para ello las siguientes bases teóricas: a) El modelo de Rash que permite, bajo un buen ajuste de los datos, la medición conjunta de personas e ítems en una misma dimensión o constructo, expresa que el estadístico de ajuste Infit para un ítem o para una persona considera desajuste en muestras con menos de 500 casos cuando los valores son superiores a 1.3. c) Para realizar un análisis factorial exploratorio, de acuerdo con Tornimbeni, S., *et. al*, (2008), para obtener resultados útiles y relativamente estables al realizar un análisis factorial que responderá a la validez de constructo, debe usar aproximadamente 300 casos.

Por tanto, es posible establecer bajo un proceso de aplicación piloto con 300 casos los análisis psicométricos necesarios para analizar la consistencia interna y validez de la Prueba diagnóstica para primer grado de primaria.

a. Software utilizado

Entorno R

R es un lenguaje de programación con un enfoque al análisis estadístico. R surgió como una reimplementación de software libre del lenguaje S, adicionado con soporte para ámbito estático. (R Core Team, 2021).

Los paquetes utilizados para realizar los análisis son:

eRm: que se comprende como la estadístico el análisis del modelo Rash, por sus siglas en inglés: *Extended Rasch Modeling*. Además, para el análisis de ajuste se utilizó: *Fits Rasch*

models (RM), en el caso del análisis logístico lineal del modelo se implementó: *linear logistic test models* (LLTM), a la vez, las medidas de: *rating scale model* (RSM), *linear rating scale models* (LRSM), *partial credit models* (PCM), y *linear partial credit models* (LPCM). Así como el estadístico para calcular la máxima verosimilitud de estimación del parámetro de Person. Esta librería también incluye: *Andersen's LR-test*, *item-specific Wald test*, *Martin-Loef-Test*, *nonparametric Monte-Carlo Tests*, *análisis infit y outfit* con sus estadísticos de habilidad, la curva característica del ítem ICC y otras gráficas.

Itm: *Latent Trait Models under IRT*. Con las librerías: análisis multivariado para ítems dicotómicos y politómicos que se pueden utilizar para los modelos de la Teoría de Respuesta al ítem, el cual incluye Rash y el análisis logístico de dos parámetros. (R Core Team, 2021)

SPSS Por sus siglas en inglés *Statistical for the Social Sciencs*

SPSS significa Paquete Estadístico para las Ciencias Sociales. Es un *software* que se utiliza para el análisis de estadísticos de datos en las ciencias sociales porque presenta un conjunto de programas que permiten ejecutar, entre otros, procedimientos para aclarar las relaciones entre variables y realizar predicciones. La facilidad en el uso, su gran flexibilidad y escalabilidad hace a SPSS accesible para los usuarios porque permite mejorar la eficiencia y minimiza el riesgo. (Souto Grande, 2015).

2. Análisis de los datos

Se resguardó la información que el CIE proporcionó y las bases de datos generadas. Es decir, las bases de datos la Prueba diagnóstica para primer grado de primaria, el Manual de aplicación y calificación de uso para el docente del aula para asegurar la confidencialidad de los datos, por ello en el presente informe se dan a conocer las propiedades psicométricas de la prueba guardando el contenido específico de los ítems en las cinco subpruebas, entregándose la totalidad de la información a quien corresponde. Asimismo, es importante hacer notar que en los análisis no se revelan los nombres de los estudiantes que participaron en la aplicación piloto.

a. Análisis factorial exploratorio

Además, para el análisis factorial exploratorio se define:

$$X_j = a_{j1}F_1 + a_{j2}F_2 + \dots + a_{jk}F_k + u_jD_j$$

Donde, X_j es la variable observada. Los pesos o coeficientes a_{ij} expresan la contribución relativa de cada variable latente j sobre cada variable observada i . u_j es el conjunto de pesos. Entonces las puntuaciones típicas, se desarrollan ecuaciones similares a las de la regresión.

b. Análisis Rash

La función logística que modela la relación de los supuestos anteriores es la siguiente:

$$\ln(P_{is} / 1 - P_{is}) = (\theta_s - \beta_i)$$

Donde, $(P_{is} / 1 - P_{is})$ es la probabilidad de una respuesta correcta y la probabilidad de una respuesta incorrecta a un ítem es una de la diferencia en el atributo entre (θ_s) el nivel de la persona y (β_i) el nivel del ítem. Además, el logaritmo natural de $P_{is} / 1 - P_{is}$ indica que la dificultad del ítem es equivalente al nivel de competencia de la persona $(\theta_s - \beta_i = 0)$ (Prieto & Delgado, 2003) por tanto, se puede inferir que si la competencia o atributo del sujeto es mayor a la requerida por el ítem mayor es la probabilidad de acierto y viceversa.

Entonces, la predicción indica que la probabilidad de responder correctamente al ítem a partir de la diferencia en el atributo o competencia entre el $\theta_s =$ nivel de la persona y el $\beta_i =$ nivel del ítem se representa:

$$P_{is} = e^{(\theta_s - \beta_i)} / 1 + e^{(\theta_s - \beta_i)}$$

Donde e es la base de los logaritmos naturales (2,7183).

Embretson & Reise (2000), manifiestan que los valores de la escala de las personas y los ítems pueden usar distintas métricas, pero la más utilizada es la escala *logit* que es el algoritmo natural de $P_{is} / 1 - P_{is}$, en otras palabras, $\theta_s - \beta_i$. Como resulta arbitrario la localización del punto 0, en la tradición de Rash, se centra en la dificultad media de los ítems. Siendo su interpretación que los valores de θ_s mayores que 0 indican que las personas tienen una probabilidad superior a 0.5 de éxito en los ítems de dificultad media.

Aunque la escala *logit* puede adoptar valores entre más y menos infinito, la gran mayoría de los casos se sitúa en el rango ± 5 . Otros usuarios del modelo prefieren, considerando los objetivos y la muestra utilizada, situar el punto 0 en la habilidad media de las personas. Asimismo, la familiaridad con la distribución normal ha llevado a multiplicar por la constante 1.7 el exponente de la ecuación $P_{is} = e^{(\theta_s - \beta_i)} / 1 + e^{(\theta_s - \beta_i)}$ para asimilar la escala *logit* a aquella. En este caso, la media y la desviación típica de la escala son similares a las de las conocidas puntuaciones típicas z (0 y 1 respectivamente). Por tanto, la casi totalidad de los casos se incluye en el rango ± 3 (Prieto & Delgado, 2003, p. 2).

c. Parámetro de dificultad del ítem

La dificultad del ítem, en la TRI indica dónde se sitúa el ítem en la escala de aptitud, es por tanto un índice de la posición o localización del ítem en la escala de medida del raso o aptitud θ (Martínez, Hernández, & Hernández, 2006). Si no existen aciertos al azar, la probabilidad de acierto en el punto de inflexión de la curva es de 0.5.

En cuanto al índice de dificultad, en TCT se ha definido como la proporción de sujetos que aciertan el ítem. La interpretación se basa en que los valores más altos indican ítems más fáciles y los más bajos indican ítems más difíciles.

$$p_i = \frac{\sum_{s=1}^N X_{si}}{N} = \frac{A_i}{N}$$

En donde, $\sum_{s=1}^N X_{si} = A_i$ es el número de aciertos y N representa la cantidad de personas que han respondido al test. Ahora bien, el parámetro b_i = dificultad del ítem en la TRI se puede establecer la relación al convertir p_i a unidades de la distribución normal estandarizada, z_i de modo que p_i es el área por encima de z_i , puede entonces aproximarse el parámetro de dificultad del ítem de la siguiente manera:

$$b_i \cong \frac{z_i}{r_{bis}}$$

d. Parámetro de discriminación del ítem

El índice de discriminación basado en proporción de aciertos responde a los supuestos de la teoría clásica de los test (TCT). Kelley (1939) indica que el índice de discriminación más sensible y estable es aquel que se obtiene formando grupos extremos definidos por los sujetos que conforman el 27% superior y el 27% inferior del grupo total. Estos grupos se definen seleccionando aquellos sujetos que puntúan por encima del percentil 73 y por debajo del percentil 27, respectivamente. Dicho índice se define como:

$$Di = P_s - P_I$$

Donde, omitiendo el subíndice del ítem, P_s representa la proporción de aciertos al ítem en el grupo superior y P_I la proporción en el grupo inferior. Entonces el estadístico Di toma valores en el intervalo [-1, +1]. Valores positivos indican que el ítem discrimina a favor del grupo superior, mientras que valores negativos representan una discriminación invertida.

Al hacer referencia a la discriminación del ítem se lleva a pensar en el parámetro a_i de la Teoría de Respuesta al Ítem (TRI), en donde los principios del índice de discriminación se aproximan de la siguiente manera:

$$\alpha_i \cong \frac{r_{bis}}{\sqrt{1-r_{bis}^2}}$$

e. Coeficiente Kappa

El análisis del coeficiente Kappa se aplicó al proceso de validación entre jueces que se realizó con directores y docentes para obtener el acuerdo entre examinadores.

El método consiste en es obtener la confiabilidad entre examinadores. Para ello es necesario aplicar un instrumento con una población seleccionada y obtener la puntuación independiente para este. Posteriormente, indica Murat (1985), se verifica el grado de acuerdo que obtienen los jueces. El índice Kappa se aplica cuando las escalas son nominales y los estadísticos Kappa modificados para variables ordinales o intervalares.

La fórmula es:

$$k = \frac{F_c - F_a}{N - F_a}$$

Donde, F_c son las frecuencias de coincidencia o número de casos en los que las clasificaciones de ambos jueces coinciden. F_a son las frecuencias de azar, o número de casos en que cabe esperar que las clasificaciones de los jueces coinciden por mero azar, se obtiene con la sumatoria de los productos de los subtotales de cada categoría sobre el número de casos.

La máxima concordancia posible a $K = 1$, mientras que 0 corresponde a la concordancia observada por el puro azar (Murat, 1985).

3. Codificación de los datos

Es importante aclarar que la Prueba diagnóstica para primer grado de primaria se construyó por el CETT para uso en el aula, esto implica que el Manual de calificación expresa el proceso de calificación con un enfoque de evaluación formativa. Ante la

situación, de que el CIE requirió conocer la pertinencia del instrumento para usarse en un proceso de evaluación estandarizada el análisis de los datos para los 32 ítems se realizó con codificación dicotómica.

La base con que se procedió a realizar los análisis establecidos para el alcance de los objetivos requirió la recodificación de las variables que conformaron la base de datos. Primero se codificaron los datos del centro educativo y estudiante con un código alfanumérico para su identificación. Segundo la codificación dicotómica de las respuestas de los estudiantes requirió, la recodificación de las variables de respuesta de a 1 como presencia de respuesta esperada y 0 ausencia de respuesta esperada y 9 para los datos perdidos que responden a que el estudiante se quedó callado.

F. Instrumento para la recolección de información

Posicionar al estudiante que egresa del Nivel de Educación Preprimaria con respecto a lo que se espera que conozca y que sea capaz de hacer en el dominio (Ravela, 2006) de la lectura y escritura emergente, es el propósito de la Prueba diagnóstica de primer grado de primaria, por ello se puede concluir que este instrumento fue diseñado con referencia criterial.

El instrumento de evaluación está conformado por cinco subpruebas, las cuales están distribuidas en seis ejercicios que han sido nombrados por su constructor inicial. En el proceso de análisis de la prueba realizado para este estudio se identificó que el nombre de los ejercicios no corresponde a la propiedad de la lectura y escritura emergente que midió en realidad el ejercicio. En la Tabla 4 se muestra el análisis realizado y el nombre que corresponde a cada subprueba según lo definido por la teoría de la lectura y escritura emergente que en adelante se utilizará para nombrar a cada una.

Tabla 4. *Determinación del nombre que corresponde a las subpruebas que contiene la Prueba diagnóstica para primer grado de primaria*

N.º de ejercicio	Constructo	Nombre del ejercicio (subprueba) en la prueba diagnóstica	Propiedad de lectura y escritura emergente	Habilidad demostrada con	Forma de aplicación y momento
1	Lectura	Comprensión oral. (Conciencia fonémica)	Desarrollo de la conciencia fonológica	Identificar el sonido inicial	Grupal / momento 1
2	Escritura	Expresión escrita. (Dictado)	No aplica*	Escribir a mano, por medio de la transcripción al dictado.	Grupal / momento 1
3	Escritura	Expresión escrita (Utilización de elementos convencionales para escribir, caligrafía)	Desarrollo de la expresión escrita	La aplicación de las convenciones de escritura al expresar las ideas por escrito. Expresar por escrito y dibujar las ideas provocadas por el estímulo presentado.	Grupal / momento 1
4	Lectura	Comprensión lectora (Decodificación de palabras/ Relación grafema-fonema)	No aplica**	Decodificar palabras y frases familiares	Individual / momento 2
5	Lectura	Convenciones de la lectura	Desarrollo de la conciencia del lenguaje escrito	Manipular un megalibro para responder oralmente a las preguntas dadas.	Individual / momento 2
6	Lectura	Comprensión lectora	Comprensión oral	Después de escuchar una	Individual /

N.º de ejercicio	Constructo	Nombre del ejercicio (subprueba) en la prueba diagnóstica	Propiedad de lectura y escritura emergente	Habilidad demostrada con	Forma de aplicación y momento
				historia responde preguntas que miden habilidades de la escucha.	momento 2

*La habilidad de transcribir o codificar que se evalúa con el dictado forma parte de la lectura y escritura inicial. En el Nivel de Educación Preprimaria se desarrollan habilidades relacionadas con la conciencia fonológica y expresión escrita para que al escribir al dictado, el niño ponga en práctica la segmentación de sonidos, es decir segmenta «el habla en fonemas y trasforma cada uno de esos fonemas en su correspondiente grafema» Treiman (1998), citado en Montenegro, R., (2020).

**La habilidad de relacionar el grafema con el fonema para decodificar el lenguaje escrito se evalúa en lectura inicial bajo el sustento de los componentes de principio alfabético y decodificación. En el Nivel de Educación Preprimaria se desarrolla la comprensión oral, lenguaje oral, animación a la lectura y escritura, conciencia fonológica y conciencia del lenguaje escrito para que estas habilidades permitan la decodificación.

Desarrollo de la conciencia fonológica

La subprueba se aplicó el primer día de aplicación (momento 1) a todo el grupo del salón y representa el ejercicio 1. Consta de cinco ítems y posee un ítem de ejemplo. El aplicador indujo a los estudiantes para que por medio de una historia que contextualizó el ejercicio el estudiante pudiera identificar la imagen que representa el sonido inicial que se preguntó. El estudiante marcó la respuesta por medio de un círculo o marca que hizo sobre la imagen elegida como correcta.

Transcripción al dictado

Esta subprueba forma parte del constructo de escritura y representa el ejercicio 2 de la prueba. Se aplicó el primer día de aplicación (momento 1) a todo el grupo del salón.

Antes del dictado que forma parte de la prueba, se dictó una palabra como ejemplo, la cual se escribió en el pizarrón para que el estudiante la copiara. Después de este ejemplo, el aplicador dictó cuatro palabras, dos de dos sílabas, una con tres sílabas y una última con 4 sílabas, todas con estructura consonante vocal.

Desarrollo de la expresión escrita

Esta subprueba es el segundo ejercicio del constructo de escritura y corresponde al ejercicio 3 de la prueba. Se aplicó el primer día de aplicación (momento 1) a todo el grupo del salón. El aplicador solicitó a los estudiantes que pensarán en una situación específica e hizo dos preguntas, indicó que podían conversar sus ideas con sus compañeros, después de un tiempo solicitó que escribieran y dibujaran sus ideas. Con este ejercicio se finalizó el primer día de aplicación.

Decodificación

La subprueba se aplicó el segundo día de aplicación (momento 2) de forma individual y representa el ejercicio 4. El aplicador mostró 7 tarjetas, cinco de ellas con una palabra y dos de ellas con una frase corta. Conforme las mostró, solicitó al estudiante que colocará el dedo para indicar lo que estaba leyendo.

Desarrollo de la conciencia del lenguaje escrito

La subprueba se aplicó el segundo día de aplicación (momento 2) de forma individual y representa el ejercicio 5 de la prueba. El aplicador hizo uso de un megalibro¹ que proporcionó al estudiante para que lo manipulará con la intención de que le mostrará seis aspectos que podría observar, identificar y mostrar en el megalibro que forma parte de las convenciones de la lectura.

Comprensión oral

La subprueba se aplicó el segundo día de aplicación (momento 2) de forma individual y representa el ejercicio 6 de la prueba. Aunque en el Manual de aplicación y calificación

¹ Libro de tamaño grande que permite apreciar sus letras e imágenes desde una distancia prudente por un grupo de estudiantes. Este megalibro forma parte de la colección *El tesoro de la lectura*, etapa emergente, publicado por la Digeduca en 2014.

de la prueba diagnóstica para primer grado de primaria se sugirió como objetivo «evaluar la comprensión lectora, aunque no se espera que el estudiante sea capaz de leer el texto por sí solo» (CETT-CA-RD, S.f.), se procedió a evaluar lo que el currículo Nacional Base del Nivel de Educación Preprimaria indica en la competencia 1 del área de Comunicación y Lenguaje para seis años el niño debe evidenciar que «responde con gestos, movimientos y oralmente a mensajes escuchados en poemas, cuentos y textos diversos de su cultura y de otras culturas» (p. 182), en otras palabras la aplicación de la sección se realizó leyéndole al estudiante una historia, al finalizar la lectura el aplicador procedió a hacerle cinco preguntas.

El tiempo promedio de aplicación en el día uno, donde se dio de forma grupal fue de 30 minutos y en el día dos, de manera individual, el tiempo promedio fue de 20 minutos por estudiante, logrando aplicar hasta un promedio de 9 estudiantes durante una jornada de clases.

1. Contextualización del manual de aplicación

Es de conocimiento que la aplicación estandarizada del instrumento es fundamental para asegurar que los datos obtenidos en campo son (Martínez, Moreno, & Muñiz, 2005) válidos y confiables. Se realizó un análisis cualitativo después de la primera aplicación para modificar el formato y datos indicados en el Manual. Por ser un documento confidencial no se adjunta al presente informe. En la tabla se evidencia de forma general las modificaciones realizadas al Manual de aplicación.

Tabla 5. Modificaciones al Manual de aplicador

No.	Observación	Modificación
1	Información innecesaria para el aplicador	Se seleccionó la información pertinente para el Manual del aplicador.
2	Falta de estructura	Organización de la información por modalidad de aplicación.
3	Falta de información	Se agregaron datos necesarios para realizar la aplicación e interpretación de las respuestas dadas por el estudiante.

Considerando que el aplicador de la prueba, en Guatemala, puede ser un agente externo al salón de clase del estudiante que se evaluara se procedió a cambiar el nombre del Manual de aplicación del docente por Manual de aplicación de la Prueba diagnóstica de lectura y escritura para el primer grado de primaria. Asimismo, se realizaron las modificaciones pertinentes de estructura y formato para que fuera funcional en el proceso de aplicación. Los cambios se sustentan en la experiencia vivida en campo al momento de aplicar la prueba diagnóstica para primero primaria.

G. Procedimientos

La metodología implementada para el cumplimiento del objetivo general contempló actividades como:

- El análisis del Currículo Nacional Base del Nivel de Educación Preprimaria para la Etapa 6 años, la teoría existente sobre la lectura y escritura emergente para comparar el instrumento elaborado por el CETT-CA-RD y lo establecido tanto en el Currículo Nacional Base que rige la educación en el contexto guatemalteco y la definición del constructo que mide la prueba. Producto de lo descrito fue una Tabla de Alcance y secuencia que refleja la pertinencia del instrumento para el contexto guatemalteco respecto al diseño de la prueba.

- Se realizó el proceso de juicio de expertos con la finalidad de obtener la validez de contenido del instrumento, los análisis de este proceso se obtuvieron con el coeficiente Kappa, calculado con el software SPSS. El proceso llevó dos momentos: a) juicio con directores de escuela y b) juicio con docentes del Nivel de Educación Preprimaria.

En la validación con directores participaron dos profesionales que poseen conocimiento en lectura emergente, con un rango entre 16 y 19 años de experiencia en el puesto de dirección del centro educativo que administran y como mediadoras pedagógicas. La validación con docentes se realizó con tres educadoras del Nivel de Educación Preprimaria que poseen entre 13 y 20 años de experiencia como docentes del nivel. Se elaboró una escala tipo Likert para obtener la validez inter-examinadores, (Tornimbeni, Pérez, & Olaz, 2008) el cual se utilizó para obtener la puntuación total asignada por los jueces a una escala donde todos los ítems se califican con un componente de subjetividad.

- Para el proceso de análisis de las propiedades de los ítems desde la Teoría de Respuesta al Ítem (TRI) se llevó a cabo el proceso de aplicación piloto de la prueba. Para ello, se visitó en dos momentos a cada uno de los establecimientos participantes que fueron descritos en párrafos anteriores en este mismo informe. La aplicación grupal fue administrada el primer día de evaluación y los ejercicios de aplicación individual se realizaron en el segundo día. Previamente a la aplicación individual se seleccionó una lectura² para alcanzar el objetivo de los ejercicios cinco y seis de la Prueba diagnóstica para primer grado de primaria.

Al tener los datos recogidos, se construyó una matriz de vaciado en la cual se trasladaron los datos brutos obtenidos en las pruebas aplicadas. A partir de esta, se generó la base de datos para ser analizada en los softwares estadísticos indicados anteriormente. Luego, se hizo un análisis de los valores de cada una de las variables que componen la

² Se utilizó el megalibro *El Tesoro de la lectura para el ejercicio 5 y 6*. La lectura se titula “Ana es mi amiga.” El texto tiene 52 palabras, con alta frecuencia de las vocales a, o, e, i, u; así como de las consonantes y, n, l, s, m.

base con el fin de identificar linealidad en las correlaciones o y casos atípicos para asegurar la fiabilidad de los resultados.

Para la construcción de la base de datos se usó Excel. Los análisis con el software r permitió obtener los parámetros de discriminación y dificultad del ítem correspondientes al modelo 2 pl. Así como datos del modelo Rash para el análisis de independencia local, asimismo, se obtuvo el análisis de dimensionalidad por medio del análisis factorial exploratorio para conocer los datos de los supuestos de la TRI y así proceder a aplicar los modelos corridos con la prueba.

VII. PRESENTACIÓN DE RESULTADOS

Los resultados obtenidos con el instrumento diseñado por el CETT «le permitirá, al usuario, planificar y desarrollar situaciones de aprendizaje que correspondan a las necesidades detectadas y ofrecer el apoyo adecuado» (CETT-CA-RD, S.f.). Con base en lo anterior, la Prueba diagnóstica para primer grado de primaria es una evaluación referida a criterio, como se ha indicado anteriormente. Diseñada con el propósito de proveer información al docente sobre algunos saberes que poseen los estudiantes con respecto a la lectura y la escritura emergente al comienzo del año escolar de primer grado. Cabe resaltar que el uso que se le desea dar a la prueba es para una aplicación estandarizada y externa por ello, la pertinencia de la ejecución del presente estudio y de sus resultados.

A. Tabla de alcance y secuencia

Se procedió a equiparar, bajo lineamientos de análisis cualitativo de la tabla de especificaciones propuesta por el CETT con el Currículo Nacional Base (CNB) del Nivel de Educación Preprimaria. Como se ha mencionado en líneas anteriores, la prueba está diseñada para ser aplicada en los primeros meses del ciclo escolar del primer grado de primaria por ello, se procedió a relacionar las competencias de la prueba con las competencias de lectura y escritura emergente establecidas en el CNB.

En la tabla de especificaciones que se construyó se evidencia la competencia y subcompetencia establecida por el CETT, además el estándar y la competencia del CNB, asimismo, la posición del ítem en la prueba y lo que mide el ítem. Es necesario aclarar que por ser información confidencial y de propiedad del CIE en la siguiente tabla se eliminaron datos específicos de los ítems.

Tabla 6. *Tabla de alcance y secuencia para equiparlo al contexto guatemalteco*

No.	Constructo	CETT		en Nacional Base, Etapa 6 años, Nivel de Educación Primaria			Prueba diagnóstica			
		Competencia	Subcompetencia	Contrastando con la teoría existente	Área de aprendizaje	Estándar (CNB)	Competencia (CNB)	No de ítem	lo que mide el ítem	
1	Lectura emergente	Comprensión oral	Conciencia fonológica	Desarrollo de la conciencia fonológica		7. Comprende que un sonido o fonema está representado por una letra o grafía, que a su vez, se combina con otros para formar palabras.	2. Utiliza nociones de la estructura de las palabras al expresar sus ideas	1	Sonido inicial	
2								2	Sonido inicial	
3								3	Sonido inicial	
4								4	Sonido inicial	
5								5	Sonido inicial	
6								12	Lectura de palabra	
7		13	Lectura de palabra							
8		14	Lectura de palabra							
9		Comprensión lectora	Decodificación de palabras/ grafema-fonema	Decodificación (lectura inicial)	No aplica	No aplica		3. utiliza estrategias para comunicarse oralmente	15	Lectura de frase
10									16	Lectura de palabra
11									17	Lectura de frase
12									18	Lectura de palabra
13									19	Con impreso 1inst
14									20	Con impreso 1inst
15			21	Con impreso 1inst						
16			22	Con impreso 3inst						
17			23	Con impreso 1inst						
18			24	Conc impreso 2inst						
19			25	Conc impreso 1inst						
20	Comprensión analítica			Comprensión oral	Comprensión oral		5. Escucha comprensivamente cinco libros recreativos, apropiados a su nivel de lectura por año.		26	Conc impreso 2inst
21									27	Interpretar
22		28							Entender y recordar	
23		29							Entender y recordar	
24		30							Localizar sucesos	
25		31							Localizar sucesos	
26		32							Localizar sucesos	
27	Escritura emergente	Expresión escrita	Dictado	Transcripción al dictado (escritura inicial)	Destrezas de aprendizaje	No aplica	No aplica	6	Dictado 1	
28								7	Dictado 2	
29								8	Dictado 3	
30		9	Dictado 4							
31		Caligrafía	Utilización de elementos convencionales para escribir	Expresión escrita	Destrezas de aprendizaje	2. Dibuja líneas y formas, siguiendo trazos.			10	Convenciones para la escritura
									11	Dibuja y escribe oraciones

B. Juicio de expertos

La validación de la prueba con juicio de expertos (externos) arrojó que la proporción de acuerdo entre ellos con respecto a la suficiencia, coherencia, claridad y relevancia, de cada una de las secciones que contiene la prueba posee una fuerza de concordancia de .75. De manera análoga con la validación con docentes (internos) se obtuvieron acuerdos de .58, lo cual determina que la fuerza de concordancia es moderada. En la tabla se muestran los porcentajes de acuerdo entre jueces.

Tabla 7. Fuerza de concordancia entre juicio de expertos y validación con docentes

FUERZA DE CONCORDANCIA		
JUICIO DE EXPERTOS		
	Valor k	Interpretación
EXTERNOS	0.75	Buena
INTERNOS	0.58	Moderada

C. Comprobación de los supuestos de la TRI

Los resultados que se presentan son análisis que se efectuaron bajo los criterios de la Teoría de Respuesta al Ítem (TRI). Primero se procedió a comprobar los dos supuestos de la TRI, es decir la dimensionalidad de la prueba y la independencia local. Luego, los parámetros de los ítems obtenidos se realizaron con el Modelo Rash y el Modelo de 2 parámetros (2PI): parámetros de discriminación (a_i) y dificultad (b_i) de los ítems, así como de las curvas características de los ítems y la función de información del ítem y del test con la intención de conocer la confiabilidad de la prueba.

1. Dimensionalidad de la prueba e independencia local

Se procedió a realizar el análisis factorial exploratorio para conocer la consistencia interna del modelo teórico que describe la Prueba diagnóstica para primer grado de primaria. La finalidad, entonces, fue comprobar la dimensionalidad de la prueba por medio de los factores que se agrupan como parte del proceso de comprobación de uno de los supuestos de la TRI.

Se obtuvo la matriz de correlaciones para determinar si el análisis factorial es apropiado. Dado que correlaciones cercanas a uno o menos uno indicarían que no procede hacer un análisis factorial. Se obtuvo en la observación de la matriz de correlación (ver Anexos), que: existe posible colinealidad (correlación de 0.8 a 0.86), entre el ítem 15 con los seis ítems restantes que conforman los ejercicios de Decodificación, es decir, los ítems 12, 13, 14, 16, 17 y 18. Lo que implica que son potencialmente redundantes, por tanto,

existe la posibilidad de eliminarse o unirse con las otras variables con las que tiene alta correlación.

El siguiente paso fue eliminar del análisis los ítems que están indicados con colinealidad, observándose en la segunda corrida que no existe en la matriz de correlación valores cercanos a 1. Por tanto, se procedió a realizar la prueba KMO, con 25 variables o ítems donde se obtuvo el valor de 0.81 siendo estadísticamente significativo, indicando que es meritorio porque está en el rango 0.8 a 0.9 además, la prueba de esfericidad de Barlett, aseguró que la matriz de correlaciones no es una matriz identidad, observándose un p-value de $4.94e^{-324}$ con un 99 % de confianza que es factible hacer una reducción dimensional. Luego, el determinante de la matriz de correlación indicó el valor 0.002 y como es distinto de cero se interpreta que las variables no son linealmente dependientes, por tanto, es factible proceder con el análisis factorial exploratorio para los ítems de la prueba.

Al correr el análisis factorial exploratorio se encontró que cinco métodos tienden por usar un factor y otro grupo de cuatro factores aconsejan generar cuatro factores. Se decidió proceder con el análisis de ambas propuestas.

Tabla 8. Índices de ajuste para un factor

<i>Factores</i>	índice	Valor	Intervalo de confianza
<i>1</i>	Tucker Lewis Index (índice de confiabilidad del ajuste del modelo)	0.685	---
	RMSEA	0.099	90%
<i>4</i>	Tucker Lewis Index (índice de confiabilidad del ajuste del modelo)	0.831	---
	RMSEA	0.062	90%

La interpretación indica que el índice de Tucker Lewis para considerarse un buen ajuste del modelo requiere estar cerca de 1, con un valor hasta 0.9 para considerar un ajuste

aceptable y el Error Cuadrático Medio de Aproximación (RMSEA) indica que para aceptar la H_0 el valor debe ser inferior a 0.08.

En la Tabla 9 se evidencia la carga factorial de la prueba con cuatro factores.

Tabla 9. Cuatro factores agrupados para la Prueba diagnóstica

	Factor 1	Carga	Factor 2	Carga	Factor 3	Carga	Factor 4	Carga
Ítem	6	0.61	30	0.34	6	0.61	21	0.68
	9	0.6	31	0.8	9	0.60	20	0.65
	8	0.38	32	0.73	26	0.50	22	0.63
	7	0.37	29	0.36	23	0.57		
	5	0.83	28	0.33	8	0.57		
	4	0.79			7	0.54		
	19	0.5			11	0.52		
	2	0.34			24	0.45		

El análisis factorial dejó sin agrupar a los ítems 27, 25, 3, 10, 1. Se resalta que el ítem 6 y 9 se encuentran tanto en el factor 1 como 3. En donde los factores que se agrupan son los ejercicios que miden el componente de Desarrollo de la conciencia fonológica (2, 4, 5) transcripción al dictado (6, 7, 8, 9) y Desarrollo de la conciencia del lenguaje escrito (19). El factor 2 agrupa ítems del componente de comprensión oral 28, 29, 30, 31, 32). El factor 3 agrupa ítems del componente de transcripción al dictado (6, 7, 8, 9), de Desarrollo de la conciencia del lenguaje escrito (23, 24 y 26) y decodificación (11) y finalmente, el factor 4 agrupa al componente de Desarrollo de la conciencia del lenguaje escrito (20, 21 y 22).

2. Parámetros de los ítems

Bajo el modelo de dos parámetros se obtuvo la discriminación y dificultad de los ítems de la prueba. Inicialmente se corrió el modelo 2pl con los 32 ítems observándose

que el 44 % de los ítems discrimina de manera óptima (>2), el 47 % posee discriminación moderada es decir tienen valores entre 0.64 y 1, y el 9 % de los ítems discriminan mal.

En cuanto a la dificultad el 100 % de los ítems son fáciles, donde el más fácil tiene el valor de -7.5 (ítem 1) y el ítem con mayor dificultad en la escala es de 0.45 (ítem 11). El primer ítem corresponde al conjunto de ítems que miden el desarrollo de la conciencia fonológica y el último al ejercicio de escritura, específicamente expresión escrita.

En esta misma tabla se presenta, de acuerdo con el análisis factorial exploratorio, una segunda corrida del modelo 2pl, en esta ocasión sin los ítems de Decodificación por presentar alta correlación en la matriz de correlaciones. Se obtuvieron los resultados siguientes: el 36 % tiene óptima discriminación, el 15 % discrimina moderadamente y solo el 4 % discrimina bajo. En cuanto a la dificultad los resultados no cambian, solo el valor, es decir, el ítem 1 continúa siendo el más fácil de la prueba y el 11 el más difícil pero dentro del rango de la dificultad media.

Tabla 10. Resultados 2pl

No.	Dificultad	Discriminación	No.	Dificultad	Discriminación
<i>Item_1</i>	-7.593	0.52	<i>Item_1</i>	-8.82	0.45
<i>Item_2</i>	-2.449	1.53	<i>Item_2</i>	-2.07	2.22
<i>Item_3</i>	-3.758	0.972	<i>Item_3</i>	-3.58	1.06
<i>Item_4</i>	-1.503	3.864	<i>Item_4</i>	-1.38	23.88
<i>Item_5</i>	-1.5	4.308	<i>Item_5</i>	-1.39	24.83
<i>Item_6</i>	-0.881	2.444	<i>Item_6</i>	-1.02	2.12
<i>Item_7</i>	-0.898	2.305	<i>Item_7</i>	-1.03	2.09
<i>Item_8</i>	-0.924	2.135	<i>Item_8</i>	-1.05	1.99
<i>Item_9</i>	-0.875	2.641	<i>Item_9</i>	-0.96	2.73
<i>Item_10</i>	-3.245	1.501	<i>Item_10</i>	-3.20	1.57
<i>Item_11</i>	0.446	2.267	<i>Item_11</i>	0.32	1.94
<i>Item_12</i>	-0.517	6.237			
<i>Item_13</i>	-0.105	16.139			
<i>Item_14</i>	-0.337	6.157			
<i>Item_15</i>	-0.082	17.473			
<i>Item_16</i>	-0.416	4.075			
<i>Item_17</i>	-0.234	5.01			
<i>Item_18</i>	-0.324	5.464			

No.	Dificultad	Discriminación	No.	Dificultad	Discriminación
<i>Item_19</i>	-2.451	1.424	<i>Item_19</i>	-1.97	2.28
<i>Item_20</i>	-2.558	0.956	<i>Item_20</i>	-2.03	1.40
<i>Item_21</i>	-1.894	0.935	<i>Item_21</i>	-1.52	1.40
<i>Item_22</i>	-2.926	0.557	<i>Item_22</i>	-2.05	0.88
<i>Item_23</i>	-1.301	1.327	<i>Item_23</i>	-1.33	1.45
<i>Item_24</i>	-1.965	1.251	<i>Item_24</i>	-1.80	1.56
<i>Item_25</i>	-2.331	0.949	<i>Item_25</i>	-1.82	1.45
<i>Item_26</i>	-0.893	1.288	<i>Item_26</i>	-0.94	1.43
<i>Item_27</i>	-0.698	1.057	<i>Item_27</i>	-0.74	1.21
<i>Item_28</i>	-1.628	1.12	<i>Item_28</i>	-1.54	1.33
<i>Item_29</i>	-1.294	1.133	<i>Item_29</i>	-1.19	1.46
<i>Item_30</i>	-1.48	0.66	<i>Item_30</i>	-1.11	1.05
<i>Item_31</i>	-1.517	0.669	<i>Item_31</i>	-1.22	0.96
<i>Item_32</i>	-0.984	0.5	<i>Item_32</i>	-0.72	0.82

a. Curvas características del ítem

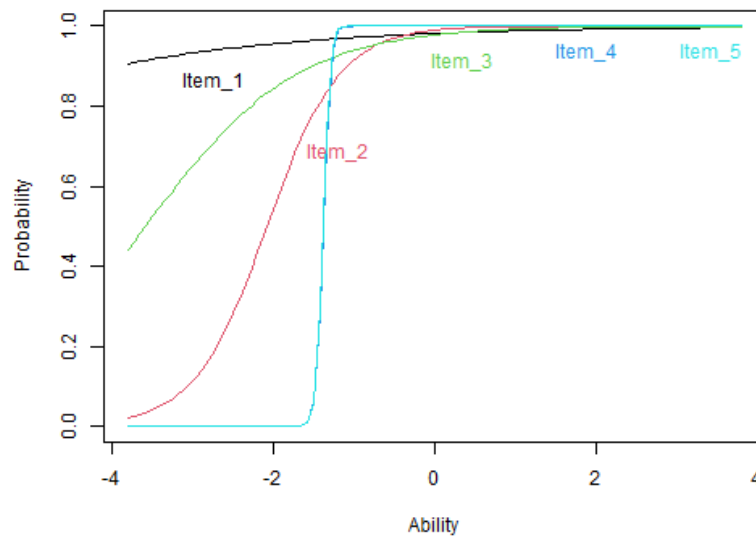
Se procedió a obtener las Curvas Características del Ítem CCI, agrupadas por ejercicio dentro de la prueba para una mejor interpretación de la relación funcional entre los valores de la variable que miden los ítems y la probabilidad de acertar estos, y de ahí que un objetivo de la TRI sea establecer la mejor función que ajuste esta relación (Chacón Moscoso & Pérez-Gil, 2008), es decir, una función que dé cuenta de la relación entre la probabilidad de acertar el ítem con la localización en el rasgo de los sujetos; en concreto, esa relación puede ser expresada mediante una función de regresión no lineal, que involucra los parámetros del ítem, que une cada valor en el rasgo con la puntuación medida condicionada en el ítem, que, en el caso de ítems dicotómicos, coincide con la probabilidad condicionada al nivel de θ de acertar el ítem.

Desarrollo de la conciencia fonológica

Se aprecia que el ítem 1 no discrimina correctamente y que es el ítem más fácil del conjunto observado, mientras que los ítems 2, 4 y 5 poseen discriminación óptima y dificultad fácil. El comportamiento de la curva no es adecuado, es decir lo que mide el ítem y la probabilidad de acierto no tiene una relación precisa.

La subprueba se compone de 5 ítems que miden el sonido inicial de palabras que escuchan para identificar la imagen que las representa.

Gráfica 1. CCI de 5 ítems de conciencia fonológica

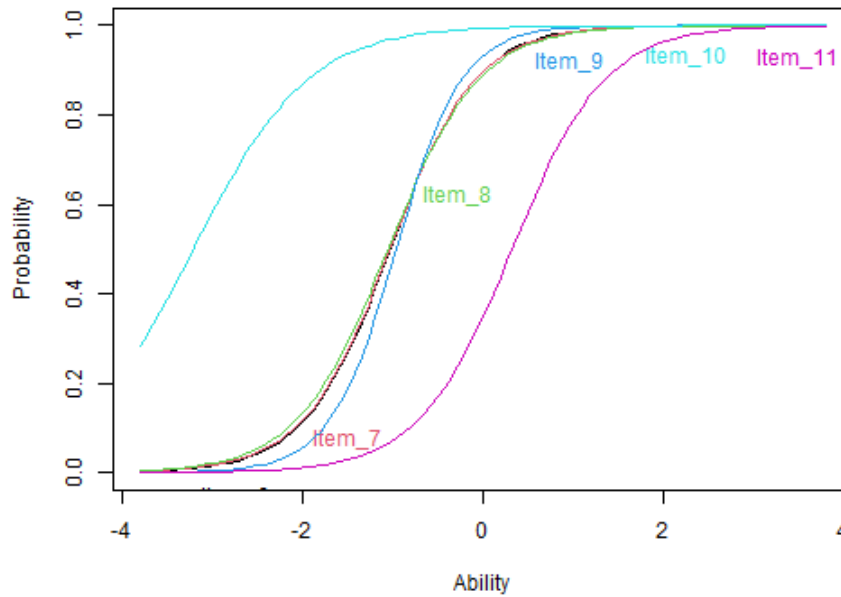


Escritura

Del conjunto observado, tres de sus ítems discriminan de forma óptima (6,7 y 9) y los otros tres, tienen discriminación moderada. Además, la dificultad de esta subprueba se ubica a la izquierda principalmente, en un rango de -3 a 0.4. El ítem más fácil de esta subprueba mide las convenciones de la escritura bajo el constructo de expresión escrita y el ítem más difícil el dibujo y la escritura de oraciones.

En la gráfica se observa su comportamiento en cuanto a lo que mide el ítem y la probabilidad de acierto, en donde el ítem 10, es que no refleja relación clara entre los criterios del ítem analizados.

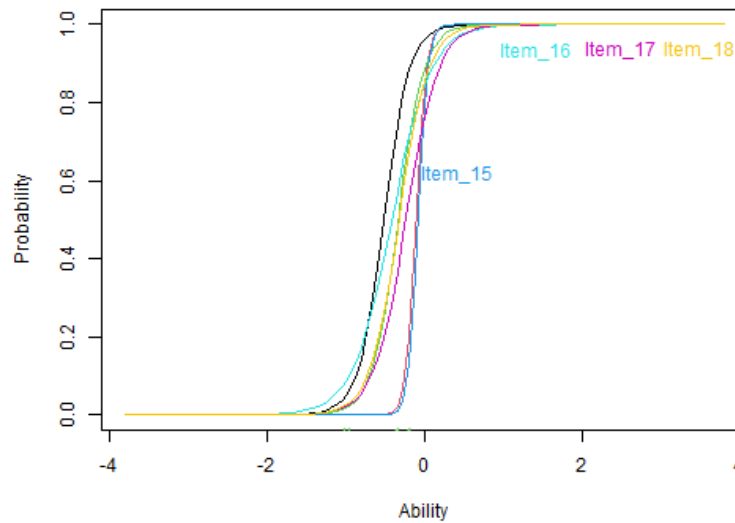
Gráfica 2. CCI de 6 ítems de escritura emergente e inicial



Decodificación

La subprueba se compone de 6 ítems que miden Decodificación posee óptimas medidas de discriminación, el rango de los valores está en >2 y sus dificultades oscilan de -0.5 a -0.09 . Son ítems fáciles, cercanos a 0 pero con curvas poco definidas para asegurar su relación clara entre la probabilidad de acierto y lo que mide el ítem.

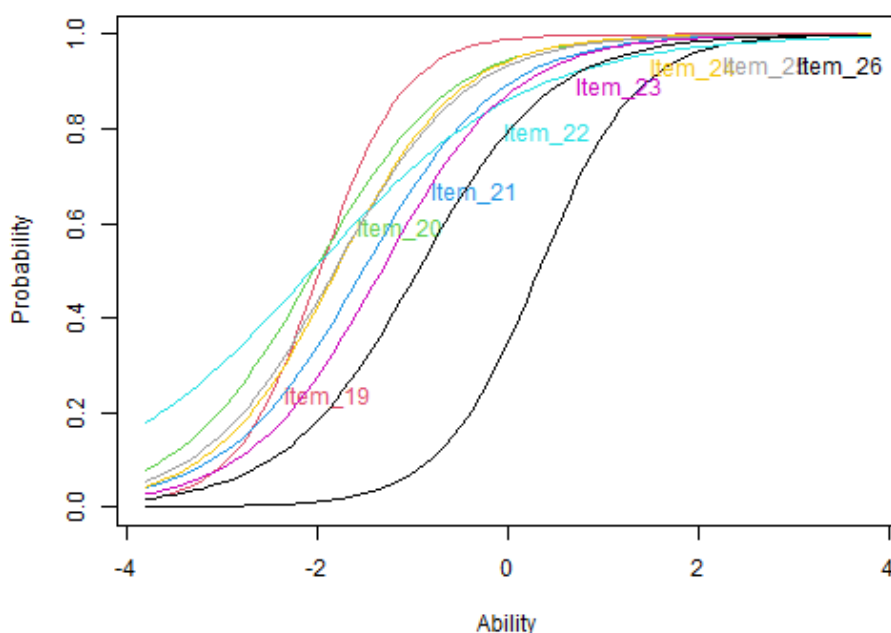
Gráfica 3. CCI de 7 ítems de decodificación



Desarrollo de la conciencia del lenguaje escrito

La Gráfica 4 ilustra ocho ítems que son el conjunto que mide las convenciones de la lectura o mejor conocido como los conceptos impresos. La discriminación de los ítems es buena, resaltando al ítem 19 que posee optima discriminación y el ítem 22 que tiene discriminación cercana a 1. En relación con la dificultad es una sección con ítems de dificultad baja, con datos que se encuentran en -2.0 y -0.9. El comportamiento de la curva, en general, es adecuado.

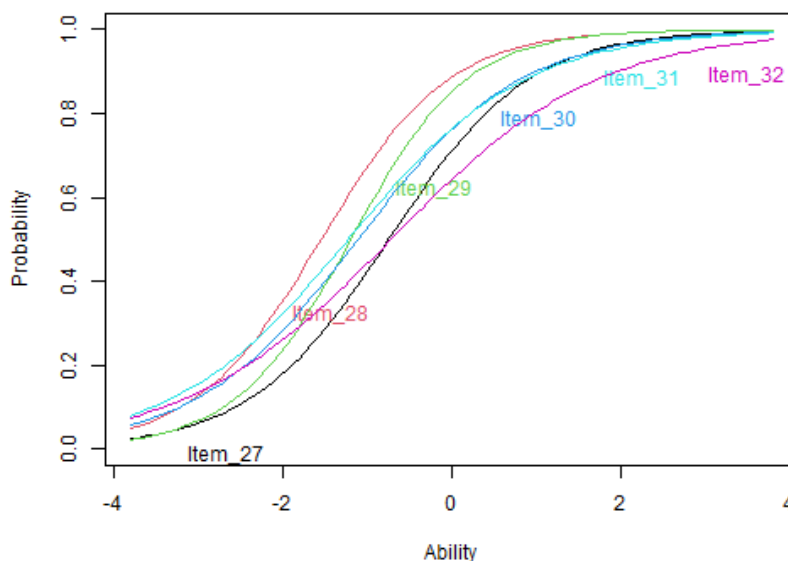
Gráfica 4. CCI de 8 ítems de desarrollo de la conciencia del lenguaje escrito



Comprensión oral

Del conjunto observado de ítems que miden comprensión oral desde las habilidades de la escucha de interpretar, entender y recordar y localizar sucesos, se resalta que la discriminación de todos es moderada. Mientras que en el tema de la dificultad se encuentran en un rango de -0.7 a -1.5, todos se encuentran en la parte fácil de la escala. El comportamiento de la curva es adecuado.

Gráfica 5. CCI de 6 ítems de comprensión oral



b. Ajuste de las respuestas a los ítems

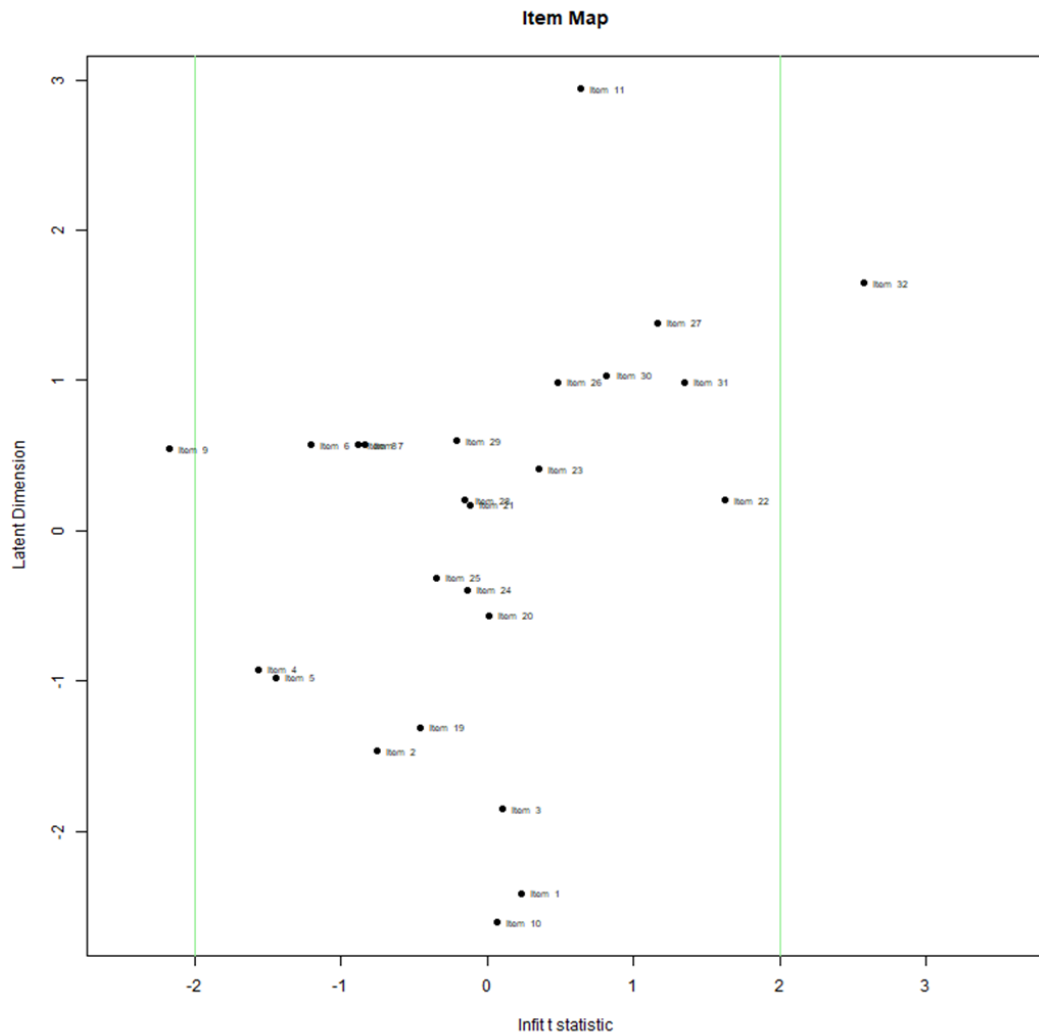
Se procedió a realizar el análisis Rash para obtener las medidas de ajuste de la media cuadrática ponderada de residuos (infit) y la suma de cuadrados de los residuales estandarizados (outfit) siendo la medida de la persona. Esto con la intención de conocer el ajuste de los ítems de la prueba y de las respuestas de los sujetos, como ajuste externo.

En este ejercicio se aplicó el modelo 1pl al conjunto de 32 ítems y al conjunto de 25 ítems, de acuerdo con las correlaciones observadas al momento de hacer el análisis factorial confirmatorio. Los resultados en la primera aplicación del modelo arrojan como resultado de la comparación de los ajustes uno a uno, que los ítems que miden decodificación no se ajustan al valor esperado de infit (rango de 0.6 – 0.5), ni a la medida del outfit (rango de -5 a -3). Además, que solo el 62 % de los ítems se encuentran dentro de los rangos de ajuste.

En la Gráfica 6 se encuentra la ilustración correspondiente a las medidas de ajuste con el modelo aplicado al conjunto de 25 ítems. Los ítems 6, 7, 8, 9, 10, 11, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 y 32 se ajustan a los parámetros del rango 0.7 - 1.3

publicados por Wright y Linacre, (1994), los cuales indican que existe buen ajuste de los ítems con respecto a la medida del modelo Rash, ahora bien, en cuanto al ajuste de las respuestas de los sujetos se encuentran los ítems 2, 3, 6, 7, 8, 10 a 31 con ajuste en el rango de -1.96 a 1.96 lo cual indica que tienen buen ajuste con respecto a la prueba.

Gráfica 6. Mapa de ítems (Infit y Outfi)



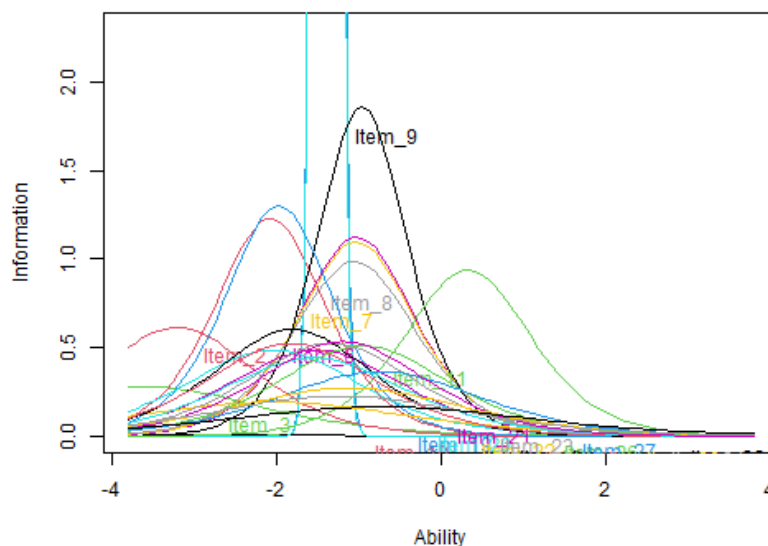
Se observó que la medida de ajuste infit es de 0.98 en la prueba por tanto el valor esperado se ajusta y el segundo análisis presentado que responde a la medida de las respuestas inesperadas de las personas también presenta un buen ajuste, es decir la prueba se ajusta al modelo y las respuestas obtenidas no provocan interpretaciones equivocadas.

c. Función de información de los ítems y del test

Posteriormente a ello, se trabajó el análisis de la función de información de los ítems y del test, siendo los resultados obtenidos los que se evidencian en las siguientes gráficas.

En la Gráfica 7 se observa que el conjunto de 25 ítems analizados arroja información de ajuste del test. Es posible ver que el comportamiento de los ítems de conciencia fonológica resalta en el test, así como el ítem de escritura que mide las convenciones de la escritura.

Gráfica 7. *Función de información del Ítem y del test*

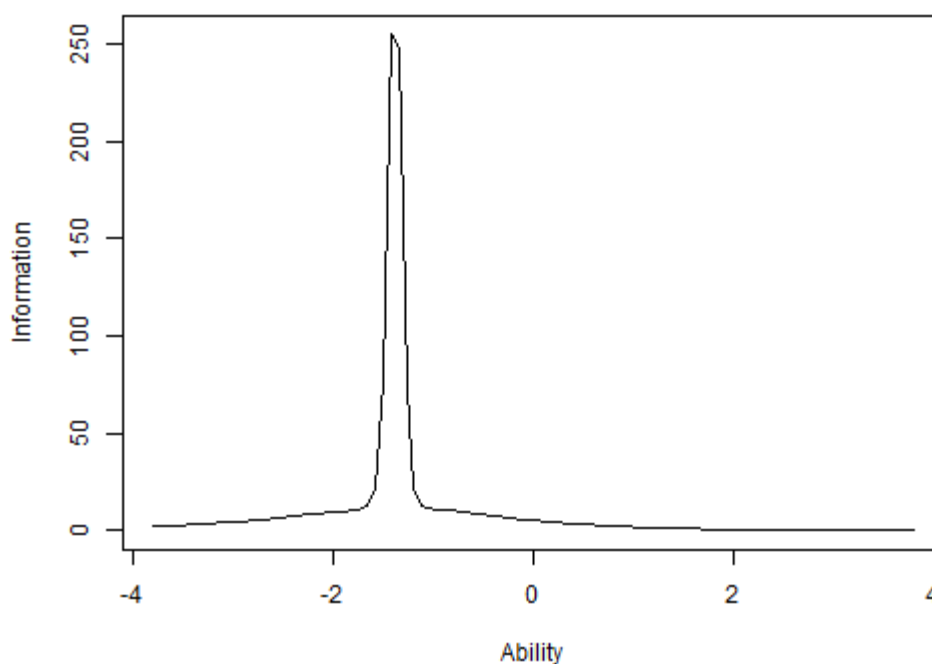


Analizando la curva de Función de información del test, puede apreciarse que la Prueba diagnóstica para primer grado de primaria proporciona su máxima información cuando la habilidad de los individuos tiende a ser menor a 0. En términos teóricos, esta curva no se aproxima a una curva esperable, es decir, una curva simétrica respecto de $\theta = 0$ y una habilidad entre -3 y 3. Sin embargo, el rango de valores en que fluctúa b sugiere que los reactivos poseen una dificultad mínima para ser respondida con acierto. En donde el parámetro de dificultad (b) fluctúa entre -8.81 y 0.32.

d. Ajuste del test

Ahora bien, con el afán de tener la interpretación del ajuste del test para identificar su confiabilidad en la medida se analizaron los distintos parámetros del ítem y del test. En la Gráfica 8 se observa que el ajuste que se encuentra en los valores que oscilan en un rango de habilidad cercano a -1.7 lo cual indica que la información del test se encuentra principalmente en este valor. No es una curva esperada la que arroja el ajuste porque es posible que existan datos atípicos, pero la prueba es consistente.

Gráfica 8. *Ajuste del test*



VIII. DISCUSIÓN DE LOS RESULTADOS / HALLAZGOS

A. Tabla de alcance y secuencia

Se ha determinado que la prueba diagnóstica para primer grado de primaria mide el constructo de lectura y escritura tanto en la etapa emergente como inicial. El peso de la etapa inicial en la prueba es del 29% en los constructos de lectura y escritura.

Desde el concepto teórico de la construcción de la prueba es posible observar que presenta ítems fáciles, medios y difíciles. Es decir, en la subprueba de desarrollo de la conciencia fonológica, se miden sonidos iniciales de vocales y de consonantes, observándose la gradualidad del aprendizaje, en el CNB del nivel sí se encuentran ejercicios que benefician de manera motora y cognitivamente el proceso del aprendizaje de este constructo, pero lo que se promueve es la identificación del sonido de palabras por sílabas y no por unidad de medida.

La decodificación, no se contempla en el CNB analizado, observándose coherencia con lo indicado en la teoría de los aprendizajes de la lectura emergente, ya que este es un elemento que se desarrolla en primer grado de primaria. Por tanto, desde el análisis cualitativo, esta sección es difícil para el estudiante que egresa del Nivel de Educación Preprimaria. Además, los ítems se observan con dificultad baja, media y alta ya que el ejercicio posee ítems donde se solicita decodificar palabras y frases de hasta dos palabras de uso frecuente en un contexto familiar y escolar.

El conocimiento de la conciencia del lenguaje escrito es un aprendizaje indispensable en la etapa emergente de la lectura y escritura. Se observa adecuada la inclusión de la medida de este elemento en la prueba. Se determinó en el análisis que tiene como producto la tabla de alcance y secuencia que hay ítems con dificultad baja, media y alta por las distintas tareas que se solicitan, pero también, se observó que tres de los 7 ítems tienen

hasta tres instrucciones que el estudiante debe ejecutar para llegar a la respuesta esperada. Situación que rompe con el criterio de medir un conocimiento claro, concreto y preciso por ítem.

Los ejercicios de comprensión oral tienen linealidad con lo establecido en el CNB del nivel, desde la concepción de la expresión oral en la competencia comunicativa, situación que es coherente con la definición del continuo en el aprendizaje del lenguaje oral, comunicación oral y comprensión oral. Siendo estas tres, tres elementos distintos con sus propias características, pero ampliamente dependientes para el logro de la competencia comunicativa en el estudiante. Al realizar el análisis se identificó que los ejercicios tienen ítems que miden tres de las cuatro habilidades de la escucha.

Finalmente, en la prueba se evalúa escritura desde la etapa emergente y también desde lo esperado en la etapa inicial. Es decir, la transcripción al dictado es un aprendizaje que debe iniciarse en primer grado con la finalidad de alcanzar la legibilidad de la escritura a mano, teniendo como aprendizajes previos lo que bien se define en los ejercicios de expresión escrita. Es necesario resaltar que en el CNB del nivel se encuentra indicado el inicio de la expresión escrita, pero no hay nada concreto que teóricamente apoye el ejercicio de expresión escrita que solicita el dibujo y la escritura de por lo menos una oración que exprese las ideas a partir de un estímulo, pero sí la medida de la expresión escrita por medio del dibujo.

Otro tema abordado en esta subprueba son las habilidades que requieren el conocimiento de las convenciones de la escritura, es decir, escribir sobre el renglón, linealidad, direccionalidad y espacios entre letras. Este constructo no se encuentra contemplado en el CNB, porque este refleja la preparación para la escritura sobre el desarrollo motor (motricidad gruesa, media y fina).

Es posible indicar ante lo expuesto que la prueba analizada mide dos constructos claramente delimitados, pero no todos aptos para diagnosticar los aprendizajes esperados

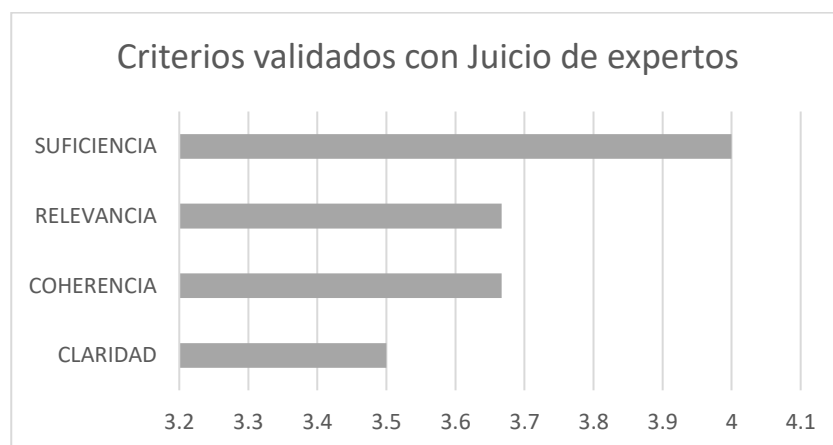
al egresar del Nivel de Educación Preprimaria según lo descrito en el Currículo Nacional Base del nivel y en la teoría existente de la lectura y escritura emergente.

B. Juicio de expertos

Para el juicio de expertos se definió el criterio de suficiencia, el cual indica que bastan los ítems que pertenecen a una misma dimensión para obtener la medición de esta. También se utilizó el criterio de relevancia que menciona que el ítem tiene relación lógica con la dimensión que está midiendo. También el criterio de coherencia que establece que el ítem se comprende fácilmente, es decir, su sintaxis y semántica o ilustración son adecuadas y finalmente el de claridad que hace referencia a que el ítem tiene relación lógica con la dimensión que está midiendo.

Los resultados obtenidos en el juicio de expertos indican que el test es suficiente en un 87.5%, que es relevante y coherente en un 91.7% y que es claro en un 100%. Dichos porcentajes se obtienen de la media entre los acuerdos de una escala entre 1 y 4. En la ilustración se evidencia las respuestas dadas por cinco expertos que participaron en el proceso de validación de contenido de la prueba.

Tabla 11. *Frecuencia de respuestas según escala utilizada*



Es importante resaltar que el resultado del juicio de expertos difiere en cierta medida con el criterio coherencia que se identificó en el análisis de la tabla de alcance y secuencia

de la prueba, porque existen dos ejercicios (subpruebas) que no pertenecen a la dimensión medida.

C. Comprobación de los supuestos de la TRI

Uno de los supuestos de la TRI es identificar la dimensionalidad de la medida y se indica en la teoría existente que este resultado será la posibilidad de concluir si las respuestas de las personas que se recogen con el instrumento pueden asegurar independencia local.

Bajo esta perspectiva se realizó el análisis factorial exploratorio, primero con un modelo de un factor y luego con el modelo de cuatro factores a razón que al hacer el análisis de factores factibles esos fueron los números posibles de analizar. Ante los resultados observados, se determina que la prueba diagnóstica para primer grado de primaria tiene buen ajuste para cuatro factores en comparación con el ajuste de un factor. Por tanto, ante esta exploración se comprende que la Prueba diagnóstica para primer grado de primaria mide más de un constructo, interpretación que se puede sustentar al contrastar este resultado con la tabla de alcance y secuencia analizada.

En otras palabras, el constructo medido es lectura y escritura desde la etapa emergente e inicial; en este se miden seis componentes que explican el constructo. Importante mencionar que la tabla de datos para realizar el análisis factorial confirmatorio, análisis que permitiría comprobar los dos supuestos de la TRI no convergió porque en el análisis de la matriz de correlación, en las variables de escritura, conciencia fonológica y decodificación se encontraron valores extremos que no permitieron proceder con el análisis, por tanto, no fue posible identificar las medidas de correlación entre las distintas subpruebas para comprobar el modelo teórico.

D. Parámetros de los ítems

Como se ha indicado anteriormente, se procedió a analizar las respuestas recogidas en campo, con el modelo de dos parámetros de la TRI, esto para obtener las medidas de discriminación y dificultad de los ítems bajo los supuestos de dicha teoría e identificar el ajuste de la prueba y reconocer la confiabilidad del instrumento por medio de la función de información del ítem y del test.

El resultado del análisis factorial exploratorio se analizó, inicialmente el conjunto de 32 ítems de la prueba para observar su comportamiento. Se encontró que la discriminación, en promedio, osciló en 3.04 con el 19 % de ítems con mal ajuste en la medida infit. En términos generales, los ítems de decodificación, en su conjunto tienen discriminación alta y poseen dificultades negativas, pero en promedio la subsección es la más difícil de la prueba diagnóstica, lo cual es lo esperado, según la teoría del constructo. Esta situación permite inferir que está recogiendo adecuadamente la información y es poco probable que un estudiante sin la habilidad adquirida responda acertadamente a estos ítems. En la Tabla 12 se observan sus parámetros.

Tabla 12. Resultados 2pl para la sección Decodificación

	Dificultad	Discriminación
<i>Item_12</i>	-0.517	6.237
<i>Item_13</i>	-0.105	16.139
<i>Item_14</i>	-0.337	6.157
<i>Item_15</i>	-0.082	17.473
<i>Item_16</i>	-0.416	4.075
<i>Item_17</i>	-0.234	5.01
<i>Item_18</i>	-0.324	5.464

Al observarse la curva de información del ítem y del test, el conjunto mostró una distribución inadecuada entre la probabilidad de acierto y la habilidad del estudiante.

Luego, se aplicó el modelo 2pl al conjunto de 25 ítems donde ya no se contó con la medida de los ejercicios de decodificación. Se encontraron cambios en el instrumento. Es

decir, ante los resultados se observó que la prueba se vuelve más fácil (-1.77), pero discrimina (3.34) mejor al eliminar los ejercicios de decodificación. Además, al compararse uno a uno los ítems, las discriminaciones mejoraron, situación que no se observó de manera constante con la dificultad.

Se calcularon los parámetros del comportamiento de los ítems y de la prueba. Ante esto, se resalta que el comportamiento del ítem 1 y 11 es constante dentro de las distintas perspectivas de análisis con el modelo 2pl, en otras palabras, son ítems extremadamente fáciles que discriminan bien, pero no se ajustan. Especialmente el ítem 1 no tiene buenos estadísticos entre lo observado con lo esperado, situación que sí cumple el ítem 11 y respecto al ajuste de las respuestas inesperadas de las personas ambos ítems no tienen el ajuste esperado.

En general, las cinco subpruebas (ejercicios de la prueba) con que se mide el constructo son fáciles a excepción de la medida de expresión escrita que posee dificultad cercana a 0.5, situación que es esperada porque la prueba diagnóstica tiene el propósito como su nombre lo indica de identificar los aprendizajes que ya han sido adquiridos por los estudiantes durante su aprendizaje en el Nivel de Educación Preprimaria, es decir la prueba cumple con el criterio de contar con ítems que identifique a los estudiantes que aún no han alcanzado el aprendizaje de lectura y escritura emergente necesario para la continuidad del aprendizaje en el Nivel de Educación Primaria.

En cuanto a las curvas características del ítem (CCI), los ítems de la subprueba (ejercicios) de desarrollo de la conciencia fonológica tienen valores extremos a la izquierda en el parámetro de dificultad, es decir son los ítems más fáciles de la prueba, con discriminación inconsistente por sus valores <0.5 y >2 . La sección requiere modificaciones para medir adecuadamente la conciencia fonológica como habilidad adquirida después de cursar el Nivel de Educación Preprimaria.

La subprueba (ejercicios) de escritura tiene comportamiento similar entre sus ítems y sus curvas indican buenas características de los ítems. En donde el ítem 11 resalta por las

propiedades analizadas en párrafos anteriores. Por otro lado, es necesario revisar los ítems de la subprueba de decodificación porque sus características de dificultad y discriminación no convergen adecuadamente. Respecto a la subprueba de desarrollo de la conciencia del lenguaje escrito se observan ítems con características moderadas en la discriminación y dificultades bajas, considerándose observar el detalle cualitativo de su construcción, como por ejemplo que en cada uno de los ítems se mida una habilidad. Finalmente, las curvas características de la subprueba de comprensión oral, todos sus ítems discriminan de manera moderada indicando que es necesario revisar la construcción de las pruebas.

En cuanto al análisis de ajuste y de la medida esperada *infit* y observada y de las respuestas inesperadas de las personas *outfit* es notable que la prueba es un instrumento que puede recoger información útil para el análisis que se requiera hacer con la población objetivo.

En este sentido, la información del ítem y del test expone que existe distribución hacia la izquierda de la habilidad de las personas, es decir confirma que la prueba diagnóstica es un instrumento que resulta ser fácil para las personas que la responden, además el ajuste de la información del test refleja que hay aglomeración en la habilidad hacia la cola izquierda de la distribución; esto indica que es necesario mejorar el instrumento de evaluación diagnóstica para acercar la medida a una distribución normal (habilidad media con distribución normal en la dificultad entre -3 y 3) y que su confiabilidad ha sido constante, no ideal según la distribución de los datos y el ajuste del modelo por sus medidas de consistencia interna son aceptables ($infit = 0.97$; $outfit = 1$).

IX. CONCLUSIONES

La prueba diagnóstica de primer grado de primaria mide dos constructos, lectura y escritura en la etapa emergente e inicial. El análisis exploratorio confirmó la existencia de por lo menos cuatro factores que se ajustan en gran parte a lo definido en el análisis de alcance y secuencia realizado a la tabla de especificaciones del instrumento observado.

El análisis de las seis subpruebas con la que se mide el constructo arroja consistencia en los distintos parámetros analizados, es decir la media del parámetro a es 3.34 donde la interpretación indica que es óptimo a partir de los valores >2 . El parámetro b es -1.77 indicando que la habilidad requerida por los sujetos para responder correctamente a los ítems no es alta. Además, que sus índices de ajuste interno son adecuados porque se encuentran dentro del rango esperado. Asimismo, las curvas características de los ítems de cada subprueba, exceptuando el ítem 1, 11 y 12 a 18 de la prueba muestra dos asíntotas horizontales en donde el punto de inflexión tiene su máxima pendiente, pudiéndose inferir la localización del sujeto en el rasgo, observándose que la estructura de la prueba mide adecuadamente lo que tiene que medir, y que la habilidad para responder los ítems al haber cumplido con los tres años de escolaridad que requiere el Nivel de Educación Preprimaria, es baja ($b = -1.7$).

En general los resultados observados con el modelo 1pl y 2pl de la Teoría de Respuesta al Ítem tienden a definirse en la cola izquierda de la curva porque esta es una prueba diagnóstica. Una prueba de este tipo pretende verificar que para los evaluados los ítems resultan fáciles, es decir, requieren habilidades bajas para responder acertadamente a los ítems, porque los aprendizajes ya fueron adquiridos. Este instrumento realiza correctamente la medida porque tiene el ajuste interno adecuado para identificar a aquellos sujetos que llegan a primer grado de primaria sin el aprendizaje esperado.

Finalmente, la validez de contenido que se realizó por medio del juicio de expertos permite inferir que quien aplica la prueba debe tener total conocimiento del proceso de aplicación y administración, es decir, llevarlo a cabo de forma estandarizada. En ambos

juicios de expertos se evidencia un índice menor al .80 que asegura que los errores de aplicación pueden afectar la confiabilidad de los datos. Por otro lado, se determina que el acuerdo entre examinadores es alto, ya que tanto los dos expertos como las tres docentes que participaron tienen una media de 3.45 sobre el acuerdo entre ellos; lo cual manifiesta, según la escala en el instrumento para este fin, que el contenido evaluado en la prueba diagnóstica es relevante para medir el dominio de lectura y escritura, según el juicio de los involucrados.

X. RECOMENDACIONES

Recomendaciones para el Centro de Investigación Educativa por ser la institución a la que le pertenece el instrumento utilizado en el estudio:

Desde la perspectiva de validez interna del instrumento es recomendable verificar los ítems que miden expresión escrita desde la expresión por medio del dibujo y la escritura de las ideas. Los ítems que miden desarrollo de la conciencia fonológica discriminan correctamente, por tanto, la medida de dificultad solo está asegurando que para aquellos estudiantes que han alcanzado el aprendizaje estos ítems no representan un reto cognitivo para que lo respondan correctamente.

De acuerdo con los resultados de la matriz de correlación, se realizaron comparaciones entre una prueba con la medida de decodificación y una sin esta medida. En este análisis se evidencio que la prueba tiene mejor ajuste interno cuando se eliminan estos ítems. Aunque la recomendación es revisarlos para identificar una forma más adecuada de medir la decodificación. Por lo tanto, se ha comprobado que la mejor medida para decodificación es una medida con pseudopalabras.

Lo antes expuesto lleva a inferir que la prueba diagnóstica para primer grado de primaria requiere ser ajustada para lograr una medida cercana a 0 que permita identificar mejor el continuo del aprendizaje en la lectura y escritura emergente, pues el marco teórico refiere a los aprendizajes que los estudiantes han adquirido antes de comenzar la educación en el Nivel de Educación Preprimaria, marco que delimita que el constructo por medir en la prueba es la etapa emergente.

XI. REFERENCIAS BIBLIOGRÁFICAS

- Aguirre International. (2005). *Un estudio reflexivo del desarrollo profesional de los docentes en los centros regionales de américa latina y el caribe para la excelencia de la capacitación a docentes*. Caribe: USAID.
- Aiken, A. (2003). *Test psicológicos y evaluación*. México: prentice Hall.
- APA. (1999). *American Psychological Association. Standards for psychological and educational test*. Washington D.C.
- Atorresi, A. (2010). *Escritura. Un estudio de las habilidades de los estudiantes de América Latina y el Caribe*. Chile: Oficina Regional de Educación para América Latina y el Caribe.
- Atorresi, A. (s.f.). *Unesco_ LLECE*. Obtenido de Competencias para la vida en las evaluaciones de lectura y escritura (SERCE-LLECE): http://www.google.com.gt/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwjW3KLrkvvKAhUDax4KHRt-BZcQFggaMAA&url=http%3A%2F%2Fwww.oei.es%2Fevaluacioneducativa%2Fcompetencias_vida_evaluaciones_lectura_escritura_atorresi.pdf&usg=AFQjCNEKeBuLGtSXrF43uILqO7
- Backhoff, E., Bouzas, A., González, M., Andrade, E., & Hernández, E. y. (2008). Factores Asociados al aprendizaje de estudiantes de 3ro primaria en México. *INEE*.
- Baker, F. B. (1985). *The basics of Ítems Response Theory*. Heinemann.
- Bandura. (2001). Guía para la construcción de escalas de autoeficacia. *Evaluar*, 2. 7-38.
- Bechger, M., Maris, G., & Verstralen. (2003). Using classical test theory in combination with item response theory. *Applied psychological measurement*, 319-334.
- Bond, L. (13 de 11 de 1999). *Norm an criterion-refenced testing*. Obtenido de Practical assessment, Research and Evaluation 2: www.ericae.net/pare/getvn.asp
- Borrero, J. (2008). *Enseñando a leer: teoría, práctica e intervención*. Colombia: Grupo Editorial Norma.
- Camargo, G., Montenegro, R., Linan-Thompson, S., & Loeza, P. (2013). *Aprendizaje de la lectoescritura*. Guatemala: Mineduc.

- Carretero-Dios, H., & Pérez, C. (2005). Normas para el desarrollo y revisión de estudios instrumentales. *International Journal of Clinical and Health Psychology*, 521-551.
- Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition*, 77-111.
- Centro de Investigaciones Educativas - CIE. (2016). *Universidad del Valle de Guatemala - UVG*. Obtenido de <https://www.uvg.edu.gt/investigacion/cie/#:~:text=Misi%C3%B3n,y%20capacitaci%C3%B3n%20en%20material%20psicom%C3%A9trico>.
- CETT-CA-RD. (S.f.). *Manual del docente. Prueba diagnóstica primer grado*. Guatemala: USAID.
- Chacón Moscoso, S., & Pérez-Gil, J. (2008). *Diseño y medición de programas de intervención neuropsicológica: aspectos fundamentales*. Sevilla: S.E.
- Charría, & González. (1993). *El placer de leer en un programa de lectura. Colección La escuela y formación de lectores autónomos*. Bogotá, Colombia: Procultura - CERLALC.
- Condemarin, M. (2001). *El poder de leer*. Chile: División de Educación General. Ministerio de Educación de la República de Chile.
- Consejo Nacional de Educación. (2010). *Ministerio de Educación*. Recuperado el 30 de julio de 2015, de Políticas públicas: http://www.mineduc.gob.gt/portal/contenido/menu_lateral/quienes_somos/politicas_educativas/pdf/Politic_Educativas_CNE.PDF
- Consoli, E. V., Canales, R., & Lingán, M. M. (2011). *PROGRAMA DE ESTIMULACIÓN DE LAS HABILIDADES PRELECTORAS EN NIÑOS Y NIÑAS DE EDUCACIÓN INICIAL DE LA PROV. CONSTITUCIONAL DEL CALLAO - PERÚ*. Perú: Investigación Educativa.
- Cortada de Kohan, N., & Cupani, M. (1998). *Teoría de Respuesta al ítem*. Paidós.
- Cotto, E. (2012). *Habilidades necesarias para aprender a leer*. Guatemala: Ajetab'al.
- Cotto, E., & Arriaga, M. (2014). *El tesoro de la lectura: material de apoyo para desarrollar la lectura emergente. Nivel de Educación Preprimaria*. Guatemala: Dirección General de Evaluación e Investigación Educativa: Ministerio de Educación.

- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Nueva York: Holt, Rinehart y Winston.
- Cronbach, L. (1985). *Fundamentos de la exploración psicológica*. Madrid: Biblioteca Nueva.
- Cuetos, F., Rodríguez, B., & Arribas, E. R. (2007). *Evaluación de la batería PROLEC-R*. España: Secop.
- Del Valle, M. J., & Cotto, E. (2015). *Modelo para explicar y predecir el aprendizaje de lectura para Guatemala*. Guatemala: Mineduc.
- Digecur. (2008). *Currículo Nacional Base. Nivel Preprimario*. Guatemala: Mineduc.
- Digecur. (2008). *Currículo Nacional Base. Primer grado. Nivel primario*. Guatemala: Mineduc.
- Digeduca. (2014). *Manual para elaboración de ítems del nivel primario*. Guatemala: Ministerio de Educación.
- Ebel, R. (1965). *Measuring educational achievement*. Englewood Cliffs: Prentice Hall.
- Egan, K. (1997). *Mentes educadas: cultura, instrumentos cognitivos y formas de comprensión*. Barcelona: Paidós.
- Embretson, S., & Reise, S. (2000). *Item response theory for psychologists*. Mahwah, N.J.: LEA.
- Ferreiro, E., & Teberosky, A. (1995). *Los sistemas de escritura en el desarrollo del niño*. México: Siglo XXI.
- Fortín Morales, Á. M. (2013). *Evaluación Educativa Estandarizada en Guatemala: Un camino recorrido, un camino por recorrer*. Guatemala: Ministerio de Educación.
- Fundar. (2011). *Pruebas de Dominio lector*. Chile: Universidad Católica de Chile.
- Gutiérrez Duarte, S., & Ruiz León, M. (2018). Impacto de la educación inicial y preescolar en el neurodesarrollo infantil. *IE Revista de investigación educativa de la REDIECH*, vol 9 núm. 17.
- Hambleton, R., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Beverly Hills: Sage.
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). *Metodología de la investigación*. McGraw Hill.
- Herrera, R. (1998). *Notas sobre psicometría*. Colombia: Inédito .

- Hulme, Snowling, & Carrol, C. &. (2005). Phonological skills are (probably) one cause of success in learning to read. *Scientific Studies of Reading*, 351-365.
- Jiménez, J., Gove, A., & Rodríguez, L. C. (2014). *Internal structure and standardized scores of the Spanish adaptation of the EGRA (Early Grade Reading Assessment) for early reading assessment*. Holanda: Psicothema.
- Jmetrik. (2015). *Sistemas de psicomedida*. Obtenido de Software y Servicios de Consultoría: https://itemanalysis-com.translate.googleusercontent.com/translate?_x_tr_sch=http&_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es-419&_x_tr_pto=sc
- Junta de Andalucía. (24 de octubre de 2015). Obtenido de Los procesos de escritura: <http://www.juntadeandalucia.es/averroes/~cepc3/competencias/lengua/aspgenerales/LOS%20PROCESOS%20DE%20ESCRITURA.pdf>
- Kelley, T. (1939). Selection of upper and lower groups for the validation of test items. *Journal of Educational Psychology*(30), 17-24.
- Kline, P. (2000). *Handbook of Psychological Testing*. Londres: Routledge.
- López, J. (Diciembre de 2005). Ítems politómicos Vs. dicotómicos: un estudio metodológico. *anales de psicología*, 21(2), 339-344.
- Martínez. (1995). *Psicometría*. Madrid: Síntesis Psicológica.
- Martínez Arias, R. (1995). *Psicometría*. Madrid: Síntesis psicológica.
- Martínez, R., Hernández, M., & Hernández, V. (2006). *Psicometría*. Madrid: Alianza.
- Martínez, R., Moreno, R., & Muñiz, J. (2005). *Análisis de los ítems*. España: Lavel S.A.
- Miller, M. D., Linn, R. L., & Gronlund, N. E. (2012). *Measurement and Assessment in Teaching (11a ed.). usa: Pearson*. EEUU: Pearson.
- Mineduc. (2012). *Plan de implementación estratégica de educación 2012-2016*. Guatemala: Usaid.
- Mineduc. (2014). *Ruta crítica*. Guatemala: Unicef.
- Mineduc. (30 de julio de 2015). *Portal del Ministerio de Educación, Digeduca*. Obtenido de <http://www.mineduc.gob.gt/digeduca/>
- Ministerio de Educación. (2010). Acuerdo Ministerial N. 560-2010. En *Reglamento interno de la Dirección General de Evaluación e Investigación Educativa (Digeduca)*. Guatemala: Ministerio de Educación.

- Montenegro, R. (2020). *Enseñanza y aprendizaje de la escritura. Libro para docentes de preprimaria y primaria*. Guatemala: USAID.
- Muñiz, J. (1997). *Introducción a la Teoría de Respuesta a los Ítems*. Madrid: Pirámide.
- Muñiz, J. (2001). *Teoría Clásica de los Test*. Madrid: Pirámide.
- Murat, F. (1985). *Evaluación del comportamiento humano*. Universidad nacional de Córdoba.
- Nunnally, J. (1991). *Teoría psicométrica*. Buenos Aires: Paidós.
- Oesterlind. (1990). Establishing criteria for meritorius test items. *Educational Research Quality*, 26-30.
- Olaz, F. (2008). *Confiabilidad*. Argentina: Paidos.
- Olilla, L. (1976). *Reading: Preparing the child. Reading: Foundations and instructional strategies*. . California: Wadsworth.
- Ollila, L. (1981). *¿enseñar a leer en preescolar*. España: NARCEA.
- Osuna, H., García, M., & Torres, M. (2004). *La segmentación en segundo grado de educación primaria*. México: Universidad Pedagógica Nacional.
- Padilla, R. D. (2007). *El dibujo del natural. En la época de la postacademia*. España: AKAL. Bellas Artes.
- Pérez, A. (S.f.). Modelos para ítems politómicos de respuestas discretas. En *Desarrollos actuales de la medición: Aplicaciones en evaluación psicológica*. (pág. 20). España: Universidad de Sevilla.
- Pérez, E., & Tornimbeni, S. (2008). *Construcción de tests*. Paidos.
- Piaget, J. (1969). *Psicología del desarrollo*. España: Mc Graw Hill.
- Popham, W. (1975). *Educational evaluation*. Englewood Cliffs, NJ: Prentice Hall.
- Prieto, G., & Delgado, A. (2003). Análisis de un test mediante el modelo de Rash. *Psicothema*, 94-100.
- psicopedagogia.com*. (02 de 2016). Obtenido de Psicología de la educación para padres y profesionales: <http://www.psicopedagogia.com/definicion/grafomotricidad>
- Quiñonez, a. (2014). *Taller del escritor. Primero primaria. Serie aprender del error. Caudernillo No. 6*. Guatemala: Ministerio de Educación.
- R Core Team. (2021). *R: A language and environment for statistical*. Obtenido de <https://www.R-project.org/>

- Ramos, R. Á. (2007). *EVALUACIÓN DE DIAGNÓSTICO PARA DETERMINAR EL NIVEL DE APRENDIZAJE DE LENGUA ESCRITA Y LAS MATEMÁTICAS. PRIMERO Y SEGUNDO GRADO. 2ª Reunión Regional Norte, Centro América y Caribe de Evaluación Educativa*. México: USEBEQ.
- Ravela, P. (2006). *Fichas didácticas. Para comprender las evaluaciones educativas*. Montevideo: PREAL.
- Real Academia Española. (2005). *Diccionario Panhispánico de dudas*. España.
- Reise, E. &. (2000). *ítem Response Theory for psychologist*. Mahwah: Erl-Baum.
- Schermelleh-Engel, K., & Moosbrugger, H. (2003). *Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures*. . *Methods of Psychological Research Online Department of Psychology*. Vol.8, No.2, pp. 23-74.
- Siqueira Loureiro, V. (s.f.). *¿Por qué y para qué enseñar vocabulario? Universidade Gama Filho / Faculdade Geremário Dantas / Centro Universitário de Volta Redonda*, 457- 465.
- Souto Grande, I. (2015). *Todo lo que necesitas saber sobre SPSS antes de utilizarlo*. Obtenido de Master en marketing digital: <https://www.uscmarketingdigital.com/todo-sobre-spss/>
- Swearingen, A., & Calder, T. (2009). La escritura necesita habilidades motrices visuales y de percepción. *Handy Handouts*, 138.
- Tabachnick, & Fidel. (2001). *Usibg multivariate statistics*. Nueva York: Harper & Row.
- Thompson, L. (febrero de 2004). *5 Components of Early Grade Reading*. Obtenido de <https://www.eddataglobal.org/video/index.cfm>
- Thorndike, R. (1989). *Psicometría aplicada*. México: Limusa.
- Tornimbeni, S., Pérez, E., & Olaz, F. (2008). *Introducción a la psicometría*. Argentina: Paidós.
- Urbina, A. A. (1998). *Test psicológicos*. México: Prentice Hall. 2da edición.
- USAID/CETT. (2009). *Manual del docente*. . República Dominicana: CETT.
- USAID/G. (1998). *Diseño de la Reforma Educativa*. Guatemala: Ministerio de Educación.
- UVG. (2015). *Centro de Investigaciones Educativas (CIE)*. Obtenido de Servicios: <http://www.uvg.edu.gt/investigacion/cie/servicios.html>

- Vigotsky, L. (1988). *El desarrollo de los procesos psicológicos superiores*. México: Grijalbo.
- Villagran, M., & et.al. (2010). velocidad de nombrar y conciencia fonológica en el aprendizaje de la lectura. *Psicothema*, 22(3), 436-442. Recuperado el julio de 2015
- Wise, & Wolfe, P. &. (2008). Phonological awareness and rapid naming skills of children with reading disabilities and children with reading disabilities who are at risk for mathematics difficulties. *Learning Disabilities Research and practice*, 125-136.
- Zeller, E. C. (1979). *Reliability and vallidity assessment*. Londres: Sage.

XII. ANEXOS

Anexo 1. Carta de solicitud para realizar aplicación en el establecimiento seleccionado para la obtención de datos.

Guatemala, 24 de agosto de 2015.

Señores (as)
Directores (as) de Educación

Estimados Señores (as) directores (as):

Reciban un cordial saludo. Hago de su conocimiento que se ha iniciado el proceso de validación de 32 ítems contenidos en una prueba de lectura y escritura diseñada para estudiantes que inician a cursar el primer grado de primaria. La prueba a validar fue elaborada bajo el contexto del programa de capacitación a docentes de Centro América y República Dominicana que es coordinado en Guatemala por la Universidad del Valle.

Al diseñar instrumentos de evaluación con el propósito de generar información que permita la toma de decisiones y la generalización de los resultados para contribuir con la mejora continua de la educación de nuestro país, es necesario realizar una aplicación piloto que arroje información pertinente que permita el análisis del funcionamiento de los ítems contenidos en la prueba de lectura y escritura emergente.

Tomando en cuenta el desempeño que los estudiantes han demostrado en años anteriores en la prueba nacional que administra el Ministerio de Educación, se ha seleccionado a los centros educativos que son parte de la muestra. Adjunto encontrará los establecimientos educativos seleccionados. Por lo expuesto, solicito su valiosa colaboración para girar instrucciones a quien corresponda para que el equipo asignado pueda realizar la aplicación de instrumentos de evaluación a los estudiantes que cursan el grado de preparatoria.

Es importante mencionar que la prueba consta de dos momentos, es decir inicialmente se hará una aplicación grupal, que tarda 20 minutos. El segundo momento dura 12 minutos por estudiante, la modalidad de aplicación es individual. Con el objetivo de propiciar un ambiente sin distracciones al momento de responder a las preguntas de la prueba, se requiere de una mesa y dos sillas ubicadas afuera del salón del estudiante.

El proceso piloto se realizará del 08 al 18 de septiembre del año en curso en el horario de la jornada de cada establecimiento. Se requiere que los estudiantes tengan a mano un lápiz y un borrador. Además, es necesario que un docente encargado del grado permanezca dentro del salón para acompañar el proceso.

La información generada será utilizada únicamente para determinar la confiabilidad y validez de los ítems por lo que será manejada de forma confidencial.

	Departamento	Municipio	Dirección	Nivel	Sector	Área	Modalidad	Jornada	Plan
1	Ciudad Capital	Zona 11	12 Avenida 12-49	Párvulos	Oficial	Urbana	Monolingüe	Matutina	Diario(Regular)
2	Ciudad Capital	Zona 1		Párvulos	Privado	Urbana	Monolingüe	Matutina	Diario (Regular)
3	Guatemala	Santa Catarina Pinula	Km. 16 Carretera A El Salvador Lote 3 Lomas De San Rafael	Párvulos	Privado	Rural	Monolingüe	Matutina	Diario(Regular)
4	Ciudad Capital	Zona 15	3a. Calle 16-53, Colonia El Maestro	Párvulos	Oficial	Urbana	Monolingüe	Matutina	Diario(Regular)
5	Guatemala	Fraijanes	1ra. Avenida 4-93 Zona 1 Aldea Lo De Dieguez	Párvulos	Oficial	Rural	Monolingüe	Vespertina	Diario(Regular)
6	Ciudad Capital	Zona 16	4a. Calle 1-24.	Párvulos	Privado	Urbana	Monolingüe	Matutina	Diario(Regular)

*Se suprimió de la tabla anterior el nombre de los establecimientos para mantener la confidencialidad de los datos.

Anexo 2. Instrumento para validación de expertos

Guatemala, _____ de _____ de 2015

Respetado (a) juez (a):

Usted ha sido seleccionada para evaluar el instrumento de **Pre lectura** que hace parte de la investigación de la **prueba diagnóstica para primer grado del** Centro de Excelencia para la Capacitación de Maestros de Centroamérica y República Dominicana (CETT-CA-RD). La evaluación de los instrumentos es de gran relevancia para lograr que sean válidos y que los resultados obtenidos a partir de estos sean utilizados eficientemente en el campo de la evaluación educativa. Agradecemos su valiosa colaboración.

Nombre del juez (a): _____

Formación académica:

(Nivel medio) _____

(Nivel superior) _____

Áreas de experiencia profesional _____

Tiempo de experiencia: _____ años _____ meses

En qué se desempeña actualmente: _____

Institución en que labora: _____

Objetivo de la investigación: análisis de reactivos y validación de la prueba diagnóstica en el Sistema Educativo Nacional

Objetivo del juicio de expertos: validar el instrumento de evaluación diagnóstica para primer grado en suficiencia, claridad, coherencia y relevancia.

Objetivo de la prueba: medir el constructo de la prueba para aportar información válida a los docentes para la toma de decisiones que marcará el rumbo para el aprendizaje de la lectura en el nivel primario.

Instrucciones:

De acuerdo con los siguientes indicadores califique cada uno de los ítems según corresponda. Escriba el número que indica la clasificación.

CATEGORÍA	CLASIFICACIÓN	INDICADOR
SUFICIENCIA BASTAN LOS ÍTEMS QUE PERTENECEN A UNA MISMA DIMENSIÓN PARA OBTENER LA MEDICIÓN DE ESTA.	1. No cumple con el criterio	Los ítems no son suficientes para medir el constructo de lectura y escritura emergente
	2. Bajo nivel	Los ítems miden algún aspecto de la lectura y escritura emergente pero no corresponden con la dimensión total
	3. Nivel moderado	Se deben incrementar algunos ítems para poder evaluar el constructo de lectura y escritura emergente completamente
	4. Cumple con el criterio	Los ítems son suficientes

CATEGORÍA	CLASIFICACIÓN	INDICADOR
COHERENCIA EL ÍTEM TIENE RELACIÓN LÓGICA CON LA DIMENSIÓN QUE ESTÁ MIDIENDO	1. No cumple con el criterio	El ítem no tiene relación lógica con la lectura y escritura emergente
	2. Bajo nivel	El ítem posee poca importancia con la lectura y escritura emergente
	3. Nivel moderado	El ítem tiene relación moderada con el constructo de lectura y escritura emergente que está midiendo
	4. Cumple con el criterio	El ítem se encuentra completamente relacionado con el constructo de lectura y escritura emergente que está midiendo

CATEGORÍA	CLASIFICACIÓN	INDICADOR
------------------	----------------------	------------------

CLARIDAD EL ÍTEM SE COMPRENDE FÁCILMENTE, DECIR SU SINTÁCTICA Y SEMÁNTICA O ILUSTRACIÓN SON ADECUADAS	1. No cumple con el criterio	El ítem no es claro
	2. Bajo nivel	El ítem requiere bastantes modificaciones o una modificación muy grande en el uso de las palabras o imágenes de acuerdo con su significado o por la ordenación de las mismas
	3. Nivel moderado	Se requiere una modificación muy específica de los términos del ítem
	4. Cumple con el criterio	El ítem es claro, tiene semántica y sintaxis o imagen adecuada

CATEGORÍA	CLASIFICACIÓN	INDICADOR
RELEVANCIA EL ÍTEM ES ESENCIAL O IMPORTANTE, POR LO TANTO DEBE SER INCLUIDO	1. No cumple con el criterio	El ítem puede ser eliminado sin que se vea afectada la medición del constructo de lectura y escritura emergente
	2. Bajo nivel	El ítem tiene alguna relevancia, pero otro ítem puede estar incluyendo lo que mide este
	3. Nivel moderado	El ítem es relativamente importante
	4. Cumple con el criterio	El ítem es relevante y debe ser incluido

Anexo 3. Matriz de correlación

	Item_1	Item_2	Item_3	Item_4	Item_5	Item_6	Item_7	Item_8	Item_9	Item_10	Item_11	Item_12	Item_13	Item_14	Item_15	Item_16	Item_17	Item_18	Item_19	Item_20	Item_21	Item_22	Item_23	Item_24	Item_25	Item_26	Item_27	Item_28	Item_29	Item_30	Item_31	Item_32
Item_1	1.000	0.307	0.504	0.051	0.054	0.043	0.102	0.043	0.045	-0.019	0.071	0.020	0.048	0.006	-0.005	0.009	-0.005	0.055	-0.034	-0.047	-0.063	-0.064	0.052	0.026	0.022	0.022	0.053	-0.064	-0.016	-0.088	-0.088	-0.060
Item_2	0.307	1.000	0.223	0.362	0.311	0.235	0.196	0.196	0.277	-0.029	0.184	0.150	0.141	0.157	0.169	0.123	0.136	0.187	0.158	0.249	0.116	0.113	0.094	0.175	0.166	0.081	0.149	0.195	0.193	0.077	0.117	0.125
Item_3	0.504	0.223	1.000	0.090	0.095	0.179	0.223	0.132	0.180	-0.024	0.117	0.180	0.141	0.195	0.136	0.120	0.136	0.151	0.039	0.002	0.069	0.116	0.052	0.053	0.048	0.140	0.108	0.017	0.085	-0.072	-0.028	-0.027
Item_4	0.051	0.362	0.090	1.000	0.725	0.357	0.357	0.357	0.383	0.163	0.235	0.405	0.392	0.393	0.287	0.330	0.440	0.368	0.211	0.222	0.080	0.125	0.065	0.177	0.143	0.295	0.252	0.322	0.255	0.148	0.129	
Item_5	0.054	0.311	0.095	0.725	1.000	0.372	0.404	0.340	0.408	0.168	0.229	0.392	0.382	0.378	0.274	0.274	0.346	0.400	0.458	0.220	0.232	0.088	0.067	0.186	0.194	0.130	0.255	0.228	0.240	0.184	0.160	0.114
Item_6	0.043	0.235	0.178	0.357	0.372	1.000	0.458	0.478	0.525	0.125	0.363	0.608	0.597	0.640	0.600	0.550	0.565	0.576	0.133	0.108	0.193	0.077	0.314	0.263	0.170	0.337	0.287	0.209	0.191	0.088	0.133	0.046
Item_7	0.102	0.196	0.223	0.357	0.404	0.458	1.000	0.458	0.505	0.253	0.330	0.571	0.548	0.587	0.514	0.550	0.565	0.524	0.242	0.219	0.216	0.099	0.251	0.237	0.195	0.337	0.184	0.165	0.191	0.033	0.078	0.046
Item_8	0.043	0.196	0.132	0.357	0.240	0.478	0.458	1.000	0.405	0.193	0.380	0.550	0.553	0.569	0.548	0.497	0.565	0.576	0.169	0.108	0.143	-0.033	0.314	0.165	0.170	0.253	0.252	0.209	0.211	0.125	0.170	0.079
Item_9	0.045	0.277	0.180	0.393	0.408	0.525	0.505	0.485	1.000	0.191	0.392	0.635	0.589	0.631	0.597	0.523	0.540	0.603	0.209	0.112	0.220	0.103	0.277	0.268	0.148	0.325	0.329	0.302	0.257	0.112	0.158	0.103
Item_10	-0.019	-0.029	-0.024	0.163	0.168	0.125	0.253	0.189	0.191	1.000	0.056	0.195	0.126	0.137	0.123	0.141	0.123	0.134	0.201	-0.043	0.013	0.012	0.003	0.121	0.034	0.039	0.076	0.082	-0.004	0.037	0.039	-0.046
Item_11	0.071	0.184	0.117	0.235	0.229	0.363	0.330	0.380	0.392	0.056	1.000	0.507	0.498	0.505	0.470	0.491	0.498	0.514	0.138	0.181	0.185	0.045	0.317	0.228	0.195	0.334	0.205	0.227	0.318	0.116	0.104	0.063
Item_12	0.020	0.150	0.160	0.405	0.392	0.608	0.571	0.590	0.635	0.155	0.507	1.000	0.916	0.936	0.913	0.791	0.768	0.822	0.190	0.157	0.219	0.091	0.433	0.250	0.223	0.431	0.361	0.312	0.225	0.116	0.160	0.063
Item_13	0.048	0.141	0.141	0.392	0.382	0.597	0.546	0.563	0.589	0.126	0.498	0.916	1.000	0.865	0.860	0.797	0.801	0.820	0.205	0.170	0.235	0.076	0.370	0.310	0.243	0.379	0.395	0.283	0.330	0.191	0.220	0.135
Item_14	0.006	0.157	0.155	0.389	0.378	0.640	0.587	0.569	0.631	0.137	0.505	0.836	0.865	1.000	0.830	0.821	0.785	0.800	0.223	0.168	0.226	0.141	0.387	0.291	0.225	0.373	0.388	0.354	0.332	0.150	0.210	0.137
Item_15	-0.005	0.169	0.136	0.383	0.374	0.600	0.514	0.548	0.597	0.123	0.470	0.913	0.860	0.830	1.000	0.778	0.824	0.800	0.230	0.233	0.205	0.122	0.357	0.322	0.212	0.379	0.338	0.272	0.283	0.162	0.205	0.118
Item_16	0.009	0.129	0.120	0.287	0.274	0.550	0.550	0.497	0.523	0.141	0.491	0.791	0.797	0.821	0.778	1.000	0.779	0.775	0.162	0.201	0.199	0.075	0.420	0.324	0.289	0.406	0.298	0.211	0.223	0.116	0.143	0.095
Item_17	-0.005	0.136	0.136	0.330	0.346	0.565	0.565	0.565	0.540	0.123	0.488	0.766	0.801	0.795	0.824	0.778	1.000	0.785	0.199	0.162	0.186	0.084	0.357	0.300	0.212	0.379	0.308	0.234	0.283	0.146	0.174	0.132
Item_18	0.055	0.187	0.151	0.410	0.400	0.576	0.524	0.576	0.603	0.134	0.514	0.822	0.820	0.800	0.800	0.775	0.785	1.000	0.219	0.211	0.238	0.096	0.378	0.307	0.218	0.378	0.330	0.249	0.375	0.156	0.152	0.111
Item_19	-0.034	0.158	0.039	0.388	0.458	0.133	0.242	0.169	0.209	0.201	0.138	0.190	0.205	0.223	0.230	0.102	0.199	0.219	1.000	0.224	0.136	0.172	-0.077	0.059	0.145	-0.042	0.183	0.225	0.275	0.254	0.289	0.248
Item_20	-0.047	0.249	0.002	0.211	0.220	0.168	0.219	0.168	0.112	-0.043	0.181	0.157	0.170	0.168	0.233	0.201	0.162	0.211	0.224	1.000	0.496	0.399	0.155	0.246	0.197	0.134	0.157	0.035	0.169	0.079	0.058	0.123
Item_21	-0.063	0.116	0.069	0.222	0.232	0.153	0.218	0.148	0.220	0.012	0.185	0.219	0.225	0.236	0.205	0.195	0.186	0.238	0.158	0.496	1.000	0.504	0.200	0.190	0.261	0.223	0.173	0.117	0.210	0.235	0.121	0.203
Item_22	-0.064	0.113	0.116	0.080	0.088	0.077	0.099	-0.033	0.103	0.012	0.045	0.091	0.076	0.141	0.122	0.075	0.084	0.086	0.172	0.399	0.504	1.000	0.080	0.129	0.200	0.135	0.203	0.184	0.138	0.167	0.135	0.176
Item_23	0.052	0.094	0.052	0.125	0.067	0.314	0.291	0.314	0.277	0.003	0.317	0.433	0.370	0.387	0.357	0.420	0.357	0.378	-0.077	0.158	0.200	0.080	1.000	0.400	0.274	0.380	0.188	0.194	0.225	0.122	0.111	0.104
Item_24	0.026	0.175	0.053	0.065	0.196	0.263	0.237	0.195	0.268	0.121	0.228	0.290	0.310	0.291	0.323	0.324	0.300	0.307	0.059	0.246	0.190	0.123	0.400	1.000	0.337	0.366	0.187	0.071	0.103	0.071	0.149	0.062
Item_25	0.022	0.166	0.048	0.177	0.184	0.170	0.195	0.170	0.148	0.034	0.195	0.223	0.243	0.225	0.212	0.189	0.212	0.218	0.145	0.197	0.261	0.200	0.274	0.337	1.000	0.344	0.190	0.200	0.216	0.266	0.153	0.259
Item_26	0.053	0.149	0.108	0.295	0.255	0.287	0.164	0.252	0.329	0.076	0.205	0.361	0.355	0.388	0.338	0.298	0.308	0.330	0.183	0.157	0.173	0.203	0.189	0.187	0.190	0.239	1.000	0.280	0.262	0.147	0.175	0.202
Item_27	-0.064	0.155	0.017	0.252	0.228	0.209	0.165	0.209	0.302	0.082	0.227	0.312	0.283	0.254	0.272	0.211	0.234	0.249	0.252	0.035	0.117	0.184	0.194	0.071	0.200	0.236	0.280	1.000	0.269	0.328	0.216	0.306
Item_28	-0.016	0.193	0.065	0.322	0.240	0.191	0.191	0.211	0.257	-0.004	0.308	0.325	0.330	0.332	0.283	0.233	0.283	0.375	0.275	0.188	0.210	0.138	0.225	0.193	0.216	0.193	0.252	0.289	1.000	0.355	0.330	0.258
Item_29	-0.088	0.077	-0.072	0.258	0.184	0.088	0.033	0.125	0.112	0.037	0.116	0.116	0.191	0.150	0.162	0.116	0.146	0.156	0.254	0.078	0.235	0.167	0.122	0.071	0.266	0.194	0.147	0.328	0.356	1.000	0.698	0.614
Item_30	-0.086	0.077	-0.028	0.148	0.160	0.133	0.078	0.170	0.158	0.039	0.104	0.160	0.220	0.210	0.205	0.143	0.174	0.152	0.259	0.058	0.121	0.135	0.111	0.149	0.133	0.045	0.175	0.216	0.330	0.696	1.000	0.615
Item_31	-0.060	0.125	-0.027	0.129	0.114	0.046	0.046	0.079	0.103	-0.046	0.063	0.063	0.135	0.137	0.118	0.095	0.122	0.111	0.248	0.123	0.203	0.176	0.104	0.062	0.259	0.100	0.202	0.306	0.258	0.614	0.615	1.000

Anexo 4. Sintaxis de r

-- Cargar de la tabla de datos --

```
file.choose()
df_uvgt <- read.csv('C:\\Users\\AGS\\Desktop\\Investigación\\Tesis
uvgt\\VF18.csv',sep=',',header=T)
dim(df_uvgt) = 300 x 49 View(df_uvgt)
```

-- Selección de las variables para el análisis --

```
md <- df_uvgt[,4:35] dim(md)= 300 x 32
```

-- Limpieza de la tabla de datos --

```
md$escritura <- 1*as.numeric(ifelse(md$escritura!='1',0,md$escritura))
md$dic_4 <- 1*as.numeric(ifelse(is.na(md$dic_4),0,md$dic_4))
```

-- Es necesario renombrar las variables, dado que es una prueba no liberada --

```
names(md) <-
c('Item_1','Item_2','Item_3','Item_4','Item_5','Item_6','Item_7','Item_8','Item_9','Item_10'
,'Item_11','Item_12',
'Item_13','Item_14','Item_15','Item_16','Item_17','Item_18','Item_19','Item_20','Item_21',
'Item_22','Item_23',
'Item_24','Item_25','Item_26','Item_27','Item_28','Item_29','Item_30','Item_31','Item_32'
)
```

```
str(md) summary(md) attach(md)
```

--- Model 2PL ---

Ajuste del modelo 2PL con la función ltm(), la cual realiza la estimación de parámetros bajo el método de la máxima verosimilitud marginal.

```

Mod_2pl <- ltm(md ~ z1)
-- ICC para los ítem's del 1 al 5 --
  png(filename = "CC1_1_5.png", width = 500, height = 400)
  plot(Mod_2pl, items = 1:5)
  dev.off()

-- ICC para los ítem's del 6 al 11 --
  png(filename = "CC1_6_11.png", width = 500, height = 400)
  plot(Mod_2pl, items = 6:11)
  dev.off()

-- ICC para los ítem's del 12 al 18 --
  png(filename = "CC1_12_18.png", width = 500, height = 400)
  plot(Mod_2pl, items = 12:18)
  dev.off()

-- ICC para los ítem's del 19 al 26 --
  png(filename = "CC1_19_26.png", width = 500, height = 400)
  plot(Mod_2pl, items = 19:26)
  dev.off()

-- ICC para los ítem's del 27 al 32 --
  png(filename = "CC1_27_32.png", width = 500, height = 400)
  plot(Mod_2pl, items = 27:32)
  dev.off()

```

```

plot(Mod_2pl, items = 0, type = "IIC")

```

```

--- ANÁLISIS FACTORIAL EXPLORATORIO pasos preparatorios ---

```

```

Descripción de los datos --> describe(md)

```

Matriz de correlación --> R <- cor(md)

Diagrama de dispersión bivariados --> ggpairs(md) *No es funcional para más de 10 ítem's*

Se retiran los ítem's del 12 al 18

```
md1 <-  
data.frame(Item_1,Item_2,Item_3,Item_4,Item_5,Item_6,Item_7,Item_8,Item_9,Item_1  
0,Item_11,Item_19,Item_20,Item_21,  
  
Item_22,Item_23,Item_24,Item_25,Item_26,Item_27,Item_28,Item_29,Item_30,Item_31  
,Item_32  
)
```

Matriz de correlación --> R1 <- cor(md1)

DETERMINACIÓN DE LA FACTORIZACIÓN

Estadística Kaiser-Meyer-Olkin -- KMO -- {KMO >= 0.60}

KMO(md1)

cortest.bartlett(md1)

det(cor(md1))

-- Número de Factores --

```
results_nfactor <- n_factors(md1, rotate = "varimax", fm = "mle", n = NULL)
```

-- Gráfica con el Número de Factores Recomendados --

```
png(filename = "Número de Factores Recomendados.png", width = 800, height = 600)
```

```
plot(results_nfactor)
```

```
dev.off()
```

```
as.data.frame(results_nfactor)
```

```
Tfactor1 <- fa(md, nfactores = 1, fm = "ml", rotate = "varimax")
```

```

print(Tfactor1,digits = 2,cut = .30,sort=TRUE)

fa.diagram(Tfactor, digits = 2, main = "Factor Diagram",
           cut = .32,
           simple = F,
           errors = T)
summary(Tfactor1, fit.measures=TRUE)
md2
data.frame(Item_1,Item_2,Item_3,Item_4,Item_5,Item_6,Item_7,Item_8,Item_9,Item_1
0,Item_11,Item_19,Item_20,Item_21,
Item_22,Item_23,Item_24,Item_25,Item_26,Item_27,Item_28,Item_29,Item_30,Item_31
,Item_32,Item_12,Item_13,
           Item_14,Item_16,Item_17,Item_18
           )
Tfactor2<-fa(md2,nfactors = 1,fm = "ml",rotate ="varimax")
print(Tfactor2,digits = 2,cut = .30,sort=TRUE)
fa.diagram(Tfactor, digits = 2, main = "Factor Diagram",
           cut = .32,
           simple = F,
           errors = T)
summary(Tfactor2, fit.measures=TRUE)
-----
Tfactor3<-fa(md1,nfactors = 4,fm = "ml",rotate ="varimax")
print(Tfactor3,digits = 2,cut = .30,sort=TRUE)
fa.diagram(Tfactor3, digits = 2, main = "Factor Diagram",
           cut = .32,
           simple = F,
           errors = T)
summary(Tfactor3, fit.measures=TRUE)

```