

UNIVERSIDAD NACIONAL AUTÓNOMA DE HONDURAS

FACULTAD DE CIENCIAS SOCIALES

ESCUELA DE CIENCIAS PSICOLÓGICAS

MAESTRÍA EN PSICOMETRÍA Y EVALUACIÓN EDUCATIVA



TESIS

Diseño de una prueba de desempeño académico en el área de Ciencias Sociales para los estudiantes de noveno grado en Honduras 2017-2018

Tesis para optar al grado de Magister en

Psicometría y Evaluación Educativa

Presentada por

Dulce Cristina Guevara Castillo

Asesor temático: PhD. **Dennis Fernando Cáceres Valladares, Ph.D.**

Asesor Metodológico: **German Edgardo Moncada Godoy, Ph.D.**

Tegucigalpa, M.D.C., abril Honduras 2020

Tegucigalpa, M.D.C., 29 de abril de 2020

Señores

Universidad Nacional Autónoma de Honduras

Coordinadores Maestría en Psicometría y Evaluación Educativa

Estimados señores:

Por este medio autorizo a la UNAH para que publique en su REVISTA DIGITAL, WEB, CD Y CUALQUIER OTRO MEDIO, y uso de beneficio sociales y académicos de la sociedad hondureña mi trabajo de tesis titulado: **“Diseño de una prueba de desempeño académico en el área de Ciencias Sociales para los estudiantes de noveno grado en Honduras 2017-2018”**, el cual he presentado como parte de los requisitos para optar por la Maestría en Psicometría y Evaluación Educativa.

Atentamente,



Dulce Cristina Guevara Castillo
Cuenta No. PSE100113

DEGT-UNAH

AUTORIDADES UNIVERSITARIAS

Doctor Francisco José Herrera Alvarado
Rector Interino

Máster Belinda Flores de Mendoza
Vice-Rectora Académica

Abogada Jessica Patricia Sánchez Medina
Secretaria General

Doctora Leonarda Andino
Directora de Docencia

Doctor Armando Euceda
Director Sistema de Estudios de Postgrado

Doctora Martha Lorena Suazo Matute
Decana Facultad de Ciencias Sociales

Máster María José Irías Escher
Coordinadora General Postgrados Facultad de Ciencias Sociales

Máster Lina María Mendoza Recarte
Coordinadora Académica Maestría en Psicometría y Evaluación Educativa

Dedicatoria

A Dios por regalarme la oportunidad de seguir avanzado en vida
A mis padres, que me han inspirado a ver el mundo de diferente manera y que me han
apoyado en cada uno de mis planes que me ha regalado la vida.

Agradecimiento

Le agradezco a mi madre; Yorna L. Castillo y a José M. Rivera por sus sabios consejos y estar pendientes de cada etapa de esta valiosa meta; reconozco a ambos por confiar en mí.

Asimismo, doy gracias a mis amigos y a mis compañeros de trabajo que estuvieron pendientes del proceso de esta tesis

Mi agradecimiento a todos los docentes de la maestría, objetivamente han cambiado la percepción de ver las cosas que me rodean; a mis compañeras de trabajo, ahora amigas; María Isabel Hernández, Keyla Rivera, Tessy Solorzano, Nusly Moreno, y a mi compañero; Carlos Aguilar que siempre nos apoyamos para alcanzar esta meta.

A mi asesor de tesis el PhD. Dennis Cáceres que desde un principio estuvo pendiente de cada uno de sus estudiantes; así mismo agradezco al PhD. Germán Moncada por cada una de sus observaciones realizadas en cada una de las presentaciones de esta tesis.

Mi agradecimiento a Msc. Lina Mendoza por su disposición a colaborar a nuestro crecimiento profesional como personal y por cada consejo brindado por ella, llego en un momento oportuno en mi vida.

Agradezco al personal administrativo y docente que labora para esta maestría y que siempre están disponible para apoyar a cada uno de los estudiantes.

Así mismo agradezco Ana Alvarado, Vilma Díaz, Jerson Cortez, Maxwell Láinez, Nancy Hernández docentes del área de Ciencias Sociales que contribuyeron a la elaboración de reactivos.

Resumen

En la actualidad la evaluación educativa juega un papel importante, ya que se encuentra vinculada con los contenidos que plantea el Currículo Nacional Básico, asimismo la práctica de la evaluación tanto en los individuos como en los centros educativos llegan a fortalecer o a enriquecer el proceso de enseñanza-aprendizaje; de esta manera se convierte en una evaluación formativa. Además, es fundamental tener en cuenta las necesidades y las características propias de cada estudiante y centro educativos.

La evaluación es un proceso que ayuda a verificar con eficacia y eficiencia los aprendizajes adquiridos por los estudiantes, en el transcurso y al final del año. Esta toma como bases los estándares educativos del país, clasificados en: estándares de contenido y estándares de desempeño.

Esta investigación se centra en la evaluación del aprendizaje de una manera estandarizada, específicamente en el área de Ciencias Sociales. Es importante mencionar que el Proyecto de Eficiencia de la Educación Primaria (PEEP) fue el pionero en realizar la evaluación estandarizada en nuestro país en el año de 1990-1994 (Evaluaciones formativas en las cuatro áreas establecida por el currículo), de igual manera la UMCE y MIDEH se han preocupado por la evaluación de los aprendizajes sin embargo solo han tomado en cuenta en dos áreas español y matemáticas.

En este estudio queda evidenciado diferentes temáticas que contribuyen al diseño de la prueba estandarizada del área de ciencias sociales de fin de grado para estudiantes de noveno, por ende, en el primer capítulo se aborda el planteamiento del problema, los objetivos de investigación y la justificación; en el segundo se describe la teoría en cual se fundamenta esta investigación, en el tercero se refiere a la evaluación de nuestro país, en el cuarto capítulo se plantea el método o la metodología que se utilizó para llevar a ejecución este trabajo, en el quinto se discute los resultados obtenidos en la aplicación del instrumento, en el sexto se presenta la discusión de los resultados realizando la comparación de los datos obtenidos.

Índice de contenido

Capítulo I: Planteamiento del Problema	1
1.1 Situación del problema/Construcción del objeto de estudio.....	1
1.2. Objetivos de la investigación	3
1.2.1. Objetivos General.....	3
1.2.2. Objetivo Específicos	3
1.3. Justificación	4
Capítulo II: Marco Teórico	5
2.1 Conceptos básicos para la elaboración de una prueba de desempeño académico	5
2.1.1. Evaluación educativa	5
2.1.2. Medición de los aprendizajes.....	6
2.1.3 Estándares Educativos	7
2.1.3.1. Estándares de contenido	9
2.1.3.2. Estándares de desempeño	10
2.1.4 Evaluación Psicométrica	11
2.1.5 Desempeño académico	11
2.2. Evaluación de los aprendizajes.....	12
2.2.1. Evaluación estandarizada de aprendizaje.	15
2.2.1.1. Evaluación estandarizada en Europa.....	20
2.2.1.2 Evaluación estandarizada en Latinoamérica.....	24
2.2.1.3. La evaluación estandarizada en Centroamérica	28
2.3. Diseño y validación de un instrumento de evaluación	30
2.3.1 Vínculo entre los objetivos y el instrumento que se elaborará o seleccionará	31
2.3.2. Tipo de Instrumentos	34
2.3.2.1 Pruebas paramétricas y no paramétricas	36
2.3.2.2. Instrumentos referidos a normas y a criterios.....	37
2.3.2.3 Instrumento de respuesta abierta o cerrada.....	38
2.3.3. Calidad de los instrumentos.....	40
2.3.4 Taxonomía de objetivos cognoscitivo	50
2.3.5. Marco conceptual y tabla de especificaciones	53

2.3.6. Elaboración de reactivos	53
2.3.7 Principios para el desarrollo de las pruebas o teoría de la medición	55
2.3.7.1 Teoría Clásica de los Test	55
2.3.7.2 Modelo lineal clásico de la TCT.....	55
2.3.7.3 Modelo de formas paralelas de la TCT	56
2.3.7.4 Supuestos de la TCT.....	56
2.3.7.5 Limitaciones de la TCT	57
2.3.7.6. Los parámetros de la TCT	58
2.3.7.2 Teoría de respuesta al ítem	59
2.3.7.3 Modelo de la TRI.....	60
2.3.7.4 Parámetro de la TRI	62
2.3.7.5 Los supuestos de la TRI	63
2.3.7.6. Curva Características del ítem	64
2.3.7.7. Diferencias de la TCT y la TRI	66
2.4 Las Ciencias Sociales en el Mundo.....	67
2.4.1 La desigualdad.....	72
2.4.2. Cambios ambientales	74
2.4.2.1. La pérdida de la biodiversidad	74
2.4.2.2. Cambio Climático.....	75
Capítulo III. Marco Contextual	75
3.1 Situación de la educación en Honduras	76
3.2 Marco legal y las reformas del sistema educativo hondureño.....	82
3.3 La educación básica en Honduras.....	86
3.3.1. Currículo Nacional Básico	88
3.4 Unidad Externa de la Calidad de la Educación (UMCE).....	91
3.5. Proyecto Mejorando el Impacto al Desempeño Estudiantil en Honduras	91
Capítulo IV: Método	92
4.1. Enfoque, alcance y diseño	93
4.2. Variables, operacionalización e indicadores.....	94
4.3 Población y muestra	96
4.4. Tipo de muestreo.....	97
4.5. Tamaño de la muestra.....	98

4.6. Diseño muestra de la prueba operativa	98
4.7. Tamaño de la muestra para la prueba operativa	99
4.8 Instrumento de medición	101
4.9 Plan de análisis	103
4.10.1 Procedimientos que se utilizó para identificar los estándares, y bloques que se evaluaron en la prueba	109
4.10.2. Diseño de los reactivos para la prueba de CCSS.....	109
Capítulo V: Resultados de la investigación.	113
5.1. Descriptivos de la muestra.	113
5.2. Resultados de ítems aplicados en la prueba Piloto.....	114
5.1.1 Digitalización de los resultados, obtenidos en aplicación de la prueba piloto.	114
5.1.2 Digitalización de los resultados, obtenidos en aplicación de la prueba operativa.....	146
CAPITULO VI: Discusión de los resultados	157
Conclusiones	160
Recomendaciones	161
Referencias bibliografía	162
Declaración de Ética de la Investigación	170
Anexos.....	171

Índice de Figuras

Figura 1: Tipo de instrumentos	35
Figura 2: Ejemplo de instrumentos paramétricas y no paramétricas	37
Figura 3: Tipo de validez	45
Figura 4: Marco teórico para analizar confiabilidad.....	46
Figura 5: Consideraciones prácticas para elaborar o seleccionar un instrumento	49
Figura 6: Modelo de la TRI	60
Figura 7: Resultados de la TRI.....	65
Figura 8: Tres dimensiones de los social.....	68
Figura 9: Tipo de desigualdad.....	73
Figura 10: Estructura general del sistema educativo hondureño	79
Figura 11: Área curriculares del nivel de educación básica.....	87
Figura 12: Instrumentos en los que se apoya el CNB.....	89
Figura 13: Estructura del CNB.....	90
Figura 14: Temática utilizada en el taller de evaluación educativa	110
Figura 15: Etapas para la construcción de una prueba.....	112
Figura 16: Verificación datos atípico de las pruebas pilotos	144

Índice de Tablas

Tabla 1:Métodos para verificar la confiabilidad	48
Tabla 2: Compendio ilustrativo de los objetivos cognoscitivos.....	51
Tabla 3: Limitaciones más importantes de la TCT	58
Tabla 4:Diferencias entre la TCT y la TRI.....	66
Tabla 5:Dimensiones conceptuales básicas de la educación para la ciudadanía mundial	71
Tabla 6: Informe de progreso educativo: Honduras, 2017.....	80
Tabla 7:Fase de implementación de las reformas educativa	84
Tabla 8:Bloques y componentes que se evaluarán	93
Tabla 9: Operacionalización de la variable	94
Tabla 10: Distribución de la muestra para la prueba piloto	96
Tabla 11:Distribución de la muestra de la prueba operativa	97
Tabla 12:Escala estandarizada.....	106
Tabla 13 :Niveles de desempeño vinculado a la escala estandarizada	107
Tabla 14: Valores de discriminación para seleccionar los ítems	114
Tabla 15: Valores de dificultad para seleccionar los ítems.....	114
Tabla 16: Análisis de los reactivos Forma “A”	115
Tabla 17: Análisis de los reactivos Forma “B”.....	129
Tabla 18: Confiabilidad de la prueba con respecto a la forma A	143
Tabla 19:Confiabilidad de la prueba con respecto a la forma B	143
Tabla 20:Confiabilidad de la prueba operativa.....	146
Tabla 21: Confiabilidad con 45 reactivo	146
Tabla 22: Parámetro estadísticos de 45 reactivos de la prueba operativa	147
Tabla 23:Análisis estadístico descriptivo por bloques del área curricular de Ciencias Sociales	154
Tabla 24: Escala de puntuación de desempeño académico.....	156

Índice de Gráficos

Gráfico 1: índice de dificultad forma A y B	145
Gráfico 2: Rango de discriminación prueba operativa.....	149
Gráfico 3: Rango de dificultad prueba operativa.....	150
Gráfico 4: Total de estudiantes por género	151
Gráfico 5: Género por centro educativo	152
Gráfico 6: Porcentaje de edad de la muestra	153
Gráfico 7: Diagrama de caja del puntaje total de 45 reactivo de la prueba de noveno grado CCSS	155
Gráfico 8: Porcentaje de estudiante por nivel de desempeño	157
Gráfico 9: Nivel de desempeño académico por centro educativo	158
Gráfico 10: Nivel de desempeño académico por género	159

Índice de anexos

Anexo 1: Tabla de especificaciones para el CCSS	172
Anexo 2: Guión metodológico para el taller de evaluación educativa y redacción de ítems.....	173
Anexo 3: Consentimiento informado para el director	179
Anexo 4: Conocimiento informado para participantes en la investigación	182
Anexo 5: Asentimiento informado	183
Anexo 6: Descriptores para el área de ciencias sociales en noveno grado.....	184
Anexo 7: Instrumentos de recolección de información	187
Anexo 8: Prueba operativa	188
Anexo 9: Manual de instrucciones para la aplicación de la prueba piloto y operativa.....	189
Anexo 10: Escala de puntuación por centro educativo	190
Anexo 11:Taller de evaluación y redacción de ítems	191
Anexo 12: Docentes del área de Ciencias Sociales validando el instrumento	191

Glosario de abreviaturas

EACEA: Agencia Ejecutiva en el Ámbito Educativo

PEEP: Proyecto de Eficiencia de la Educación Primaria

SE: Secretaría de Educación

MIDEH: Proyecto Mejorando el Impacto del Desempeño Estudiantil de Honduras

USAID: Agencia de los Estados Unidos para el Desarrollo Internacional

UMCE: Unidad de Medición de la Calidad de la Educación

DCNB: Diseño Curricular Nacional Básico

CNB: Currículo Nacional Básico

PREAL: Programa de Promoción de la Reforma Educativa de América Latina y el Caribe

IEA: Asociación Internacional para la Evaluación del Logros Educativos

OREAL: Oficina Regional de Educación para América Latina y el Caribe

TCT: Teoría Clásica de los Test

TRI: Teoría de Respuesta al Ítem

CINE: Clasificación Internacional Normalizada de la Educación

FAREMA: Fundación para la Educación Ricardo Ernesto Maduro Andreu

INIEES: Instituto de Investigación y Evaluación Educativa

EFA:(Por sus siglas en inglés): Educación para Todos

CCSS: Ciencias Sociales

CISD: Consejo Internacional de Ciencias Sociales

UNESCO: Organización de las Naciones Unidas para la Educación, Ciencia y Cultura

IIMEC: Instituto para el Mejoramiento de la Educación Costarricense

Capítulo I: Planteamiento del Problema

1.1 Situación del problema/Construcción del objeto de estudio

En Honduras existe una preocupación por las bajas calificaciones que obtienen los estudiantes de los centros de educación básica, especialmente en los grados de 7mo, 8vo y 9vo; siendo este último donde se presentan más estudiantes con bajo rendimiento. Según el Informe Nacional de Desempeño Académico de la Secretaría de Educación y Proyecto Mejorando el Impacto al Desempeño Estudiantil de Honduras (2017), “en el nivel insatisfactorios el promedio de los estudiantes de los centros de educación básica es ligeramente más bajo en 7mo y 8vo grado con el 1% al contrario de 9vo grado con el 5%” (Secretaría de Educación y Proyecto Mejorando el Impacto al Desempeño Estudiantil de Honduras, 2017, pág. 54) independientemente del centro educativos al que estén matriculados los estudiantes de 9vo grado al menos seis de cada diez estudiantes se encuentran en los niveles de insatisfactorio y debe mejorar.

Sin embargo en 1990-1994 se aplicaron pruebas formativas estandarizadas a los primeros cinco grados en las cuatro asignaturas básicas (Matemáticas, Español, Ciencias Naturales y Ciencias Sociales) en diez (10) municipios, gracias al Proyecto de Eficiencia de la Educación Primaria (PEEP) que funcionó hasta 1995; a partir de 1997 hasta el 2004 se realizaron pruebas estandarizadas de español y matemáticas, a diferencia que en el año 2004 se aplicó prueba a la asignatura de Ciencias Naturales (Secretaría de Educación, 2009, pág. 27). Por otro lado, la SE- MIDEH mencionan que en el año 2016 se aplicaron por sexto año consecutivo pruebas para estudiantes de primero a noveno grado en las áreas de matemáticas y español, desde hace 19 años la SE, ha priorizado las pruebas estandarizadas en estas dos áreas curriculares.

Hasta este momento en Honduras no se cuenta con una prueba estandarizada que mida el desempeño académico en el área de Ciencias Sociales en los estudiantes del tercer (III) ciclo, el diseño de esta prueba se enfocará en los estudiantes de noveno grado ya que son los que presentan más dificultad en las áreas curriculares de matemáticas y español. El informe mundial sobre ciencias sociales, publicado con la colaboración de las Naciones Unidas para la Educación, la Ciencia y la Cultura y el Consejo Internacional de Ciencias Sociales menciona la ausencia y vacíos de investigaciones en el área de las ciencias sociales (Costa, 2017, pág. 6).

De acuerdo Moncada & Abuslim (2016) en nuestro país, las evaluaciones de los aprendizajes se empezó con el Proyecto PEEP-USAID, seguido por la UMCE y finalmente por MIDEH, en los dos últimos casos han venido evaluando las mismas área curriculares (español y matemáticas) en los centros de educación básica dejando a un lado las otras dos (Ciencias Naturales y Ciencias Sociales), que están establecidas en el Diseño Curricular Nacional Básico (Moncada & Abuslim, 2016, pág. 304) ; se está preparando a los estudiantes para que sean competitivos en dos áreas curriculares (matemáticas y español) cuando la educación es permitir el desarrollo de los individuos que la sociedad requiera, con una formación académica integral para que los jóvenes del mañana desarrollen una serie de destrezas y habilidades que se deben ampliar en todas las áreas del conocimiento (Reyes, 2016, pág. 43).

Por otro lado Cueto (2009) menciona que el fin de la educación es un esfuerzo programado desde la sociedad para actualizar el potencial humano de todos de una manera integral, en otras palabras es incluir conocimientos, valores y conductas vinculadas a los diferentes aspectos del currículo, más adelante agrega que la educación debería integrar diversos conocimientos para el desarrollo de la educación ciudadana o democrática, relacionado con el desarrollo individual y de los grupos humanos, este último abarcaría las Ciencias Sociales, pues este sería la guía al uso de habilidades académicas, siempre y cuando el desarrollo de la educación esté vinculada con la instrucción individual, pero también social, y de este modo contribuir a la formación de sociedades más justas (Cueto, 2009, pág. 40)

La problemática antes planteada es una realidad que involucra a todos los hondureños, es preciso integrarse para mejorar la calidad de la educación del país, a medida que se afronta el reto en participar en las evaluaciones estandarizadas en español y matemática; así mismo inculcar las evaluaciones estandarizadas en las Ciencias Sociales de una forma periódica; una ventaja es que se cuenta con las experiencias de las dos áreas antes mencionadas, y de esta manera se fomentará una cultura de evaluación que será significativa para el sistema de educación en el nivel básico, siendo esto un precedente para diseñar una prueba piloto en el área curricular de las Ciencias Sociales.

1.2. Objetivos de la investigación

1.2.1. Objetivos General

Desarrollar una prueba de fin de grado que evalúe el desempeño académico de los estudiantes de noveno grado en el área curricular de Ciencias Sociales, de Tegucigalpa 2018.

1.2.2. Objetivo Específicos

1. Clasificar los estándares de contenido según los bloques y componentes del Diseño Curricular Nacional Básico en el área de Ciencia Sociales para medir el desempeño académico de los estudiantes de noveno grado.

2. Diseñar los ítems de una prueba que cumpla con las características psicométricas requeridas que evalúe el desempeño académico de los estudiantes de noveno grado en área de Ciencias Sociales.

3. Determinar las propiedades psicométricas de la prueba que evalúe el desempeño académico de los estudiantes de noveno grado en el área de Ciencias Sociales.

4. Describir el desempeño académico de los estudiantes de noveno grado en el área de Ciencias Sociales.

1.3. Justificación

La necesidad de proporcionar información pertinente para ayudar a que la Secretaría de Educación pueda planificar posibles mejoras en el proceso de enseñanza – aprendizaje y así contribuir a que la educación de Honduras sea de calidad. *La importancia de este estudio* es el diseño y la validación de una prueba de desempeño académico en el área de Ciencias Sociales para los estudiantes de noveno grado. Se conoce que este tipo de evaluación es un tema de interés y que en la actualidad se realiza hasta el nivel básico (Primero a noveno grado). Es necesario mencionar que la Secretaría de Educación con la ayuda de organizaciones internacionales, han realizado importantes esfuerzos para participar en las evaluaciones del desempeño académico, dichos avances se han elaborado en el área curricular de español y matemáticas.

La relevancia social del diseño y la validación de esta prueba es que en el país no se cuenta con una evaluación del desempeño académico en el área de Ciencias sociales, esta área también debe ser incluida en el proceso de una evaluación externa ya que posee autonomía e independencia en el Diseño Curricular Nacional Básico, analizar el desempeño académico en esta área es un desafío para los encargados de la evaluación ya que se precisa saber con exactitud qué es lo que aprende el estudiante en esta área. Para esto es necesario conocer cuánto aprenden los estudiantes en esta asignatura y cuáles son las condiciones en que lo hacen. Es importante que esta área cuente con una evaluación externa ya que las Ciencias Sociales son fundamentales en el desarrollo del ser humano.

El aporte teórico que de este estudio es brindarle a la sociedad hondureña una prueba que cumpla con los criterios de calidad: pertinencia, confiabilidad y validez, además la prueba diseñada ayudará a resolver uno de los problemas que presenta la Secretaría de Educación ya que hasta el momento no cuenta con una prueba estandarizada para medir el desempeño académico de los estudiantes de noveno grado en Ciencias Sociales. En cuanto a lo metodológico el estudio aportará un nuevo instrumento de medición en el desempeño académico orientada al noveno grado de educación básica, la cual estará a disponible para todas las personas que deseen conocer el desempeño de los estudiantes en esta área que se está analizando.

Capítulo II: Marco Teórico

En el presente marco teórico se les proporcionará una amplia información sobre los conceptos básicos para la elaboración de una prueba de desempeño académico (evaluación educativa, medición de los aprendizajes, estándares de contenido, estándares de desempeño, evaluación psicométrica y el desempeño académico), también se encontrará la evaluación de los aprendizajes dentro de este, incluirá las evaluaciones estandarizadas en el mundo y los subtemas que estará compuesto por la evaluación estandarizada en Europa, evaluación estandarizada en Latinoamérica y evaluación estandarizada en Centroamérica; otro tema que abordaremos en este apartado es el diseño y validación de un instrumento de evaluación (vínculo entre los objetivos y el instrumento que se elaborará o seleccionará, tipo de instrumentos, calidad de los instrumentos, taxonomía de objetivos cognoscitivo, taxonomía de objetivos afectivos y psicomotrices, marco conceptual y tabla de especificaciones, elaboración de los reactivos del test), principios para el desarrollo de las pruebas o teoría de la medición (teoría clásica de los test y teoría respuesta al ítem); después se contará con el apartado de las Ciencias Sociales y el mundo (se abordará la desigualdad y los cambios ambientales).

2.1 Conceptos básicos para la elaboración de una prueba de desempeño académico

Para mejor dominio de las generalidades que se tienen que tener en cuenta para la elaboración de una prueba, se comenzará brindando algunos significados sobre los siguientes enunciados: evaluación educativa, medición de los aprendizajes, estándares educativos (está compuesta por dos: estándares de contenido y estándares de desempeño), evaluación psicométrica y desempeño académico.

2.1.1. Evaluación educativa

La evaluación educativa está vinculada con los procesos de enseñanza-aprendizaje, establecer la evaluación como una temática demandante y en particular en un debate educativo, pero las afirmaciones son muy variadas con respecto a las dificultades que presenta el campo educativo que se dirige al hecho de evaluar y es aquí donde se presenta diversas opiniones o significado acerca de la evaluación, con respecto a sus propósitos, planificación y su práctica. Es importante mencionar que todas las prácticas de la evaluación

de los individuos y centros educativos vienen a ratificar nuevas maneras de enseñar ya que esta se convierte en una evaluación formativa la cual enriquece el proceso de enseñanza-aprendizaje del individuo, considerando las necesidades y las características propias de cada sujeto (Escorcía, 2008, pág. 9).

Por otro lado, recapacitar sobre la evaluación educativa es un trabajo continuo e invariable, que produce de pronto diversos dictámenes y debates del mismo. La responsabilidad de la persona encargada de evaluar tiene el trabajo más significativo ya que es el que dirige los procesos de formación del sujeto que se encuentran en las instituciones educativas. De acuerdo con Escorcía (2008) donde expresa “(...) todo proceso de aprendizaje va ligado íntimamente a procesos evaluativos, que se convierten en obligatorios en la medida en que se exigen lineamientos, objetivos y estándares propios de los procesos de enseñanza” (Escorcía, 2008, pág. 1). Por consiguiente, podríamos decir que enseñar involucra evaluar y viceversa.

Para Islas (2010) la evaluación es “el procedimiento sistemático y comprensivo en la cual se utilizan múltiples estrategias (...) por lo tanto es un conjunto de estrategias destinadas a la mejorar de la calidad de la enseñanza” (Islas, 2010, pág. 2) . Sin embargo, Escorcía (2008) lo considera como sinónimo de “valorar, estimar, examinar, calcular, acreditar, ponderar, apreciar, puntuar, criticar, ajustar, tallar, medir, graduar, calificar y hasta juzgar; sin, embargo, estas consideraciones inmediatas se hacen más claras y reveladoras a la luz de los procesos de enseñanza propios.” (Escorcía, 2008, pág. 2)

A pesar de que la autora lo considera un sinónimo todos buscan el mismo objetivo; valorar el conocimiento, las habilidades que han logrado los estudiantes en el desempeño académico en el curso que han sido matriculado. Por lo tanto, es necesario ver a la evaluación como una parte integral del proceso de enseñanza aprendizaje. Concuerdo con Murillo & Román (2008) cuando expresan que “la evaluación es el único instrumento que tenemos para verificar el cumplimiento de ese derecho superior, internacionalmente reconocido y que supone el acceso igualitario al conocimiento y a las oportunidades disponibles en la sociedad” (Murillo & Román, 2008, pág. 2)

2.1.2. Medición de los aprendizajes

Ya se ha conocido un poco lo que es la evaluación educativa; en este enunciado brindaremos algunos significados del término de medición de los aprendizajes. En la cual

empezaremos citando a Ayala (2006) ella nos menciona que cuando medimos establecemos valores a los datos o cualidades observadas, a la vez nos expresa que medición se refiere a la representación numérica de datos o cualidades (Ayala, 2006, pág. 1)

Es decir que la medición es brindarle un número a una variable (el aprendizaje es una variable), en cambio la calificación numérica pertenece a la puntuación obtenida por un individuo en los diferentes tipos de test diseñados para medir el resultado del aprendizaje, la medición puede ser dura si es ejecutada en un ambiente controlado y es a la vez una descripción rigurosa e irrefutable que posee un significado constante en el ambiente científico y profesional. En otras palabras, medir significa reemplazar los datos por números.

Por otro lado, Arredondo & Diago (2006) expresan que la evaluación es muy amplia y que engloba a los dos términos (la calificación y medición), pero también hacen la aclaración que no se identifican con ellos, como todos saben que para evaluar se necesita medir, para que se pueda realizar la medición se necesita recoger todos los datos que sean necesario y de la manera más objetiva posibles para que los datos sean confiables sin embargo, no es suficiente solo medir ya que la medición tiene que ser interpretada a la hora de evaluar por lo tanto los datos tiene que contar con criterios de evaluación para cada objetivo. (Arredondo & Diago, 2006, pág. 388)

2.1.3 Estándares Educativos

La palabra estándar nació en 1862 en Inglaterra, cuando se estableció una ley donde mencionaba que se debía prever el pago de subsidios a cada escuela en función a los logros que obtenía los estudiantes. Para que el propósito se cumpliera se le asigna supervisores a cada centro educativo, dicho supervisores se fijaban seis niveles diferentes de logro, denominados “estándares”, del 1 al 5, lo supervisores eran los encargados de clasificar a los estudiantes en exámenes anuales de lectura y cálculo, con los resultados obtenidos en estas dos áreas, se les asignaban un estándar a las escuelas que deberían de corregir, esta leyes se eliminó formalmente en 1895, sin embargo la palabra estándar se siguió utilizando en Gran Bretaña. El concepto de estándar se entiende como norma, pauta, tipo o regla (Camilloni, 2009, pág. 56).

Al pasar de los tiempos la palabra estándar se fue aplicando a gran diversidad de cosas como: bienes y servicios productivos, se aplicó también a las instituciones, a las personas, los conocimientos de los aprendizajes y las competencias (Camilloni, 2009, pág. 56). Sin

embargo, el Ministerio de educación de Guatemala/USAID (2007) Ministerio de Educación de Guatemala/USAID (2007) expresa que:

Los estándares son enunciados que establecen criterios claros, sencillos y medibles que los maestros y maestras deben considerar como meta del aprendizaje de sus estudiantes, que se traducen en lo que debe saber y saber hacer. Dicen lo que se espera lograr en el proceso de enseñanza-aprendizaje. (Ministerio de Educación de Guatemala/USAID, 2007, pág. 6)

Es decir que los estándares es un modelo de esperanza que el estudiante salga con sobresaliente, ya que proponen metas que desafían al estudiante, por medio de la investigación y el desarrollo del conocimiento, además ayudan a establecer vínculos entre conocimientos previos y los conocimientos en proceso de enseñanza-aprendizaje. Por otro lado, la Secretaría de Educación (2011) expresa que los estándares son objetivos educativos que señalan lo que los estudiantes tienen que saber (se refiere a conocimiento) y saber hacer (se refieren a las habilidades y a las destrezas) independientemente del contexto geográfico, cultural o social. Según la misma Secretaría los estándares deben de contar con las siguientes características:

Definen los que los estudiantes deben saber y saber hacer, deben estar en consonancia con las aspiraciones nacional en materia educativa, deben de ser sencillo, claros, priorizado, alcanzables y medibles, deben ser entendibles sobre todo por docentes, estudiantes y padres de familia, deben ser de alta calidad y comparables con estándares internacionales. (Secretaría de Educación, 2011, pág. 2)

Los puntos expuestos por la Secretaría de Educación de Honduras son muy importantes, sin embargo, en la característica que expresa que deben ser entendibles sobre todo por docentes, estudiantes y padres de familia, están fallando ya que los estándares solo son conocidos por las autoridades de los centros educativos, docentes y estudiantes.

Además de las características la Secretaría de Educación nos proporciona la finalidad de estos estándares las cuales son las siguientes:

Coordinar todos los elementos del sistema educativo, servir como norma para orientar la labor de los docentes, servir como base en el establecimiento de criterios para el diseño de pruebas, orientar la formación inicial y la capacitación permanente de docentes en servicio, comunicar a los distintos grupos (directivos, docentes, padres de familia) lo que los estudiantes deben saber y saber hacer, permitir contrastar el rendimiento entre instituciones

o regiones del país, al establecer criterios de evaluación precisos para realizar evaluaciones internas y externas (Secretaría de Educación, 2011, pág. 2)

Una vez explicado lo anterior, es necesario mencionar que en Guatemala se cuentan con tres tipos de estándares: de contenido, desempeño y de oportunidad, este último describe las condiciones necesarias para el aprendizaje tales como: ambientes escolares apropiados, maestros capaces, materiales que ayuden al proceso de enseñanza-aprendizaje, y el tiempo de aprendizaje en la escuela y en el hogar (Ministerio de Educación de Guatemala/USAID, 2007, pág. 6). Sin embargo, en Honduras, Nicaragua, El Salvador, Panamá, solo existen dos tipos de estándares: los estándares de contenidos y los estándares de desempeño (Secretaría de Educación, 2011, pág. 1) (Ministerio de Educación de Nicaragua, 1999, pág. 2) (Programa de Promoción de la Reforma Educativa en América Latina y el Caribe, 2002, pág. 16) (Sistema Nacional de Evaluación de la Calidad de los Aprendizajes, 2007, pág. 16)

2.1.3.1. Estándares de contenido

Los estándares pueden ser definidos como criterios claros y públicos que establecen los parámetros de lo que los alumnos pueden, deben saber y hacer en cada una de las asignaturas de los planes de estudios correspondientes a los diferentes niveles por lo tanto se podría decir que los estándares son una meta para que todas las instituciones del sistema educativo cuenten con una educación de calidad en el cual se espera que ofrezcan a los estudiantes y los resultados que deben alcanzar para lograr dicha calidad (Cabrero, 2010, pág. 21).

Para Cabrero (2010) y Lansdale, Vukmirovic, & Estrada (2012) coinciden que los estándares de contenido se refieren al conocimiento y las habilidades que los estudiantes deben adquirir en una asignatura particular. En otras palabras, los estándares de contenido son todos aquellos conocimientos que dichos estudiantes deben saber y ponerlos en práctica (Cabrero, 2010, pág. 21) (Lansdale, Vukmirovic, & Estrada, 2012, pág. 153). Por otro lado, el Ministerio de Educación de Guatemala /USAID (2007) menciona que los estándares de contenido son los que describen los conocimientos y destrezas que deben tener los estudiantes, es decir el “que” de lo que los estudiantes deben saber y poder hacer (Ministerio de Educación de Guatemala/USAID, 2007, pág. 4)

2.1.3.2. Estándares de desempeño

Al momento de referirse a los estándares de desempeños los mismos autores citados anteriormente expresan: “Los estándares de desempeño los constituyen ejemplos concretos y definiciones explícitas de lo que los estudiantes tienen que saber y ser capaces de hacer para demostrar su pericia en las habilidades y el conocimiento que están delineados en los estándares de contenido” (Cabrero, 2010, pág. 4). Sin embargo, Lansdale, et al. (2012) opinan que “Los estándares de desempeño establecen los niveles de aprendizaje logrado por los estudiantes. Las evaluaciones alineadas a los estándares sirven para determinar el nivel de dominio de los estándares por los estudiantes.” (Lansdale, Vukmirovic, & Estrada, 2012, pág. 153) . Los niveles de desempeño se refieren al grado en que los estudiantes dominan los estándares de contenido. Además, agregan:

Definir estándares de desempeño es un procedimiento para conceptualizar los niveles de desempeño que se usarán para evaluar los aprendizajes: se trata de decidir acerca del número y el propósito de los niveles de desempeño, escoger los nombres de los niveles de desempeño (Etiquetas de Niveles de Desempeño) y desarrollar descripciones generales y específicas de los niveles de desempeño (Descriptorios de Niveles de Desempeño) (Lansdale, Vukmirovic, & Estrada, 2012, pág. 157)

Los estándares de desempeño muestran que tanto los estudiantes deben aprender y hacer lo que está determinado en los estándares de contenido, estos vienen siendo los indicadores de calidad que detallan cómo debe ser las competencias de los estudiantes; en este apartado se enuncian en forma precisa y sintetizada los niveles de desempeño que deben ser alcanzados además se ofrece una escala cualitativa y cuantitativa que permite interpretar el desempeño académico de los estudiantes en el alcance de los estándares de contenido (Ministerio de Educación de Guatemala/USAID, 2007, pág. 6).

En conclusión, los estándares de desempeño dependen de que tanto contenido lograron aprender los estudiantes, para poder ubicarlo en el nivel que le corresponde estar; para que se obtenga una excelente puntuación el docente tiene que brindar los contenidos de manera objetiva y con distintos métodos para que todos los estudiantes logren la adquisición de ese contenido.

2.1.4 Evaluación Psicométrica

El término “psicometría” procede del griego “psykhé” (mente) y “metron” (medida), lo que quiere decir literalmente medir los rasgos psicológicos de una persona. La evaluación psicométrica resultante se realiza principalmente en forma de test estandarizados (Central Test, 2013, pág. 5). Se entiende por Psicometría el conjunto de modelos formales que establecen las bases para que se lleve a cabo de manera adecuada la medición de variables, con el objetivo de proporcionar los métodos para poder transformar los hechos en datos, mediante la asignación de valores numéricos tanto a las respuestas dadas por los sujetos como a los estímulos presentes en la situación de prueba, y poder interpretar esos números en función de la teoría sobre la que la prueba descansa (Aragón, 2015, pág. 56).

Los términos psico y metría, significan: medida de los fenómenos psíquicos. Por tanto, los tests psicométricos intentan medir habilidades cognitivas o rasgos de personalidad, en las personas estudiadas. Los tests psicométricos son instrumentos estructurados, es decir, en ellos la persona tiene que escoger, entre alternativas de respuestas posibles, aquella que considera se ajusta mejor en su caso particular (González, 2007, pág. 8)

2.1.5 Desempeño académico

El desempeño académico es una problemática que preocupa a los padres de familia, estudiantes, docentes y a las autoridades de nuestro país, también es una preocupación en varios países latinoamericanos. El significado de desempeño académico es muy complejo, empezando desde su definición, de acuerdo con Lamas (2015) en su artículo sobre el rendimiento académico el expresa que “el desempeño académico en ocasiones se le denomina aptitud escolar y rendimiento académico, es necesario aclarar que las diferencias de conceptos solo se explican por cuestiones de estudio ya que se utilizan como sinónimos” (Lamas, 2015, pág. 315), sin embargo, Delgado & Palos (2003) mencionan que el desempeño puede ser expresado por medio de una calificación proporcionada por el docente o el promedio que ha logrado el estudiante además agrega que el promedio es el resumen del desempeño académico. (Delgado & Palos, 2003, pág. 6)

Los autores antes mencionados concuerdan que el desempeño académico es el resultado de múltiples factores que influyen en el proceso de aprendizaje que se les atribuye a los logros de los estudiantes teniendo como intervención a los docentes, además el desempeño académico no sólo se trata de obtener una nota por parte de los estudiantes también aumentan

el grado de satisfacción psicológica de los individuos involucrados. Por lo tanto, el desempeño académico viene siendo un indicador de las capacidades de un estudiante que brinda una rendición de cuenta de todo lo que aprendido a lo largo del proceso de enseñanza aprendizaje ya que este tiene que responder a los estímulos brindados por el sistema educativo.

Para Guzmán (2012) “el desempeño académico es el resultado de un proceso educativo que puede ser medido en términos cuantitativos y cualitativos, que sirve para retroalimentar a los individuos y a las instituciones sobre el logro de los objetivos que se establecieron previamente” (Guzmán, 2012, pág. 57)

De acuerdo con Molinare, Meriño, & Barro (2014) “El rendimiento académico es un problema educativo que afecta al desarrollo de la sociedad y que ha generado preocupación en nuestro país, especialmente por los bajos rendimientos en las evaluaciones estandarizadas” (Molinare, Meriño, & Barro, 2014, pág. 154). También concuerdo con Martínez (2013) al expresar que el desempeño académico cuenta con tres dimensiones fundamentales en el estudiante, la primera Aptitudes (el poder) es decir, la capacidad que tiene el estudiante para obtener una buena nota; la segunda dimensión es Actitud (es el querer) cuanto está dispuesto hacer es estudiante para obtener es buena nota y la tercera dimensión es la metodología (Saber hacer) estas son las habilidades y destrezas que posee el estudiante para realizar eficientemente las asignaciones por el docente (Martínez, 2013, págs. 3-4)

2.2. Evaluación de los aprendizajes

Una de las tareas más importantes que efectúa el docente en el proceso de enseñanza-aprendizaje es la evaluación que aplica a cada uno de los estudiantes. Es preciso que la sociedad esté reflexiva sobre cómo es entendida la evaluación hoy, ya que los docentes y los encargados de perpetrar las evaluaciones estandarizadas se plantean las siguientes preguntas ¿para qué sirve?, ¿Qué evaluar?, ¿Cuándo y cómo se evalúa?; estas preguntas se las plantean con el fin de alcanzar los aprendizajes que están plasmados en el DCNB.

Para Rivera (2007) “La evaluación permite detectar si los problemas en el aprendizaje son generados por el propio docente, los estudiantes, los contenidos, las condiciones de enseñanza que se están propiciando” (Rivera, 2007, pág. 21) . Es trascendental expresar que, sin la información alcanzada mediante la evaluación, no poseeríamos fundamentos para realizar mejorar en el proceso de enseñanza-aprendizaje.

La evaluación de los aprendizajes se ha convertido en uno de los procesos más difíciles de abordar en la actualidad, y no sólo por los profesores sino también para los centros educativos, este proceso se ha tornado muy complejo debido a las desiguales características que poseen un estudiante en el salón y además por las diferentes tipologías que poseen planes de estudios o el CNB, en los cuales se plantea los contenidos de aprendizajes por áreas del saber y se enfatiza en la adquisición de conocimientos conceptuales, procedimentales y actitudinales. (Guadrón, Matos, Castro, & Mendi, 2013, pág. 94).

La realización de la evaluación de los aprendizajes se ha vuelto una parte fundamental en las instituciones educativas, los docentes efectúan una evaluación al terminar cada semestre algunos la ejecutan de manera intermedia y otra parte de los docentes al final. Como es de conocer en el año 2013 la Secretaría de Educación realizó una Evaluación de aprendizajes a los docentes del todo país, Honduras empieza a recorrer un largo camino como es la evaluación, por lo tanto, es necesario conocer diferentes conceptos de lo que es evaluación de los aprendizajes, lo cuales se brindaran a continuación.

Para Gil (2007) “La evaluación de los aprendizajes es un instrumento para ejercer el poder y la intimidación” (Gil, 2007, págs. 212-213). Por otro lado, Castillo (2009) expresa que “La evaluación de los aprendizajes es un proceso constante de producción de información para la toma de decisiones, sobre la mejora de la calidad de la educación en un contexto humano social, mediante sus funciones diagnóstica, formativa y sumativa” (Castillo, 2009, pág. 1).

Sin embargo, Pereira (2015) expresa que la evaluación de los aprendizajes es:

La valoración de los procesos de enseñanza y aprendizajes mediante el diálogo entre los participantes del hecho educativo para determinar si los aprendizajes han sido interiorizados y tienen sentido y valor funcional. Además, lleva a la reflexión sobre el desarrollo de las competencias y los logros alcanzados. Es un proceso pedagógico. Sistemático, instrumental, analítico y reflexivo. (Pereira, 2015, pág. 408)

De acuerdo con Guzmán (2010) menciona que evaluar es emitir un juicio de valor de acuerdo a ciertas normas establecida por la persona encargada de realizar la evaluación de los aprendizajes el autor lo expresa de la siguiente manera:

evaluar implica emitir juicios de valor, juzgamos si algo es “bueno”, “mejor” o “no satisfactorio” o “deficiente”, conforme a ciertos valores o parámetros; por eso, conlleva siempre la comparación de lo encontrado o existente con valores, normas o criterios de lo

que se considera adecuado o digno de lograr, la evaluación debe realizarse de manera sistemática, rigurosa y con recopilación de información ha de valerse tanto de lo cuantitativo como lo cualitativo y cumplir con los requisitos técnicos de tener validez, confiabilidad, representatividad, poder discriminatorio y factibilidad. (Guzmán, 2010, pág. 3)

La evaluación de los aprendizajes se entiende como aquella acción dentro del proceso de enseñanza y aprendizaje donde los y las docentes de manera particular o los organismos privados y públicos habilitados para ello determinan a través de pruebas relativamente objetivas, unas más que otras, el nivel de logro en términos de adquisición de contenido, o si se prefiere de competencia, por el estudiante luego de haber sido expuesto a las prácticas y actos pedagógicos diseñados para él (Delgado & Rivilla, 2014, pág. 24).

Con los conceptos presentados anteriormente de distintos autores, se han clasificado algunas palabras de dichos conceptos ya que para nuestro parecer son claves para realizar una evaluación de manera eficiente por lo tanto diremos que: la evaluación del aprendizaje es un proceso pedagógico que se da de manera instrumental, sistemática analítica y reflexiva ya que por medio de esta se recopila y se produce información de manera cuantitativa y cualitativa, siempre y cuando la evaluación sea elaborada con los requisitos técnicos como ser la validez, confiabilidad, poder discriminatorio y factibilidad, todo esto con el fin de mejorar la calidad de la educación.

La importancia de la evaluación de los aprendizajes ha sido un tema de interés colectivo, sin embargo, muchos autores manifiestan que la importancia radica en informar a partir de los resultados obtenidos, y por otro lado permite dar la pauta para orientar la toma de decisiones hacia la búsqueda de la mejora de la calidad de la educación. Es así, que coinciden en que el fin primordial de las evaluaciones de los aprendizajes es la “utilización de los resultados”, siendo este el concepto que predominó en las primeras etapas del desarrollo histórico de la evaluación (Dirección Nacional de Información y Evaluación de la Calidad Educativa / Ministerio de Educación, 2009, pág. 78). Hay que tener en claro que la evaluación es un proceso de investigación continuo, sistemático y participativo que tiene como propósito la obtención, análisis e interpretación de información para establecer juicios de valor que sirven de base en la toma de decisiones con el fin de lograr el perfeccionamiento del objeto y adoptar decisiones que conduzcan a la mejora (Guevara, Hernández, & Cabrera, 2014, pág. 15).

De acuerdo con Córdova & Lozano (2015) la evaluación de los aprendizajes se ha vuelto un interés a nivel nacional e internacional no solamente en los centro educativos, actualmente se efectúan procesos de evaluación estandarizadas a nivel mundial, por ejemplo, PISA (Programa Internacional para la evaluación de estudiantes), MIDEH (Mejorando el Impacto al Desempeño Estudiantil en Honduras), SERCE (Segundo Estudio Regional Comparativo); así como estos existen diferentes organizaciones que están interesado en la evaluación de los aprendizajes (Córdova & Lozano, 2015, pág. 133)

Normalmente las prácticas de evaluación de los aprendizajes tradicionalmente se limitan a examinar y brindar una calificación, pero no ven más allá de eso; no comprenden las causas y consecuencias de las dificultades en las que se encuentran los estudiantes en el proceso de aprendizaje. Medina, García, Martínez, & Abundez (2011) ven a la evaluación como:

Un potencial único para mejorar el desempeño de los alumnos. Su importancia es tal que puede ser el eje de toda la enseñanza, pues a partir de ella se podría definir qué, cómo y cuándo se enseña. No obstante, para que la evaluación explote su potencial se requiere que esté estrechamente ligada a los procesos de enseñanza y aprendizaje. De esa forma, los diferentes actores que intervienen en el quehacer educativo (docentes, alumnos, padres de familia y la escuela como institución) se verán beneficiados. (Medina, García, Martínez, & Abundez, 2011, pág. 11)

De algo si hay que estar seguros que las evaluaciones de los aprendizajes han tenido un auge significativo en la actualidad y que nos ayuda hacernos una idea del avance que han tenido los estudiantes. Una vez proporcionada algunas definiciones de evaluación de los aprendizajes, en este apartado hablaremos de la evaluación estandarizada en el mundo, desglosados en los subtemas siguientes: Evaluación estandarizada en Europa, Evaluación estandarizada en Latinoamérica, Evaluación estandarizada en Centroamérica.

2.2.1. Evaluación estandarizada de aprendizaje.

En los últimos años el sistema educativo se ha preocupado por evaluar el aprendizaje a nivel nacional e internacional; otorgando a la evaluación el camino para ayudar a que la educación sea de calidad. Es por ello, que en la actualidad ha surgido el interés de realizar evaluaciones estandarizadas de los aprendizajes en el campo educativo; permitiendo de esta manera que llegue al alcance de los políticos, económicos y sociales; para saber si los estudiantes están obteniendo un aprendizaje significativo, el principal indicador es el

desempeño académico y a la vez este indica el funcionamiento del sistema educativo (López & Lever, 2016, pág. 132).

Sin embargo López & Lever (2016) expresan que “en el ambito internacional, diversas organizaciones instrumentan evaluaciones, que involucran la aplicación de pruebas objetivas, enfocandas en la medición de conocimientos, habilidades, dominios específicos y competencias de estudiantes” (López & Lever, 2016, pág. 133) Estos autores mencionan que las evaluaciones más destacadas internacionalmente son:

➤ El Estudio Internacional de Tendencias en Matemáticas y Ciencias (TIMSS) y el Estudio Internacional de Progreso en comprensión Lectora (PIRLS): TIMSS (por sus siglas en inglés) se aplicó en 1995 por primera vez y se realiza cada cuatro años, en cambio PIRLS (por sus siglas en ingles) ha sido el tercer estudio internacional, después de PIRLS 2001 y 2006. La primera se centra en valorar los conocimientos en matemática y ciencias a los estudiantes de los grados de cuarto y octavo de educación básica; el segundo evalúa las tendencias en el rendimiento en comprensión lectora de los estudiantes de cuarto grado de educación básica, es importante mencionar que los dos proyectos son instrumentado por la Asociación Internacional para la Evaluación del Logro Educativo (IEA por sussiglas en inglés), en estos estudios han participado más de 50 países (López & Lever, 2016, pág. 133) .

➤ Programa para la Evaluación Internacional de Estudiantes (por sus siglas en PISA): Este programa se ha condesado en evaluar una área en específica cada tres años, en estudiantes de 15 años; dichas áreas son: lectura (2000), matemáticas (2003) y ciencias (2006), llevando a cabo su segunda etapa de evaluación en: lectura (2009), matemáticas (2012) y ciencias (2015); este programa cuenta con el apoyo de la Organización para la cooperación y el Desarrollo Económico (OCDE), el objetivo de esta evaluación es identificar el nivel de habilidades necesarias que han logrado los estudiantes para participar completamente en la sociedad. Lo que busca PISA en los estudiantes es verificar si todo lo aprendido por los estudiantes pueden reproducirlo en la realidad; las competencias que evalua el programa son: capacidad de análisis, razonar y comunicar sus ideas de manera efectiva (López & Lever, 2016, pág. 133) .

➤ Estudios Comparativos y Explicativos (en Latinoamérica): Están compuesto por PERCE (Primer Estudio Regional Comparativo y Explicativo) y SERCE (Segundo Estudio Regional Comparativo y Explicativo): En el primer caso (PERCE), se aplicó por primera vez en 1997 y se evaluó matemáticas y lectura en tercer y cuarto grado de educación básica. En el caso de SERCE se aplicó en el 2006, nueve años más tardes; las áreas que se evaluó fueron matemáticas y lectura en tercer y sexto grado, para sexto grado se agregó Ciencias Naturales. El tercer estudio se realizó en el año 2013 evaluando las mismas áreas y grados. Estos estudios han sido realizado por el Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE), se fundó en 1994 como una red de unidades de medición y evaluación de la calidad educativas de la región y técnicamente fue coordinado por el equipo de la Oficina Regional de Educación para América Latina y el Caribe (OREAL/UNESCO). En estos estudios se emplean las pruebas estandarizadas de evaluación de aprendizaje en las diferentes áreas curriculares a estudiantes de educación básica. Es importante mencionar que LLECE brinda información y conocimiento que retroalimentan la política educativa, por medio de los informe que manifiestan el estado de la calidad de la educación; En estos estudios participaron 15 países: Brasil, Argentina, Colombia, Chile, Guatemala, Ecuador, México, Nicaragua, Honduras, Paraguay, República Dominicana, Uruguay, Perú, Costa Rica, Panamá (López & Lever, 2016, pág. 133).

Además de los programas mencionados anteriormente, es importante mencionar que en la región existen otras instituciones encaminadas en la evaluación educativa las cuales destacan las siguientes:

➤ El Grupo de Trabajo Estándares y Evaluaciones (GTEE): Este grupo se encuentra conformado por el Programa de Promoción de la Reforma Educativa en América (PREAL), que desde 1999 ha realizado un seguimiento sistemático del desarrollo de los sistemas educativo nacionales e internacionales de evaluación de aprendizaje en América Latina. Así mismo el grupo se ha responsabilizado de efectuar el monitoreo permanente de los avances en las políticas curriculares, principalmente en el desarrollo de estándares de currículo (Ferrer, 2007, pág. 3).

- Instituto de Evaluación y Asesoramiento Educativo (IDEA): En este instituto participan los siguientes países: Puerto Rico, Brasil, México, Chile y se creó con el objetivo de colaborar con los centros educativos de los países partícipes en el conocimiento y la mejorar de la calidad en procesos de enseñanza aprendizaje, lo cual se realiza una evaluación independiente y externa.

Por otro lado SE/MIDEH (2016) en su Informe Nacional de Desempeño Académico 2016 Español y Matemáticas. 1ro a 9no grado, nos expresa que se ha desarrollado en el sistema nacional de evaluación estandarizada con el objetivo de conocer el desempeño y el resultado de aprendizaje en el sistema educativo. Así mismo hacer referencia a las evaluaciones internacionales estandarizadas tales como el Programa para la Evaluación Internacional de Alumnos (PISA), Programa Internacional del Progreso en Comprensión Lectora (PIRLS), Programa Internacional de Tendencias en Matemáticas y Ciencias (TIMSS), Tercer Estudio Regional Comparativo y Explicativo (TERCE), El Consorcio de África Meridional y Oriental para el Monitoreo y Supervisión de la calidad Educativa (SACMEQ).

2.2.2. La evaluación estandarizada en el mundo

En los últimos años la evaluación estandarizada se ha convertido en una pieza clave para satisfacer las diferentes necesidades del sistema educativo. Meliá (2017) menciona que “la estandarización es el procesamiento de datos de todos los elementos de una aproximación a la recolección de información e interpretación de manera que se utilicen los mismos instrumentos con los criterios de corrección, de análisis para realizar la interpretación de la misma” (Meliá, 2017, pág. 5). En otras palabras, las evaluaciones estandarizadas son pruebas que se administran y se les asigna una calificación siguiendo un procedimiento estándar. Existen dos tipos primordiales de pruebas estandarizadas: Las pruebas de aptitud y las pruebas de logro; la primera hace referencia a la capacidad o el desempeño del individuo en alguna área o nivel educativo.

Estoy de acuerdo con Camilloni (2009) cuando nos menciona que “los test estandarizados se diferencian de las pruebas informales preparadas por el docente, por el proceso de su construcción” (Camilloni, 2009, pág. 59) estos se diferencian porque en las pruebas estandarizadas se delimitan con precisión, los conocimientos y destrezas a evaluar y se realiza una prueba piloto antes de incorporar los ítems a la prueba operativa Se administra la prueba a una muestra representativa de la población, se perfecciona hasta lograr la definitiva. Las

evaluaciones estandarizadas son todas aquellas pruebas que han sido normalizadas o aprobadas por una población determinada (estudiantes, docentes o dependiendo del estudio que se realizará), estas pruebas tendrán que cumplir con algunos principios generales como: la confiabilidad y la validez de cada test elaborado.

Sin embargo, el Ministerio de Educación de España (2010) en su informe titulado “Pruebas nacionales de evaluación del alumnado en Europa: objetivos, organización y utilización de los resultados”, expresa lo siguiente:

Los procedimientos que rigen la elaboración de su contenido, su administración y su valoración, así como su interpretación y la utilización de sus resultados se establecen a nivel central. Las pruebas nacionales se llevan a cabo bajo la autoridad de un órgano nacional o centralizado y las condiciones de ejecución de las mismas son lo más similares posibles para todos los que las realizan. La evaluación nacional del alumnado, entendida como “la administración nacional de pruebas estandarizadas y de exámenes establecidos a nivel central”, es uno de los instrumentos para la medición y el seguimiento sistemáticos del rendimiento tanto de los alumnos a nivel individual, como de los centros y de los sistemas nacionales de educación. Estas pruebas nacionales se han venido conformando y han evolucionado de acuerdo con las agendas políticas y los contextos estructurales nacionales, y a menudo están vinculadas con otras formas de evaluación (Ministerio de Educación de España, 2010, págs. 7-11).

Es necesario tener en cuenta que la evaluación estandarizada ayuda a visualizar los resultados educativos del conjunto de estudiantes, entregan información sobre el real acceso al conocimiento y a las capacidades que alcanzan los alumnos, más allá de la cantidad de años de estudio que tengan. Las evaluaciones estandarizadas ayudan a hacer visibles un conjunto de aspectos centrales de la labor educativa, aportan información sobre: en qué medida los alumnos están aprendiendo lo que se espera de ellos al finalizar ciertos grados o niveles; cuál es el grado de equidad o inequidad en el logro de dichos aprendizajes; cómo evolucionan a lo largo de los años, tanto los niveles de logro como la equidad en el acceso al conocimiento por parte de los diversos grupos sociales. (Ravela , Arregui, & Wolfe, 2008, pág. 47). Ya conocemos algunas de las definiciones de evaluación estandarizada, ahora procederemos a conocer como son estas evaluaciones en algunos países del mundo.

2.2.1.1. Evaluación estandarizada en Europa

La mejora de la calidad de la educación ocupa un importante lugar en el debate político a nivel nacional, así como en europeo, además contar con información fiable con respecto al desempeño de los estudiantes resulta esencial a la hora de aplicar con éxito las políticas de educación que persiguen objetivos concretos. Las pruebas para los estudiantes en Europa están adquiriendo una creciente importancia ya que de esta forma mide y monitorea la calidad de la educación, así como el sistema educativo europeo. Las pruebas nacionales (Como se le conocen en Europa), son una de las formas operables de evaluación de los estudiantes si se le define como “prueba estandarizada y establecidos a nivel central y administrados a nivel nacional y central”, es un instrumento para medición y el seguimiento sistemático del desempeño de los estudiantes de manera individual, así como de los centro educativos; los responsables de realizar el proceso de estandarización de la prueba son las autoridades educativa sin embargo, en el caso de Bélgica España y Alemania las encargadas son las autoridades educativas de mayor rango (equivalente a nivel central) (EACEA, 2010, págs. 3-4).

Fowler en el 2004 nos proporciona un ejemplo con respecto a Inglaterra, donde los estudiantes de dieciséis años de edad son expuesto a una evaluación estandarizada con el objetivo de determinar cuáles son los estudiantes que obtendrán una posibilidad de entrar al nivel avanzado o “A-Level”; todo esto sucede en la educación secundaria, una vez finalizada la educación secundaria los estudiantes se tienen que someter de nuevo a una evaluación estandarizada para conocer a que universidades pueden asistir. Por otro lado Rotberg en el 2006, nos expresa otro ejemplo: esta vez es el caso de Turquía dada la sobredemanda que existe de estudiantes con la finalidad de ingresar al círculo universitario, en este país solo el 20% del población pueden ingresar a la universidad debido a cuestiones de cupos, y es por medio de las evaluaciones estandarizadas donde la población ingresa a la universidades (Barrenechea, 2010, pág. 3).

En cuanto a evaluación de estudiantes se refiere, los centros educativos de los Países Bajos tienen un alto grado de autonomía, ya que muchos de ellos utilizan objetivos y pruebas intermedios para evaluar el progreso de los estudiantes. Según la EACEA en el 2010 uno de los seguimientos del estudiante y de la enseñanza es el (Leerlingen Onderwijsvolgsysteem, LVOS) es la prueba de acceso (Entreetoets) que realiza la medición del progreso de los

estudiantes y los resultados de la enseñanza en las áreas de lenguaje, aritmética/matemáticas y en las técnicas de estudio, también realizan un examen final de la educación primaria (Eindtoets Basisonderwijs) esta es una prueba nacional sin embargo no es obligatoria realizar, pero la mayoría de estudiantes se someten a esta prueba, el objetivo de esta prueba es recoger información independiente para fundamentar las recomendaciones que el centro educativo le hace a los padres de familia a la hora de escoger el tipo de educación secundaria que más le favorece a su hijo. (EACEA, 2010, pág. 13)

El objetivo principal de los centros educativos es conseguir un marco integrado para que el sistema de seguimiento de los estudiantes mediante las pruebas mencionadas anteriormente. Asimismo se está ejecutando un estudio de seguimiento académico de Cohortes (Educational Careers Cohort Survey-Cool) en la cual valoran el progreso de los estudiantes desde los 5 años hasta los 18 años con respecto a su desarrollo cognitivo, social y emocional, para poder hacer este seguimiento de los ámbitos mencionados el estudiantes se someten diariamente a diferentes pruebas y cuestionarios en toda su trayectoria educativa, permitiendo de esta manera que todo quede documentado constantemente.

De igual manera, en Escocia, los diferentes tipos de evaluación se encuentran articulados en el marco de un sistema único y coherente, contribuyendo al apoyo del aprendizaje en todos los niveles: estudiantes de manera individual, centros educativos, autoridades y sistemas educativo escocés en su totalidad. En este país se apuesta por la evaluación formativa a hora de evaluar el aprendizaje y el proceso de cada estudiante de manera individual, así como la información brinda por la evaluación cuando se trata de evaluar y mejorar la calidad del aprendizaje en el sistema educativo; con respecto a la evaluación sumativa del aprendizaje de cada estudiante a nivel individual depende del criterio profesional que el docente emplee a la hora de evaluar el trabajo de los estudiantes a lo largo del curso. (EACEA, 2010, pág. 14)

En este contexto los resultado de las evaluaciones estandarizada en lengua inglesa y matemáticas se pueden utilizar para verificar las evaluaciones de los trabajos elaborados por los estudiantes en el aula de clases durante todo el curso, en cambio el Estudio Nacional de Rendimiento (Scottish Survey of Achievement, SSA), esta prueba es utilizada para evaluar cada cuatro años el desempeño en inglés/lectoescritura (Español), matemáticas/ cálculo, ciencias naturales y ciencias sociales de una muestra de estudiantes a nivel nacional. Por lo

tanto, las pruebas estandarizadas están integradas en una política más amplia de evaluación para el aprendizaje, como aprendizaje y del aprendizaje, este último aspecto su función es limitada y hay que considerarla junto con otras actividades importantes para la evaluación. (EACEA, 2010, pág. 14)

Las evaluaciones estandarizadas de los estudiantes en Europa son relativamente nuevas, la introducción y el uso de las evaluaciones fue un inicio lento y ocasional y se incrementó hasta la década de los 90. En la actualidad algunos países están iniciando a introducir la evaluación estandarizada. Los primeros en introducir las pruebas estandarizadas se encuentran todos aquellos que desarrollaron los instrumentos con el objetivo de tomar decisiones sobre el trayecto académico de los estudiantes. En 1946 Islandia creó exámenes finales coordinado de ámbito nacional para el nivel CINE 1 (Clasificación Internacional Normalizada de la Educación) más tardes estas evaluaciones serias sustituida por pruebas similares para los niveles CINE 1 y 2, estos exámenes se utilizaron para decidir si los estudiantes debían cursar al siguiente grado (EACEA, 2010, págs. 14-15)

En el caso de Portugal, los exámenes nacionales surgieron en 1947 para los estudiantes en los niveles CINE 1 y 2 esta sirvió como base para el avance de los estudiantes al nivel siguiente y para obtener un certificado (hasta su derogación en 1974) (EACEA, 2010, pág. 15). En este mismo año (1947) se introdujo en Irlanda del Norte (en Reino Unido) exámenes de selección para la educación post primaria en virtud de la ley de educación. En el año 1962 se empezó en Escocia (también en el Reino Unido) los exámenes certificadorios Ordinary Grade (Nivel CINE 2), que consistía en obtener un certificado a los 16 años. Los exámenes de acceso en Lexemburgo empezaron en 1968 siendo reemplazado en el año 1996 por las evaluaciones estandarizados encuadrados en el procedimiento de orientaciones y eran utilizadas para promover de primaria a secundaria. De igual modo, en los Países Bajos el examen final al término del nivel CINE 1, que se realizó por primera vez en 1970 (EACEA, 2010, pág. 15)

En Malta y Dinamarca las evaluaciones estandarizadas se empezaron en 1975; en Malta se inició bajo el formato de exámenes anuales para los centros educativos de primaria y secundaria, en cambio en Dinamarca se introdujeron como exámenes finales al finalizar la enseñanza secundaria. Cinco países iniciaron relativamente pronto las evaluaciones estandarizadas cuyos objetivos no estaban relacionados con la toma de decisiones sobre el

trayecto a seguir en lo académico individual del estudiante, estas se administraban una muestra representativa de la población, estos cinco países eran: Francia, Hungría, Irlanda Reino Unido y Suecia, en este último la pruebas estandarizadas se iniciaron en el año 1962 en el CINE 2 con el propósito de ayudar a los docentes a comparar los resultados de los estudiantes con un estándar nacional, en cambio en Irlanda se empezaron las pruebas estandarizada en 1972, con las pruebas de lectura en inglés con nivel CINE 1 esto fue posible por el resultado de un debate público sobre el nivel de los estudiantes en esta materia (EACEA, 2010, pág. 15)

Cinco años más tarde se introdujeron las evaluaciones estandarizadas tras la reforma educativa ocurrida en el año 1977, iniciando en primer lugar con los centros de primaria y más tarde con los de secundaria, con la finalidad de mejorar las evaluaciones del sistema educativo. En el año 1978 se creó en Inglaterra, Gales e Irlanda del Norte (en el Reino Unido) la Unidad de Evaluación del Rendimiento (Assessment of Performance Unit, APU). Esta unidad realizó varias investigaciones en muestras de estudiantes de 11 y 15 años de edad generalmente y a veces a estudiantes de 13 años, con la intención de identificar las diferencias significativas en el desempeño de los estudiantes conexas al contexto de aprendizaje, incluida los acontecimientos de bajo desempeño, y brindar conclusiones obtenidas en estas investigaciones y ponerlas a disposición de las autoridades. Por otro lado, en Hungría se inició a efectuar trabajos regulares de “seguimientos” en los niveles CINE 1 y 2 en el año de 1986 con la entrada en vigencia de su Ley de Educación en 1985, donde menciona que tanto el contenido como los instrumentos para el seguimiento, la evaluación y el control de la calidad de la educación pública (EACEA, 2010, pág. 15).

Según EACEA (2010) en los años noventa se supuso el esparcimiento de las evaluaciones estandarizada para los estudiantes, debido a que varias regiones o países la comenzaron a utilizar como instrumento de evaluación. En España, por ejemplo, con la Ley de Ordenación General del Sistema Educativo (LOGSE) en 1990 logrando cambios significativos en el sistema de evaluación, incluida su aplicación en el año 1994 la primera evaluación estandarizada sobre los resultados de los estudiantes en el nivel primaria. En ese mismo año la Comunidad Francesa en Bélgica inició a realizar pruebas estandarizadas sobre los resultados de los estudiantes en educación primaria y después en secundaria, con la finalidad de recopilar información sobre el desempeño de los estudiantes. En los países de Letonia y

Estonia se iniciaron las evaluaciones un poco después de la recuperación de la independencia en 1991 con el progreso del sistema nacional para evaluar estudiantes de los niveles cine 1 y 2, que respectivamente se pusieron en práctica por primera vez en los años de 1994 y 1997, mientras que, en Rumania, la pruebas se comenzaron en 1995 al final de la educación primaria (EACEA, 2010, págs. 15-16).

En la actualidad son 11 los países que se han introducido en las pruebas estandarizada, aunque no se están aplicando completamente en todos los países; En Bélgica en la Comunidad Flamenca, Lituania y Polonia se comenzó su aplicación desde 2002, en el caso de Noruega, inicio desde 2004, Eslovaquia en el 2003 emprendió un proyecto de evaluación de los estudiantes por medio de las pruebas estandarizadas al finalizar su educación secundaria. Mientras que Austria y Alemania comenzaron la evaluación del estudiante basado en estándares educativos nacionales en el 2003 y 2005 sin embargo en el 2008 en Austria el gobierno de este país dio enmienda a la Ley de Educación Escolar, donde hace constar que la base jurídica para la introducción de los estándares educativos, las evaluaciones estandarizadas con bases a estándares estaban prevista para los años de 2011, 2012 y 2013 (EACEA, 2010, pág. 16).

En Alemania los estándares educativos adaptados en el 2004, se aplicaron a través las evaluaciones estandarizadas en educación secundaria en todo el país en el año del 2005 y 2006. El propósito de estas evaluaciones es efectuar comparaciones a nivel central del país. En el 2009 en cada país se empezó a realizar pruebas comparativas en base a estándares educativos. En cambio, en el 2006 Bulgaria introdujo las evaluaciones estandarizadas a finales del grado de educación primaria. De igual manera en Chipre el Ministerio de Educación empezó a realizar estas evaluaciones desde el 2007 en todo el país, la prueba es anual y se aplica al final de educación primaria. El objetivo por el cual se realizan esta prueba es para detectar al estudiante en riesgo de desarrollar un analfabetismo funcional al finalizar la educación obligatoria, de manera que se le pueda brindar un apoyo adicional en la educación secundaria (EACEA, 2010, pág. 16).

2.2.1.2 Evaluación estandarizada en Latinoamérica

En los últimos años, los sistemas de evaluación de la calidad de la educación de todos los países de América latina han creado diversos mecanismos de medición de resultados; al mismo tiempo, la mayoría de los países se han incorporado a las evaluaciones de carácter

internacional. Los sistemas de evaluación de la calidad de la educación han convertido los resultados en un debate público por medio de informes de las pruebas ya sean estas nacionales o internacionales, por otro lado, la mayoría de los docentes que son evaluados con estos instrumentos (un ejemplo de esto es la prueba PISA), ven este mecanismo como una amenaza a su desempeño. Cerón y Cruz en 2013 prácticamente todos los países de Latinoamérica han fundado un sistema nacional de la evaluación para la calidad de la educación, la cual se han convertido en instalaciones de diversos mecanismos de intervención de resultado; de igual modo, la mayoría se han unido a las evaluaciones de carácter internacional (Cerón & Cruz, 2013, pág. 98)

Además, los sistemas de evaluación han convertido los resultados en un debate público, a través de los informes que se brinda al finalizarse las pruebas nacionales como internacionales. En los países de Latinoamericanos se ha observado un inusual y metódico interés por la evaluación, estrechamente asociada a la calidad de la misma, como el resultado de esto se ha incrementado la aplicación de diferentes evaluaciones. La PREAL en unos de sus informes ubica a México y a Chile entre los países que más experiencias poseen en evaluación. Los especialistas consideran que Chile posee uno de los sistemas de evaluación más amplio y mejor administrados del Latinoamérica; por otro lado, México es el país de américa latina que más experiencia posee con respecto a las evaluaciones, pero se identifica porque las autoridades renuncian a la divulgación de los resultados. Es necesario mencionar que una de las herramientas que contribuye a fortalecer el sistema de evaluación de la enseñanza en todo el mundo es el TIMSS (Third International Mathematics and Science Study), en la cual compara y explica los aprendizajes en ciencias y matemáticas en 41 países (Dirección Nacional de Información y Evaluación de la Calidad Educativa / Ministerio de Educación, 2009, págs. 17-18).

Por otro lado, Cerón & Cruz (2013) señalan que la the Evaluation of Educational Achivement es una de las organizaciones que está preocupada por las políticas de evaluación ellos lo expresan así:

La International Association for the Evaluation of Educational Achivement (IEA) es la que lidera la preocupación por las políticas de la evaluación, sin embargo, en el caso de Latinoamérica este fenómeno es impulsado y apoyado por la Organización para la Cooperación y el Desarrollo Económico (OCDE) por medio del Programme Cooperación y

el Desarrollo Económico (OCDE) por medio del Programme for International Student Assessment (PISA), prueba que se le aplica a jóvenes de 15 años cada 3 años a los países que son miembros de la OCDE. (Cerón & Cruz, 2013, pág. 100)

Para Gallegos (2008) las evaluaciones estandarizadas en esta región, son coordinadas y organizada por la Oficina Regional de Educación para Latinoamérica y el Caribe (OREALC) de la UNESCO y los proyectos dirigido por el Laboratorio Latinoamericana de la Evaluación de la Educación (LLECE) que incluyeron aspectos referidos a la calidad educativa de los sistemas educativos, el desempeño académico y los factores asociados (Gallegos, 2008, pág. 16). Según Arregui y otros en el 2003 mencionan que durante los años noventa América Latina además de incorporar las evaluaciones internacionales también estableció procedimientos nacionales de evaluación al interior del sistema educativo de cada país. Por otro lado, Arregui, y otros, (2008) en su artículo las evaluaciones educativas que América latina necesitan mencionan que estas iniciaron con mayor auge en los años 90 y que desde esa fecha algunas naciones han continuado con las evaluaciones, aunque se originaron cambios en los sistemas de educativos, sin embargo, en otros países se discontinuo la importancia de las evaluaciones como consecuencia de esa decisión ha iniciado de cero en varias ocasiones. Estos mismos autores nos expresan lo siguiente:

Actualmente son dieciséis países que han participado en el Segundo Estudio Regional en 3 y 6 grado de primaria que lleva adelante la OREALC/UNESCO: Argentina, Chile, Colombia, Costa Rica, Cuba, Ecuador, El Salvador, Guatemala, México, Nicaragua, Panamá, Paraguay, Perú, República Dominicana y Uruguay. Seis de estos países participaron en PISA en el año 2006 y otros tres se incorporaron en el año 2009 al ciclo PISA. Algunos países de la región han participaron o están participando en los estudios de Matemática y Ciencia (TIMSS), Lectura (PIRLS) y educación cívica que la coordina la IEA. En los 12 años los países de Latinoamérica han estado activo en la red de los sistemas de evaluación organizada por el Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE) de OREALC/ UNESCO, en la cual casi todos los países de la región participan. Esta red realiza reuniones cada dos veces al año en la cual se intercambian experiencias y formación en el aspecto de la evaluación (Arregui, y otros, 2008, pág. 49).

Las evaluaciones no solo han crecido en esta región sino también están mejorando la medición de los resultados, algunos cambios y mejorar que se han observado en los últimos

años se encuentran las siguientes: a) *Mayor transparencia en la difusión de los resultados*, al parecer se ha superado la etapa en la cual las autoridades de diferentes países intervenían en la divulgación de los resultados de las evaluaciones cuando estos no eran favorables para la región. b) *Tendencias crecientes a pasar de pruebas normativas*, estas tienen el objetivo principal establecer comparaciones entre los estudiantes a pruebas de criterios, que se centra en qué es lo que los estudiantes saben y son capaces de hacer, asimismo en esta prueba de criterio se presenta una definición de cuál es el resultado que deberían de alcanzar los estudiantes para que su desempeño sea considerado satisfactorio. c) *Mejoría de las capacidades técnico-metodológicas para la construcción de pruebas y para el procesamiento de los datos*, el esfuerzo por desarrollar prueba para evaluar el rango de conocimiento y capacidades más amplias que contengan preguntas de respuesta breve o construida (Arregui, y otros, 2008, págs. 49-50)

Por otro lado, se han incorporado metodología más sofisticada para el análisis de datos como la Teoría de Respuesta al ítem. d) *Creciente atención a la difusión y uso de los resultados*, se ha comprendido que no es suficiente implementar una evaluación y publicar informe de resultado, también es necesario desarrollar estrategias de divulgación y realizar un reporte individual a cada centro educativo a la que se le quiere hacer llegar la información. e) *Mayor preocupación por la investigación de los factores que inciden sobre los aprendizajes*, si bien es cierto que falta mucho que hacer con respecto a los términos metodológico e interpretación para realizar buenas investigaciones, en los diferentes países coinciden que es preciso avanzar en las investigación en los sistemas educativos, las prácticas del proceso de enseñanza-aprendizaje y las decisiones de política educativa, esta última influye en las oportunidades de aprendizaje de los jóvenes (Arregui, y otros, 2008, págs. 49-50).

A pesar de las mejoras mencionadas anteriormente, todavía persiste una serie de debilidades que según los autores es preciso mencionar: en primer lugar es usual que a nivel político se demanda la aplicación de grandes pruebas operativas sin ningún propósito y con recursos insuficientes para hacerlo bien; segundo, todavía no se tiene claro que no cualquiera evaluación sirve para cualquier objetivo, por ende es necesario diseñar una evaluación que cumpla los requisitos para que se pueda usar a largo plazo; tercero, hace falta discutir los contenidos que se deben de enseñar en el aula de clases y que están establecidos en el

currículo y que los estudiantes deben de saber al finalizar el año escolar; cuarto, en cuanto a investigación se refiere es necesario mejorar los instrumentos que son complementarios y que se aplican con las pruebas y que el diseño de investigación sea más ambicioso por ejemplo: estudios de factores asociados; quinto, el Ministerio de Educación ha tenido una restringida capacidad para crear, formular e implementar políticas que brinden respuestas a los problemas presentado en la evaluación (Arregui, y otros, 2008, pág. 50)

Sexto, las acciones para la divulgación y el uso de los resultados de las evaluaciones son todavía insuficientes, para conseguir que los docentes comprendan y manejen los resultados para que los incorporen en el aula de clases; séptimo, en varios países persiste significativas debilidades técnicas en cuanto a diseño de evaluación se refiere, por ejemplo, existe excesiva simplicidad para elaborar preguntas difíciles ya que estas evalúan las capacidades cognitivas complejas, enfocando las evaluaciones en los logros académicos de un determinado grado, lo que dificulta saber qué aprendieron y quienes todavía no lograron lo esperado en ese grado específico; Octava, no hay suficientes profesionales certificado para coordinar y diseñar este tipo de evaluaciones, lo que empeora la continuidad de los equipos técnico de los países, la mayoría de veces por razones políticas. Esto impide la acumulación de nuevos conocimientos y experiencias en esta región que da como resultado que los países inicien de cero con su sistema de evaluación, cuando en años anteriores han tenido un buen funcionamiento (Arregui, y otros, 2008, pág. 50)

2.2.1.3. La evaluación estandarizada en Centroamérica

La educación en Centroamérica es una de las claves primordiales para prosperar en el desarrollo de estos países. Esta región está conformada por siete países Costa Rica, Nicaragua, Panamá, Honduras, El Salvador, Guatemala y Belice, sin embargo, solo seis de ellos forman parte de las evaluaciones estandarizadas (Costa Rica, Nicaragua, Panamá, Honduras, El Salvador y Guatemala), estos países han situado, con distinto énfasis, a la educación como un eje central de su agenda, persuadidos que la calidad, la pertinencia, la equidad y la eficiencia que se logre depende el futuro de la integración de Centroamérica, así como su nivel de competitividad y viabilidad como países, con respecto a Belice todavía no se encuentran informes o investigaciones de que haya participado en una evaluación estandarizada

Según la UNESCO en 1997 menciona que el sistema nacional de medición de logro educativo en Costa Rica se ejecutó en tres etapas. La primera etapa inicio desde 1986 a 1988 ha iniciativa del Instituto para el Mejoramiento de la educación costarricense (IIMEC) de la Universidad de Costa Rica. En esta etapa se elaboraron evaluaciones con referencia a criterios de carácter diagnóstico y censales para los grados de tercero, sexto, noveno y onceavo en las áreas de español, matemáticas, ciencias y estudios sociales; ya en el año 1988 se originaron las variantes siguientes: las evaluaciones fueron administradas a una muestra nacional, no se aplicó a 110 grado y además incluyeron las áreas de inglés y francés para noveno grado. En la segunda etapa se desarrolló y se administró las evaluaciones con referencia a normas con carácter sumativo para onceavo grado (Se empezaron aplicar a partir de 1988) y se le conocieron como pruebas de bachillerato y en sexto grado se empezó aplicar en 1989. La primera prueba la elaboró el IIMEC en base a un temario que fue admitido por el Consejo Superior de educación, mientras que la segunda fue preparada por cada región educativo con guía del Departamento de Evaluación del Ministerio Pública.

La tercera etapa, se originó en 1993 dándole seguimiento a las evaluaciones sumativa y añadiendo pruebas con referencia a criterios en los grados de tercero y noveno, estas evaluaciones fueron de carácter diagnóstico y fue aplicada a una muestra nacional de este país. Por otro lado, Costa Rica, este país cada año realiza informe sobre los avances o desafío que han tenido en educación, dichos informes tienen como propósito promover conocimientos actualizados sobre el desempeño académico del sistema de educación de esta región, de esta manera induce un debate entre los ciudadanos para que ayuden a la toma de decisiones en este aspecto. A la vez estos informes procuran establecer cuanto se alejan o se acerca a las aspiraciones de oportunidad que la población y el país aspira tener. Sin embargo, este país empezó la divulgación de resultado en educación en el 2005 con la publicación de su primer informe, esto gracias a la iniciativa del Consejo Nacional de Rectores (Conare) con el objetivo de proporcionarles a la población un análisis actualizado del sistema educativo del país. En el año 2006 en Conare acordó dar seguimiento permanente a este esfuerzo por mejorar la calidad de educación, este trabajo se lo encargo al Programa de Estado de la Nación de Costa Rica. Desde eso años este Programa ha realizado investigación hasta la actualidad (Programa Estado de la Nación, 2013, pág. 472).

2.3. Diseño y validación de un instrumento de evaluación

Según Covacevich (2014) la ejecución de prácticas y políticas educativas está muy relacionada a la evaluación de los aprendizajes de los estudiantes, porque permite monitorear avances y logros, mejorar la enseñanza en el aula de clases, perfeccionar las políticas y evaluar la efectividad de programas que realizan dichas evaluaciones. Para que la evaluación logre sus propósitos, es fundamental hacer una adecuada elección de los instrumentos de evaluación del aprendizaje que serán utilizados. Es importante destacar que en los últimos años las evaluaciones han tenido un auge, debido que éstas buscan medir el resultado de un programa realizando una comparación entre una institución que cuenta con los beneficios de ese programa y otra institución que no posee dicha ayuda. Asimismo, se evalúa aprendizajes para la rendición de cuentas, verificación de competencias obtenidas en determinado grado (Primero grado, segundo grado, etc.) de enseñanza, comprobación del nivel de conocimientos de un estudiante para brindarle la educación más apropiada con el fin de mejorar la calidad de la educación. (Covacevich, 2014, pág. 1)

Para que una evaluación tenga efectividad dependen de la calidad y adecuaciones de los instrumentos de evaluación de los aprendizajes que se utilicen, de acuerdo con Covacevich (2014) “decidir cómo evaluar el aprendizaje se deben de considerar diversos factores, analizar los pros y contras de cada instrumento disponible, y tomar la decisión cuál es el más adecuado para la situación que se pretender evaluar” (Covacevich, 2014, pág. 1). De acuerdo con Vergara & Quiles (2016) un instrumento, una prueba “es una manera de medir el comportamiento o ejecución de un individuo en una situación determinada que permite recoger información para usos variados como ser: de tipo diagnóstico o de alcance de logros” (Vergara & Quiles, 2016, pág. 3).

Es de suma importancia que al momento de la elaboración de un instrumento tengamos en cuenta los siguientes tópicos que abordaremos en este apartado tales como ser: *el vínculo entre los objetivos y el instrumento que se elaborará o seleccionará, tipo de instrumento, taxonomía de los objetivos cognoscitivos, tablas de especificaciones, la preparación de los reactivos de un test, teoría, de respuesta al ítem, teoría clásica de los test* entre otros temas que explicaremos a continuación.

2.3.1 Vínculo entre los objetivos y el instrumento que se elaborará o seleccionará

Para escoger un instrumento de evaluación de los aprendizajes es importante considerar que el propósito de la evaluación sea adecuado. En este apartado se describirá algunos posibles objetivos de evaluación, para después continuar con los elementos que debe poseer un instrumento para identificar si son coherentes con los objetivos de evaluación. Es necesario que nos preguntemos ¿Para qué queremos evaluar? Con respecto a esto Toranzos (2014) expresa lo siguiente:

¿Qué evaluar? es la pregunta por el objeto de la evaluación que, si bien parece simple, en la práctica no resulta de fácil resolución. En el espacio educativo se tiende a pensar en objetos múltiples de evaluación: los alumnos y estudiantes, el docente, los aprendizajes, las estrategias de enseñanza, la gestión pedagógica, los recursos empleados, la institución, un programa, etc. Cuando esto sucede y ni el objeto ni los sujetos de la evaluación resultan bien definidos se suele producir una reducción notable de las reflexiones sobre la evaluación y sus efectos a cuestiones meramente instrumentales o técnicas. Esto, en general, trae como consecuencia una mirada que empobrece el valor de los procesos de evaluación como instancias de construcción colectiva y producción de conocimiento (Toranzos, 2014, pág. 4).

Si bien es cierto que las preguntas planteadas por los dos autores son importantes, se han olvidado de la comunicación de los resultados tal y como lo plantean Valcárcel, Cabellos, & Avilés (2012) “la base necesaria para tomar las decisiones adecuadas sobre ¿para qué evaluar?, ¿qué evaluar? ¿cómo?, ¿cuándo?, ¿con qué? y ¿cómo? comunicar los resultados.” (Valcárcel, Cabellos, & Avilés, 2012, pág. 7). Todas estas preguntas hay que realizarse para evaluar los aprendizajes de los estudiantes, pero siempre hay que tener en cuenta comunicar los resultados de dichas evaluaciones ya que de esta manera se seguirá mejorando la calidad de la educación. Determinar el objetivo o propósito de evaluación implica hacerse las preguntas ¿para qué queremos medir aprendizaje?; ¿qué es lo que queremos medir? y ¿a quiénes queremos evaluar? (Covacevich, 2014, pág. 8). Son muchos los motivos para decir evaluar aprendizajes. A continuación, se presentan algunos de los objetivos que se tienen que tener en cuenta al momento de realizar una evaluación en el sistema educativo.

a) *Obtener información a nivel del sistema educativo*: Algunas maneras de obtener información de los sistemas educativos es tomar una fotografía de cómo está un sistema educativo, para poder a tomar decisiones de políticas educativa, ya sea de manera nacional o municipal, ya que estos pueden necesitar información sobre de cómo los estudiantes están logrando los objetivos de aprendizaje propuestos. Lo correcto es diseñar una prueba que evalúe el currículo nacional básico en la asignatura y grado que se consideren. Y es normal que las pruebas se apliquen al final de año escolar, pero en algunas ocasiones es necesario hacer la medición intermedia para saber el grado de avances que se ha logrado en el proceso de enseñanza aprendizaje y de esta manera realizar una retroalimentación a tiempo. (Covacevich, 2014, pág. 3)

Covacevich en el (2014) expreso lo siguiente:

Si se desea obtener información agregada de lo que está sucediendo con el conjunto de escuelas o estudiantes, basta con realizar una evaluación muestral, es decir, no es necesario aplicar las pruebas a todas las escuelas ni estudiantes sino solo a una muestra representativa de ellos, lo que hace la evaluación más barata y simplifica los procesos logísticos. Si también se desea información específica a nivel de escuela o de estudiante, y no solo del sistema educativo en su totalidad, la evaluación debe ser censal, es decir, aplicada a todos los estudiantes y escuelas. (Covacevich, 2014, pág. 3)

b) *Compararse con otros países*: Hoy es importante contar con información de cuanto están aprendiendo los estudiantes del país que se está comparando con referencia a otros países. Tal y como nos señala Covacevich (2014):

(...) Un país o subsistema nacional puede participar de los estudios internacionales tales como PISA, TIMSS, PIRLS y las pruebas PERCE, SERCE y TERCE del Laboratorio Latinoamericano de Evaluación de la Calidad Escolar -LLECE. Todos estos estudios evalúan aprendizajes de estudiantes a nivel escolar en un conjunto de países (o subsistemas nacionales), lo que permite compararlos contra un marco de evaluación conocido y acordado por los países participantes. (p.p. 8-9). Es necesario aclarar que cada estudio evalúa diferentes asignaturas y grados y con objetivos diferentes, la misma autora nos brinda un ejemplo de esto "(...) PISA como TIMSS evalúan matemáticas y ciencias, pero mientras TIMSS lo hace en estudiantes de 4º y 8º grado y con un enfoque curricular, PISA lo hace en estudiantes de 15 años, con un enfoque de habilidades para la vida" (Covacevich, 2014, págs. 3-4).

c) Medir la evolución del sistema educativo a lo largo del tiempo: Las evaluaciones nacionales como internacionales permiten obtener una información más amplia de cómo están los logros del aprendizaje en un determinado tiempo, pero a la vez puede servir para monitorear el avance del proceso de enseñanza- aprendizaje a lo largo del tiempo. Estas pruebas se aplican de forma periódica, de manera anual; en el caso de las evaluaciones nacionales. Sin embargo, en las evaluaciones internacionales se administra en ciclos de tres o más años. Esto permite observar avances en el aprendizaje promedio de cada país y también ayuda a identificar otras debilidades, como ser la disminución de género, si los estudiantes han mejorados en alguna asignatura (Covacevich, 2014, pág. 4).

d) Brindar información a nivel de escuela: Son muchos los objetivos para obtener información de los aprendizajes a nivel de cada escuela, uno de ellos puede de la retroalimentación al equipo docente y administrativo, la rendición de cuentas entre otras. Es importante, tener en cuenta que las pruebas nacionales están diseñadas para entregar información por escuela, en otros centros educativos puede ser necesario diseñar pruebas especiales, ya sea porque el objetivo de evaluación es diferente a lo que se desee medir (Covacevich, 2014, pág. 5).

e) Obtener información a nivel de estudiante: Este podría ser uno de los más importantes ya que son los estudiantes que nos brindarán la información para el proceso de mejoramiento continuo del sistema, para poder identificar las debilidades de un estudiante y que su padre y docente puedan apoyar a que el estudiante convierta esa debilidad en fortaleza (Covacevich, 2014, pág. 5).

Según Muñiz & Fonseca (2008) la evaluación se puede dar de tres formas: la evaluación externa, cuando las personas evaluadoras no pertenecen a la institución; la evaluación interna cuando la persona pertenece al centro educativo, y mixta cuando se mezclan ambas. Además, expresa que no existe regla universal para poder utilizar estos tres modelos o formas, por lo tanto, son legítimos y dependerá del investigador que modelo utilizará (Muñiz & Fonseca, 2008, pág. 15). Para que todo resulte, los responsables de coordinar estas evaluaciones tienen que tener en cuenta lo siguiente:

a) Alinear los objetivos de evaluación y el instrumento: Según la Center for Assessment and research, James Madison University, 2014 un elemento fundamental al momento de escoger un instrumento de evaluación de aprendizajes es que sea adecuado para los objetivos

de la evaluación. Este paso es crucial porque se relaciona con la utilidad de la información que se obtendrá. Para que los resultados de la evaluación no se obtengan baja o poca información, se debe de revisar los objetivos, los contenidos y la población declarados por el instrumento y asegurarse que el propósito y la evaluación estén alineados. (Covacevich, 2014, pág. 5)

b) Alinear los propósitos: Para poder realizar una prueba lo primero que hay que preguntarse es el objetivo para diseñarla y si alcanza la evaluación hay que tener en cuenta que los objetivos o los propósitos de los instrumentos pueden ser varios, tales como; medir el logro académico de una determinada población, un instrumento diagnóstico para conocer la realidad de las instituciones educativas. Algunos instrumentos son diseñados para realizar evaluaciones diagnósticas, formativas, o sumativa (Covacevich, 2014, pág.6).

Las evaluaciones diagnósticas (se dan al inicio de clases de un Centro educativo) son aquellas que evalúan las debilidades y fortalezas de un grupo de estudiantes puede ser individual o grupal. Esta evaluación incluye varios ítems que profundizan en un solo tema, con el fin de identificar con precisión las dificultades de aprendizajes de los estudiantes, y de esta manera buscar una solución. La formativa, es el tipo de evaluación que se presenta durante un año escolar y el objetivo es supervisar el avance que ha obtenido el estudiante durante un periodo de clases determinado. La evaluación sumativa se utiliza al final de año escolar y está diseñada para medir los logros de cada estudiante (Covacevich, 2014, pág. 6).

c) Alinear los contenidos: De acuerdo con Covacevich (2014) La segunda pregunta que hay que plantearse es si el instrumento mide lo que se quiere medir. Hay que analizar en detalle el contenido del test, cuáles son los niveles de aprendizaje que cubre el grupo al que se le aplicará el test y realizar una comparación con los objetivos de evaluación para asegurarse que los ítems de la prueba están acordes con lo enseñado en un aula de clases.

d) Alinear los objetivos con la población: Es importante tener en cuenta un tercer elemento es la coherencia que tiene que existir entre la población y lo que se desea evaluar, y si se elaboran pruebas basadas en normas también se tienen que revisar para que población fueron construidas estas normas (Covacevich, 2014, pág.6).

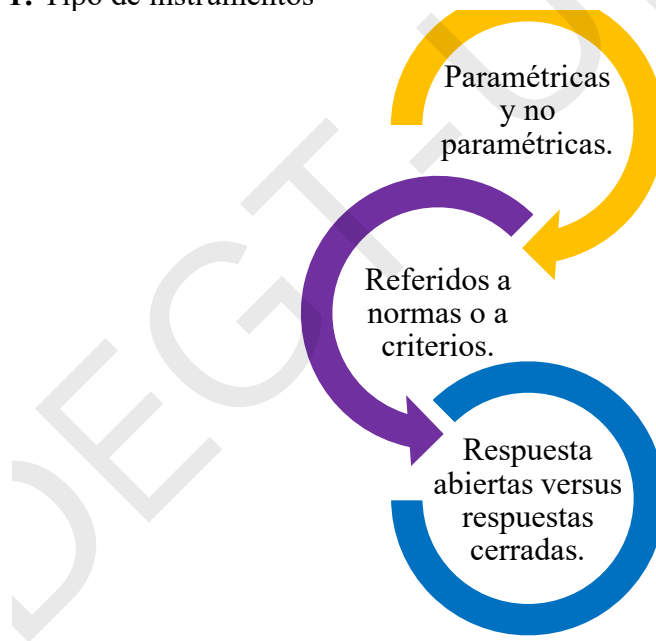
2.3.2. Tipo de Instrumentos

Hay diversas formas de clasificar los instrumentos de evaluación de aprendizajes. Sin embargo, Cueto, (2007) propone una forma para clasificar los sistemas de evaluación, es

según las implicancias de sus resultados; es decir, que existen los sistemas de bajas implicancias, esta proporciona información para fines formativos, no cuenta consecuencias directas para los actores implicados; a diferencia de las altas repercusiones, que usa los resultados de pruebas para diversos objetivos, en este si tiene consecuencias para las personas involucradas (Covacevich, 2014, pág. 20)

Los instrumentos también se pueden clasificar según sus características técnica; en este trabajo nos enfocaremos en este tipo de clasificación. Según Covacevich (2014), una manera de clasificar bajo este enfoque es según si son medidas directas o indirectas del aprendizaje. La primera se refiere a las observaciones de un producto del trabajo del estudiante, como ser una tarea realizada en casa, un examen de conocimiento, en cambio la segunda las percepciones de estudiantes, docentes, padres de familia, directores (Covacevich, 2014, pág. 20). A continuación, explicaremos cada uno de los tipos de instrumentos que existen en la Figura 1 se mencionan los tipos de instrumentos para después ser explicados

Figura 1: Tipo de instrumentos



Fuente: Elaboración propia con base en Covacevich (2014).

El instrumento de tipo paramétricas con las evaluaciones que realiza el gobierno al finalizar el año escolar y no paramétricas son aquellas que se elaboran para determinado contexto.

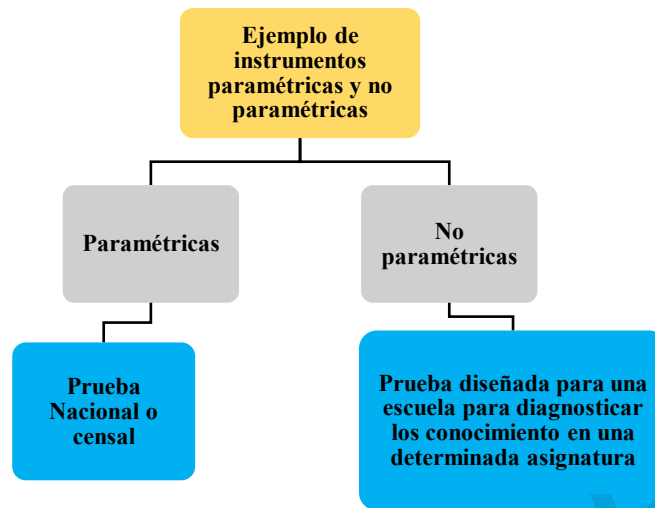
Los instrumentos referidos a normas o a criterios, la primera se refiere a al puntaje de un estudiante la cual se le compara con el de un grupo, la segunda es la que compara el puntaje con un criterio o estándar de aprendizaje establecido.

Las respuestas abiertas versus respuestas cerradas son el tipo de preguntas que se utiliza para obtener información del estudiante.

2.3.2.1 Pruebas paramétricas y no paramétricas

Las pruebas paramétricas están elaboradas para representar a una población en general, en cambio las pruebas no paramétricas están diseñadas para un determinado grupo de la población. De acuerdo con Cohen, Manion y Morrison (2000) las pruebas paramétricas normalmente se encuentran disponibles comercialmente, ya sea por medio de internet o una librería, esto se debe a que dichas pruebas han sido piloteados y estandarizados en una población completa y brindan muchos datos sobre su muestreo, validez y confiabilidad (Covacevich, 2014, pág. 20).

Los autores antes mencionados expresan que las pruebas paramétricas cuentan con dos grandes supuestos: la primera es que hay una curva normal de distribución en la población, y la segunda es que presenta un cero real e intervalo continuos y equivalentes entre los puntajes, en cuanto a las pruebas no paramétricas hacen poco o ningún supuesto sobre la distribución de la población y no asumen una distribución normal, debido a esto no se puede realizar una comparación con la población general. En la Figura 2 se presenta un ejemplo de este tipo de instrumento.

Figura 2:Ejemplo de instrumentos paramétricas y no paramétricas

Fuente: Elaboración propia con base en Covacevich (2014)

Este ejemplo ayuda para orientar a cada uno de estos instrumentos y nos facilita al momento de decidir a qué contexto deseamos aplicar el instrumento.

2.3.2.2. Instrumentos referidos a normas y a criterios

Se empezará explicando los instrumentos referidos a normas, estas evaluaciones son todas aquellas donde el individuo obtiene un puntaje y este es comparado con el puntaje logrado por un grupo, permitiendo de esta manera ubicar el rendimiento de un estudiante o grupo de estudiantes frente a otro, es decir, si un estudiante está arriba de la media o bajo de la media. De acuerdo con Cohen y Swerdlik (2009), se entiende por “norma” al comportamiento usual, habitual y esperado, en cambio en el contexto psicométrico, las normas son datos sobre el desempeño de un grupo específicos que se utilizan como referencia para interpretar o evaluar el puntaje individual de un estudiante (Covacevich, 2014, págs. 21-22).

En otras palabras, la puntuación referida a norma se basa en comparar a un individuo con su grupo normativo. Este tipo de evaluaciones cuentan con tres tipos de puntuaciones normativas y que son fundamentales: la puntuación de percentiles, la puntuación cronológica y la puntuación típica. Los percentiles muestran los porcentajes de los sujetos que puntúan por debajo de la puntuación obtenida por el grupo normativo, sin embargo, la puntuación cronológica muestra la relación de la puntuación de un test con la edad cronológica del individuo, mientras que la puntuación típica indica la distancia que separa a un sujeto de la media del grupo normativo, dichas distancias se expresan en unidades de desviación típica.

El siguiente instrumento es el referido al de criterio, tal y como se describió en la evaluación referida a normas, un modo de brindarle significado al puntaje de un individuo es compararlo con el de un grupo. La otra manera es comparar el puntaje con respecto a algún criterio o estándar de aprendizaje predeterminado (Covacevich, 2014, pág. 22). Este tipo de evaluación es aquella que les brinda un significado a los resultados de un puntaje de una persona comparándolo con un estándar, ya sea de conocimiento, habilidades o de competencias. El punto significativo de esta puntuación es el desempeño del individuo evaluado, lo que le interesa es qué puede o no hacer, que aprendió, si cumple los criterios esperado para el grupo.

2.3.2.3 Instrumento de respuesta abierta o cerrada

Los instrumentos que se utilizan para evaluar el aprendizaje pueden dividirse en dos: las de respuestas abiertas y las de respuesta cerrada. Estos últimos consiste en elegir alternativas dentro de una serie de opciones facilitadas. Las más conocidas de este tipo de respuesta (cerrada), son las pruebas estandarizadas de ítems de selección múltiple, en cambio las respuestas abiertas son las evaluaciones que se necesitan crearlas, un ejemplo de esta son los ensayos. Están con conocidas como preguntas abiertas o de desarrollo.

Para el Center for Assessment and Research, James Madison University (2014), las ventajas de los instrumentos de selección múltiples se pueden administrar de manera rápida y fácil a grandes números de estudiantes, abordando diferentes temas, pero la desventaja que brinda es: que ofrecen menos profundidad, menos riqueza; se enfocan en la capacidad de recordar y otras habilidades cognitivas que son bajas. Las ventajas de las respuestas abiertas o construidas son las que proporcionan información más profunda de lo que los estudiantes saben y pueden hacer, las desventajas de este tipo de instrumentos es que son caros porque deben ser puntuados por personas capacitadas especialmente para este tipo de respuesta, además requiere bastante tiempo de aplicación. De acuerdo con Covacevich (2014), con respecto a las respuestas abiertas “si la puntuación no es efectuada correctamente y de la misma manera por todos los correctores, la comparación de los puntajes se verá afectada, lo que no sucede en el caso de las respuestas cerradas” (Covacevich, 2014, pág. 24).

Es importante expresar, que la cantidad de tiempo y esfuerzo invertido en elaborar un test psicológico o educativo varia con el tipo de prueba y con los propósitos para los cuales se crea. De acuerdo con Aiken (2003) “La elaboración de un test requiere de las consideraciones

cuidadosas de sus propósitos específicos, el proceso de elaboración varía en cierto grado de acuerdo con el propósito que se pretenda lograr” (Aiken, 2003, pág. 18). Al mismo tiempo nos mencionan los tipos de pruebas que existen y los pasos que hay que seguir para poder elaborar un instrumento ya sean psicológico o educativos, a continuación, explicaremos los tipos de pruebas que existen según este autor son las siguientes:

a) *Pruebas de Observación*: Según este autor, la elaboración de una prueba de aptitud para observar a solicitantes de un trabajo en particular comienza con el análisis de todas las actividades que compone dicho empleo. El análisis de las tareas o del trabajo consiste en especificar los componentes del mismo de manera que pueda construirse los reactivos más apropiados para predecir el desempeño de los empleados que se someterán a dicha prueba. Esta prueba evalúa a la observación como una variable psicológica, esta condiciona los procesos de estudiar, aprender y pensar, ya sea en un contexto educativo o laboral, la observación está relacionada con la atención, la percepción visual y con la rapidez de ejecución. Esta evaluación solo trata de mostrar los comportamientos más, importantes relacionado con el trabajo (Aiken, 2003, pág. 18).

b) *Pruebas de inteligencias*: La prueba de inteligencia se elabora como cualquier otra, se reúne un conjunto de ítems o reactivos que supuestamente miden el constructo (Inteligencia), dichos reactivos pueden ser elaborados de acuerdo a una teoría específica de la conducta de inteligencia o haciendo referencia sólo a los tipos de tareas que las personas muy inteligentes pueden realizar de manera más efectiva que las personas menos inteligentes. En la actualidad las pruebas de Binet ya no son la única ocupación de los especialistas en psicología aplicada, pero la evaluación de aptitudes cognoscitivas todavía forma parte de las actividades de los psicólogos en los ámbitos clínicos, educativos y empresariales (Aiken, 2003, pág. 19).

El motivo, por el cual se realizan estos tipos de prueba es porque “la sociedad valora altamente la inteligencia de una persona y además la consideran un factor decisivo para los logros de la educación, el éxito del trabajo y para el desarrollo socioeconómico de una población” (Ardila, 2011, pág. 97). De acuerdo con Ardila (2011) sobre la definición de inteligencia la cual expresa lo siguiente:

Inteligencia es un conjunto de habilidades cognitivas y conductuales que permite la adaptación eficiente al ambiente físico y social. Incluye la capacidad de resolver problemas, planear, pensar de manera abstracta, comprender ideas complejas, aprender de la experiencia.

No se identifica con conocimientos específicos ni con habilidades específicas, sino que se trata de habilidad cognitiva general, de la cual forman parte las capacidades específicas (Ardila, 2011, pág. 100).

c) Pruebas de inventarios y escalas de personalidad: Según Aiken al diseñar una prueba de inventario de personalidad y escala de calificación sea empleado varios enfoques, algunos basado en el sentido común, otros en la teoría de la personalidad y otros más en el procedimiento estadístico (Aiken, 2003, pág. 19).

d) Pruebas de rendimiento: Se les ha brindado más atención a los procedimientos utilizados para diseñar pruebas de rendimiento académicos que a los otros tipos de prueba antes mencionadas. A pesar del uso generalizado de las pruebas de rendimiento, la mayoría de los docentes que supuestamente están familiarizados con su área de estudio, no le dedican el tiempo suficiente a la evaluación del proceso de los estudiantes, ya que la mayoría de veces los profesores consideran que las pruebas son algo desagradable para la enseñanza, en lugar de verlos como parte integral y formativa del proceso educativo (Aiken, 2003, pág. 19).

Aiken (2003), aconseja que las personas que realizarán este tipo de pruebas deberán de empezar por atender con cuidado las siguientes preguntas:

¿Cuáles son los temas y materiales sobre los que se examinará a los estudiantes? ¿Qué tipo de preguntas deben elaborarse? ¿Qué formatos o esquemas de reactivos y pruebas deben utilizarse? ¿Cuándo, dónde y cómo debe administrarse la prueba? ¿Cómo debe calificarse y evaluarse la prueba resuelta? (Aiken, 2003, págs. 19-20)

En la prueba de rendimiento se tiene que partir de una taxonomía de objetivos cognoscitivos, objetivos afectivos y psicomotrices y con una tabla de especificaciones (Que se explicara más adelante).

2.3.3. Calidad de los instrumentos

Para Covacevich, (2014) otro aspecto que nace al analizar un instrumento es la calidad técnica que este posea, entre más calidad conserve un instrumento, más útil será. Además, menciona que los dos elementos principales que se dan en la calidad de un instrumento son su validez y su confiabilidad, estandarización y su importancia para la validez y confiabilidad.

a) Validez: Según Covacevich, (2014) y Prieto & Delgado (2010) en los últimos años, el concepto de validez a sufrido varias transformaciones (Covacevich, 2014, pág. 10; Prieto &

Delgado, 2010, pág. 70), sin embargo, Prieto & Delgado explican que estas transformaciones son provocadas por los distintos objetivos a los que se han destinado los tests (Prieto & Delgado, 2010, pág. 70). Para Barajas (2011) y Prieto & Delgado, (2010) la validez es fundamental para el desarrollo y evaluación de una prueba, sin embargo, Barajas expresa:

La validación de una prueba es el proceso de acumulación de evidencias que apoyen tales interpretaciones, y su objetivo es determinar qué tan apropiadas, significativas y útiles resultan las inferencias específicas que se hacen a partir de las mediciones realizadas mediante la prueba en función del uso específico para el cual se diseñó. (Barajas, 2011, pág. 17)

Por su parte Covacevich (2014) nos comparte las opiniones que tienen los expertos en evaluación, así mismo nos hace referencia del tipo de población a la que está dirigida el instrumento que en su artículo expresa lo siguiente.

Los especialistas en evaluación han considerado que la validez no es una propiedad fija e inherente del instrumento, sino que es un juicio, basado en evidencia, sobre qué tan apropiadas son las inferencias realizadas o acciones implementadas a partir de los puntajes de una prueba en un determinado contexto. Es importante tener en cuenta, que un test es válido para un grupo o población determinada. Por ejemplo, si un test de razonamiento espacial diseñado para estudiantes con un nivel de lectura de 8^a grado, es aplicado a un estudiante con habilidades lectora de 4^a grado sus resultados reflejan tanto su nivel lector como su capacidad de razonamiento espacial. (Covacevich, 2014, pág. 8)

Más adelante Covacevich en su artículo cómo seleccionar un instrumento para evaluar aprendizajes estudiantiles menciona que la validez relacionada con los programas de evaluación, y que además es necesario especificar el programa que se evaluará, cuál será el contexto en que se implementará, también se tiene que tener en cuenta descartar variables externas que puedan afectar los resultados del test. En la actualidad y tradicionalmente, la validez de un instrumento se ha entendido como hasta qué punto el instrumento efectivamente mide lo que se quiere medir.

Una de las prioridades que se debería tener en cuenta al momento de diseñar o seleccionar un instrumento para la evaluación de aprendizajes es la validez, estar pendiente de la validez, de cómo esta puede verse amenazada nos permitirá a tomar decisiones sobre qué evaluaciones son adecuadas, y cuál vale la pena hacer y qué uso se les puede dar a estas

evaluaciones. Además, la evaluación permite hacer juicios sobre el progreso de los estudiantes que sean robustos y útiles, y que tengan consecuencias positivas en la evaluación (Covacevich, 2014, pág. 9).

Según Darr 2005 (citado en Covacevich, 2014) “la validez está referida a inferencias y decisiones hechas por un grupo específico en un contexto específico, para juzgar la validez de un instrumento se requiere reunir mucha información”. Existen diferentes tipos, los más conocidos son: validez de constructo, contenido y criterio, tal y como lo expresa Wilson, 2005 (citado en Covacevich (2014) “Clásicamente, se habla de validez de constructo, contenido y criterio. Distintos autores hacen distintas clasificaciones de los tipos de validez que se deben considerar, por ejemplo, validez de constructo, contenido, ítem, predictiva, “face”, relativa a criterios, concurrente (Covacevich, 2014, pág. 10). A continuación, se explicará cada uno de los tipos de validez:

a) *Validez basada en contenidos*: Según Prieto & Delgado (2010), esta validez “se extendió desde el análisis del criterio al de validez de los tests predictores: Una prueba no puede considerarse válida si los ítems que la componen no muestran adecuadamente el contenido a evaluar” (Prieto & Delgado, 2010, págs. 70-71), más adelante estos dos autores hacen mención que este tipo de validez es especialmente fértil cuando las circunstancias pueden identificarse y definirse claramente. Con respecto a las pruebas dirigidas a evaluar el rendimiento académico Prieto y Delgado aconsejan que pueden especificarse en función de los objetivos de la instrucción (Conceptos y habilidades que un estudiante ha de poseer). Sin embargo, Darr 2005 (citado en Covacevich, 2014) expresa lo siguiente:

En los tests no se pueden evaluar todos los conocimientos de los estudiantes, sino solo una muestra de ellos, por lo tanto, es muy importante que esta sea una muestra adecuada del área de aprendizaje que interesa evaluar. Si esto se logra, aumenta nuestra posibilidad de hacer inferencias válidas sobre los logros de aprendizaje en un cierto dominio. (Covacevich, 2014, pág. 11)

De acuerdo con el Center for assessment and research, James Madison University (2014) si realmente se quiere saber qué es lo que mide un test, no es suficiente guiarse por el nombre, es primordial observar detalladamente los ítems que componen el instrumento. Cada ítem se puede analizar en relación con el dominio del contenido que los estudiantes han visto en

clases, otra opción sería solicitar la opinión a un experto. Así como lo afirma Prieto & Delgado (2010):

La metodología de validación descansa fundamentalmente en la evaluación de expertos acerca de la pertinencia y la suficiencia de los ítems, así como de la adecuación de otras características de la prueba como las instrucciones, el tiempo de ejecución, etc. Sin embargo, especificar con precisión el contenido de las manifestaciones de constructos como la extraversión, la memoria de trabajo o la motivación de logro es una tarea más difícil. (Prieto & Delgado, 2010, pág. 71)

De acuerdo con Covacevich (2014) “Si un instrumento es bueno, tendrá ítems que evalúen diferentes aspectos del tema evaluado, y expertos en el área, que no están familiarizados de antemano con los ítems, estarán de acuerdo en qué evalúa cada ítem” (Covacevich, 2014, pág. 11). Tanto Prieto y Delgado y Covacevich coinciden que se debe contar con la presencia de un experto en el área a evaluar. Covacevich nos advierte de los dos riesgos que deben ser evitados en la validez de contenido: “1) la subrepresentación del constructo, es decir que elementos importantes del constructo que se quiere evaluar no estén siendo evaluado. 2) la varianza relacionada con constructos que son irrelevantes para lo que está midiendo” (Covacevich, 2014, pág. 11).

Como ya se sabe que los test se componen de un conjunto de ítems o preguntas, en la cual se elaboraran para obtener una puntuación que representará el nivel de una persona en un contenido determinado (español, ciencias sociales, matemática entre otros), las instrucciones, los ejemplos de práctica que se brindan en los tests, el material de la prueba, el tiempo de ejecución son parte de la validez por contenido. Así como lo expresa Prieto & Delgado (2010):

Los tests están compuestos por un conjunto de ítems destinados a obtener una puntuación que represente el nivel de una persona en un constructo (...) son objeto de la validez de contenido las instrucciones, los ejemplos de práctica, el material de la prueba, el tiempo de ejecución (Prieto & Delgado, 2010, pág. 72).

Usualmente, para evaluar la calidad del contenido se consulta a un experto principalmente en el ámbito educativo. Los procedimientos estandarizados de consulta proveen la obtención de datos cuantitativos indicativos del porcentaje de ítems de calidad, el porcentaje de las facetas del dominio suficientemente evaluadas, así como el de los jueces que han valorado

positivamente la calidad de los materiales y la concordancia entre los expertos (Prieto & Delgado, 2010, pág.72).

b) Validez basada en el proceso de respuesta: Con el surgimiento de la ciencia cognitiva (Conocimiento), la validación de las pruebas de inteligencias, aptitudes y rendimiento académico, es importante incluir el análisis de los procesos, las estrategias que utilizan los estudiantes para la resolución de problemas y además las representaciones mentales que utilizan los participantes para resolver cada ítem del test (Covacevich, 2014, pág. 11).

De acuerdo con Prieto & Delgado (2010) “se obtendrá evidencia de validez cuando los procesos utilizados se ajusten a los que se postulan en las teorías relativas al constructo. En este enfoque predomina más el diagnóstico cognitivo” (Prieto & Delgado, 2010, pág. 72). Con respecto a las teorías que menciona los autores se refieren a la TCT o a la TRI, a la vez nos hacen referencia que la metodología para este tipo de validez, la entrevista a los colaboradores formulándole un problema de la vida diaria, el tiempo en que se tardan en contestar las repuestas; este enfoque tiene una orientación cognitiva. Para Covacevich (2014) “esta validez entrega información entre estos procesos y los constructos que se desea evaluar” (Covacevich, 2014, pág. 11).

c) Validez basada en la estructura interna: Según Prieto & Delgado (2010), “(...) el análisis de la estructura interna persigue verificar empíricamente si los ítems se ajustan a la dimensionalidad prevista por el constructor de la prueba” (Prieto & Delgado, 2010, pág. 72). Para Covacevich, (2014) “busca recoger evidencia sobre el grado en que las relaciones entre el ítem de un test y sus componentes se adecúan al constructo que supuestamente busca evaluar, el que puede implicar una sola dimensión, o varias” (Covacevich, 2014, pág. 11).

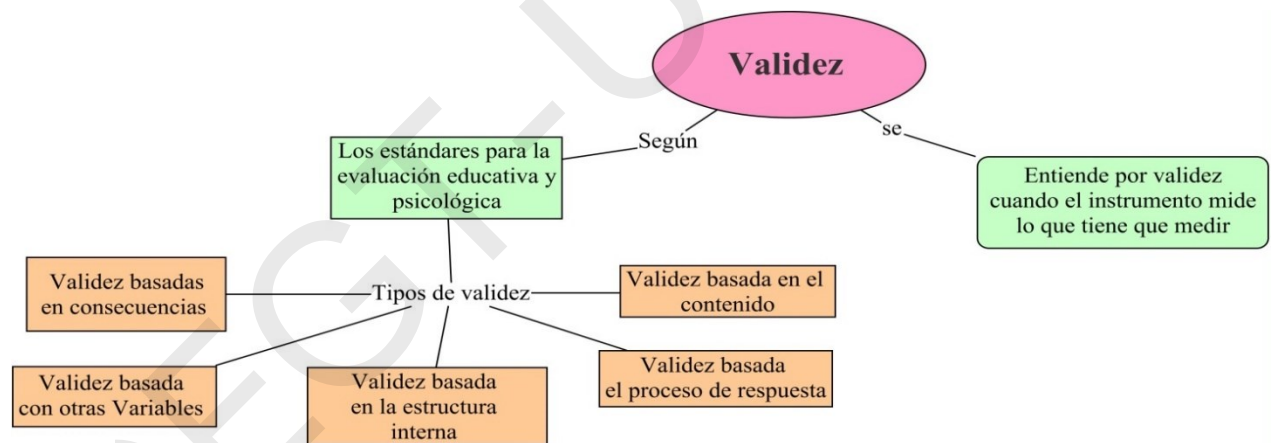
d) Validez basada en relaciones con otras variables: De este tipo se desglosa en validez convergente y discriminatoria. La validez convergente demuestra que los constructos que se espera que estén relacionados de hecho lo están, mientras que la validez discriminatoria de muestra lo contrario, es decir, prueba que los constructos que no deben tener ninguna relación de hecho no la tengan. Tal y como lo expresa Covacevich (2014):

Este tipo de evidencia se desglosa en validez convergente y discriminatoria. La evidencia relativa a la validez convergente implica comparar los resultados obtenidos en un determinado test con los obtenidos por los mismos estudiantes en tests que midan el mismo constructo, o constructos similares. Se espera que los puntajes de un cierto instrumento se

correlacionen con otros que declaran medir constructos iguales o parecidos. La evidencia relativa a la validez discriminatoria se obtiene comparando los resultados obtenidos en el test con otras evaluaciones que midan constructos opuestos o diferentes. (Covacevich, 2014, pág. 12)

e) *Validez basada en consecuencias*: El uso de las evaluaciones en particular tienen o pueden tener consecuencias negativas, o las consecuencias de usar esos resultados pueden estar en contra de los objetivos educativos. Esta es una consideración se debe tomar en cuenta al momento de elaborar o de seleccionar un instrumento esta decisión dependerá de la persona que está realizando la investigación. Otro punto importante que es necesario analizar es si las consecuencias indeseables se deben al constructo que se quiere medir o al instrumento que se está utilizando (Covacevich, 2014, pág.12). En manera de resumen la figura 3 presentará de manera sintetizado lo antes explicado:

Figura 3: Tipo de validez



Fuente: Elaboración propia con base en Covacevich (2014)

En la figura 3 proporciona de manera sintetizada todo lo explicado anteriormente. De esta manera se puede comprender los tipos de validez con lo que se cuenta para poder seleccionar al momento de crear o elaborar nuestro instrumento.

b) *Confiabilidad*: Según Cohen y Swerdlik (2009) “La confiabilidad se refiere a la consistencia con que el instrumento mide, o visto de otro modo, al grado de error presente en la medida, un instrumento confiable mide de la misma manera” (p. 13). Sin embargo, Haertel

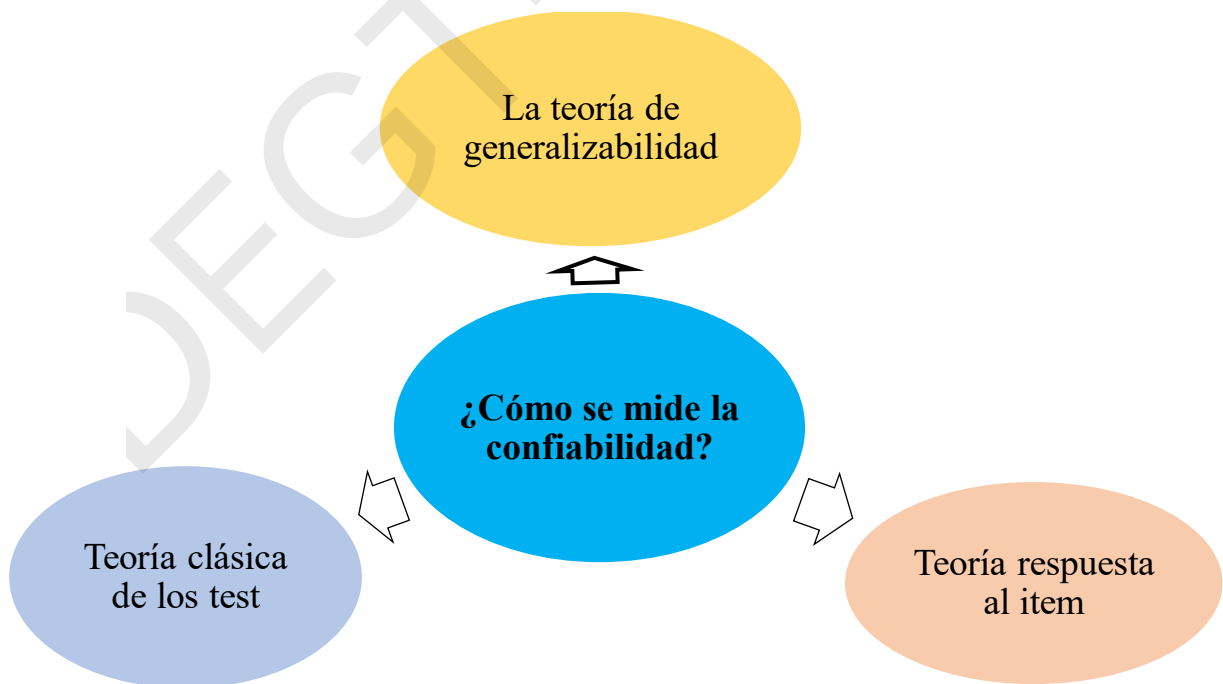
(2006) “La confiabilidad al igual que la validez, se tiene que entender en contextos y propósitos evaluativos determinado” (Covacevich, 2014; p. 13).

Para Prieto & Delgado (2010) “La fiabilidad se concibe como la consistencia o estabilidad de las medidas cuando el proceso de medición se repite” (Prieto & Delgado, 2010, pág. 67). Los individuos responsables de que los instrumentos posean confiabilidad son todas aquellas que diseñan el manual de aplicación para los instrumentos o las instituciones que se encargan de publicar los instrumentos en páginas web. De acuerdo con Covacevich (2014) los responsables de brindar la información sobre la confiabilidad deberán de brindar lo siguiente:

La información que debe presentarse es la identificación de las principales fuentes de error, resúmenes estadísticos que cuantifiquen el tamaño de estos errores, y el grado de generalizabilidad de los puntajes entre distintas formas, puntuadores, administradores, y otras relevantes. También una descripción de la población con el que fueron hechas estas estimaciones. (Covacevich, 2014, pág. 12)

En la Figura 4 se menciona el marco teórico con el cual se mide la confiabilidad. Para proseguir a explicarla y tengamos unas ideas más claras de cómo podemos conocer la confiabilidad.

Figura 4:Marco teórico para analizar confiabilidad



Fuente: Elaboración propia con base en Covacevich (2014)

En la figura 4 se puede observar las teorías de como analizar la confiabilidad, sin embargo, la confiabilidad se calcula de manera diferente todo depende del enfoque o teoría que se vaya a utilizar. También dependiendo del tipo de confiabilidad que se utilice reportará información de diversas maneras: como ser las desviaciones estándares, el coeficiente, el chi cuadrado entre otras.

a) La teoría de generalizabilidad: Esta teoría utiliza coeficientes que permiten especificar y estimar los diversos componentes de la verdadera varianza del puntaje, la varianza del error, y varianza del puntaje observado. Esta misma autora más adelante menciona que esta teoría utiliza dos tipos de estudio, de generalizabilidad (G-Study) y de decisión (D-Study). Habitualmente las herramientas de análisis utilizadas es ANOVA, asimismo como el programa computacional GENOVA (Covacevich, 2014,pág.14).

La teoría clásicas de los test (TCT): esta teoría analiza los siguientes tipos de confiabilidad: Coeficiente derivado de la administración de formas paralelas en sesiones independientes, coeficiente obtenido por la administración del mismo instrumento en ocasiones separadas (Se conoce también como test retest) y el coeficiente basado en la relación entre puntajes derivados de ítem individuales o subtest dentro de un test, dicha información es obtenida de la misma administración (conocida como coeficiente interno, o inter ítem), esta teoría el coeficiente ,as usado en el Alpha de Cronbach (Covacevich, 2014,pág14).

La teoría de respuesta al ítem (TRI): A partir del año 90 esta teoría ha sido utilizada en la mayoría de las evaluaciones estudiantiles. La TRI es una familia de modelos estadísticos usados para analizar los datos de ítems de tests entregando un proceso estadístico unificado para estimar características de los ítems y los individuos examinados y definir como estas características interactúan en el desempeño en los ítems y el test (Covacevich, 2014, pág. 14). Así como existen teorías para conocer la confiabilidad, también se cuentan con métodos para verificar la confiabilidad en la tabla 1 se le presenta a cada uno de estos métodos, las dimensiones, cantidad de vez que se tiene que aplicar y los estadísticos a utilizar.

Método	Dimensión de la confiabilidad	N° de sesiones de administración de la prueba	Estadístico
Test retest	Estabilidad	2	R de Pearson
Formas paralelas	Equivalencia	2	R de Pearson
Partición en mitades	Consistencia interna	1	R y formula de correlación Spearman-Brown
Coefficiente alfa	Consistencia interna	1	Alfa Kuder-Richardson
Acuerdo entre examinadores	Confiabilidad entre examinadores	1	Kappa W de Kendall
			Coefficiente de correlación intraclase

Tabla 1: Métodos para verificar la confiabilidad

Fuente: Elaboración propia con base en los contenidos del módulo ámbitos de la Psicometría y evaluación educativa impartida por Zarko Vukmirovik y Mike Fast (2016)

Esta tabla facilita la comprensión y método a utilizar para el tipo de instrumento que se usará en la recolección de datos de nuestra investigación.

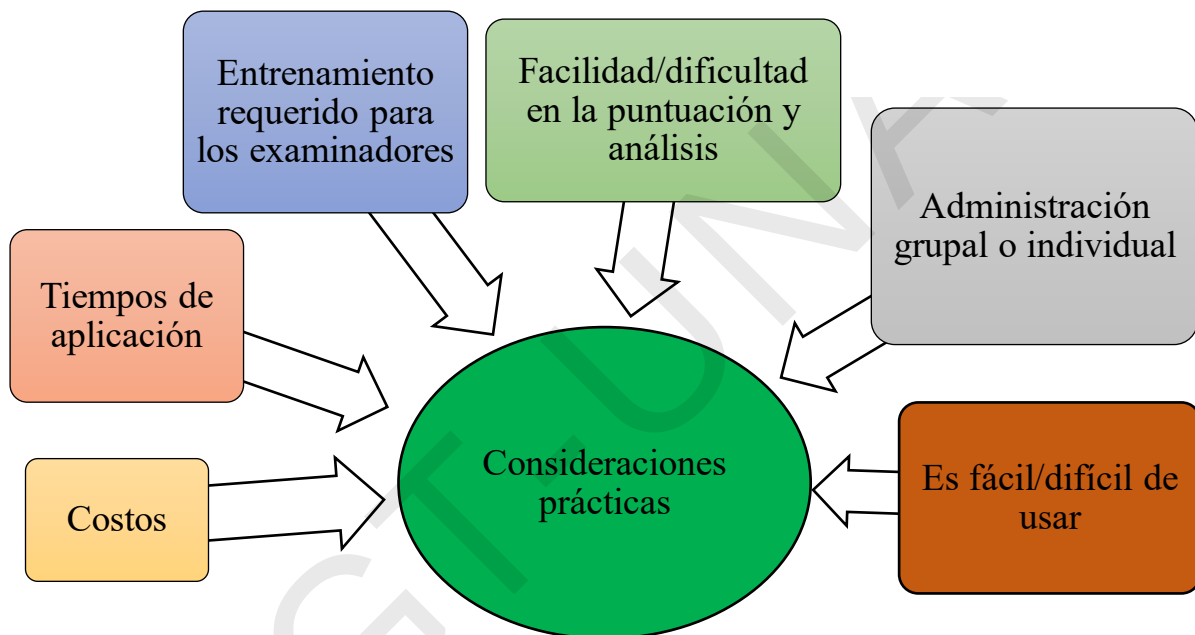
c) Estandarización y su importancia para la validez y confiabilidad: Es importante saber que la forma de cómo se analice y se implemente un instrumento puede afectar la validez y la confiabilidad. De acuerdo con McCallin (2006) citado por Covacevich (2014):

Para que un test sea válido, no basta con que las características técnicas del instrumento lo sean, también es fundamental que todos los instrumentos hayan sido administrados bajo las mismas condiciones estandarizadas de aplicación. Esto significa que las instrucciones, el contexto de aplicación y los procedimientos de puntuación han sido exactamente los mismos para todos los examinados. Eso asegura que los datos puedan ser adecuadamente, interpretados, comparados, y usados de acuerdo al principio de que cada usuario fue tratado de manera justa. Por lo tanto, cualquier alteración en la estandarización de la aplicación afecta la comparabilidad y la validez de la prueba (Covacevich, 2014, pág. 15).

En este punto empezamos hablar de la estandarización y para que un instrumento sea estandarizado dependerá de las condiciones de la aplicación (si las instrucciones son iguales,

si las condiciones de administración y si posee el mismo puntaje para todos los examinados). La estandarización de un instrumento es significativa sin interesar el tipo de instrumento ya sea este de criterio o de norma. Para que el instrumento sea de calidad tal y como se explicaron anteriormente, ahora se presentará una serie de consideraciones que es necesario tomar en cuenta al momento de elaborar o seleccionar un instrumento. En la figura 5 se presentan las consideraciones que se deben de tener en cuenta.

Figura 5: Consideraciones prácticas para elaborar o seleccionar un instrumento



Fuente: Elaboración propia con base en Covacevich (2014).

Estas consideraciones ayudan a realizar un presupuesto y a tomar las decisiones correctas para iniciar una investigación y de esta manera todo sea un éxito.

Además de las consideraciones técnicas (calidad de los instrumentos) antes explicada, también existen una serie de consideraciones prácticas que son importantes al momento de decidirse o elaborar un instrumento. Tal y como se observa en la figura 5, a continuación, se realizará una breve explicación.

a) *Los costos:* el investigador antes de proceder a desarrollar un instrumento tiene que hacer una planificación o presupuesto que tendrá que llevar a cabo al momento de iniciar el proceso. Ya es importante cuantos recursos se necesita para realizar la evaluación en su totalidad.

b) Tiempo de aplicación: Es necesario crear un manual de instrumento donde especifique el tiempo que durará el test. En algunos casos los instrumentos toman en consideración los bloques escolares para que las respuestas de los estudiantes sean más confiables y no afecten los resultados de los instrumentos. De acuerdo con Covacevich, (2014) “(...) Hay que evaluar los tiempos de aplicación contra el tiempo disponible y el tipo de estudiante al que se le aplicará el test, para evaluar si es adecuado o no al contexto particular” (Covacevich, 2014, pág. 18)

c) Entrenamiento requerido por los examinadores: Es de suma importancia capacitar al personal que aplicará las pruebas en los centros educativos, ya que son ellos los que responsable de que los estudiantes no copien al momento de responder las prueba. Además, será encargado de brindar las instrucciones pertinentes para que los estudiantes contesten el test de manera confidencial y honesta posible.

d) Facilidad/ dificultad en la puntuación y análisis: Tenemos que saber que no todos los test son iguales, algunas pruebas se requieren tiempo o un entrenamiento especial para ser analizados o puntuados los datos, lo que implica contratar y entrenar a una persona para que lo realice.

e) Administración grupal o individual: Esto dependerá del tipo de investigación o el enfoque el investigador le quiera dar al estudio, algunas pruebas se requieren de una aplicación uno a uno, es decir de manera individual, mientras otros instrumentos son administrados de forma grupal. Según Covacevich (2014) “(...) lo más fácil es ocupar un instrumento de aplicación grupal, es decir, que implica tener solo uno o dos examinadores por sala o aula” (Covacevich, 2014, pág. 19). En cuanto a la administración individual es más utilizado cuando se evalúa niños pequeños.

f) Es fácil/difícil de usar: Un instrumento debería ser fácil de utilizar, ya que el desempeño de los estudiantes puede salir afectados sino comprendieron las instrucciones, y más si son niños que con facilidad pierden la concentración, por ellos es necesario, elaborar un instrumento que se comprensible para todos los estudiantes.

2.3.4 Taxonomía de objetivos cognoscitivo

Desde mediados de la década de 1950 se ha prestado mucha atención a los sistemas formales y estándar de la clasificación de los objetivo cognoscitivos, afectivos y psicomotrices de la instrucción, así como diseñar una prueba de observación para la selección

de personal requiere un análisis preliminar del trabajo a desempeñar, la preparación de una prueba para medir objetivos instruccionales específicos es más afectiva cuando las conductas que van a evaluarse se definen claramente al inicio (Aiken, 2003, pág. 20). En la tabla 2 se les brindará algunos compendios sobre este tema.

Tabla 2: Compendio ilustrativo de los objetivos cognoscitivos

Autor y año	Objetivo Cognitivo	Definición	Verbos relacionados
Bloom y Krathwohl (1956)	Conocimiento	La RAE (2014), conocimiento es acción y efecto de conocer, y el significado de conocer Según la RAE “Averiguar por el ejercicio de las facultades intelectuales de la naturaleza, cualidades y relaciones de las cosas”, es decir, que el conocimiento es recordar hechos específicos sobre algún tema o experiencias vividas.	✓ Definir ✓ Identificar ✓ Mencionar ✓ Nombrar ✓ Escribir ✓ Reproducir ✓ Numerar ✓ Leer ✓ Etiquetar
	Comprensión	Para Aguilera, Barbán, Cabrera, & Segura (2011) definen a la comprensión como: Un proceso que requiere la sucesión progresiva de etapas para lograr el fin deseado, así como también de algunas categorías filosóficas presentes en el proceso de comprensión del significado: causa-efecto, esencia-fenómeno, general-particular. Es la facilidad con la que el ser humano cuenta para entender el significado o el propósito de algún tema que ha aprendido.	✓ Clasificar ✓ Citar ✓ Convertir ✓ Describir ✓ Discutir ✓ Explicar ✓ Exponer ✓ Resumir
	Aplicación	De acuerdo a la REA (2014), proviene de la	✓ Usar ✓ Carcular

	acción y efecto de aplicar, la cual esta última significa según la RAE: “Emplear, administrar o poner en práctica un conocimiento, medida o principio, afin de obtener un determinado efecto o rendimiento en alguien o algo” en el caso de educación serian en los estudiantes	✓ Construir ✓ Determinar ✓ Solucionar ✓ Resolver
Análisis	Según Aiken, (2003), “es descomponer algo para reverlar su estructura y las interrelaciones que hay entre sus partes” (Aiken, 2003, pág. 21)	✓ Analizar ✓ Discriminar ✓ Categorizar ✓ Comparar ✓ Contrastar ✓ Analizar
Síntesis	De acuerdo con René Descartes, la síntesis consiste en ordenar los pensamiento o idea de manera sencilla y fácil ascendiendo poco a poco en otras palabras de lo sencillo a lo complejo. Sin embargo, para Aiken (2003), es combinar varios elementos o artes en un todo estructural.	✓ Crear ✓ Adaptar ✓ Anticipar ✓ Planificar
Evaluación	Para Castillo (2009), la evaluación es: Un campo disciplinar que se define como un proceso complejo de producción de información de naturaleza cuantitativa como cualitativa, con el propósito de tomar decisiones.	✓ Valorar ✓ Justificar ✓ Juzgar ✓ Decidir ✓ Concluir

Fuente: Adaptado de herramienta para formular objetivos de la Universidad Interamericana para el desarrollo (2003) y Test psicológico y evaluación de Aiken (2003).

2.3.5. Marco conceptual y tabla de especificaciones

Marco conceptual: Al momento que uno decide elaborar una prueba lo primero que realizar es un marco conceptual (En inglés Test Framework), este marco conceptual nos ayudará a definir la estructura organizadora y las características claves de una evaluación y esto nos permitirá construirla, además este marco definirá el contenido a evaluar. La función de este es asegurar coherencia y consistencia a en la construcción de las pruebas de evaluación, la cual abarca los siguientes aspectos:

1. Detalles precisos de la población meta que se desea evaluar.
2. Explicación precisa del propósito de la evaluación, que decisión se tomará con la información obtenida, y como se ayudará a resolver los problemas que sean planteados por el estudio.
3. Explicar el contenido a evaluar.
4. Brindar información contextual que permita entender el motivo y la forma de evaluación.
5. Tabla de estándares de contenido que representa un análisis del contenido a evaluar.

Creación de un Marco operativo o tabla de especificaciones para una evaluación (en inglés test Blueprint): El momento de tomar la decisión de elaborar o diseñar una prueba, es necesario empezar por planificar en construir una tabla de especificaciones para el diseño de los ítems, esta tabla debe de ser razonable y detallada en cuanto a término del conocimiento y habilidades que esperan que los examinados demuestren. Esta tabla nos ayudará elaborar los reactivos de una manera sencilla. Además, nos brindará una relación de los ítems que aparecerá en la forma operativa o final, ya que sería arriesgado elaborar los que uno quiere para ensamblar la forma operativa, por ello se recomienda elaborar más ítems de lo necesario, ya que en las pruebas pilotos uno que otro puede fallar en cuanto a la confiabilidad, validez, discriminación.

2.3.6. Elaboración de reactivos

Una vez que se cuenta con la tabla de especificaciones, está servirá de guía para la elaboración de los reactivos con los que se van a evaluar los objetivos, ya con la tabla de especificaciones diseñada, se procederá a la creación de reactivos, es recomendable que para las pruebas objetivas se preparen alrededor del 20% de los ítems, de lo que realmente se

necesitaran, de manera que se cuente con una cantidad adecuada de buenos reactivos para el ensamble final de la prueba. Todos los reactivos incorporan procedimientos para brindar información acerca de los individuos, pero dicha información varía según la naturaleza o el objetivo con el cuál fue creada el reactivo. (Aiken, 2003, pág. 20)

El término reactivo es el planteamiento de un problema para que el individuo brinde una respuesta a esas preguntas o reactivos con el fin de conocer el nivel de dominio que posee en ese tema. En una prueba estandarizada se pueden encontrar diferentes tipos de reactivo tales como: reactivos de ensayo y reactivos objetivo (reactivos, de respuesta corta, de verdadero y falso de apareamiento, reactivos de opción múltiples).

Los reactivos de ensayo: A estos tipos de reactivos también se le conoce como respuesta breve en este tipo de pregunta se puede evaluar las habilidades personales que posee el individuo para: relacionar, organizar y comunicar ideas sobre determinado tópico. Estos tipos de reactivos no requieren mucho tiempo para su elaboración y además disminuye la probabilidad que los estudiantes contesten de manera correcta los ítems por medio de la adivinanza. Sin embargo, estos tipos de reactivos son tan generales que se interpretan de manera distinta por diferentes individuos y el tiempo de aplicación no es suficiente para determinar el conocimiento que tiene una persona sobre el tema. De acuerdo con Aiken (2003) quien expresa “que no se debe usar reactivos de ensayo cuando el mismo conocimiento o habilidad puede ser evaluados con reactivos objetivos” (Aiken, 2003, pág. 26)

Reactivos Objetivos: Este tipo de reactivos puede clasificarse de forma fácil e imparcial y se requiere menos tiempos para contestar cada reactivo. Los reactivos más conocidos de este tipo están: las de verdadero y falso, respuesta corta o de completación, apareamiento y opción múltiple. Al momento de elaborar los reactivos objetivos se debe de cuidar que los reactivos sean claros, precisos y gramaticalmente correctos, además se tienen que escribir de acuerdo al lenguaje de los estudiantes a las que se dirige la prueba. Las personas encargadas de elaborar este tipo de reactivos deben tener cuidado de no incluir claves para la respuesta correcta, además tiene que evitar reactivos entrelazados o interrelacionados esto quiere decir que dos preguntas están entrelazadas cuando es preciso conocer la respuesta de una de las preguntas para poder contestar la otra; y dos ítems están interrelacionados cuando el planteamiento de uno provee una señal para la respuesta del otro (Aiken, 2003, pág. 27)

2.3.7 Principios para el desarrollo de las pruebas o teoría de la medición

Para una comprensión y manejo adecuado de las propiedades psicométricas de los test es preciso ir más allá del puro cálculo empírico, y conocer los fundamentos científicos en los que se basan estos. Para poder comprender y manejar adecuadamente estas propiedades psicométricas, se tratarán dos teorías que nos permiten analizar los datos, estas dos teorías son: La teoría clásica de los test (TCT) y la teoría de respuesta al ítem (TRI), también se conocerá los modelos, las limitaciones de cada una de ella, sin embargo, en la TRI, se mencionaran las diferencias entre estas dos teorías.

2.3.7.1 Teoría Clásica de los Test

Según Muñiz (2010) esta teoría inicio en los “trabajos pioneros de Spearman a principio del siglo xx (...) Serán Lord y Novick (1968) quienes lleven a cabo una reformulación de la teoría clásica” (Muñiz, 2010, pág. 59). Por otro lado, Attorresi, Lozzia, Abal, Galibert, & Aguerri, (2009) expresan “Desde su primera formulación (Spearman 1904), la teoría clásica de los test (TCT) ha servido como modelo para dar una interpretación a los puntajes de las personas en los tests” (Attorresi, Lozzia, Abal, Galibert, & Aguerri, 2009, pág. 179) . Prieto & Delgado (2003) mencionan que “La teoría clásica de los test ha sido el principal modelo psicométrico empleado en la construcción y análisis de los test” (Prieto & Delgado, 2003, pág. 94)

2.3.7.2 Modelo lineal clásico de la TCT

Según Moscoso & Pérez-Gil (2008) los trabajos presentados por Spearman en el cual planteo viene a ser el clásico ““Modelo Lineal de Puntuaciones” denominado teoría débil de la Puntuación verdadera o simplemente como la conocemos hoy teoría clásica de los tests” (Moscoso & Pérez-Gil, 2008, págs. 39-40). Estos mis autores Moscoso & Pérez-Gil (2008) mencionan que:

La tesis central de Spearman era formalizar un modelo estadístico que pudiese tener en cuenta las estimaciones de los errores de media inherentes a todo proceso de medición. De este modo se posibilitaría una fundamentación adecuada para las puntuaciones de los tests en el campo de la psicología. (Moscoso & Pérez-Gil, 2008, pág. 40).

Por otro lado, Muñiz (2010) expresa que si una persona obtiene en un test 70 puntos de una puntuación empírica, este modelo permite saber cual es la puntuación verdadera ni el error contenido en esa puntuación. Solo contamos con una puntuación y con dos incógnitas, la puntuación verdadera (v) y la del error (e), hasta estos momentos no se avanzado nada (Muñiz, 2010, pág. 60). Sin embargo Muñiz (2010) y Moscoso & Pérez-Gil (2008) coinciden que para avanzar hay que auxiliarse de los tres supuestos de esta teoría.

2.3.7.3 Modelo de formas paralelas de la TCT

Spearman asume que se puede construir dos más test que midan lo mismo, aunque con distintos ítems, es decir dos formas equivalentes de mismo test a lo cual Spearman los denomino test formas paralelas o equivalentes (Moscoso & Pérez-Gil, 2008; pág. 44). Es decir, dos conjuntos de puntuaciones X y X' , y se dice que son equivalentes o paralela si ambas tienen la misma puntuación verdadera, $X = V + e$ y $X' = V + e'$ y, además, ambas poseen la misma varianza de error. A partir de los supuestos, en el caso de que se cumplan, se consigue establecer de manera simple e intuitivo que para el número (k) de tests paralelos, y desde el punto de vista de sus valores paramétricos, las medias de la población son idénticas, cada tests paralelo presenta la misma varianza poblacional y las intercorrelacionan entre cada par de ellos son iguales (Moscoso & Pérez-Gil, 2008, págs. 41-42).

Los supuestos (que se explica a continuación) hacen referencia a las puntuaciones empíricas, sí se pueden comprobar y dado que la TCT toma el modelo probabilístico y que los estadísticos obtenidos en muestras bastante amplias constituyen buenos estimadores de los parámetros poblacionales, las deducciones expuestas pueden ser fácilmente sometidas a contrastación. De las situaciones de paralelismo se deducen múltiples consecuencias, las más importante es que la correlación entre las dos medidas equivalentes expresa la proporción de varianza en puntuaciones empíricas que es debida a la varianza de las puntuaciones verdaderas (Moscoso & Pérez-Gil, 2008; p. 45).

2.3.7.4 Supuestos de la TCT

El primer supuesto de esta teoría Moscoso & Pérez-Gil (2008) lo expresan de la siguiente manera se define el concepto de puntuación verdadera de un sujeto como la esperanza matemática de la puntuación empírica observada en infinitos ensayos de medidas independientes realizados con ese individuo, y, considerando el error de medida como un componente aleatorio que simplemente se suma a la puntuación verdadera del sujeto en el

test, es decir, la puntuación en el test libre del componente de error. (Moscoso & Pérez-Gil, 2008, pág. 42)

Es decir que este primer supuesto establece que la puntuación verdadera, coincide con el valor esperado de la puntuación empírica, y como la puntuación verdadera es un concepto matemático se representa así $V = E(X)$ (Muñiz, 2010 y Moscoso & Pérez-Gil, 2008).

El segundo supuesto que planteó Spearman fue que no existe relación entre la puntuación verdadera de una persona y el tamaño de los errores que afectan a esa puntuación; es decir que el valor de la puntuación verdadera de un individuo no tiene nada que ver con el error que afecta esa puntuación, en otras palabras, puede haber puntuaciones verdaderas altas con errores altos o bajos y no hay conexión entre el tamaño de la puntuación verdadera y el tamaño de los errores (Muñiz, 2010). (Muñiz, 2010, pág. 61)

El tercer supuesto según Muñiz (2010) es el siguiente:

Establece que los errores de medida de las personas en un test no están relacionados con los errores de medida en otro test distinto, es decir, no hay ninguna razón para pensar que los errores cometidos en una ocasión vayan a covariar sistemáticamente con los cometidos en otra ocasión. (Muñiz, 2010, págs. 61-62)

En otras palabras, los tres supuestos o principios básicos de la TCT estarían sintetizados de la siguiente manera:

1. La puntuación observada (X) es una combinación lineal del componente verdadero (V) y el componente de error (e).
2. La puntuación verdadera puede ser conceptualizada como la puntuación media que una persona obtendría en el mismo test en una cantidad infinita de sesiones de evaluación repetida.
3. V es hipotético porque no es posible obtener una cantidad infinita de puntajes del test, sin embargo, es el aspecto central de la TCT.

2.3.7.5 Limitaciones de la TCT

A pesar de lo simple y fácil adaptación a diversos campos, la TCT ha recibido críticas o limitaciones brindadas tanto a los conceptos que maneja como los supuestos en los que se fundamenta. En la siguiente tabla aparecen las limitaciones más importantes de esta teoría

Tabla 3: Limitaciones más importantes de la TCT

1. En la TCT las características de los ítems o prueba dependen de las características de la muestra que contesta al test, por lo tanto, también las características de la población de referencia.
2. Las características del test dependen además de la muestra, del número de ítems que compone el test.
3. La estimación de la puntuación verdadera depende de las características de la muestra y del propio test.
4. La definición de tests paralelo es muy limitada.
5. Homocedasticidad de los errores de medida

Fuente: Obtenido Gómez-Benito & Hidalgo-Montesinos (2003) y en Terrón (2010).

En la tabla 3 se evidencia las dificultades que presentan la TCT, es necesario que al momento de utilizar se tenga en cuenta.

2.3.7.6. Los parámetros de la TCT

Los parámetros con los que cuenta esta teoría son: El parámetro “a” que representa la dificultad clásica del ítem y el parámetro “b” que representa la discriminación clásica del ítem:

El parámetro a: Dificultad clásica del ítem- valor P-value: En los ítems de opción múltiples representa la cantidad de estudiantes que contestaron el ítem correctamente y en las respuestas escritas el valor de P ajustado es la medida de los ítems dividido por el puntaje máximo posible. La dificultad del ítem en la TCT, representa que tan fácil es contestar ese ítem ya que representa la proporción de estudiantes que respondieron correctamente. El rango del valor es de 0 a 1 con valores arriba de 0.9 indican ítems demasiado fácil y valores bajo de 0.3 indicando ítems demasiado difíciles.

Parámetro b: Discriminación clásica del ítem: correlación ítem – puntaje total: La correlación ítem-puntaje total indica el grado de acuerdo entre los puntajes individuales de los ítems y el puntaje total de la prueba. En cambio, la correlación punto biserial es un tipo específico de correlación usados para ítems dicotómicos o ítems de opción múltiple, esta indica el acuerdo entre escoger cada opción (u obtener un punto en ítems de respuesta breve) y el total de la prueba. Por otro lado, una correlación ítem-puntaje total muestra que los

estudiantes que contestaron correctamente un ítem o que recibieron un puntaje mayor en un ítem de respuesta escrita también recibieron un alto puntaje en la prueba y viceversa. Los valores arriba de 0.3 son deseables y los valores debajo de 0.25 deber ser investigado.

2.3.7.2 Teoría de respuesta al ítem

Varios autores han coinciden que la TRI representa una opción a la TCT, que además ha brindado las excelentes respuestas a las dificultades que muestra el modelo tradicional y simboliza una innovación al paradigma de las propiedades de la teoría de los tests. Según Muñiz (2010) la TRI expresa “que el primero surgimiento de esta teoría se pude atribuir a los trabajos pioneros de Thurstone en 1925, y que Lawley en los años 1943, 1944 y Tucker en 1946 continuaron aportando a esta teoría” (Muñiz,2010, pág. 63). Hidalgo-Montesinos & French (2016) expresan “El análisis usando TRI, a través de modelos matemáticos, nos proporciona una visión de la relación entre el nivel en el rasgo de un individuo (por ejemplo, nivel de depresión) y las características de los ítems” (Hidalgo-Montesinos & French, 2016, pág. 14). Para Attorresi et al., (2009) mencionan que:

El objetivo sustancial de la TRI es la construcción de instrumentos de medición con propiedades invariantes entre poblaciones. Si dos individuos presentan idéntico nivel de rasgo medido ambos tendrán igual probabilidad de dar la misma respuesta, independientemente de la población de pertenencia. (Attorresi, Lozzia, Abal, Galibert, & Aguerri, 2009, pág. 180)

Por otro lado, Hidalgo-Montesinos & French (2016) nos mencionan que una de las características importante de la TRI es:

Proporciona estimaciones invariantes de las propiedades psicométricas de los ítems, así como de las características de los sujetos, es decir, que los parámetros que caracterizan al ítem y al test son menos dependientes de la muestra particular de sujetos utilizada y que los parámetros que caracterizan al sujeto no dependen de la muestra particular de ítems utilizada. (Hidalgo-Montesinos & French, 2016, pág. 14)

Bajo la TRI se intenta relacionar el rasgo latente que se desea medir en las personas con la probabilidad de contestar correctamente al ítem. En esta teoría la relación se representa en la Curva Característica del ítem donde:

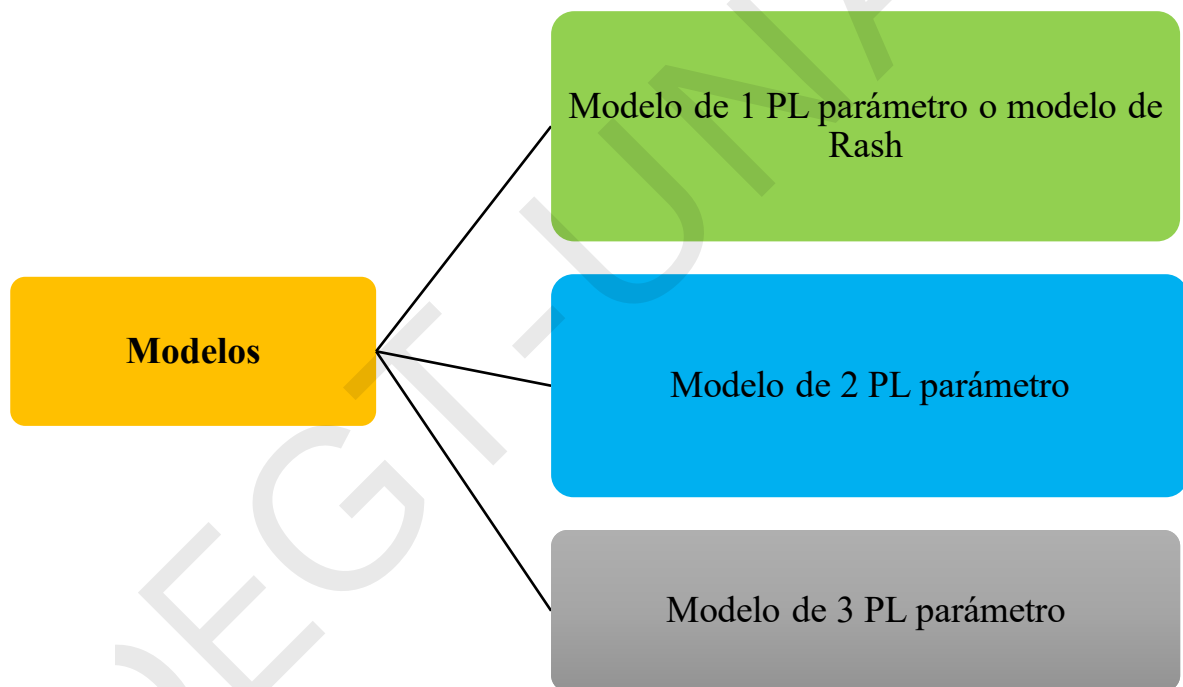
1. La escala de habilidad o rasgo latente adquiere valores prácticos de -3.0 a +3.0 y se representan con la letra griega Θ

2. La probabilidad de contestar al ítem correctamente tiene un rango de 0.0 a 1.0 y se representa con $P(\Theta)$

2.3.7.3 Modelo de la TRI

La TRI facilita una extensa escala de modelos que ayudan a trabajar con pruebas que sean unidimensionales o multidimensionales y con diferente formato de respuesta (Poliatómicos, dicotómicos, continuos). Los modelos de la TRI se diferencian en función del número de parámetros que contiene dependiendo de los supuestos que subyacen a los datos (Hidalgo-Montesinos & French, 2016: p. 14). La cual esta conformada por los siguiente modelos.

Figura 6: Modelo de la TRI



Fuente: Elaboración propia

El modelo de 1 PL parámetro o modelo de Rasch: Georg Rasch contribuyó a la fundación del campo de la TRI hace más de 50 años es por ello que este modelo lleva su nombre, y el cual está caracterizado por el uso de un solo parámetro para modelar la relación entre la dificultad del ítem y la habilidad de una persona este modelo es considerado por mucho tiempo como el estándar de excelencia en la medición objetiva. Según Prieto & Delgado (2003) Este modelo posee los siguientes supuestos en lo cual se fundamenta:

1. El atributo que se desea medir puede representarse en una única dimensión en la que se situaría conjuntamente las personas y los ítems.
2. El nivel de la persona es el atributo y la dificultad de ítem determinan la probabilidad de que la respuesta sea correcta.
3. Las *ventajas* de este modelo para Prieto & Delgado (2003) son las más relevantes es: Medición conjunta, objetividad específica, propiedades de intervalo y especificidad de error típico de medida. *La medición conjunta:* representa que los parámetros de los individuos y de los ítems se expresan en las mismas unidades y se localizan en el mismo continuo. Es decir que esta ventaja confiere a este modelo un carácter más realista que la de la TCT, además esta ventaja permite analizar las interacciones entre las personas y los ítems (Prieto & Delgado, 2003; pág. 95).
4. *Objetividad específica:* Una medida solo puede ser considerada valida sino depende de las condiciones específicas con que han sido obtenidas. En otras palabras, la diferencia entre dos individuos en un atributo no debe depender de los reactivos específicos con los que sea estimada y de igual manera la diferencia entre dos ítems no debe depender de los sujetos específicos que se manejen para cuantificar los datos (Prieto & Delgado, 2003; pág. 95).
5. *Propiedades de intervalo:* Es importante mencionar que las discrepancias en la escala es la misma a lo largo de la propiedad medida. En otras palabras, las diferencias iguales entre una persona y un reactivo le corresponden probabilidades idénticas de una respuesta correcta. La métrica intervalar tiene gran importancia, ya que es una condición necesaria para utilizar con rigor en los análisis paramétricos más frecuentemente empleados en las ciencias sociales (Prieto & Delgado, 2003; pág. 96).
6. *Especificidad del error típico de medida:* La objetividad específica no implica que la precisión de las estimaciones de los parámetros sea similar en diferentes conjuntos de ítems y de individuos. Si los reactivos son fáciles, se estimarán con más precisión los parámetros de las personas de bajo nivel. De manera similar, si los sujetos son de alto nivel, se estimarán con mayor precisión los parámetros de los ítems difíciles. El modelo de Rasch permite: (a) medir la cantidad de información con la que se calcula cada punto de la dimensión y (b) seleccionar los ítems que permiten incrementar la información en regiones del atributo previamente especificadas. Este último es de

sumo interés en los tests referidos a criterio, en los que interesa maximizar la fiabilidad en torno a los puntos de corte (Prieto & Delgado, 2003; pág. 96).

Modelo de 2 PL parámetro o modelo de Birnbaum: El modelo de 1 PL antes explicado admite que todos los ítems discriminan de igual manera entre las personas con diferente capacidad latente. Sin embargo, Sánchez (2004) “esta hipótesis es poco realista, ya que, en la práctica, lo habitual es que la discriminación entre los individuos de reducida capacidad latente y aquellos otros de elevada capacidad difiera de unos ítems a otros” (Sánchez, 2004, pág. 8)

Modelo de 3 PL parámetros: Este modelo es muy popular ya que se utiliza con ítems de respuesta dicotómicas. Estos tipos de reactivos son muy utilizados en tests de desempeño académico, donde cada ítem se formula con dos o más opciones de respuesta donde hay una respuesta correcta y las demás incorrectas (Hidalgo-Montesinos & French, 2016; pág.14).

2.3.7.4 Parámetro de la TRI

A igual que la TCT, la TRI también cuenta con parámetros que se basan en los modelos antes explicados, los parámetros son:

Parámetro a: Representa a la discriminación del ítem: la discriminación indica que tan bien un ítem diferencia a los examinandos, cuyas habilidades están por debajo del punto de ubicación de la Curva Característica del Ítem (CCI) de aquellos examinandos cuyas habilidades están por arriba del punto de ubicación de la CCI. Este parámetro se relaciona con la pendiente de la CCI, cuanto más inclinada sea la curva, mayor será el valor del parámetro y mostrara una mejor discriminación del ítem. El valor de este parámetro siempre es positivo. La capacidad discriminatoria de da para los valores que tenga la habilidad θ en torno al parámetro b. Para efectos de la construcción de pruebas los valores aceptables son aquellos mayores a 0.5 (Attorresi, Lozzia, Abal, Galibert, & Aguerri, 2009, págs. 182-183)

Parámetro b: Indica la dificultad del ítem, este parámetro muestra en que parte de la escala el ítem funciona mejor. Es decir, un ítem fácil es útil para identificar a un individuo de baja habilidad mientras que un ítem difícil es útil para identificar sujetos con alta habilidad. Según Attorresi, et al., (2009) “el índice de dificultad o parámetro de localización del ítem concuerda con el valor necesario para tener probabilidad 0.5 de contestar la respuesta clave al ítem” (Attorresi, Lozzia, Abal, Galibert, & Aguerri, 2009, pág. 182). En este sentido el

parámetro b nos brinda la ubicación del ítem en la escala de habilidad; para efecto de la construcción de pruebas los valores aceptables son aquellos entre -2.5 a 2.5.

Parámetro c : Representa el azar pseudoadivinanza, es el índice de pseudo-azar del ítem, indica la probabilidad de acertar de las personas que desconocen la respuesta correcta. Attorresi, et al., (2009) expresan “el valor de la asíntota a izquierda, es decir cuando θ tiende a $-\infty$. En ítems que miden habilidades refleja la probabilidad que tienen los individuos con muy bajo nivel de rasgo de responder correctamente” (Attorresi, Lozzia, Abal, Galibert, & Aguerri, 2009, pág. 183).

2.3.7.5 Los supuestos de la TRI

La aplicación de los modelos se fundamenta sobre un conjunto de suposiciones que deben suponerse de partida. Asumir que la conducta de un sujeto ante un reactivo persigue un modelo probabilístico con una forma determinada es en sí mismo un supuesto de la TRI. El conjunto de rasgo latente que intervienen al momento de responder un reactivo establece una importante distinción clasificatoria de los modelos (Attorresi, et al., 2009).

Los autores Muñiz, (1997a/2010) y Terrón (2010) coinciden que el supuesto clave del modelo de la TRI es que existe una relación funcional entre los valores de una variable no observada, latente, que miden los ítems y la probabilidad de acertar estos, denominando a dicha función Curva Característica del ítem (Muñiz, 1997a/ 2010, pág. 64) (Terrón, 2010, pág. 2).

El supuesto de esta teoría es la Unidimensionalidad, en otras palabras, el ítem o el tests solo mide un rasgo, Attorresi, et al., (2009) nos mencionas que “este supuesto es difícil que se satisfaga totalmente en la práctica, debido a que existen diversos factores que pueden afectar la respuesta del ítem” (Attorresi, Lozzia, Abal, Galibert, & Aguerri, 2009, pág. 182). Muñiz, (2010) nos aconsejas que antes de utilizar este modelo hay que verificar que los datos cumpla con los requisitos (Muñiz, Las teorías de los tests: Teoría clásica y teoría de respuesta a los ítems, 2010, pág. 64).

El tercer supuesto de esta teoría es la denominada independencia local; Terrón (2010) ,Attorresi, et al., (2009) y Muñiz, (2010) expresan que en este tipo de modelo los ítems tienen que ser independientes unos de otros, se supone que este requisito conocido como el nivel Θ de un individuo, las respuestas a cualquier subconjunto de reactivos no proporcionan ninguna información para el cálculo de probabilidad de la respuesta de un reactivo en particular, en

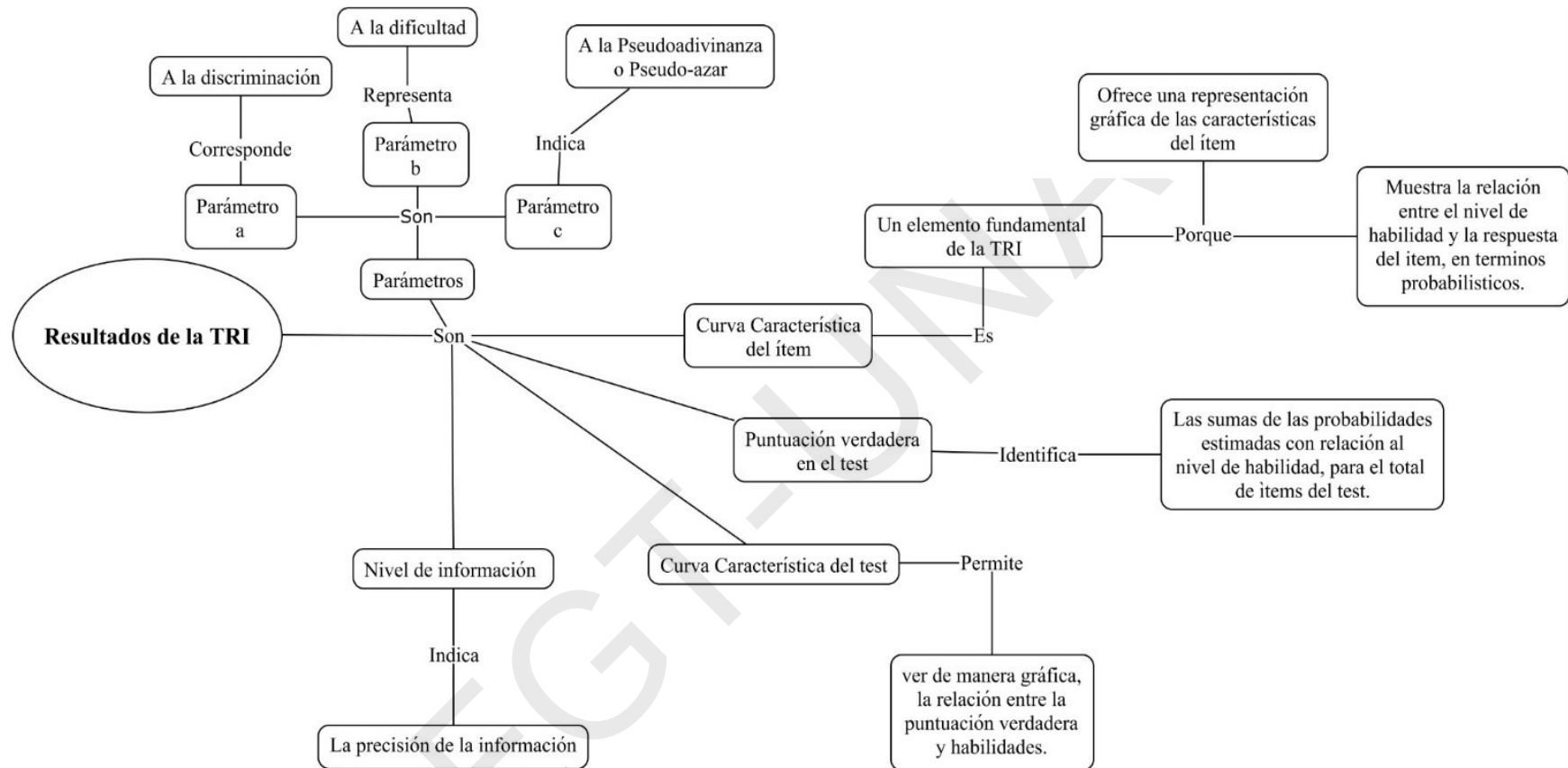
otras palabras las diferentes respuestas de los reactivos son estadísticamente independientes. La unidimensionalidad del rasgo latente y la independencia local en realidad si se cumple la una también se cumple la otra, es decir son dos caras de un mismo requerimiento.

2.3.7.6. Curva Características del ítem

La TRI asevera que, si una persona posee un nivel n de conocimiento sobre matemáticas, por ejemplo, poseerá un valor inscrito a una probabilidad de que responda correctamente a una pregunta de matemáticas, es decir que, si un sujeto contesta correctamente a un reactivo, es posible suponer que este individuo posee un determinado nivel de conocimiento (Terrón, 2010; pág. 3).

En tal sentido, Attorresi, et al., (2009) explica que en el marco de esta teoría “se postula la existencia de una relación directa entre el comportamiento de un individuo frente a un ítem y el rasgo que posee esta conducta” (Attorresi, Lozzia, Abal, Galibert, & Aguerri, 2009, pág. 181). Además, agrega que dicha relación acoge la forma de una función matemática que se vincula con la probabilidad de proporcionar una determinada respuesta a un reactivo con cada nivel de rasgo latente. Según Terrón (2010) “En la práctica la CCI se suele representar en una escala situada entre -3 y $+3$ puntos de habilidad. En el eje “ x ” se sitúan los valores de la habilidad y en el eje “ y ” las probabilidades asociadas” (Terrón, 2010, pág. 3). La curva característica del ítem depende de los tres parámetros anteriormente explicados.

La manera concreta de la CCI será determinada por el valor que adquieran los tres parámetros: a , b , c . De acuerdo a los valores que tomen los parámetros se proporcionaran diferentes formas de curvas. En manera de síntesis en la ilustración 10 se proporcionará un resumen de los resultados de esta teoría.

Figura 7: Resultados de la TRI

Fuente: Elaboración propia con base en Terrón (2010)

En la figura 7 se sintetiza la información sobre la teoría de respuesta al ítem y es importante que al momento de decidir qué teoría utilizar se tenga en cuenta

2.3.7.7. Diferencias de la TCT y la TRI

En la siguiente tabla se proporcionará las diferencias entre estas dos teorías.

Tabla 4:Diferencias entre la TCT y la TRI

Aspecto	TCT	TRI
Modelo	Lineal	No lineal
Asunciones	Débiles (fáciles de cumplir por los datos).	Fuerte (difícil de cumplir por los datos)
Invarianza de las mediciones	No	Si
Invarianza de las propiedades del test	No	Si
Escala de las puntuaciones	Entre cero y la puntuación máxima en el test	Entre $-\infty$ y $+\infty$
Énfasis	Test	Ítem
Relación Ítem-Test	Sin especificar	Curva Característica del Ítem
Descripción de los ítems	Índices de Dificultad y de Discriminación	Parámetros a, b, c
Errores de medida	Error típico de medida común para toda la muestra	Función de Información (varía según el nivel de aptitud)
Tamaño Muestral	Puede funcionar bien con muestras entre 200 y 500 sujetos aproximadamente	Se recomiendan más de 500 sujetos, aunque depende del modelo

Fuente: Obtenida de Muñiz (1997a/2010)

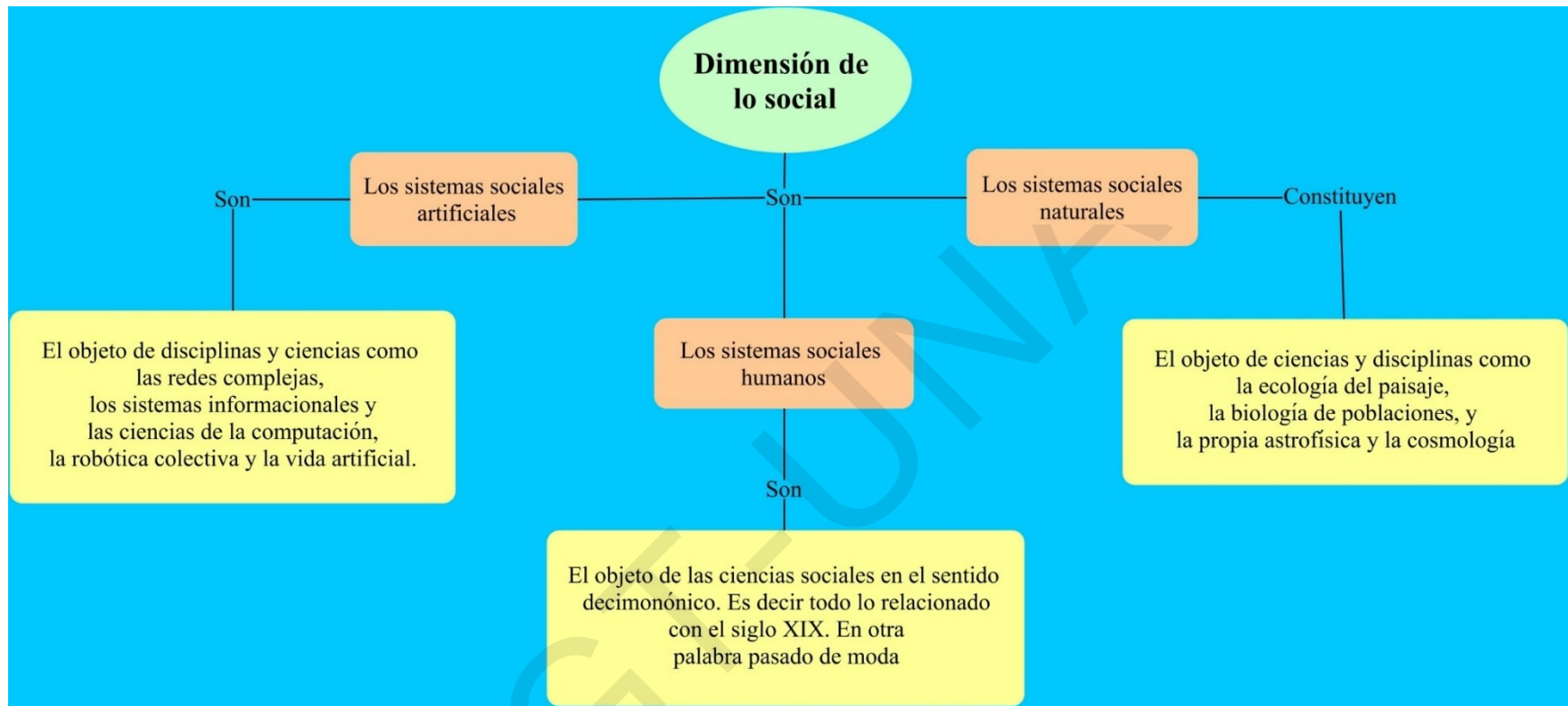
En la tabla 4 se realiza una comparación sobre ambas teorías explicadas anteriormente.

2.4 Las Ciencias Sociales en el Mundo

Según la UNESCO y Foro Consultivo (2011) las Ciencias Sociales iniciaron en la época de la ilustración europea de siglo XVII, cuando surgieron las ideas de la religión, la razón, la humanidad y la sociedad, estas ideas se unieron proporcionando de esta manera una visión del mundo que hizo énfasis en los derechos humanos, el individualismo y el constitucionalismo que hasta la fecha sigue vigente. Sin embargo, Maldonado (2009) expresan:

Las ciencias sociales, nacieron posteriores al desarrollo de las ciencias exactas, físicas y naturales. Este hecho hizo que el espíritu, las aspiraciones, el lenguaje y numerosos métodos y aproximaciones de las ciencias sociales y humanas nacieran y permanecieran, durante mucho tiempo, a la luz o a la sombra; según se mire de las ciencias físicas y naturales. (Maldonado, 2009, pág. 2)

Lo importante del surgimiento de estas ciencias es que nos conversan sobre nuestros intereses de la comunidad a la que pertenecemos, también nos platica sobre el sentido de identidad y de los objetivos que tenemos o no alcanzamos. Para Barros (2012) “Las Ciencias sociales comprende un conjunto de disciplinas que se ocupan de diferentes aspectos relacionados con los grupos humanos y la sociedad, en general, estudian las manifestaciones materiales e inmateriales de la humanidad” (Barros, 2012, pág. 60), es decir que las ciencias sociales buscan expresar lo que somos, como somos, en el contexto de la sociedad y la cultural. Además, contribuye a brindarnos conocimiento acerca de cómo podemos y debemos vivir (Maldonado, 2009, pág. 148). Según Maldonado existen tres dimensiones o modos de lo social lo cual se explicara a continuación:

Figura 8: Tres dimensiones de lo social

Fuente: Elaboración propia con base en Maldonado (2009).

Para el Consejo Internacional de Ciencias Sociales CICS/UNESCO (2015) ve a las ciencias sociales como una ciencia que se encarga de estudiar a la sociedad y su comportamiento como ser humanos, a la vez esta es llamada a recapacitar y a expresar leyes sociales que normalicen el comportamiento de la sociedad para compensar los diferentes problemas que afectan al planeta y al ser humano. Por eso es importante que tanto la comunidad como el gobierno deben reconocer el impacto que tiene el cambio climático en el planeta. Según CICS/UNESCO (2015) los seres humanos deberían de plantearse esta simple pregunta:

¿Podemos decir que más cemento, más edificios, más automóviles, más carreteras y más industrias constituyen verdaderamente el mejor modelo de que disponemos para el desarrollo? Si existe un modelo mejor, el reto de los especialistas en ciencias sociales será ayudar a definirlo y comprenderlo, y aportar conocimientos para modificar el comportamiento humano y la práctica social con miras a adoptar un modelo de desarrollo y un estilo de vida que dejen una huella de carbono mucho más ligera y, como es de esperar, un mundo mucho más verde

La Ciencias Sociales contribuyen con herramientas que permiten analizar e interpretar el universo social, donde intervienen, decide, cambia y busca medios que permitan transformarlo; en cuanto a la educación en cuanto a la educación esta ciencia se ha convertido en una parte primordial, permitiéndole a los sujetos que conozcan de donde proceden sus orígenes y a la vez brindándole la oportunidad a que crean su propio criterio sobre el universo en que habitan. Hoy la desigualdad y los cambios ambientales se ha convertido en punto de analices para los científicos ya que de una forma u otra afecta a la sociedad.

Según Consejo Internacional de Ciencias Sociales CICS /UNESCO (2015) expresan que los especialistas en Ciencia Sociales deben de estar seguro que los esfuerzos realizado por es CICS han logrado tres cosas la cual se les da a conocer a continuación:

1. el actual modelo de desarrollo es simplemente insostenible, porque los recursos se agotan y no se ha estructurado un plan para contrarrestar este problema.
2. el comportamiento humano es el principal factor para alcanzar progresos significativos, esto significa crear nuevos estilos de vida, crear conciencia de los efectos negativos de los cambios ambientales para generar propuestas de transformación y evitar la crisis mundial.

3. los especialistas en las ciencias sociales ocupan una posición única para contribuir al cambio del actual modelo de desarrollo por un proceso más sostenible, al comprender e influir en el comportamiento humano, las instituciones y los sistemas culturales en las que emergen y encuentra su expresión

Por otro lado, la UNESCO (2015) expresa que la sociedad debe poseer una ciudadanía mundial la cual se define de la siguiente manera “se refiere a un sentido de pertenencia a una comunidad más amplia y a una humanidad común. Hace hincapié en la interdependencia política, económica, social y cultural y en las interconexiones entre los niveles local, nacional y mundial” (UNESCO, 2015, pág. 14); más adelante la UNESCO 2015 agrega:

La educación para la ciudadanía mundial aspira a ser un factor de transformación, inculcando los conocimientos, las habilidades, los valores y las actitudes que los educandos necesitan para poder contribuir a un mundo más inclusivo, justo y pacífico. La educación para la ciudadanía mundial adopta “un enfoque polifacético, utilizando conceptos y metodologías que ya se aplican en otros ámbitos, entre ellos la educación para los derechos humanos, la educación para la paz, la educación para el desarrollo sostenible y la educación para el entendimiento internacional” y procura que se alcancen sus objetivos comunes. (UNESCO, 2015, pág. 15)

Siguiendo la lectura de la Unesco 2015 menciona que la educación para construir una ciudadanía mundial aspira a que los educandos puedan (UNESCO, 2015, pág. 16):

- Comprender las estructuras de gobernanza mundial, los derechos y las responsabilidades internacionales, los problemas mundiales y las relaciones entre los sistemas y procesos mundiales, nacionales y locales.
- Reconocer y apreciar la diferencia, así como las identidades múltiples, por ejemplo, en materia de cultura, lengua, religión, género, humanidad común y adquirir aptitudes para vivir en un mundo cada vez más diverso.
- Adquirir y aplicar competencias críticas para el conocimiento cívico, por ejemplo, indagación crítica, tecnología de la información, competencias básicas en medios de comunicación, pensamiento crítico, adopción de decisiones, solución de problemas, negociación, consolidación de la paz y responsabilidad personal y social.

- Reconocer y examinar creencias, valores, así como la manera en que las percepciones acerca de la justicia social y el compromiso cívico influyen en la adopción de decisiones políticas y sociales.
- Desarrollar actitudes de interés y empatía respecto al prójimo, el medio ambiente; también de respeto por la diversidad.
- Adquirir valores de equidad, justicia social y capacidades para analizar críticamente las desigualdades basadas en el género, la condición socioeconómica, la cultura, la religión y la edad.
- Interesarse en las cuestiones mundiales contemporáneas en los planos local, nacional y mundial para aportar contribuciones propias de ciudadanos informados, comprometidos, responsables y reactivos

Pero para que esto se pueda llevar a cabo los educandos deben de poseer las siguientes dimensiones de aprendizaje; estas dimensiones son propuesta por la UNESCO (2015):

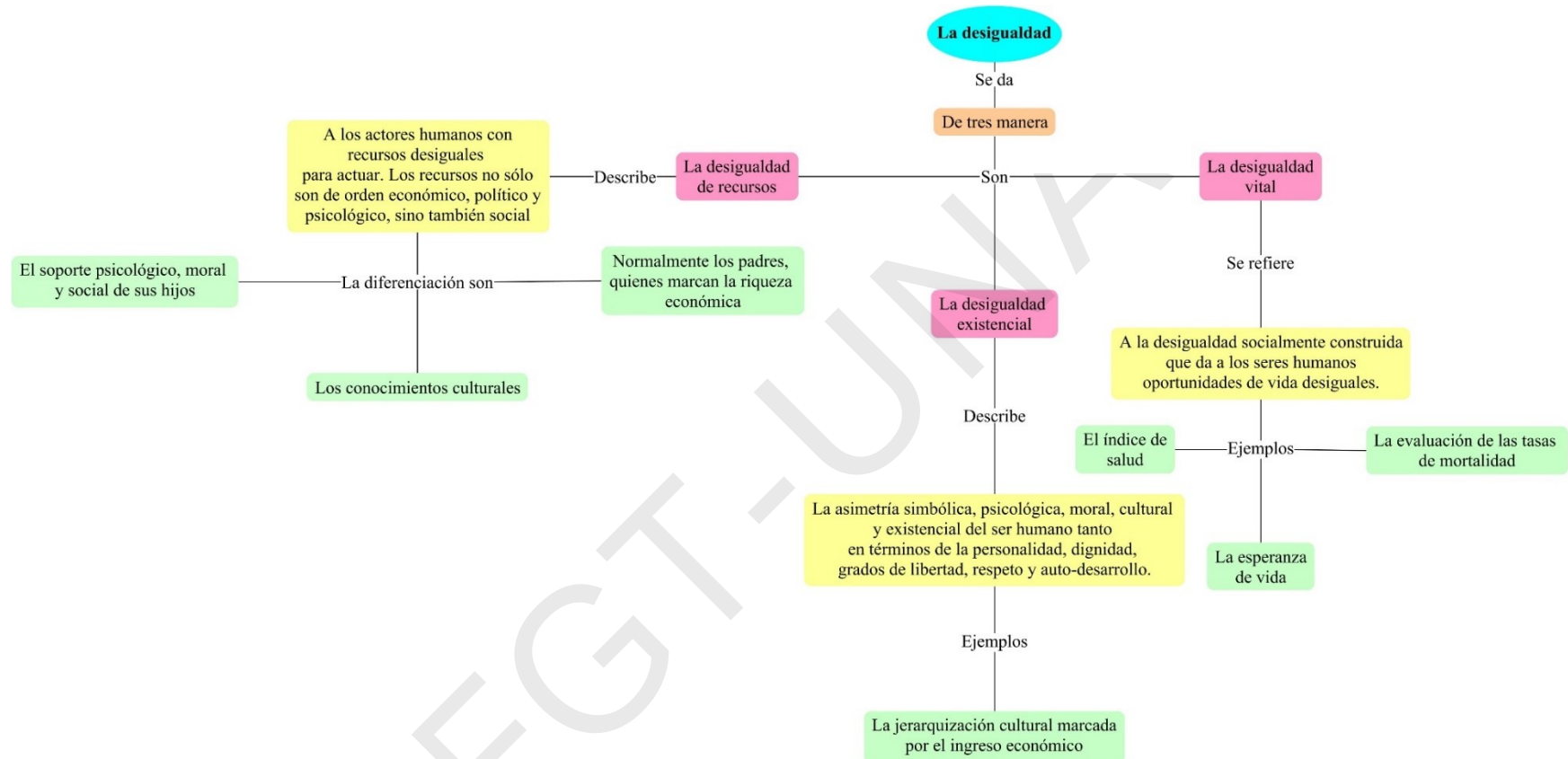
Tabla 5: Dimensiones conceptuales básicas de la educación para la ciudadanía mundial

Nº	Dimensiones	Descripción
1.	Cognitivo	Adquisición de conocimientos, comprensión y pensamiento crítico acerca de cuestiones mundiales, regionales, nacionales y locales, así como de las interrelaciones y la interdependencia de diferentes países y grupos de población.
2.	Socioemocional	Sentido de pertenencia a una humanidad común, compartiendo valores y responsabilidades, empatía, solidaridad y respeto de las diferencias y la diversidad.
3.	Conductual	Acción eficaz y responsable en el ámbito local, nacional y mundial con miras a un mundo más pacífico y sostenible.

Fuente: Adaptado de UNESCO (2015)

2.4.1 La desigualdad

La desigualdad está en todas partes, es difícil de eliminarla, y varias personas la consideran injusta. En la realidad en la que nos encontramos se ha convertido en un debate o discusión de los políticos sin embargo estos no llegan a ningún acuerdo entre ellos. Realmente todos hablamos que vivimos en una desigualdad, pero nos hemos preguntado ¿Qué tipo de desigualdad?, porque es imposible que todos los ciudadanos de un país estemos estancados en un solo tipo de desigualdad de ingreso, desigualdad en la educación. Sin embargo, Martin, Moore, & Schindler, (2016) estamos acostumbrados a relacionar la desigualdad con el acceso al sueldo, normalmente esta es definida por una composición de indicadores económicos referidos a los ingresos y la riqueza (Martin, Moore, & Schindler, 2016, pág. 30), pero me quedo con la definición que brinda Nieto (2016) en la que expresa que “la desigualdad es una violación a la dignidad humana. Es la negación de la posibilidad para que cada ser humano desarrolle sus capacidades” (Nieto, 2016, pág. 273). Como podemos observar no solo es desigualdad de los ingresos, también tiene es una situación sociocultural que para la mayoría de las personas disminuyen las capacidades de trabajar como seres humanos. En la figura 9 podremos conocer los diferentes tipos de desigualdad que puede haber.

Figura 9: Tipo de desigualdad

Fuente: Elaboración propia con base en Nieto (2016).

2.4.2. Cambios ambientales

Todos saben que en el mundo en el cual vivimos estamos sometidos a diferentes cambios, así como constantes crisis; el medio ambiente es uno de ellos. Los cambios ambientales es un problema tanto social, económico y político, así como lo es la desigualdad y la pobreza; es por ello que se precisa con urgencia la protección del planeta con el fin de asegurar la equidad, bienestar y la dignidad del ser humano. Cuando se refiere a cambios ambientales se describe la pérdida de la biodiversidad, cambios en los recursos hídricos, así como en los ciclos de nitrógeno y el fósforo y el cambio climático siendo este uno de los problemas más grave que nos enfrentamos en la actualidad. Hoy la crisis del cambio ambiental es un problema que requiere de esfuerzo compartidos (Hackmann & Moser, 2013).

Es por ello que este apartado se abordará la pérdida de la biodiversidad, el cambio climático para que podamos entender un poco estos tópicos.

2.4.2.1. La pérdida de la biodiversidad

Se le conoce como biodiversidad a la diversidad biológica y el Convenio sobre la Diversidad Biológica la define como:

La variabilidad de organismos vivos de cualquier fuente, incluidos, entre otras cosas, los ecosistemas terrestres y marinos y otros ecosistemas acuáticos y los complejos ecológicos de los que forman parte; incluye la diversidad dentro de cada especie y de los ecosistemas. (Secretaría del Convenio sobre la Diversidad Biológica., 2010, pág. 15)

Para Badii, Rodríguez, Aguilar, & Acuña (2015) nos mencionan que la biodiversidad alcanza tres niveles los cuales son: a) el genético estudia los genes entre especies e individuos de la misma especie. b) el taxonómico, estudia la diversidad de especies, género. c) el ecológico este investiga la variedad a nivel superior como lo son los ecosistemas (Badii, Rodríguez, Aguilar, & Acuña, 2015, pág. 157).

Una vez definido lo que es biodiversidad, es importante hablar de la pérdida de la misma, en el año 2002 los gobiernos de todos los países acordaron que para el año 2010 lograrían una reducción significativa de la pérdida de la biodiversidad dicha meta no se alcanzó. Las investigaciones realizadas entre los años 1970 y 2006 demuestran que las especies de vertebrados se redujo en un tercio, en cambio las especies que están en riesgo de extinción se ha evaluado corren cada vez más peligro y los que están más amenazados son los anfibios

y el estado de las especies de coral se está deteriorando más rápidamente (Secretaría del Convenio sobre la Diversidad Biológica., 2010, pág. 94)

2.4.2.2. Cambio Climático

El cambio climático es una de las preocupaciones de la actualidad y por sus implicaciones, económicas, sociales y ambientales todos los ciudadanos tienen que comprometerse a mejorar el futuro de los países. El deterioro del medio ambiente con el constante cambio climático es una bomba de tiempo que se debe de detener si queremos seguir en el planeta tierra, para que el cambio climático no siga avanzando se debe empezar educando a la población sobre el mismo y el gobierno tiene que poner de su parte creando políticas que se cumplan para el beneficio de nuestro planeta. ¿Qué es cambio climático? La literatura consultada nos expresa que el cambio climático ocurre cuando se le atribuye de manera directa o indirecta a las diferentes actividades humanas que descompone a la atmósfera y se manifiesta a través del calentamiento global; que es el aumento de las temperaturas tanto terrestres como marítimas, y esta es producida por la abundancia de los gases de efecto invernadero, pero ¿Qué son los gases de efecto invernadero?

Martón (2008) define a los gases invernadero o gases de efecto invernadero cuando existe presencia de los gases en la atmósfera y esto hace posible que dicho fenómeno ocurra también menciona que la mayoría de los gases procede de la quema de combustibles fósiles y se agrupan de la siguiente manera (están ordenados de acuerdo al orden de importancia): 1. Vapor de agua, 2. Dióxido de carbono, CO₂, (75%), 3. Metano CH₄, (18%), 4. Oxidos de nitrógeno, NO_x (95), 5. Ozono, O₃ 6. Clorofluorocarburos (artificiales) el incremento de estos gases han colaborado al aumento de la temperatura provocado por el efecto invernadero. (Martón, 2008, pág. 13)

Capítulo III. Marco Contextual

En este apartado conversaremos un poco sobre la realidad de nuestro país Honduras, conoceremos como está conformado la educación básica, la reformas que han surgido en la educación y las organizaciones nacionales que se encarga de evaluar el desempeño académico en las dos áreas que se han evaluado hasta el momento (español y matemáticas).

3.1 Situación de la educación en Honduras

De acuerdo con los autores Alas & Moncada (2011) cuando expresan que:

En la sociedad contemporánea, la educación se consigna como un derecho universal. Para muchas personas, la educación puede ser la única forma de escapar de una vida de pobreza; para las sociedades, es crucial para su desarrollo socioeconómico. Se espera que la educación apoye el desarrollo de valores tales como honestidad, equidad, justicia social, etc., y el sentido de responsabilidad compartida que es clave para el tejido social y las formas de gobierno democráticas. Pero también se espera que desarrolle competencia. (Alas & Moncada, 2011, pág. 234)

Y es por ello que Honduras ha avanzado en el ámbito de educación en los últimos años, pero hay que tener en cuenta que estos avances han sido lentos pero que han sido significativo para la sociedad en la que vivimos, es necesario mencionar que el modelo con el cual se está trabajando en los centros educativos de Honduras es un modelo adaptado de otros países que de una manera u otra ha brindado algunos resultados tal y como nos orienta Méndez (2013):

“El modelo educativo hondureño es el resultado de la sucesiva incorporación de ideas desarrolladas en otras sociedades con procesos de adaptación que no siempre ha producido los efectos esperados. En el periodo colonial la provincia de Honduras era una de las más atrasadas, la educación estaba a manos de las órdenes religiosas con predominio en el escolasticismo. Durante el periodo independentista nacieron los primeros lineamientos para la educación primaria”. (Méndez, 2013, pág. 13)

La educación hondureña se ha caracterizado por la implementación de una serie de leyes para lograr un sistema de educación de calidad. La educación en Honduras está dirigida por la Secretaría de Educación y la Universidad Nacional Autónoma de Honduras. Tal como lo expresa la República de Honduras “El sector educativo hondureño está liderado por el consejo Nacional de Educación que preside el presidente de la República y coordina la acción de las dos entidades rectoras, a saber: Secretaria de Educación y Universidad Nacional Autónoma de Honduras” (República de Honduras/Sector Educativo, 2004, pág. 5).

En los últimos veinte años, el sistema educativo hondureño ha sufrido varias reformas con el objetivo de mejorar la calidad de la educación. Al tener conocimiento sobre la demanda actual que posee la formación y el desarrollo humano y productivo, se puede identificar los problemas con los que cuenta la educación nacional. La Secretaría de Educación menciona

que: “Al reconocer la sobre demanda actual que tiene la educación y el desarrollo humano y productivo, cobran importancias algunos elementos que revelan los déficits de la realidad de la educación Nacional” (Secretaría de Educación, 2003, pág. 8). La Secretaría de Educación en el año 1999, presento la propuesta de Transformación de la Educación Nacional, en esta propuesta se identifica que existe la necesidad de formar a cada uno de los ciudadanos y ciudadanas para que responda a las demandas que la sociedad exige.

En la propuesta de Transformación de la Educación Nacional, presentada por la Secretaría de Educación (1999), expresa que existe necesidad de formar integralmente al ciudadano y ciudadana para que responda con eficacia y eficiencia a las demandas de una sociedad que exige recursos humanos aptos para incorporarse a un mundo competitivo. (Secretaría de Educación, 2003, pág. 8)

Con la implementación de las nuevas tecnologías los egresados de los centros educativos se encuentran en desventajas ya que no están formados para el manejo de las mismas, en la actualidad las mayorías de las empresas están utilizando las tecnologías para la producción del país. Hay que mencionar que en educación el currículo que se implementan en los centros educativos se encuentra desactualizado, por lo tanto, no está dando respuesta a las demandas de la sociedad.

Las y los egresado de los centros educativos no están formados para el manejo de las nuevas tecnologías relacionadas con el proceso productivos del país. Existe un currículo que no responde a las necesidades del alumnado, de la familia, de la comunidad y del país. No forma para el futuro. La falta de fundamentos para el desempeño de las instituciones educativas y la falta de desempeños del personal docente, además es importante un buen sistema de evaluación del rendimiento académico del estudiante para evaluar el proceso de enseñanza aprendizaje (Secretaría de Educación, 2003, pág. 9)

Uno de los problemas más importantes que presenta la Secretaría de Educación es la falta de educación de calidad, ya es por medio de la educación donde se evidencia los aprendizajes de cada sociedad, lastimosamente en Honduras se identifica más la deserción, repitencia y reprobación de los estudiantes. (Secretaría de Educación, 2003, pág. 9). Algunos estudios reflejan que el sector público resulta menos eficiente que el privado; la tasa de repitencia en niñas (6-15 años) es del 8% y en los niños es de 10.7% (6-15 años) en las escuelas públicas

en cuanto a la deserción en las escuelas públicas en niñas fue 47.9% y en los niños 52.1%. (UNICEF, 2011, pág. 20)

Según la Secretaría de Educación (2003) La tasa de repitencia del nivel primario fue de 9.7% en el sector público y en las escuelas privada solo es del 2.9%. En cuanto a la educación secundaria fue del 10.3% en colegios públicos público y el 80.7% en colegios privados, y respecto a la educación superior fue de 17% en el sector público el doble del 8.8% del sector privado (Secretaría de Educación, 2003, pág. 9). Asimismo, expresa que los factores que influyen en el nivel de escolaridad y en la eficiencia terminal de los estudiantes de educación básica son: la repitencia, deserción y la sobre edad. A pesar de las dificultades que, a presentados la educación hondureña, el gobierno ha estado cumpliendo con su parte. Actualmente, en Honduras el sistema educativo labora con dos tipos de calendario académicos, el de febrero a noviembre y el de septiembre a junio (este último es calendario anglosajón). (Secretaría de Educación, 2016, pág. 10).

Para que los procesos de educación se realicen de manera transparente la Secretaría de Educación en el año 2013 crea el Sistema Nacional de Información Estadística de Honduras (SINIEH) en la cual realiza recolección y procesamiento de la información educativa dicha información es brinda por cada centro educativo a nivel nacional. La recolección de la información es de manera cualitativa y cuantitativa.

La Secretaría de Educación en el año 2013 oficializa el Sistema Nacional de Información Estadística de Honduras (SINIEH), sistema que permite la recolección y procesamiento de información educativa, cuantitativa y cualitativa, generada desde los centros educativos y con procesamiento en los niveles municipal o distrital, departamental y nacional. (Secretaria de Educación, 2016, pág. 6).

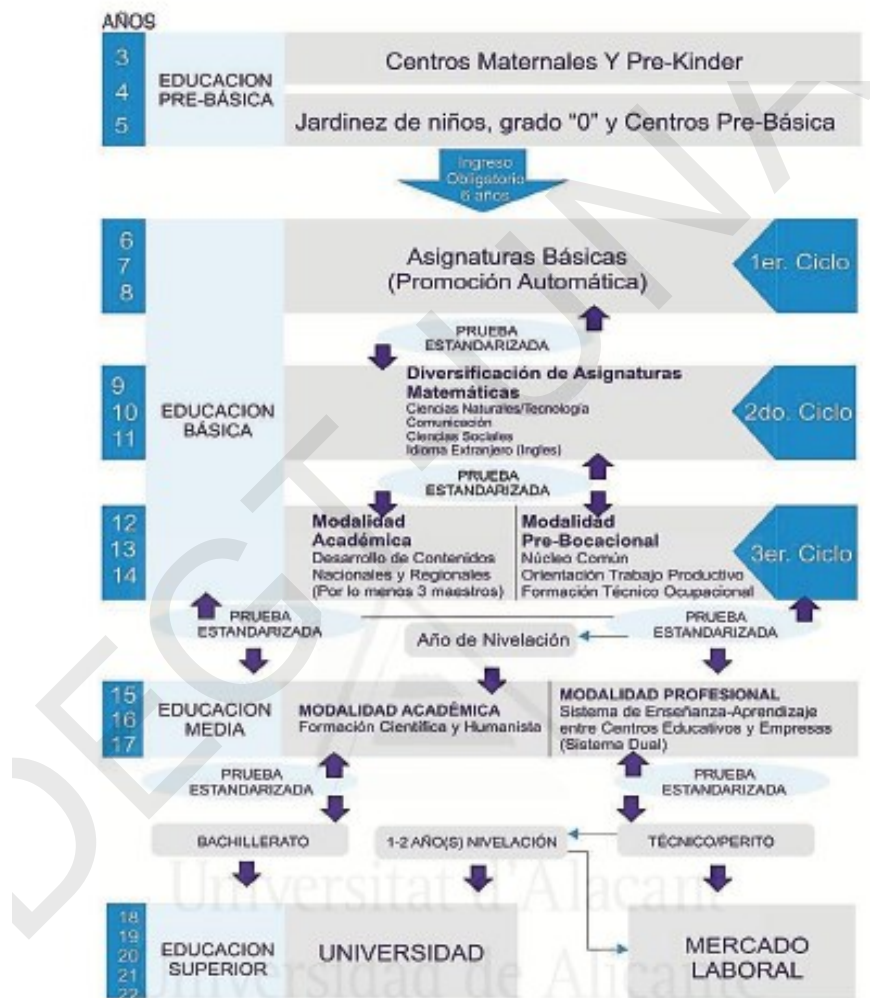
Según la Secretaría de Educación (2016) que la información proporcionada por el SINIEH tenga confiabilidad y transparencia dicho sistema puede contar con los apoyos del: Sistema Administrativo de Centro Educativos (SACE), Sistema de Planificación de la Infraestructura Educativa (SIPLIE), Sistema Integrado de Administración de los Recursos Humanos docentes (SIARHD), Sistema de Capacitación Docente (SICAVI), Sistema de Estadística Educativas (SEE), Sistema de Evaluación de Rendimiento Académico (ERA), Sistema Monitoreo de resultados para la Equidad (MORES). El Sistema Administrativo de Centros

educativos es uno de los primordiales que le brinda información al SINIEH, ya que le SACE cuenta con toda la información de los centros educativos ya sean públicos o privados.

El SACE es uno de los principales sistemas de información del SINIEH y la principal herramienta tecnológica de gestión de la información que manejan los centros educativos a nivel nacional, tanto del sector público como del privado y del cual se extraen la mayoría de las estadísticas educativas del país (Secretaría de Educación, 2016, pág. 6).

Hoy el sistema educativo de Honduras está estructurado de la siguiente manera:

Figura 10: Estructura general del sistema educativo hondureño



Fuente: Obtenida de Secretaría de Educación. (2008) .

Por otro lado, el FEREMA (Fundación para la Educación Ricardo Ernesto Maduro Andreu), publicó su informe titulado "Educación: Una deuda pendiente. Informe de progreso educativo Honduras" en el año 2017 mencionan que "la disponibilidad periódica de datos

sobre logros de aprendizajes ha ayudado a identificar, con mayor precisión, los avances y desafíos en el servicio educativo hondureños” (FAREMA, 2017, pág. 9). Además, nos proporciona una tabla con los avances y desafíos que todavía presenta la educación en Honduras, a continuación, se las presentamos.

Tabla 6: Informe de progreso educativo: Honduras, 2017

ÁREA	CALIFICACIÓN	TENDENCIA	COMENTARIOS
Cobertura	C	↔	Modestos avances al inicio, pero al final del período los datos se estancaron en algunos niveles. Brechas altas en tercer ciclo y media.
Eficiencia	D	↑	Tanto en edad oportuna como en edades múltiples, las tasas de graduación han mejorado ligeramente en el período. La deserción interanual continúa siendo elevada y en general la eficiencia interna es baja.
Calidad	D	↔	Ligera mejoría en el período, pero resultados muy bajos aún, tanto en evaluaciones nacionales como internacionales.
Equidad	D	↔	Los indicadores educativos por quintil de ingreso, zona de residencia y pertenencia a grupo étnico, continúan reflejando marcadas inequidades.

Autonomía y rendición de cuentas	C	↑	Avances importantes en gobernabilidad. Existe un marco legal que favorece la descentralización, la autonomía y rendición de cuentas, pero aún se presentan limitaciones en su aplicación
Estándares y sistemas de evaluación	B	↑	Se ha avanzado significativamente en evaluación de los aprendizajes y un poco menos en evaluación del desempeño docente. Los resultados de las evaluaciones no se utilizan en la toma de decisiones de política educativa.
Profesión docente	C	↔	Se ha iniciado la formación inicial docente en el nivel universitario. Existe actualmente un entorno favorable para introducir cambios fundamentales tanto en formación inicial y permanente, como en el ingreso a la carrera docente.
Financiamiento	D	↓	Aun cuando Honduras invierte más en educación que el promedio regional tanto como porcentaje del

PIB, como en proporción del gasto nacional, en el período analizado se presenta una reducción importante en el presupuesto asignado y baja eficiencia en el gasto educativo.

ESCALA A. Excelente
 B. Bueno
 C. Regular
 D. Deficiente
 E. Muy deficiente

Progresando
 Regular Sin tendencia definida
 Deficiente Retrocediendo



Fuente: Obtenida de FEREMA (2017)

3.2 Marco legal y las reformas del sistema educativo hondureño

El sistema educativo hondureño se respalda en la constitución de la república y la ley fundamental de educación en la cual el poder Judicial de Honduras (1982) en la Constitución de la República de Honduras en el capítulo VIII de la educación y Cultura en el artículo 151 señala que “La educación es función esencial del Estado para la conservación, el fomento y difusión de la cultura, la cual deberá de proyectar sus beneficios a la sociedad sin discriminación de ninguna naturaleza”. En este mismo artículo también se explica que la educación es laica y que se basará en los principios esenciales de la democracia, y que se inculcará se fomentará en los estudiantes un profundo sentimientos de hondureñistas, así como se involucraran en el proceso de desarrollo económico y social del país.

En los artículos 153 ,157, 160, 171 del capítulo VIII de la constitución de la república expresa que el Estado tiene la obligación de desarrollar la educación básica del pueblo, y desarrollar todos los niveles des sistema educativo formal, excepto el nivel superior, ya que será supervisada, organizada y dirigida exclusivamente por el Poder Ejecutivo por medio de la Secretaría de Educación. La Universidad Nacional Autónoma de Honduras es una institución autónoma del Estado, goza de la exclusividad de organizar, dirigir y desarrollar

la educación superior y profesional. La educación impartida será oficialmente será gratuita y la básica será, además, obligatoria y totalmente costeadas por el estado.

La República de Honduras (2012), en la ley fundamental de educación en el capítulo I del objeto y finalidad, derecho y obligaciones garantías y jerarquías normativas en el artículo 1, garantiza el derecho humano a la educación y establece los principios, garantías, fines y lineamientos generales de la educación nacional. Reconoce al educando como titular del derecho y actor principal; establece que el fin primordial de la educación es el desarrollo al máximo de sus potencialidades y su personalidad, en este sentido es deber del Estado tutelar la educación, ya que esta tiene la finalidad de garantizar el acceso equitativo de todas las personas a una educación integral de calidad, sin discriminación.

Según la Ley Fundamental de Educación, del capítulo 1 del objeto y finalidad, derecho y obligaciones garantías y jerarquías normativas en el artículo 8 de los niveles obligatorios explica que el estado está obligado a brindar la educación pública al menos desde un (1) año de educación prebásica hasta el nivel medio. Más adelante en el capítulo II de los principios, valores y fines de la educación Nacional en el Título II de la estructura del sistema Nacional de educación en el artículo 15 expresa; el sistema nacional de educación está conformado por los componentes siguientes: Educación Formal, Educación no Formal y Educación Informal.

En el capítulo II de los principios, valores y fines de la educación nacional, en el título II de estructura del sistema nacional de educación, capítulo II de la educación formal sección primera de los niveles de la educación formal, sección tercera (De la educación Básica) en el artículo 22, de la ley fundamental de educación expresa que la educación básica, se orienta hacia la formación integral de los educandos en sus dimensiones físicas, afectiva, cognitiva, social, cultural, moral y espiritual, al igual que la educación Prebásica, este nivel de educación es gratuita y obligatoria. Consta de nueve (9) años con edades de referencia desde los seis (6) hasta los catorce (14) años y se divide en tres (3) ciclo secuenciales y continuos de tres años cada uno.

En cuanto a reformas se refiere, estas vienen implementándose desde los años noventa permitiendo de esta manera la modernización de gobierno. Desde ese momento han asumido una línea continua, aunque se le han proporcionado otro nombre, por miedo a las protestas por parte de los docentes. Estas reformas se planifican tomando como guía las concepciones de calidad, eficiencia y equidad, estas transformaciones han brindado un impacto en el

sistema educativo hondureño ya que ha mejorado la formación académica de los docentes y ha realizado modificaciones es el sistema educativo (Mejía, 2008).

Las reformas educativas surgen cuando se observa una necesidad de organizar un cambio en las estructuras del sistema educativo; un ejemplo de esta es cuando se modifica el diseño curricular de alguna carrera o los cambios que la sociedad experimenta (Mejía, 2008, pág. 58). De acuerdo con López (2008) reforma “se refiere a las intervenciones de políticas propuestas y conducidas a nivel macro del sistema, por los estados o gobiernos y los organismos internacionales, independientemente de las denominaciones que estas adopten en los distintos países” (López, 2008, pág. 44). Asimismo, concuerdo con Mejía (2008) cuando expresa que:

Las reformas educativas son hechos que aparecen como prácticas sociales privilegiadas de proyectos políticos, y forman parte de las políticas públicas del Estado. Las reformas son uno de sus principales medios para incorporar, modificar, cambiar, mover y experimentar estrategias que afectan directamente el proceso de enseñanza aprendizaje y en los contenidos curriculares. (Mejía, 2008, pág. 60)

Explicado lo que es reforma procederemos a conocer las reformas que ha sufrido el sistema educativo en Honduras. Que según Mejía (2008) las implementaciones de la misma se dieron en tres fases la cual se conocerán a continuación:

Tabla 7:Fase de implementación de las reformas educativa

Fase	Características	Comentario
I	La obligatoriedad de la educación básica de nueve grados.	Inicio formalmente en 1994, con el gobierno del Doctor Carlos Roberto Reina. En estas características sobre salen tres aspectos: la obligatoriedad de la educación básica, la cobertura de la enseñanza y la descentralización o desconcentración del sistema educativo.
	El inicio de programas educativos alternativos para jóvenes y adultos (como EDUCATODOS, Programa de Alfabetización y Educación de Adultos (PRALEBAH), financiados por la cooperación externa).	La descentralización trajo consigo la reducción del aparato público y del gasto en la educación en lo que respecta a la creación de nuevos centros educativos y equipamiento de estos.
	La inserción de los valores como ejes transversales en el currículo nacional.	

	Desconcentración de la Secretaría de Educación.	
	Reestructuración de la Secretaría de Educación.	
	Constitución de las Direcciones Departamentales de Educación.	
	Creación de la Unidad de Medición de la Calidad Educativa (UMCE)	
II	En esta etapa se elimina la Unidad Coordinadora de las direcciones departamentales de educación. Se dividió a los departamentos en distritos educativos a cargo de directores distritales. Transformación del Currículo Nacional Básico. Legalización y obligatoriedad de la Educación Básica hasta el noveno grado. Formación Inicial de los docentes. Formación permanente de los docentes en servicio Educación bilingüe para las culturas étnicas. Se define la multiculturalidad en la educación. Autogestión escolar mediante la participación comunitaria, a través del Programa Hondureño de Educación Comunitaria (PROHECO).	Esta segunda etapa empieza en 1996 con la emisión del Decreto de Reformas Educativas No 162-96, de fecha del 22 de octubre de 1996. En esta fase sobresalen la calidad y equidad en la educación
III	Conversión de las normales Reformas en educación media Creación de dos bachilleratos: Bachillerato Científico Humanista y Bachillerato Técnico Profesional Diseño del Currículo Nacional para Educación Básica Elaboración de la Ley Marco que reestructura el Sistema de Formación Profesional Elaboración y reforma de Leyes educativas	Esta reforma se inició en el 2008, esta misma se centran en la autonomía de la escuela y en la denominada descentralización pedagógica, que promueven la transferencia de decisiones pedagógicas y curriculares desde gobierno central a los centros educativos. El punto central de estas reformas se enfoca en el proceso de enseñanza-aprendizaje.

Fuente: Elaboración propia con base en Mejía (2008).

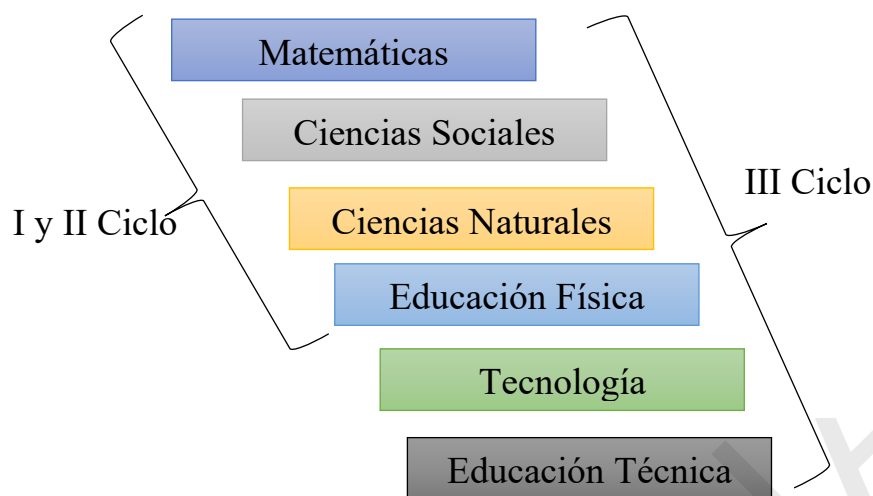
3.3 La educación básica en Honduras

La Secretaría de Educación (2004) y Paz (2014) mencionan que la educación Básica inicia a los seis años de edad, en esta etapa los niños adquieren contenidos básicos de la ciencia, la cultura y tecnología. En este nivel lo que se busca es crear en el estudiante su creatividad, el sentido crítico y autocrítico de su entorno, el pensamiento lógico, racional y reflexivo. Esta modalidad comprende tres ciclos de tres años de estudio, cada uno posee sus objetivos, perfiles y currículo. A continuación, se explicará cada uno del ciclo que comprende la educación básica de acuerdo a estos dos autores:

Primer ciclo: Esta está conformada del primer al tercer grado, con niños/as de los 6 a 8 años de edad, es aquí donde se inicia a desarrollar las destrezas instrumentales, como ser el lenguaje y el cálculo matemático. Durante este ciclo se construirán los fundamentos para la comprensión de su entorno natural y social.

Segundo ciclo: Este ciclo está atendido por el cuarto al sexto grado, su objetivo es profundizar el desarrollo de las destrezas instrumentales, así como incrementar la formación mediante procesos mentales y actitudinales, de esta manera los niños adquirirán una visión más clara de su entorno y social.

Tercer ciclo: Del séptimo a noveno grado, aquí se realiza una retroalimentación y ampliación de los contenidos curriculares de los ciclos antes mencionados, además adquieren un carácter científico y tecnológico. Este período es de carácter vocacional, de orientación profesional para la educación media. A continuación, se le presentará las áreas curriculares con las que cuentan la educación básica. En la siguiente figura 11 daremos a conocer las áreas curriculares que se dan en este nivel

Figura 11: Área curriculares del nivel de educación básica

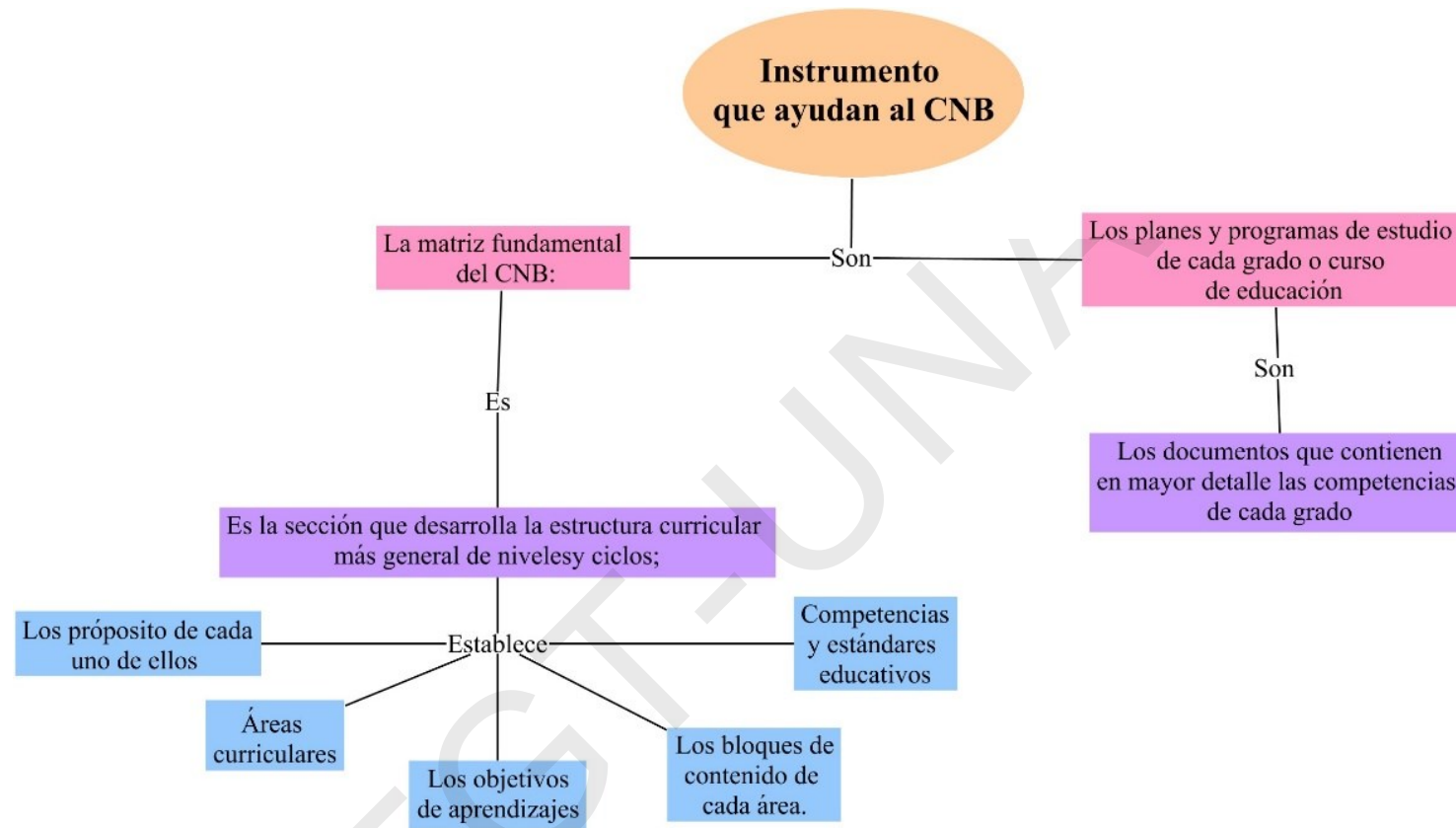
Fuente: Obtenida de Secretaría de Educación.(2004) Currículo Nacional Básico.

De acuerdo con la SE-MIDEH (2017) en la educación básica existe una de las características en específico que debe de ser analizada, y es el alto porcentaje de los estudiantes de sobre edad, siendo este más evidente en el segundo y tercer ciclo. En el 2016 los estudiantes con sobre edad aumento al 20% en quinto grado y al final del tercer ciclo lo valores fueron del 29.2%, lo que significa que al comenzar el tercer ciclo los estudiantes con sobre edad aumenta un 2 % en 8^{vo} grado y un 3% en 9^{vo} (SE/MIDEH, 2016, págs. 56-57).

Los principales factores que enfrenta la educación básica en estos días es la deserción, la sobre edad tal y como lo expresa MIDEH, la repitencia y que causa podríamos atribuirles a estos factores; uno de estos es el ausentismo debido a la pobreza que enfrenta Honduras y por ende los niños tiene que trabajar para contribuir económicamente con su familia, otro el nivel de calificaciones que han obtenido a lo largo del año escolar y esto hace que vuelvan a repetir. En la actualidad todavía existen las escuelas de multigrado en docente brinda clases a 3 o más grados esto debido a que no se invertido suficiente en la creación de nuevas aulas de clases en las zonas rurales. Para que la educación básica sea de calidad docentes se guían de Currículo nacional básico (CNB).

3.3.1. Currículo Nacional Básico

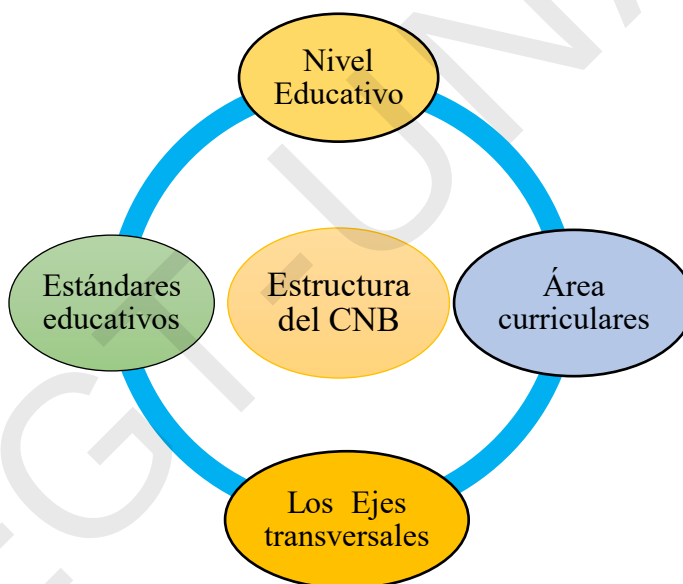
Algunas preguntas que se realizan las personas que no trabajan en el ámbito de la educación básica son ¿Qué es el CNB? ¿Para qué sirve? el CNB según la Secretaría de Educación (2003) es una herramienta o instrumento legal en el cual describe las capacidades, los conceptos, las competencias, las habilidades, las destreza y actitudes que los estudiantes tienen que lograr en cada nivel del sistema educativo. Este se encuentra organizado por áreas curriculares que se vinculan con los objetivos de cada nivel o grado. Además, este currículo toma en cuenta los cuatro saberes que establece la UNESCO: Saber ser, saber hacer, saber conocer y el saber convivir. Es por ello que le sirve como guía al docente en el aula de clases. El CNB cuenta con el apoyo de dos instrumentos más los cuales se presenta en la siguiente figura 12

Figura 12:Instrumentos en los que se apoya el CNB

Fuente: Elaboración propia en base en la Secretaría de Educación (2003).

En la Figura 12 podemos observar que CNB se auxilia de otros instrumentos para poder asegurar la calidad de la educación en el sistema educativo hondureño y de esta manera contribuir con jóvenes que estén a servicio de la sociedad. Además, la Secretaría de Educación (2003) expresa “el CNB tiene sustentos fundamentales que sistematizan los aportes provenientes de distintas disciplinas, a estos sustentos se les llama fundamentos curriculares, los cuales se apoyan en principio, fines, objetivos y fuentes que delimitan los alcances del Currículo Nacional Básico” (Secretaría de Educación, 2003, pág. 13) . Esto nos evidencia una vez más que el CNB es un instrumento confiable que los docentes tiene que saber manejar en el aula de clases. La estructura del CNB está compuesto por cuatro elementos.

Figura 13: Estructura del CNB



Fuente: Elaboración propia con base en la Secretaría de Educación (2003).

En la figura 13 se observa que dependiendo del nivel educativo en el que se encuentre el estudiante el docente brindará las áreas que establece el CNB de acuerdo a las edades del individuo, después continua los ejes transversales y estos se encarga de ayudar al sujeto a que se formen y aprendan a descubrir los conocimientos que ellos necesitan y que la sociedad demanda para que pueda ponerlo en practico a beneficio de la misma sociedad, el CNB establece los siguientes ejes transversales: Identidad, trabajo y democracia participativa. Los

estándares educativos, están clasificado de acuerdo a cada área curricular y consiste en guiar al docente que conocimientos deben dominar los estudiantes al finalizar el año escolar.

3.4 Unidad Externa de la Calidad de la Educación (UMCE)

Honduras en los últimos años se ha preocupado por la calidad de la educación, empezó a realizar evaluación estandarizada con el proyecto PEEP en los años 90, sin embargo, en el año de 1995 se crea un convenio entre la Secretaría de educación y la UPNFM, donde establecen la creación de la Unidad Externa de la Calidad de la Educación en el cual el objetivo medir los avances, logros y dominio de los aprendizajes en los centros educativos. En la actualidad UMCE funciona como una Unidad Técnica del INIEES (Instituto de Investigación y Evaluación Educativa) de la UPNFM efectuando la evaluación externa del desempeño académico de los estudiantes en los diferentes niveles de la educación y además realizan factores asociados que puedan afectar el desempeño de los estudiantes.

3.5. Proyecto Mejorando el Impacto al Desempeño Estudiantil en Honduras

El Proyecto Mejorando el Impacto al Desempeño Estudiantil (MIDEH) en Honduras empezó a funcionar en el año 2004 y finalizó en el 2018, fue ejecutado por American Institutes for Research y estuvo financiado por el Pueblo y Gobierno de los Estados Unidos por medio de USAID (Agencia de los Estados Unidos para el Desarrollo Internacional) apoyando el proceso educativo específicamente en educación básica. Según Hernández (2018)

MIDEH ha apoyado técnicamente a la Secretaría de Educación con el desarrollo de Estándares de contenidos (Español, Matemáticas, Ciencias Sociales y Ciencias Naturales 1ero -11mo grado) y de Desempeño para los estudiantes, Programaciones Educativas Mensuales, Pruebas Formativas Mensuales y sus respectivos instructivos, estándares de contenidos para uso de padres de familia y pruebas de fin de grado todos los cuales constituyen materiales educativos alineados al Diseño Curricular Nacional Básico (DCNB) (Hernández, 2018, pág. 2).

Durante estos 14 años que MIDEH funcionó, logró cuatro grandes resultados: 1. Mejorar la capacidad técnica nacional para alcanzar las metas EFA bajo el liderazgo fortalecido de la

Secretaría de Educación .2. Fortalecer las capacidades nacionales en evaluación estandarizada. 3. Fortalecer la participación de la Sociedad Civil en apoyo a la mejora de la calidad educativa. 4. Desarrollar investigación para proveer información que apoye la calidad educativa (Hernández, 2018, pág. 1)

Capítulo IV: Método

En la actualidad se ha venido evidenciando una preocupación por la calidad de educación en todos los niveles (I ciclo, II ciclo y III ciclo) en el país, como una respuesta a esta preocupación de la calidad educativa se ha dado importancia a la evaluación externa en educación. Tomando en consideración que lo más importante en los procesos educativos es el aprendizaje. Es de aquí donde se desprende la variable dependiente donde se centra la investigación. En este apartado se dará a conocer lo que han aprendido y cuanto han aprendido los estudiantes y de esta manera medir las diferencias entre un grupo y otro respecto al aprendizaje. La tarea primordial es conocer el funcionamiento del sistema educativo.

Detallaré los parámetros y procedimiento que se aplicarán para conocer los resultados de rendimiento académico de los estudiantes de noveno grado en la educación hondureña mediante la aplicación de una prueba psicometría en noviembre 2018. En el siguiente trabajo se identificará los resultados que proporcionen los estudiantes de dicho grado en el área curricular de ciencias sociales. El test se basa en los contenidos del DCNB, seleccionando los estándares educativos de contenido usados en el nivel; los estándares evaluados se dividen en bloques y componentes del contenido que presenta el DCNB. Ciencias Sociales cuenta con los siguientes bloques: La persona y su ser social, las sociedades y los espacios geográficos, las sociedades y el tiempo social, las sociedades organizadas y las actividades humanas. En la siguiente tabla se proporciona los componentes de cada una de los bloques

Tabla 8: Bloques y componentes que se evaluarán

Bloque	Componente	Grado
La persona y su ser social	Relación interpersonal	9no
Las sociedades y los espacios geográficos	Espacio geográfico	
Las sociedades y el tiempo social	Tiempo histórico	
Las sociedades organizadas y las actividades humanas		
	Acción social del ser humano	

Fuente: Obtenido de República de Honduras, Secretaría de Educación, (2009)

De los bloques y componentes presentados en la tabla 8 se procederá a crear los reactivos para ser aplicados a los estudiantes.

4.1. Enfoque, alcance y diseño

Este trabajo tendrá un *enfoque* de carácter **cuantitativo** ya que se basa en el uso de los métodos estadísticos con el objetivo de conocer aspectos de interés sobre la población que se está analizando; tal y como lo plantean López (2011), Sampieri, Collado, & Lucio (2014) la metodología cuantitativa maneja la recolección y el análisis de datos para responder problemas de investigación y además probar hipótesis que son determinadas anticipadamente, ya que se apoya en la medición numérica y el uso de estadística para crear pautas del comportamiento de una población que ha sido analizada. A la vez contará con un *estudio descriptivo* ya que deseamos brindar información de manera independiente o relacionada sobre las variables que deseamos describir, según Cortés & León (2004) expresan que los estudios descriptivos indagan sobre las características, especifican las propiedades, de un grupo, individuos o comunidades u otro fenómeno que sea objeto de análisis.

Este trabajo contará una *temporalidad transversal* ya que se recopilarán los datos en un solo momento. La *finalidad* de la misma es de tipo **aplicada** ya que en esta investigación se pretende encontrar conocimientos adquiridos y por medios de los resultados saber si es útil para la población.

4.2. Variables, operacionalización e indicadores

Para esta investigación se analizará la variable de desempeño académico y para esto se partirá de los bloques su componente antes mencionada y en este apartado se conocerá los indicadores que se evaluará. Para que el estudiante pueda contestar las respuestas deben de poseer las habilidades cognitivas necesarias para poder obtener un buen resultado.

Tabla 9: Operacionalización de la variable

Variable	Definición conceptual	Definición operacional	Dimensiones	Componentes	Indicadores
Desempeño académico	El desempeño académico es el resultado de un proceso educativo que puede ser medido en términos cuantitativos y cualitativos, que sirve para retroalimentar a los individuos y a las instituciones sobre el logro de los objetivos que se establecieron previamente (Guzmán M. P., 2012)	Por medio del diseño de una prueba estandarizada.	La persona y su ser social	Convivencia armónica	Conocen los antecedentes que determinan la razón de ser de los Derechos Humanos, identificando el carácter utilitario de sus regulaciones.
			La Sociedad y los espacios geográficos	Espacio Geográfico	Reconocen la importancia actual del estudio de la Geografía Económica y Política, analizando la condición internacional de Honduras en el orden económico, político, diplomático y social.
			Las sociedades y el tiempo social	Tiempo histórico	Conocen y manejan la división política de cada continente, relacionando las características físico-geográfico y la distribución de la población a nivel mundial.
					Establecen relaciones entre las diferentes formas de periodizar la Historia Mundial, identificando los aportes de las civilizaciones en el desarrollo de la humanidad.
					Comprenden las grandes revoluciones sociales, identificando causa y consecuencias en el mundo actual.
					Comprenden e identifican los diferentes conflictos mundiales, por

			medio de conocimiento de sus origen, antecedentes y lecciones aprendidas.
La sociedad y las actividades humanas	Acción social del ser humano	del	Comprenden las diferentes acciones y actividades económicas que realiza la población mundial, por medio de la identificación y caracterización de los sectores de la economía.
			Comprenden los indicadores de Desarrollo Humano (IDH), analizando comparativamente la situación en Honduras y estableciendo relaciones con las políticas y sociales actuales.

Fuente: Elaboración propia

4.3 Población y muestra

La evaluación estandarizada se aplicó en noviembre en dos momentos; el primero se realizó la prueba piloto en el centro educativo. Instituto Gubernamental Saúl Zelaya con 158 estudiantes de ambas jornadas en el departamento de Francisco Morazán en la ciudad de Tegucigalpa. La población antes descrita fueron estudiantes de noveno grado de educación básica, la edad con la que contaban los estudiantes es de los 14 años hasta los 16.

Tabla 10: Distribución de la muestra para la prueba piloto

Distribución de la muestra para la prueba piloto		
Centro Educativo	Sección	N° de estudiantes
Instituto Gubernamental Saúl Zelaya	1	26 estudiantes
	2	23 estudiantes
	3	27 estudiantes
	4	20 estudiantes
	5	18 estudiantes
	6	20 estudiantes
	9	24 estudiantes
Total		158

Fuente: Elaboración propia.

Los datos obtenidos presentado en la tabla 10 se consiguieron de la aplicación de la prueba piloto

El segundo momento de la aplicación se dió en los siguientes centros educativos: 1) Instituto Gubernamental Héctor Pineda Ugarte (IHPU) 2) Instituto Gubernamental Hibueras 3) Instituto Gubernamental Abelardo Fortín con los estudiantes de noveno grado en las áreas de ciencias sociales. De las siguientes edades desde los 13 años hasta los 19, algunos están en el rango de edad establecida por el DCNB sin embargo algunos sobrepasa las edades.

Tabla 11:Distribución de la muestra de la prueba operativa

Distribución de la muestra de la prueba operativa		
Centro Educativo	Sección	Cantidad de estudiante
Instituto Gubernamental Héctor Pineda Ugarte	1	39 estudiantes
	2	17 estudiantes
	3	25 estudiantes
	4	27 estudiantes
Instituto Gubernamental Hibueras	1	28 estudiantes
	2	27 estudiantes
	3	28 estudiantes
	4	27 estudiantes
	5	24 estudiantes
	6	24 estudiantes
	7	22 estudiantes
Instituto Gubernamental Abelardo Fortín	1	20 estudiantes
	2	24 estudiantes
	3	23 estudiantes
	4	19 estudiantes
	5	24 estudiantes
Total		398

Fuente: Elaboración propia.

4.4. Tipo de muestreo.

Para la prueba piloto la muestra es de tipo no probabilístico ya que el centro educativo fue seleccionado por conveniencia y por la cantidad de estudiantes que contaba dicha institución. Tal y como lo plantea menciona Otzen & Manterola (2017) “Este tipo de muestreo permite seleccionar aquellos casos accesibles que acepten ser incluidos. Esto, fundamentado en la conveniente accesibilidad y proximidad de los sujetos para el investigador” (Otzen & Manterol, 2017, pág. 230).

Es decir que los resultados obtenidos con este tipo de muestreo no implican un procedimiento aleatorio ya que no generaliza a una población. El motivo por el cual se realizó de esta manera fue por lo siguiente: 1. Fue el único centro educativo que nos permitió el acceso a la institución; 2. Porque fue la primera aplicación de los reactivos, es decir que en la prueba piloto se estaba analizando el funcionamiento de los ítems y no el desempeño académico de los estudiantes. La aplicación de la prueba piloto nos permitió conocer el comportamiento que cada reactivo desempeña en el campo de aplicación, con la finalidad de

probar algunas situaciones que pueden influir de manera negativa en la aplicación del instrumento, asimismo se obtuvieron las propiedades psicométricas para validar los ítems estadísticamente y seleccionar los mejores reactivos que conformaran la prueba operativa.

Para que la prueba piloto cumpliera su función se tomaron los siguientes criterios:

1. Los estudiantes que participaron en la prueba, estar debidamente matriculado en el grado correspondiente en el centro educativo seleccionado.
2. Las personas con necesidad especiales no participaron en la aplicación de la prueba.

Para la prueba operativa se tomaron los mismos criterios.

4.5. Tamaño de la muestra

Como se explicó anteriormente la muestra de la prueba piloto fue de 158 estudiantes del centro educativo Saúl Zelaya, el cual estuvo conformada con el 42% del género femenino y el 58% del género masculino. Esto sucedió porque el instituto es un técnico. Vale mencionar que la participación de los estudiantes fue de manera voluntaria, y al momento de aceptar ser parte del estudio se comprometieron a responder la prueba estandarizada de fin de año en el área curricular de Ciencias sociales para noveno grado.

Al momento de aplicar la prueba de desempeño académico se presentaron 110 estudiantes del Instituto Gubernamental Abelardo Fortín; 108 del Instituto Gubernamental Héctor Pineda Ugarte y 180 del Instituto Gubernamental Mixto Hibueras, siendo un total de 398 estudiantes. En la prueba operativa estuvo conformada 53% del género femenino y el 47% del género masculino.

4.6. Diseño muestra de la prueba operativa

Para realizar la administración de la prueba operativa se procedió a solicitar información a la Unidad del Sistema Nacional de Información Educativa de Honduras (USINEIH). Con la información brindada por USINEIH se pudo acceder a la matrícula de los estudiantes para el año 2018 de noveno grado tanto de los centros Gubernamentales como no Gubernamentales, pero en este caso solo se tomó la matrícula de los centros educativos gubernamentales que contaban con el grado a evaluar. Con esta información se pudo seleccionar a los centros con base a un procedimiento científico.

La selección de la muestra fue de manera probabilística ya que este tipo de procedimientos permite que cada centro educativo de la población cuente con la misma probabilidad de ser incluido en el estudio. Tal y como lo expresa Sampieri, Collado , & Lucio (2014) “En las

muestras probabilísticas todos los elementos de la población tienen la misma posibilidad de ser escogidos y es por medio de una selección aleatoria en donde se toman las unidades de análisis para realizar el estudio” (Sampieri, Collado, & Lucio, 2014, pág. 176).

Por otro lado, la técnica que se utilizó para realizar las unidades de análisis del estudio fue un muestreo de aleatorio simple Otzen & Manterol (2017) lo explica de la siguiente manera:

“El aleatorio simple garantiza que todos los individuos que compone la población tienen la misma oportunidad de ser incluidos en la muestra. Significa que la probabilidad de selección de un sujeto es independiente de la probabilidad que tiene el resto de los sujetos que integran a la población”. (Otzen & Manterola, 2017, pág. 228)

Mediante un procedimiento multiétapico se inició a seleccionar los elementos de la muestra: El primer paso fue seleccionar los centros educativos que formarían parte de la investigación, solo se tomó en cuenta el Distrito Central de Tegucigalpa.

Con el programa de Excel y de acuerdo al número de distrito de Tegucigalpa; utilizando el aleatorio simple se eligieron los distritos que formarían parte del estudio, en este caso los distritos seleccionados fueron el 5 y el 9.

La segunda etapa consistió en seleccionar los centros educativos que pertenecen a los dos distritos previamente seleccionados al azar, para seleccionar los centros educativos se realizó el mismo procedimiento que utilizamos para elegir los distritos. Los cuales fueron seleccionados los siguientes: del distrito 5 el Instituto Héctor pineda Ugarte y del distrito 9 Instituto Abelardo R. Fortín, y el instituto Mixto Hibueras.

4.7. Tamaño de la muestra para la prueba operativa

Una vez seleccionado los distritos y los centros educativos se procedió a sacar la muestra del estudio la cual se basó a la siguiente fórmula.

$$n = \frac{Z^2 N \sigma^2}{Z^2 \sigma^2 + E^2 N}$$

Donde:

N = Tamaño de la población

e = Error muestral máximo aceptable (E m)

Z = Nivel de confianza

P = Probabilidad que ocurra el evento

q= Probabilidad de que no ocurra el evento

n = Tamaño de la muestra

Tamaño de la muestra	55190
Error muestral máximo aceptable	5%
Nivel de confianza	95% (1.96)
Probabilidad que ocurra el evento	50%
Probabilidad de que no ocurra el evento	50%
Tamaño de la muestra	382

$$n = \frac{((55190 * (1.96)^2) * 50 * 50)}{((5)^2 * (55190-1)) + ((1.96)^2 * 50 * 50)}$$

$$n = \frac{530,044,760}{1389,254}$$

$$n = 382$$

$$n = 382$$

Con la fórmula aplicada se seleccionó la muestra de los estudiantes que participarían en la investigación la cual es de 382. Sin embargo, se sobre paso la cantidad de estudiantes debido a la matricula con que contaba los institutos.

4.8 Instrumento de medición

Para evaluar los contenidos de los aprendizajes de los estudiantes establecidos en el DCNB se utilizará los supuestos de ambas teorías; teoría de la respuesta ítem teoría y clásica de los test. Las evaluaciones estandarizadas miden el mismo contenido por medio de ítems iguales o equivalentes en parámetro como: el nivel de dificultad, nivel de discriminación y la pseudoadivinanza. El proceso de la elaboración de la prueba se inició con la descripción así como las especificaciones establecidas por el DCNB en el área de ciencias sociales y con el respectivo grado que analizaremos. Dichas especificaciones contienen la cantidad de ítems que debe poseer cada bloque de contenido, componente y el estándar, de tal forma que el test ofrezca información confiable (SE-MIDEH, 2017)

A partir de esa información, maestros hondureños que estén laborando, especializados en el área de Ciencias Sociales, anticipadamente capacitados para la construcción de reactivos, redactaron los ítems para el área y el grado solicitado. Con dichos reactivos se procederá al análisis de jueces expertos para después desarrollar el pilotaje de la prueba para empezar a crear el banco de ítems en esta área. El instrumento a utilizar será de opción múltiple ya que en esta técnica las calificaciones de los ítems son menos afectadas por la adivinación y se puede obtener información de diagnóstica útil a partir de los análisis de los distractores o respuesta incorrecta seleccionada por el estudiante (Aiken, 2003; pág 31).

. 4.8.1 Instrumento piloto e instrumento operativo

Instrumento piloto: La prueba estandarizada estuvo estructurada de dos manera: Forma A y Forma B y cada una de ella estaba conformada con un total de 46 reactivos don 36 ítems era unicos y 10 ítems eran anclas. El instrumento es de selección única, el cual consistía en brindarle un enunciado al estudiantes con cuatro opciones de respuesta siendo verdadera una opcion y tres distractores; la duración de la prueba fue de una hora.

Características de los reactivos: fueron creados según la taxonomía de Bloom, asimismo fueron redactados en los diferentes niveles cognitivos; conocimiento, comprensión y aplicación.

Aplicación de la prueba piloto: Para llevar a cabo la administración de la prueba piloto se le solicitó colaboración a personas externas al centro educativo, dichas personas fueron instruidas previamente para poder supervisar la aplicación de la misma. Una vez en el aula de clases se distribuyeron las sillas en filas equivalentes para repartir la misma cantidad de pruebas forma A y B.

Una vez organizado el espacio en el cual se desarrollaría la prueba piloto se les brindaron las instrucciones: Contestar en lápiz grafito, a los estudiantes que no andaban lápiz; el aplicador le prestaría uno, borrador, sacapunta. Dadas las instrucciones generales se procedió a brindarles el cuadernillo de manera intercalada al estudiante 1 se le dio la forma A al estudiante 2 la forma B y así sucesivamente donde se encontrarían los reactivos. Después se dio inicio al llenado de la caratula y la manera en que se contestaría la prueba.

Calificación de instrumento: Para cuantificar los datos de la prueba aplicada se le asignó una puntuación a cada respuesta del estudiante de 0 y 1; donde 0 es la respuesta incorrecta y 1 la respuesta correcta.

Validez del instrumento: Para validar el instrumento se utilizó la validez de contenido ya que los reactivos fueron elaborados bajo un modelo teórico ya establecido.

Instrumento para la prueba operativa: Este instrumento contaba con 50 ítems de tipo selección única y consistía en presentarle al estudiante un enunciado y cuatro opciones una verdadera y tres distractores; a diferencia de la prueba piloto esta prueba si media el desempeño académico de los estudiantes.

Características de los reactivos: Una vez analizados los ítems de la prueba piloto se seleccionaron los reactivos que discriminaran bien y con un índice de dificultad aceptable, y que estadísticamente no fueran negativo.

Aplicación de la prueba operativa: Para llevar a cabo la administración de la prueba operativa se le solicito colaboración a personas externas al centro educativo, dichas personas fueron instruida previamente para poder supervisar la aplicación de la misma. Una vez en el aula de clases se distribuyó las sillas en filas equivalentes para repartir la misma cantidad de pruebas. Para esta prueba se le solicito permiso a los directores por medio de una nota donde se le explicaba el propósito de la investigación por el cual se eligió el centro educativo. Por razones de ética se le solicito a estudiante que firmara un asentamiento informado, esto para

verificar que la participación del mismo es voluntaria y que al momento de aceptar participar en el estudio contestaría la prueba de fin de grado.

La duración de la prueba fue de una hora y se aplicó de manera individual en formato de papel y de esta manera asegurar la confiabilidad de los datos. A la hora de la administración de la prueba se organizó el espacio en el cual se desarrollaría la prueba operativa y se les brindaron las instrucciones: Contestar en lápiz grafito a los estudiantes que no andaban lápiz el aplicador le prestaría uno, borrador, sacapunto. una vez dada las instrucciones generales se procedió a brindarles el cuadernillo donde se encontraría los reactivos. Después se dio inicio al llenado de la caratula se socializo el asentimiento informado y luego se procedió a llenar los datos que se le pedían, seguidamente se les dio las instrucciones de forma guiada sobre cómo llenar los datos de la caratula de la prueba.

Calificación para instrumento de la prueba operativa: al igual que la prueba piloto el puntaje fue de 0 y 1, cero para las repuestas incorrectas y uno para la repuesta correcta. Tal y como lo plantea Abad, Garrido, Olea , & Ponsoda (2006), “los ítems de cuestionarios de rendimiento se cuantificará con 1 el acierto y con 0 el error, de tal manera que la puntuación directa de un sujeto en un cuestionario será igual al número de ítems que ese sujeto acierta” (Abad, Garrido, Olea , & Ponsoda, 2006, pág. 12)

Para realizar los estadísticos se utilizó el programa SPSS y se realizó el análisis de cada pregunta que conformaba la prueba estandarizada, a la vez se utilizó Excel para sacar la discriminación la dificultad y otros estadísticos.

4.9 Plan de análisis

Plan de análisis prueba piloto: Después de la aplicación de la prueba piloto se tabularon los datos en Excel para después ser importados al programa SPSS con las repuestas proporcionadas por los estudiantes. Donde se le asignó un puntaje a cada respuesta (0 y 1), en SPSS. Para el análisis de los datos se tomaron en cuenta los supuestos de la TRI y la TCT. Para generar el índice de discriminación, el índice de dificultad y el índice de funcionamiento diferencial de los ítems se utilizó la TCT. Una vez revisadas las características psicométricas de cada reactivo, se seleccionaron los ítems que mejor índice de discriminación e índice de dificultad poseía para que conformara la prueba operativa o la prueba final. Vale resaltar que en la prueba piloto también se generó el alfa de Cronbach

Plan de análisis prueba operativa: Una vez administrada la prueba operativa se procedió a ingresar las respuestas a la plantilla de Excel para luego ser importado para el programa SPSS una vez realizado este paso se generó el estadístico de tablas cruzadas para verificar que no había casos perdidos. Después genero el estadístico del alfa de Cronbach para confirmar que la prueba es confiable y valida, así mismo se creó el índice de dificultad, discriminación.

Para analizar los datos obtenidos en esta prueba es de mucha importancia tal como lo menciona SE-MIDEH, (2017) “Dada la gran importancia de los resultados de aprendizaje de los alumnos de cualquier sistema educativo, se han desarrollado varios procedimientos para analizar y presentar los resultados de las evaluaciones estandarizadas” (SE-MIDEH, 2017, pág. 33). Asimismo, los organismos encargados de realizar las pruebas de fin de grado presentan los resultados de desempeño académico en las dos áreas evaluadas hasta el momento (español y matemática), por lo tanto, tomando como bases las dos métricas (niveles de desempeño y escala de puntuación estandarizada) que utiliza la Secretaría de Educación y la que utilizaba MIDEH, se ha decidido utilizar el mismo esquema para analizar los datos generado en la prueba de ciencias sociales para noveno grado.” SE/MIDEH (2016) agrega que “La construcción de Niveles de Desempeño es un modelo para hacer más comprensible y relevante la información aportada por una evaluación” (SE/MIDEH, 2016, pág. 25) .

Para comprender los resultados de esta prueba se utilizó la métrica de niveles de desempeño antes explicada. A la vez se tomaron las etiquetas ya establecidas por la Secretaría de Educación, para cada etiqueta se crearon descripciones para cada una de los niveles, tanto a nivel general como a nivel específico. Por lo tanto, los niveles de desempeño utilizados por la Secretaría de Educación/ MIDEH y que se utilizó en esta prueba son:

Avanzado	En este nivel los estudiantes manifiestan un desempeño excepcional en los contenidos evaluados
Satisfactorio	En este nivel los estudiantes han logrado cumplir con el objetivo educativo y tiene las habilidades y destrezas para continuar en el siguiente grado
Debe Mejorar	En este nivel los estudiantes no poseen las habilidades ni destrezas mínimas en los contenidos evaluados por lo tanto no pueden avanzar al siguiente grado.
Insatisfactorio	En este nivel los estudiantes no poseen las habilidades ni destrezas mínimas en los contenidos evaluados por lo tanto no pueden avanzar al siguiente grado.

De acuerdo con la Secretaría de Educación y MIDEH (2016) cuando plantea que el propósito de la evaluación de los aprendizajes es “Clasificar a los estudiantes en función de su desempeño en una prueba” (SE/MIDEH, 2016, pág. 25), y para que esto tenga validez en la prueba se clasificó a los estudiantes en diferentes categorías de acuerdo al nivel de desempeño y esto se logró gracias a los puntos de corte ya establecidos por la Secretaría de Educación que se han utilizado en la dos áreas evaluadas en los años anteriores.

Para establecer los porcentajes de estudiantes que conformarían cada nivel de desempeño se utilizó la escala estandarizada que consiste en adaptar los puntajes directos de cada individuo. Para determinar la escala estandarizada se apoyó de la siguiente fórmula:

$$Ti = a * zi + b$$

Donde;

a = Desviación estándar

zi = Puntaje Z

b = Media

Se utilizó la media teórica nacional de 300 y la desviación teórica de 50. Tal y como se presenta a continuación en la tabla 12.

Tabla 12: Escala estandarizada

PUNTAJE TOTAL 45						
		Frecuencia	Frecuencia Acumulada	Percentiles	Puntaje Zi Media: 25.0829 Desviación: 6.93	Escala de puntuación Estandarizada = (50* Zi) +300
Válido	9.00	2	2	1	-2.3	184
	10.00	1	3	0	-2.2	191
	11.00	3	6	1	-2.0	198
	12.00	6	12	2	-1.9	206
	13.00	5	17	1	-1.7	213
	14.00	11	28	3	-1.6	220
	15.00	10	38	3	-1.5	227
	16.00	13	51	3	-1.3	234
	17.00	11	62	3	-1.2	242
	18.00	11	73	3	-1.0	249
	19.00	19	92	5	-.9	256
	20.00	19	111	5	-.7	263
	21.00	24	135	6	-.6	271
	22.00	17	152	4	-.4	278
	23.00	13	165	3	-.3	285
	24.00	17	182	4	-.2	292
	25.00	28	210	7	.0	299
	26.00	16	226	4	.1	307
	27.00	16	242	4	.3	314
	28.00	16	258	4	.4	321
	29.00	22	280	6	.6	328
	30.00	24	304	6	.7	335
	31.00	16	320	4	.9	343
	32.00	13	333	3	1.0	350
	33.00	15	348	4	1.1	357
	34.00	15	363	4	1.3	364
	35.00	8	371	2	1.4	372
	36.00	9	380	2	1.6	379
	37.00	9	389	2	1.7	386
	38.00	4	393	1	1.9	393
	39.00	2	395	1	2.0	400
	41.00	2	397	1	2.3	415
	42.00	1	398	0	2.4	422
	Total	398		100.0		

Fuente: Elaboración propia

Los niveles de desempeño están vinculados con la escala de puntuación, la ventaja de esta escala es; que admite realizar comparaciones entre un grupo de individuo de la misma aplicación y entre un grupo de participante a través del tiempo. Para expresar el desempeño académico de los estudiantes en una escala numérica se usó la escala de 100 a 500 puntos. Tal y como se observa en la tabla 13:

Tabla 13 :Niveles de desempeño vinculado a la escala estandarizada

Niveles de desempeño	Escala estandarizada
Avanzado	394-500
Satisfactorio	300- 393
Debe Mejorar	200-299
Insatisfactorio	100 – 199

Nota: Adaptado de MIDEH (2017)

Los puntajes que se encuentra a la par de cada nivel de desempeño son porque sus puntajes verdaderos se encuentran en eso el rango.

4.10: Procedimiento

De manera general se explicará la fases o etapa que se desarrolló para diseñar la prueba estandarizada de CCSS para noveno grado y poder medir el desempeño académico. Se desarrollaron las siguientes actividades:

1. **Seleccionar los estándares que se utilizarán para la elaboración de los reactivos:**
Del Currículo Nacional Básico se eligieron los estándares educativos correspondientes al grado que se está evaluando.
2. **Crear una tabla de especificaciones:** En esta tabla se detallaron los bloques, los estándares y la cantidad de reactivos que se tenía que elaborar. Ver anexo 1
3. **Seleccionar expertos en el área de Ciencias Sociales para la creación de ítems:**
Se convocó a docentes que estuvieran laborando en un centro educativo, que fueran docentes de Ciencias Sociales para capacitarlos en la redacción de ítems. Es importante realizar este paso ya que los docentes son los expertos en el área que se desea evaluar.

4. **Realizar un taller sobre evaluación educativa y redacción de reactivo de selección única:** Este paso es importante para el diseño de los ítems ya que es aquí donde se les brinda las directrices a los docentes sobre la formulación de los enunciados.
5. **Seleccionar los expertos o jueces para la validación de reactivos:** Una vez creado los reactivos se procedió a seleccionar docentes que no hayan participado en la redacción de ítems para obtener una criterio valido y confiable.
6. **Elaboración de la prueba piloto:** Se creó la prueba para ser aplicada a cada uno de los centros educativos y al grado seleccionado.
7. **Aplicación de la prueba piloto:** En este apartado consiste en recolectar la información que los estudiantes nos proporcionen en cada reactivo que se le presente a cada estudiante.
8. **Tabulación de la prueba:** En este inciso se digitalizó cada respuesta de los estudiantes para que sean analizados.
9. **Analizar cada uno de los reactivos con el programa SPSS:** Se precedió a sacar los estadísticos (Frecuencias, Alfa de Cronbach, etc.) para ser la interpretación de los mismo.
10. **Presentar los hallazgos sobre el comportamiento psicométrico de los reactivos después de la aplicación de la prueba piloto:** Se presentan los resultados de cada pregunta basándose en la TCT.
11. **Seleccionar los ítems con las propiedades psicométricas aceptables:** Esta etapa es crucial porque es a partir de este análisis que se seleccionan los reactivos que serán incluidos en la prueba operativa tomando como base la teoría que se está empleando.
12. **Elaboración de la prueba operativa:** Con los ítems estadísticamente aceptables se procedes a ensamblar los reactivos. Se inicia con los reactivos que estadísticamente son fáciles al reactivo más difícil; después se procede a generar los estadísticos pertinentes para conocer el desempeño académico de los estudiantes.
13. **Presentar los resultados de la prueba operativa:** Presentar los resultados obtenidos durante la investigación y con ellos realizar las recomendaciones y conclusiones pertinente.

4.10.1 Procedimientos que se utilizó para identificar los estándares, y bloques que se evaluaron en la prueba

Para poder llevar un orden e identificar los bloques y estándares se diseñó un marco operativo donde se encuentran los contenidos, bloques, componentes y los estándares a la vez cuenta con los niveles de conocimientos y el número de reactivo que se debe de redactar para conformar la prueba al marco operativo también se le conoce como tabla de especificaciones.

Según lo explicado por Fast & Vukmirovic (2016), que al momento de estar elaborando la tabla de especificaciones y estableciendo la cantidad de reactivos a redactar estos se multiplicaban por 3 o 4, dependiendo del nivel de confianza que se tenga de la calidad de los reactivos a pilotear a la vez los recursos disponibles para el diseño y la revisión de las preguntas, ya que sería peligro diseñar solo los ítems que se van a necesitar en la prueba final, porque es posibles que los poco reactivos que se construyan no cumplan estadísticamente los criterios psicométricos para ser aceptados. Es por ello que en la tabla de especificación se elaboraron más reactivos para asegurar la confiabilidad y la validez de toda la prueba tanto piloto como operativa.

4.10.2. Diseño de los reactivos para la prueba de CCSS

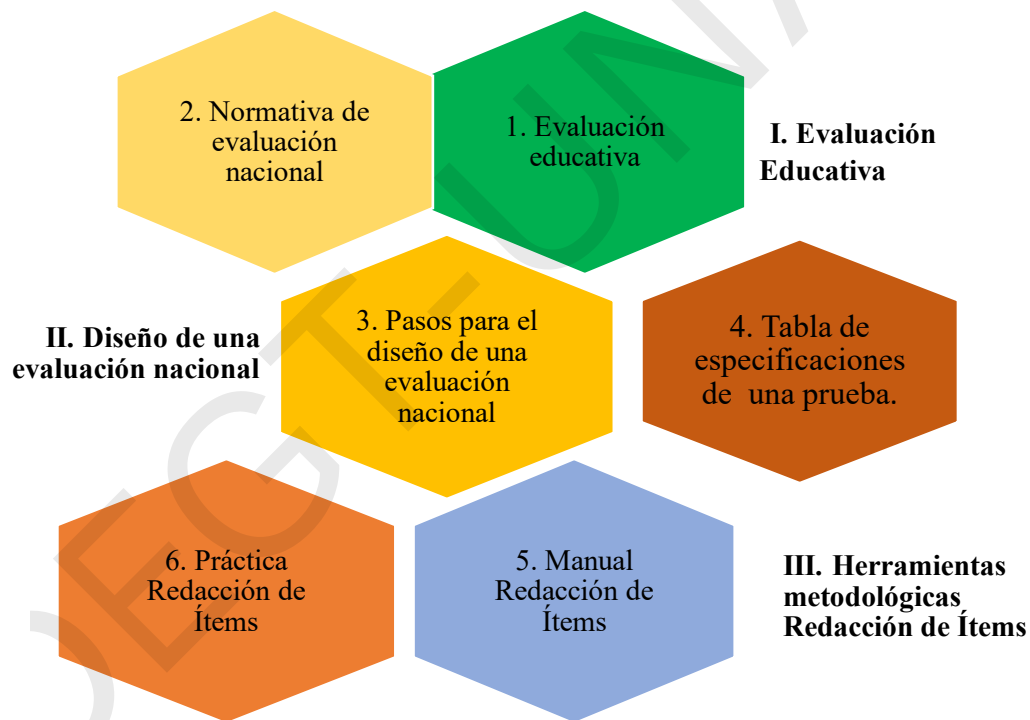
Para el diseño de los reactivos se preparó la ejecución de un taller con los docentes de las cuatro áreas de educación básica (Español, Matemáticas, Ciencias Naturales y Ciencias Sociales), el cual tenía como finalidad; fortalecer los procesos de actualización y formación docente en competencias acerca de la evaluación de logros de los aprendizajes y también tenía el propósito de proporcionar a los docentes elementos de carácter teórico-metodológico que les faciliten la elaboración de ítems para conformación de instrumentos de evaluación educativa válidos y confiables.

El taller titulado “Taller de Evaluación Educativa y Redacción de Ítems”, tuvo lugar en el Centro Regional Centro Sur- Oriente, Tegucigalpa M.D.C. (antiguas instalaciones de la Normal Pedro Nufio), con una duración de tres días, las fechas establecidas fueron del 16 al 18 de mayo del 2018, tuvo la participación de 80 docentes de las áreas antes mencionadas: y los cuales tenía como objetivos específicos los siguientes:

- ✓ Capacitar a docentes de diferentes centros educativos en los procesos de evaluación educativa en el aula conforme a los nuevos lineamientos del Acuerdo No. 1796-SE-2017.
- ✓ Validar las tablas de especificaciones a través de la revisión de los docentes especialistas de las áreas del conocimiento de español, matemáticas, ciencias naturales y ciencias sociales.
- ✓ Contribuir en la formación de un banco de ítems de opción múltiple selección única para la construcción de pruebas estandarizadas del nivel básico y medio.

El taller se desarrolló en dos maneras: en la primera se abordó de manera general y en la segunda parte de forma específica. Tal y como se puede observar en la siguiente figura 14:

Figura 14: Temática utilizada en el taller de evaluación educativa



Fuente: Elaborado por el equipo de taller de evaluación y redacción de ítems (2018)

El taller se orientó a la evaluación estandarizada y a la redacción de ítems.

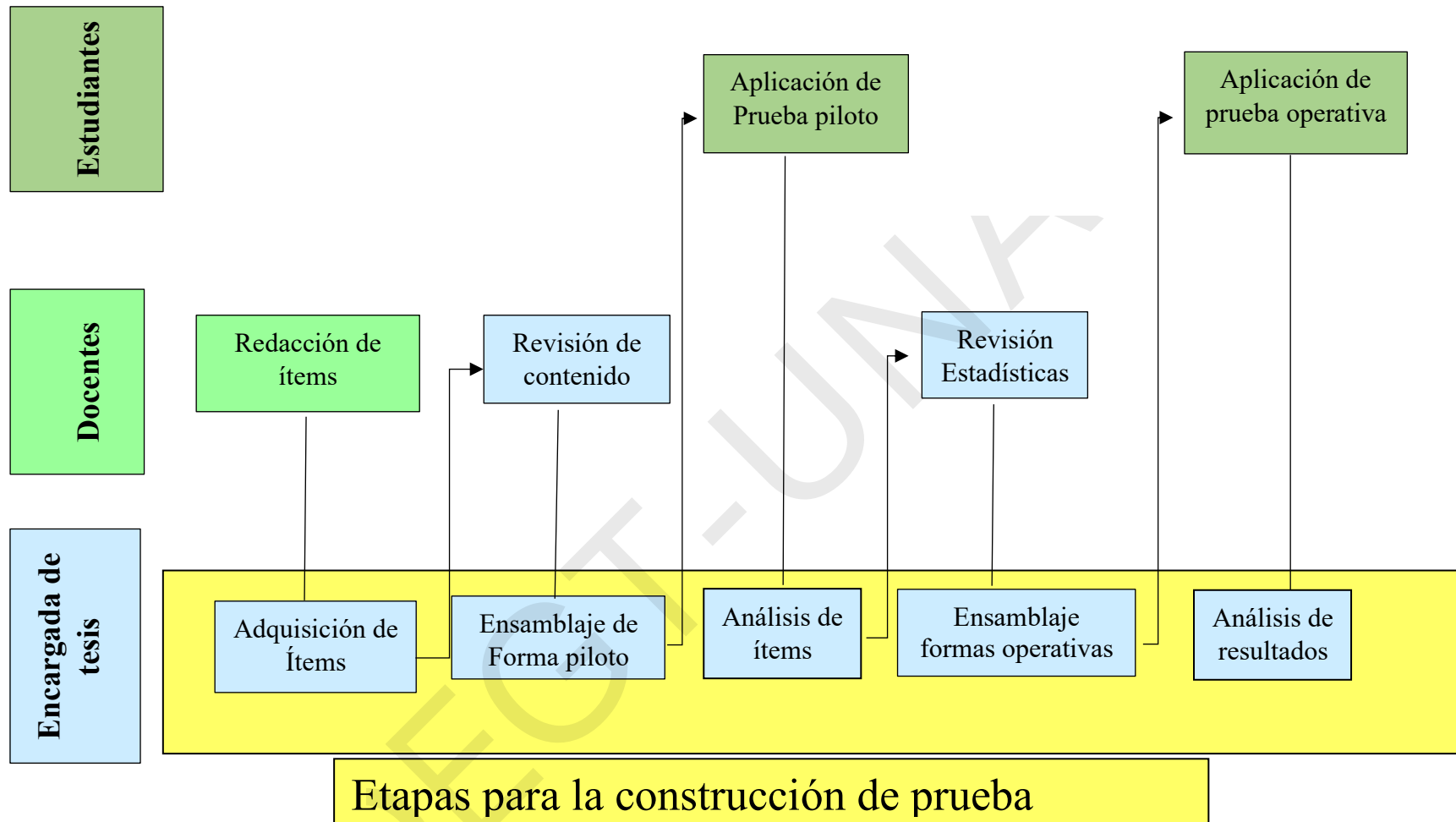
En la segunda parte se abordó la redacción de los reactivos; consistió en que los participantes especialistas en las áreas curriculares crearan ítems guiados por la tabla de especificaciones. Para que todo esto se ejecutara, se elaboró un guion metodológico para orientar todas las actividades realizadas en la misma, ver anexo 2.

Para poder elaborar los reactivos del área de Ciencias Sociales de noveno grado se, capacitaron a 6 docentes de dicha área ya que uno de los requisitos fue; que el docente estuviera laborando en las áreas solicita. En el taller se construyeron 200 ítems sin embargo no todos cumplían con los requisitos explicados en el taller, por consiguiente, se procedió a solicitar ayuda a nuevos docentes y capacitarlos en la redacción de ítems lo cual se logró en tiempo y forma. Asimismo, la prueba piloto contenía 72 reactivos, 26 era funcionales en cada forma y 20 era ítem anclas.

Para poder realizar todo lo explicado anteriormente se sintetiza en la siguiente figura



Figura 15: Etapas para la construcción de una prueba



Fuente: Información proporcionada por el Dr. Mike Fast en el módulo de Psicometría y evaluación I (2016), y por el Dr. Mario Alas (2017)

Capítulo V: Resultados de la investigación.

Para responder la siguiente pregunta ¿Qué se está evaluando?, es pertinente señalar que este trabajo tiene como objetivo principal *desarrollar una prueba de fin de grado que evalúe el desempeño académico de los estudiantes de noveno grado en el área curricular de Ciencias Sociales, de Tegucigalpa 2018*, por tanto, los contenidos que se está evaluando corresponden al espacio curricular de Ciencias Sociales, dichos contenidos se encuentran contemplados en el Currículo Nacional Básica.

Para evaluar qué y cuánto aprendizaje han adquirido los estudiantes de los centros educativos, es necesario la aplicación de pruebas estandarizadas en el área curricular de Ciencias Sociales, ya que cuentan con reactivos equivalentes para que toda la población estudiantil que se encuentre en noveno grado puedan responderla. Por ende, esta investigación se centró en detallar el rigor de los procedimientos y parámetros utilizados en el proceso de construcción y aplicación de la misma.

Dando seguimiento al tercer objetivo específico de la investigación el cual apunta a *Determinar las propiedades psicométricas de la prueba que evalúe el desempeño académico de los estudiantes de noveno grado en el área de Ciencias Sociales*. Es por ello que este capítulo se evidencia los resultados obtenidos mediante el análisis estadístico ejecutados en cada una de los ítems diseñados para la recolección de la información.

5.1. Descriptivos de la muestra.

Para la validación del instrumento se aplicó una prueba piloto los estudiantes pertenecían al Instituto Gubernamental Técnico Saúl Zelaya, cuya muestra se seleccionó de manera intencionada con 158 estudiantes de ambas jornadas, de ambos sexos, con edades entre 14 y 17 años (120 jornada matutina y 38 jornada vespertina). En cuanto al género, el 58.20% son masculino y el 41.40% del género femenino. La prueba constaba de dos formas; en la forma N.º 1, el 48% de los estudiantes es del género masculino y el otro 38% es de género femenino, en cambio en la forma N.º 2 el género masculino fue contestada por el 44% y el otro 28% es género femenino, esto es debido a que el centro educativo es Técnico y por su naturaleza el género masculino es mayor.

5.2. Resultados de ítems aplicados en la prueba Piloto.

5.1.1 Digitalización de los resultados, obtenidos en aplicación de la prueba piloto.

Una vez realizada la aplicación de la prueba piloto, se digitalizaron los datos en Microsoft Excel para después ser exportados al programa de SPSS. En el primero (Microsoft Excel) se diseñó el índice de discriminación, el índice de dificultad de cada ítem. Para seleccionar los reactivos más pertinentes, se basó en los valores de 0.23 teniendo en cuenta lo que expresa la teoría que se deben considerar arriba de 0.20. Con este mismo criterio se seleccionaron los reactivos de ambas formas. Como lo afirma Guevara (2017) los valores de Küder Richardson para las variables de discriminación son las siguientes:

Tabla 14: Valores de discriminación para seleccionar los ítems

Valor	Criterio
0.40 o más	Buen índice de discriminación
0.30 a 0.39	Razonable índice de discriminación
0.20 a 0.29	Regular índice de discriminación
0.10 o menos	Deficiente índice de discriminación

Fuente: Elaboración propia con base en (Guevara R. S., 2017)

En cambio, para el índice de dificultad, Guevara (2017) considerando la formula de Küder Richardson describe que los valores deben oscilar entre:

Tabla 15: Valores de dificultad para seleccionar los ítems

Valor	Criterio
0.81 o más	Muy fácil
0.61 a 0.80	Relativamente fácil
0.51 a 0.60	Dificultad adecuada
0.31 a 0.50	Relativamente difícil
0.11 a 0.30	Difícil
0.11 o menos	Muy difícil

Fuente: Elaboración propia con base en (Guevara R. S., 2017)

Una vez definido los valores del índice de discriminación y del índice de dificultad para la selección los reactivos; en la siguiente tabla se presenta los valores y el análisis de mismo.

A continuación, se analizan los resultados obtenidos de la aplicación de la prueba piloto dicho análisis se sustenta en la Teoría Clásica de los Test

Tabla 16: Análisis de los reactivos Forma “A”

Código de los ítems en el test mapa	Forma	Índice de Discriminación	Índice de Dificultad	Análisis	Criterio de selección		
					Descartar	Revisar	Conservar
CCSS18N001F1O	A	0.32	0.56	Tomando como base los valores antes mencionados se puede observar que este reactivo cuenta con un índice de discriminación de 0.32 por tanto, es razonable y discrimina de los estudiantes que si saben el contenido de los que no. Con respecto al índice de dificultad podemos observar que se encuentra en el nivel adecuado ya que el 0.56 de los participantes lograron contestar correctamente.			X
CCSS18N002F1O	A	0.41	0.85	En este ítem el índice de discriminación es de 0.41 por lo tanto, es un buen índice ya que si logra diferenciar a los estudiantes que si poseen los conocimientos de los estudiantes que no. Con respecto a índice de dificultad es un reactivo muy fácil con el 0.85, es decir que casi todos los participantes pueden contestar correctamente este ítem,			X
CCSS18N003F1O	A	0.41	0.50	Con respecto a esta pregunta el índice de discriminación es bueno porque cumple los criterios			

				establecidos, al igual que el anterior discrimina muy bien a los estudiantes que sí sabe los contenidos y de los que no, sin embargo, en el índice de dificultad se encuentra en relativamente difícil ya que el 0.50 de los estudiantes la pueden contestar correctamente.	X
CCSS18N0 04F1A	A	0.27	0.45	En este reactivo el índice discriminatorio es regular porque se encuentra en los límites contemplados en la fórmula Küder Richardson y posee un índice de dificultad relativamente difícil para las repuestas de los estudiantes.	X
CCSS18N0 05F1O	A	0.32	0.74	De acuerdo a los datos de este reactivo se puede analizar que con el 0.32 del índice discriminatorio se distingue que estudiante conocen los contenidos y quienes no. Sin embargo, en el índice de dificultad es un ítem relativamente fácil de contestar correctamente para la mayoría de los estudiantes.	X
CCSS18N0 06F1O	A	0.23	0.69	En relación a este ítem el índice discriminatorio es regular porque se encuentra en los límites contemplados en la fórmula considerada, discrimina muy poco a los estudiantes que sí saben de los que no, y en cuanto al índice de dificultad es	X

				relativamente fácil de responder para los participantes.	
CCSS18N0 07F1O	A	0.18	0.23	En este reactivo, el estadístico evidencia que la discriminación es regular, pero no es significativa, no evidencia quienes posee los conocimientos sobre el tema y quienes nos, sin embargo, el índice de dificultad refleja que es difícil de responder para los estudiantes con el 0.23.	X
CCSS18N0 08F1A	A	0.09	0.78	De acuerdo con los estadísticos de este reactivo se puede observar que en el índice discriminatorio es de 0.09, no se puede distinguir si lo estudiantes tiene conocimiento sobre el tema, pero en el índice de dificultad proporciona lo contrario, es decir que el ítem es relativamente fácil de responder, aunque no está discriminando.	X
CCSS18N0 09F1A	A	0.09	0.10	El índice de discriminación de este reactivo es de 0.09 y con un índice de dificultad del 0.10 es decir que este reactivo no discrimina lo suficiente para distinguir de los estudiantes que, si sabe de los que no, además es muy difícil para que los participantes puedan contestarlo.	X
CCSS18N0 10F1O	A	0.68	0.47	Con respecto a esta pregunta el índice de discriminación es bueno, porque se puede diferenciar quien tiene conocimiento sobre el tema y	X

				quien no, el índice de dificultad es relativamente difícil, es decir que la mayoría puede contestar correctamente es reactivo.	
CCSS18N0 11F1O	A	0.45	0.73	Con respecto a este ítem se puede observar que cuenta con una discriminación considerablemente buena con el 0.45, es decir que discrimina de las personas que si saben y de las que no. Con respecto al índice de dificultad queda evidenciado que es relativamente fácil responderlo, por lo tanto, se conserva para la prueba operativa.	X
CCSS18N0 12F1O	A	0.41	0.67	El índice de dificultad de este reactivo es de 0.41 por los que su discriminación es considerablemente buena, se distingue quien sabe o quién no sobre la temática. Por otro lado, el índice de dificultad es de 0.67 lo que su dificultad es relativamente fácil de responder para los estudiantes.	X
CCSS18N0 13F1O	A	0.41	0.53	El estadístico de este reactivo nos muestra que su índice de discriminación al igual que al anterior es considerado como bueno, con respecto al índice de dificultad a diferencia que el anterior este reactivo es el adecuado ya que el 0.53 de los estudiantes pueden contestar correctamente este reactivo.	X

CCSS18N0 14F1O	A	0.32	0.65	El índice de discriminación de este ítem es de 0.32, por lo que se encuentra en la categoría de razonable, es decir, que solo discrimina el 0.32 de los estudiantes que poseen los conocimientos de los estudiantes que no lo tienen. Sin embargo, el índice de dificultad nos evidencia que el reactivo es relativamente fácil de responder por los estudiantes.	X
CCSS18N0 15F1O	A	0.41	0.28	El estadístico de este reactivo muestra que su índice de discriminación es bueno según la fórmula utilizada en este estudio, con respecto al índice de dificultad es un ítem difícil de contestar para los participantes. Es decir que el ítem posee la cualidad de discriminar a quienes cuentan con las habilidades cognitivas y quienes no, de los participantes solamente 0.28 contestaron correctamente.	X
CCSS18N0 16F1A	A	0.27	0.17	El índice de discriminación de este reactivo es del 0.27, se puede considerar en el rango de regular, solo discrimina en 0.27 a los estudiantes que cuentan con los conocimientos de aquellos que no, también el índice de dificultad es de 0.17, este se encuentra en la categoría de difícil, es por ello que solo el 0.17 de los participantes logran contestar correctamente el reactivo.	X

CCSS18N0 17F1A	A	0.23	0.55	El estadístico acumulado en este reactivo es del 0.23, se encuentra en la categoría de regular, es decir que solo distingue al 0.23 de los estudiantes que conocen el contenido de los que no, por otro lado, el índice de dificultad es el adecuado porque el 55% de los participantes logran responder correctamente este ítem.	X
CCSS18N0 18F1A	A	0.36	0.57	De acuerdo con los datos obtenidos en este estadístico, el índice de discriminación se considera regular debido a que distingue el 0.36 de los participantes que poseen el conocimiento de los que no y con respecto al índice de dificultad, se encuentra en la categoría adecuada debido a que solo el 0.57 de los participantes lo contestan correctamente.	X
CCSS18N0 19F1O	A	0.00	0.60	El índice de discriminación de este reactivo es deficiente no se identifica que estudiantes posee los conocimientos de los participantes que no poseen los saberes, con respecto al índice de dificultad se encuentra en el nivel adecuado, por lo tanto, se debe revisar.	X
CCSS18N0 20F1A	A	0.05	0.29	De acuerdo al estadístico de este reactivo, el índice de discriminación es deficiente, es decir, que no logra diferenciar adecuadamente a los estudiantes	X

				que saben de aquellos que no poseen los conocimientos, en cambio en el índice de dificultad se evidencia que el reactivo es difícil de responder para los participantes. Por lo tanto, no se tomó en cuenta para formar la prueba operativa.	
CCSS18N0 21F1A	A	0.59	0.38	Con respecto al índice de discriminación en este reactivo, encaja en la categoría de bueno con 0.59 ya que distingue entre los estudiantes que poseen los conocimientos de aquellos que no, por otra parte, el índice de dificultad es relativamente difícilmente para los estudiantes, solo el 0.38 logró contestar correctamente esta pregunta.	X
CCSS18N0 22F1A	A	0.27	0.37	El índice de discriminación de este reactivo se puede encontrar en la categoría regular ya que solo puede diferenciar al 0.27 de los estudiantes que tiene conocimiento del tema de lo que no, el índice de dificultad se encuentra en el nivel de relativamente difícil, son pocos los estudiantes logran responder el reactivo correctamente.	X
CCSS18N0 23F1O	A	0.59	0.65	De acuerdo al estadístico de este ítem se puede observar que su índice de discriminación está en la categoría de bueno con 0.59 distingue a las personas que saben de las que no. Sin embargo, en	X

				cuanto al nivel de dificultad se considera relativamente fácil de responder con un 0.65 para los estudiantes investigados,	
CCSS18N0 24F1A	A	0.41	0.47	El índice de discriminación de esta pregunta entra en la categoría de bueno con 0.41 porque si diferencia a los estudiantes que si poseen los conocimientos de los que no, además en el índice de dificultad se considera relativamente difícil con un 0.47 de estudiantes que pueden responderlo de forma acertada.	X
CCSS18N0 25F1A	A	0.36	0.30	El índice de discriminación se encuentra en la categoría razonable con 0.36, permite distinguir que estudiante sabe del que no, sin embargo el reactivo considerado como difícil, solo el 0.30 de los estudiantes respondieron correctamente este ítem.	X
CCSS18N0 26F1A	A	0.59	0.66	El ítem evidencia una discriminación que se puede categorizar como buena en 0.59, permite discernir entre los estudiantes que saben de lo que no, por otra parte, en el índice de dificultad se encuentra en la categoría de relativamente fácil, es decir, solo el 0.66 de los estudiantes saben responder correctamente este reactivo.	X
CCSS18N0 27F1O	A	0.55	0.67	El estadístico de este reactivo arroja que el índice de	

				discriminación es bueno con el 0.55, se deduce que permite distinguir entre los investigados que si saben de aquellos que no, en cambio, el índice de dificultad muestra que es relativamente fácil de responder acertadamente para los estudiantes en un 0.67.	X
CCSS18N0 28F1O	A	0.36	0.52	En este reactivo se evidencia que se encuentra en la categoría razonable en un 0.36, logra diferenciar a los estudiantes que saben de lo que no, en cambio el índice de dificultad es de 0.52 y se encuentra en el nivel adecuado, la mitad de los estudiantes sabe contestar correctamente este reactivo.	X
CCSS18N0 29F1O	A	0.50	0.76	El índice de discriminación de este reactivo se encuentra en la categoría de bueno con 0.50, identifica a los estudiantes que poseen los conocimientos de lo que no, en cambio el índice de dificultad es 0.76, es relativamente fácil de responder, casi todos los estudiantes pueden contestarlo acertadamente.	X
CCSS18N0 30F1A	A	0.14	0.38	La discriminación de este reactivo es deficiente en 0.14, no permite una adecuada diferenciación de aquellos que saben el contenido de los que no saben, en cuanto al índice de dificultad se refleja que es un ítem relativamente difícil de	X

				responder con un 0.38, es decir que son poco los estudiantes que lo puedes responder correctamente.	
CCSS18N0 31F1A	A	0.50	0.67	El estadístico de este reactivo refleja que el índice de discriminación es bueno con 0.50, discierne entre los estudiantes que saben de los que no, en cuanto al índice de dificultad, es relativamente fácil de responder con 0.67, es decir, que la mayoría de los estudiantes saber contestar correctamente este reactivo.	X
CCSS18N0 32F1O	A	0.73	0.62	El índice de discriminación de este reactivo se encuentra en la categoría de buena en discriminación 0.73, se evidencia que estudiantes saben y cuáles que no, pero su índice de dificultad es relativamente fácil con 0.62, casi todos los estudiantes pueden responder correctamente el reactivo.	X
CCSS18N0 33F1O	A	0.14	0.33	En este reactivo el índice de discriminación es deficiente en 0.14, diferencia poco a los estudiantes que, si saben de los que no poseen los conocimientos, por otro lado, el índice de dificultad es 0.33, se categoriza como relativamente difícil, son poco que lo saben responder correctamente.	X

CCSS18N0 34F1A	A	0.36	0.28	El índice de discriminación de este reactivo es razonable con 0.36, sabe distinguir un poco de los estudiantes que saben de los que no, en cambio el índice de dificultad es encuentra en la categoría de ser difícil con 0.28, es decir que solo unos cuantos estudiantes lo responden correctamente.	X
CCSS18N0 35F1O	A	-0.18	0.14	El estadístico de este ítem se evidencia que el índice de discriminación es deficiente con -0.18 no distingue bien al estudiante que sabe del que no conoce el contenido que se evaluó, y el índice de dificultad es categorizado como muy difícil en 0.14, son pocos los estudiantes que lo pueden responder correctamente.	X
CCSS18N0 36F1O	A	0.68	0.76	El índice de discriminación de este reactivo se encuentra en la categoría de muy bueno con 0.68, sabe diferenciar entre los estudiantes que, si cuenta con el conocimiento de los que no, por otro lado, el índice de dificultad evidencia en 0.76, el ítem es relativamente fácil ya que casi todos los estudiantes pueden responder correctamente.	X
CCSS18N0 37F1O	A	0.50	0.63	El índice de discriminación de este reactivo se encuentra en la categoría de bueno en 0.50, sabe distinguir bien de los estudiantes	X

				que saben de lo que no, por otro lado, el índice de dificultad se refleja que es un ítem relativamente fácil en 0.63, es decir que casi todos los estudiantes pueden responder bien este reactivo.	
CCSS18N0 38F1O	A	0.59	0.70	El estadístico de este ítem nos demuestra que el índice de discriminación es bueno en un 0.59 porque diferencia muy bien las personas que cuentan con los conocimientos de las personas que no los poseen, así mismo este estadístico identifican el índice de dificultad el cual se encuentra en la categoría de relativamente fácil en 0.70 casi todos los estudiantes saben responderlo correctamente.	X
CCSS18N0 39F1A	A	0.05	0.35	El índice de discriminación de este reactivo es deficiente en un 0.05, discrimina muy poco a los estudiantes que sabe de los que no, sin embargo, el índice de dificultad refleja que se encuentra en el nivel ser un ítem relativamente difícil en 0.35 por lo tanto se tiene que revisar para el siguiente pilotaje.	X
CCSS18N0 40F1O	A	0.32	0.62	De acuerdo al estadístico de este ítem se puede observar que el índice de discriminación es razonable en 0.32 porque sabe distinguir entre los participantes que tienen conocimientos de los que no, por otro lado, el índice de	X

				dificultad se encuentra en el nivel de relativamente fácil en 0.62, casi todos los participantes pueden responder correctamente este reactivo.	
CCSS18N0 41F1A	A	0.27	0.56	El índice de discriminación de este reactivo es regular ya que solo diferencia el 0.27 de los participantes que poseen los conocimientos de los que no. En cambio, el índice de dificultad se encuentra en la categoría adecuada en 0.56, la mitad de los estudiantes puede contestar correctamente el reactivo.	X
CCSS18N0 42F1A	A	0.09	0.33	El índice de discriminación de este ítem es deficiente en 0.09, no discrimina, no se evidencia de los participantes que si saben de los que no saben. Por otro lado, el índice de dificultad se encuentra en la categoría relativamente difícil en 0.33, no todos los estudiantes logran contestarlo correctamente.	X
CCSS18N0 43F1A	A	0.14	0.15	El índice de discriminación es deficiente en 0.14, discrimina muy poco, no diferencia muy bien a las personas que saben de las que no saben, por otro lado, en cuanto al índice de dificultad es difícil con un 0.15, debido a que pocos son los estudiantes que saben responder correctamente este reactivo.	X

CCSS18N0 44F1O	A	0.14	0.44	El estadístico de este reactivo se evidencia que el índice de discriminación es deficiente en 0.14, discrimina muy poco, no se diferencia muy bien a los participantes que saben de las que no, en cambio en el índice de dificultad refleja que los participantes se encuentran en el nivel adecuado para responder al reactivo.	X
CCSS18N0 45F1O	A	0.18	0.35	El estadístico de este reactivo se evidencia que el índice de discriminación es deficiente en 0.18, discrimina muy poco porque no se diferencia muy bien de las personas que saben de las que no poseen los conocimientos, en cambio el índice de dificultad se encuentra en el rango de relativamente difícil en 0.35, una buena parte de los estudiantes se les dificultada responderlo.	X
CCSS18N0 46F1A	A	0.14	0.17	El índice de discriminación es deficiente en 0.14 discrimina muy poco, no se diferencia muy bien de las personas que saben de las que no. Por otro lado, el índice de dificultad es difícil en 0.17 son pocos los estudiantes que sabes responder correctamente este reactivo.	X

Fuente: Elaboración propia basándose en los valores de Kunder Richardson

Tabla 17: Análisis de los reactivos Forma “B”

Código de los ítems en el test mapa	Forma	Índice de Discriminación	Índice de Dificultad	Análisis	Criterio de selección		
					Descartar	Revisar	Conservar
CCSS18N001F2O	B	0.11	0.36	El índice de discriminación de este reactivo es deficiente porque discrimina poco no se diferencia muy bien que participante cuentan con los conocimientos y quienes no cuenta con los conocimientos. Con respecto al índice de dificultad es relativamente difícil porque son poco los estudiantes que pueden contestar correctamente la pregunta	X		
CCSS18N002F2O	B	0.33	0.41	El índice discriminatorio de este ítem es razonable ya que distingue a los participantes que posee lo conocimientos de los que no; con respecto al índice de dificultad el reactivo es relativamente difícil es decir que no todos los estudiantes pueden contestar correctamente esta pregunta			X
CCSS18N003F2A	B	0.06	0.36	Con respecto a la discriminación de este reactivo se puede evidenciar que es deficiente es decir que no se distingue que participantes tiene conocimiento sobre el contenido y que participantes no poseen los conocimientos, El índice de dificultad muestra que	X		

				es una pregunta relativamente difícil no todos los participantes lo responde correctamente.	
CCSS18N004F2O	B	0.06	0.74	Al igual que el reactivo anterior este reactivo el índice de discriminación es deficiente, no evidencia muy bien que estudiante posee los conocimientos de lo que no saben, sin embargo, el índice de dificultad evidencia que es un reactivo relativamente fácil, es decir que la mayoría de los estudiantes logran contestar correctamente este ítem	X
CCSS18N005F2O	B	0.22	0.38	El índice de discriminación de esta pregunta es regular por distingue a los participantes que saben el contenido de los que no sabe, pero el índice de dificultad es de razonable porque la mayoría de los estudiantes logran responder correctamente el reactivo,	X
CCSS18N006F2A	B	0.06	0.13	Con respecto a este reactivo el índice discriminatorio es deficiente ya que no se diferencia bien a los estudiantes que tienen conocimiento de los que no saben sin embargo el índice de dificultad es difícil es decir que no todos los estudiantes logran responderlo bien	X
CCSS18N007F2A	B	0.28	0.14	El índice discriminatorio de es reactivo es regular porque	X

				diferencia a los participantes que cuentan con el conocimiento de lo que no, sin embargo, el índice de dificultad evidencia que es un reactivo difícil no todos los participantes logran responder correctamente el reactivo	
CCSS18N008F2A	B	0.72	0.57	El índice discriminatorio de este reactivo es un bueno porque sabe diferenciar que estudiante poseen los conocimientos de lo que no, y con respecto a la dificultad se encuentra en el nivel adecuado.	X
CCSS18N009F2O	B	0.56	0.61	Este reactivo cuenta con una discriminación buena porque diferencia muy bien a los estudiantes que poseen los conocimientos de los que no, en cuanto a la dificultad se evidencia que el reactivo es relativamente fácil porque la mayoría de los estudiantes saben responder correctamente este reactivo	X
CCSS18N010F2O	B	0.56	0.33	El estadístico de este ítem evidencia que la discriminación es buena, porque sabe distinguir que estudiante sabe los contenidos y quienes no. El índice de dificultad es relativamente difícil son poco los estudiantes que logran responder correctamente este reactivo	X

CCSS18N011F2A	B	0.11	0.22	La discriminación de este reactivo es deficiente no distingue que participantes cuentan con los conocimientos de los que no, sin embargo, el índice de dificultad evidencia que es un reactivo difícil son poco las personas que responde correctamente el reactivo.	X
CCSS18N012F2O	B	0.28	0.50	Con respecto a la discriminación de esta pregunta se puede observar que se encuentra en el nivel regular por qué diferencia un poco que estudiante posee los conocimientos de los que no, por otro lado, el índice de dificultad es relativamente difícil no son todos los estudiantes que logran responder correctamente este reactivo	X
CCSS18N013F2O	B	0.44	0.54	La discriminación de este reactivo es muy buena porque sabe distinguir muy bien de los estudiantes que si cuentan con el conocimiento de lo que no y la dificultad de este reactivo es la adecuada porque la mayoría de los estudiantes saben responder correctamente esta pregunta.	X
CCSS18N014F2O	B	0.56	0.64	El estadístico de esta pregunta nos evidencia que la discriminación del mismo es muy buena y nos permite	X

				diferenciar que estudiantes tiene conocimiento sobre el contenido y quien no, sin embargo, la dificultad demuestra que es una pregunta relativamente fácil es decir que hasta los estudiantes con un mínimo de conocimiento lo está respondiendo correctamente.	
CCSS18N015F2O	B	0.50	0.68	La discriminación de este reactivo es buena porque distingue al participante que conocen los contenidos de los que no sabe, con respecto al índice de dificultad es relativamente fácil es decir que hasta los estudiantes con un mínimo de conocimiento lo está respondiendo correctamente.	X
CCSS18N016F2A	B	0.33	0.28	En el estadístico de este reactivo se puede observar que el índice de discriminación es razonable porque permite diferenciar que estudiante posee los conocimientos sobre el contenido de lo que no sabe, en el caso del índice de dificultad se puede observar que el reactivo se encuentra en el nivel difícil son pocos los estudiantes que logran contestar correctamente.	X
CCSS18N017F2O	B	0.33	0.39	En este reactivo se puede observar que el índice de discriminación es razonable porque permite distinguir que	X

				participantes posee los conocimientos sobre el contenido de lo que no sabe, en el caso del índice de dificultad se observar que el reactivo se encuentra en el nivel relativamente difícil ya que son pocos los estudiantes que logran contestar correctamente.	
CCSS18N018F2O	B	0.00	0.17	El estadístico de este reactivo con respecto al índice de discriminación nos proporciona que el ítem no diferencia que estudiantes poseen los conocimientos y quienes no, por otro lado, el índice de dificultad evidencia que la pregunta es difícil de responder por casi todos los estudiantes.	X
CCSS18N019F2A	B	0.67	0.50	La discriminación de esta pregunta se observa que es buena porque permite conocer que estudiante saben los contenidos y quienes no, y a la vez cuenta con un índice de dificultad adecuada es decir que la mayoría de los estudiantes saben responder correctamente este reactivo	X
CCSS18N020F2O	B	0.28	0.25	La discriminación de este reactivo es regular porque no se distingue con exactitud que estudiantes cuentan con los conocimientos y quienes no, con respecto al índice de dificultad es considerado difícil,	X

				no todos los estudiantes pueden responderlo correctamente.	
CCSS18N021F2O	B	0.33	0.35	En este reactivo se puede observar que el índice de discriminación es razonable porque permite distinguir que participantes posee los conocimientos sobre el contenido de lo que no sabe, en el caso del índice de dificultad se observar que el reactivo se encuentra en el nivel relativamente difícil ya que son pocos los estudiantes que logran contestar correctamente.	
CCSS18N022F2O	B	-0.22	0.15	En el caso de este reactivo se puede observar su discriminación es negativa es decir es deficiente, ya que su discriminación no se diferencia quien tiene conocimiento de lo que no saben y la dificultad está en el nivel difícil son pocos los estudiantes que los saben responder.	X
CCSS18N023F2A	B	0.67	0.63	El índice discriminatorio de este reactivo es reactivo es bueno, porque distingue que participante saben los contenidos de los que no sin embargo el índice de dificultad es relativamente fácil porque la mayoría de los estudiantes lo respondieron correctamente.	X
CCSS18N024F2A	B	0.67	0.61	El índice discriminatorio de este reactivo es reactivo es bueno,	

				porque distingue que participante saben los contenidos de los que no sin embargo el índice de dificultad es relativamente fácil porque la mayoría de los estudiantes lo respondieron correctamente	
CCSS18N025F2A	B	0.06	0.44	La discriminación de esta pregunta es deficiente discrimina poco de los participantes que cuenta con los conocimientos de los que no, la dificultad es relativamente difícil es decir que la mayoría de los estudiantes pueden responder correctamente este reactivo.	X
CCSS18N026F2A	B	0.61	0.35	Con respecto al índice de discriminación se evidencia que este ítem cuenta con una buena discriminación es decir que, si se puede evidenciar que estudiante posee los conocimientos de quienes no, en cuanto a la dificultad esta pregunta es relativamente difícil son pocos los estudiantes que pueden contestarlo correctamente.	X
CCSS18N027F2O	B	0.11	0.19	En este ítem la discriminación es deficiente ya que no discrimina lo suficiente para distinguir que participantes tiene los conocimientos de los que no lo tienen, la dificultad de este reactivo es muy difícil ya	X

				que son pocos los participantes que logran responderlo	
CCSS18N028F2O	B	0.33	0.31	La discriminación de este reactivo es razonable, porque sabe reconocer que estudiantes poseen el conocimiento de quienes no, con respecto a la dificultad se considera que es un reactivo relativamente difícil porque no todos los estudiantes logran responder correctamente.	X
CCSS18N029F2O	B	0.06	0.07	El índice de discriminación de este reactivo es deficiente ya que no se identifica que estudiantes cuenta con los conocimientos y quienes no cuenta con los mismos, asimismo el índice de dificultad evidencia que la preguntas es muy difícil de responder por los participantes.	X
CCSS18N030F2O	B	-0.11	0.21	El estadístico de esta pregunta evidencia que el reactivo no está discriminando absolutamente nada, es decir que no se sabe que estudiantes poseen los conocimientos y quienes no saben. Con respecto al índice de dificultad evidencia que es un ítem difícil de responder por los participantes	X
CCSS18N031F2A	B	0.00	0.24	El índice de discriminación de este reactivo es deficiente porque no se distingue que participantes cuenta con los	X

				conocimientos y quienes no, al igual que reactivo anterior este es una pregunta difícil de contestar por los estudiantes	
CCSS18N032F2O	B	0.56	0.54	La discriminación de esta pregunta es un buen índice porque identifica muy bien que participante poseen los conocimientos y quienes no, asimismo la dificultad es la adecuada porque identifica que estudiantes contestan correctamente la pregunta y quien no.	X
CCSS18N033F2A	B	0.56	0.64	La discriminación de este reactivo es muy buena porque distingue bien que participantes cuentan con el aprendizaje y quienes no, por lado la discriminación es relativamente fácil porque la mayoría de los estudiantes logran responderlo correctamente.	X
CCSS18N034F2O	B	0.28	0.36	Con respecto al índice de discriminación de este reactivo el estadístico evidencia que su índice de discriminación es regular por identifica poco que participante tiene conocimiento sobre el contenido y quienes no, por otro lado, el índice de dificultad es relativamente difícil no son todos los estudiantes que logran responderlo correctamente.	X

CCSS18N035F2A	B	0.39	0.28	La discriminación de este ítem es razonable porque diferencia un poco más que estudiantes cuentan con los conocimientos y quienes no sin embargo su índice de dificultades evidencia que es un reactivo difícil no todos los participantes logran responder correctamente.	X
CCSS18N036F2O	B	0.61	0.46	El estadístico de esta pregunta nos evidencia que la discriminación es buena idéntica muy bien que estudiante sabe el contenido de quienes no, la dificultad es relativamente difícil porque no todos los estudiantes logran responderlo correctamente.	X
CCSS18N037F2O	B	0.17	0.26	La discriminación de este reactivo es deficiente porque discrimina muy poco que estudiantes poseen los conocimientos sobre el contenido de quienes no saben. La dificultad de este reactivo se considera que es difícil porque no todos los estudiantes logran responderlo.	X
CCSS18N038F2O	B	0.00	0.04	En este estadístico se evidencia que esta pregunta no está discriminando nada no se identifica que estudiantes saben el contenido y quienes no lo saben; asimismo el índice de dificultad se observa es un	X

				reactivo muy difícil para ser respondido correctamente.	
CCSS18N039F2A	B	0.17	0.22	La discriminación de este reactivo es deficiente porque no diferencia bien que estudiantes poseen los conocimientos sobre el contenido de quienes no lo poseen. La dificultad de este reactivo se considera que es difícil porque no todos los estudiantes logran responderlo	X
CCSS18N040F2A	B	0.00	0.36	El estadístico de este reactivo se observa que este ítem no está discriminando nada de los participantes que saben sobre el contenido de quienes no saben el contenido, por otro lado, su índice de dificultad evidencia que es una pregunta relativamente difícil puede ser contestada correctamente por casi todos los participantes.	X
CCSS18N041F2A	B	0.00	0.24	El estadístico de este reactivo se evidencia que este ítem no está discriminando sobre los participantes que saben sobre el contenido de quienes no lo saben, por otro lado, su índice de dificultad se observa que es una pregunta difícil puede ser contestar correctamente por casi todos los participantes.	X
CCSS18N042F2A	B	0.56	0.44	La discriminación de este reactivo es buena porque evidencia muy bien que estudiantes manejan el	X

				contenido de quienes no, con respecto al índice de dificultad se observa que es un ítem relativamente difícil no todos los estudiantes logran responderlo correctamente.	
CCSS18N043F2O	B	0.44	0.43	El estadístico de este ítem se observa que la discriminación es buena porque identifica que participante cuentan con los conocimientos de quienes no, la dificultad es relativamente difícil porque son pocos los estudiantes que logran responderlo correctamente.	X
CCSS18N044F2O	B	0.56	0.51	Con respecto al índice de discriminación de este reactivo se puede observar que es un buen ítem porque sabe diferenciar que estudiantes manejan el contenido de quienes no; asimismo el índice de dificultad es el adecuado porque da conocer que la mayoría de los estudiantes pueden responder correctamente el reactivo.	X
CCSS18N045F2A	B	0.17	0.18	En este reactivo se evidencia que tanto el índice de discriminación como el índice de dificultad son muy bajo es decir que en el primero es deficiente no identifica muy bien qué estudiantes maneja correctamente los contenidos de quienes no lo manejan, en	X

				cambio en el segundo caso el ítem es difícil no todos los estudiantes logran responderlo correctamente.	
CCSSI8N046F2O	B	0.39	0.58	La discriminación de este reactivo es razonable porque distingue que participantes saben de quienes no; por otro lado, el índice de dificultad es la adecuada porque da conocer que la mayoría de los estudiantes pueden responder correctamente el reactivo.	X

Fuente: Elaboración propia basándose en los valores de Kuder Richardson

En tabla anterior se detallan los lineamientos utilizados en la selección de los reactivos que conforman la prueba operativa. Los reactivos seleccionados de la forma A son los siguientes: 1-5, 6, 10-18, 21, 22-29, 31-34, 36-38, 40-41 (total de reactivos aceptados en de la forma A es de 32). Para la forma B se seleccionaron los siguientes: 2, 5, 7-10, 12-17, 19-24, 26, 28, 32-36, 42-44, 46 (un total de 17 reactivos); algunas de las preguntas no fueron seleccionadas ya que se encuentran en la forma A porque fueron reactivos anclas. Según la literatura de la Teoría clásica de los Test los criterios de calidad de los ítems dependen de la muestra de participantes que los han contestado

Para verificar la confiabilidad de la prueba se utilizó el Alfa de Cronbach tanto para la forma A como para la forma B, los estadísticos obtenidos son los siguientes (Forma A)

Tabla 18: Confiabilidad de la prueba con respecto a la forma A

Estadísticas de fiabilidad	
Alfa de Cronbach	N de elementos
0.733	46

Fuente: Elaboración propia

En la forma A, de 46 reactivos se obtuvo una confiabilidad de 0.733, es una confiabilidad alta en cambio, en la forma B se calculó el índice de confiabilidad de la siguiente forma:

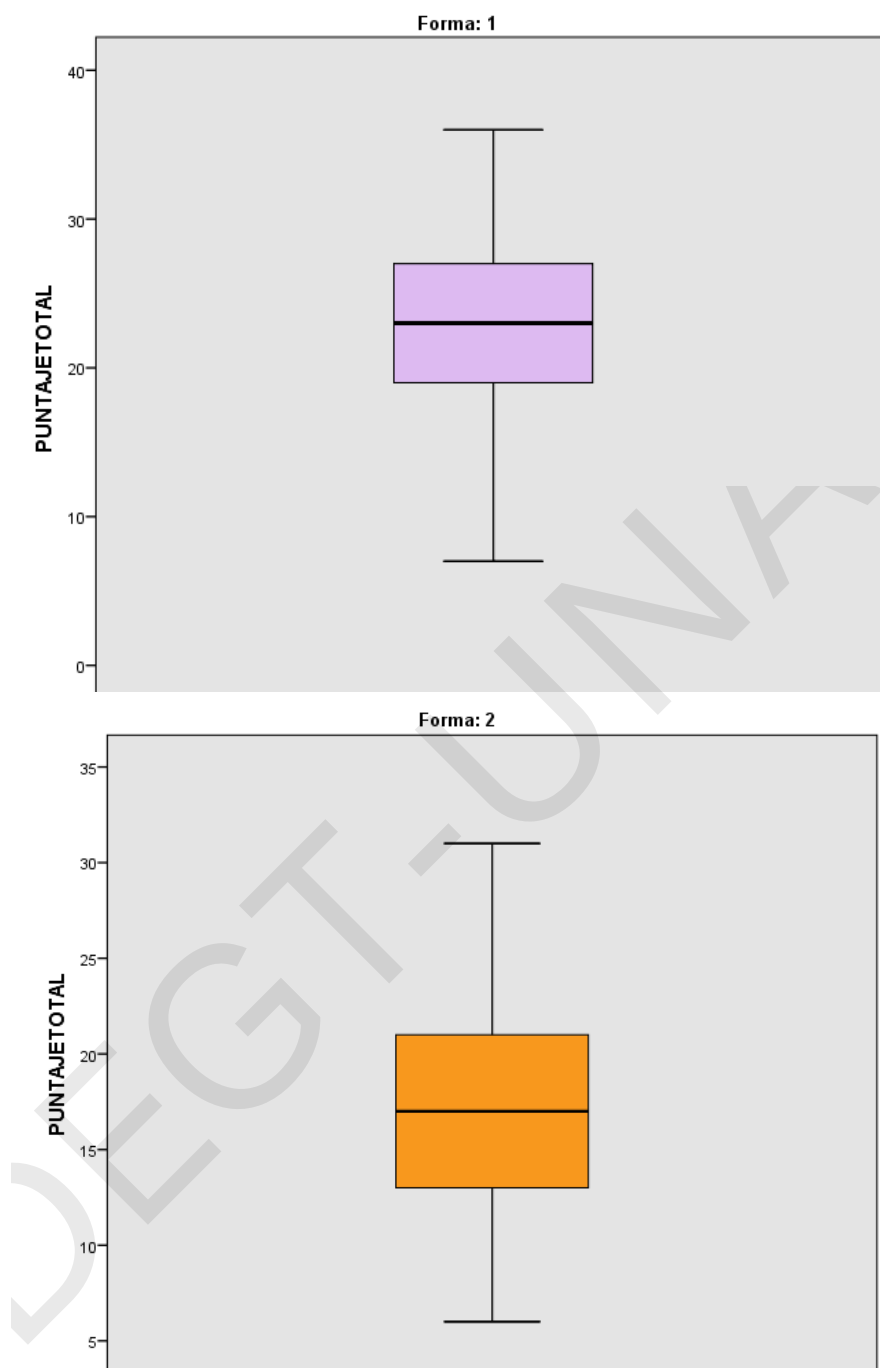
Tabla 19: Confiabilidad de la prueba con respecto a la forma B

Estadística de fiabilidad	
Alfa de Cronbach	N de elementos
0.493	26

Fuente: Elaboración propia

En la forma B de 26 reactivos se obtuvo una confiabilidad de 0.493, los demás ítems son anclas, como lo expresa Reidl-Martínez (2013), “La confiabilidad de una prueba se refiere a la consistencia de las calificaciones obtenidas por las mismas personas en ocasiones diferentes o con diferentes conjuntos de reactivos equivalentes” (Reidl-Martínez., 2013, pág. 109).

Para verificar que no existiera datos atípicos en ambas formas se elaboró un gráfico de caja de bigotes teniendo como resultado lo siguiente:

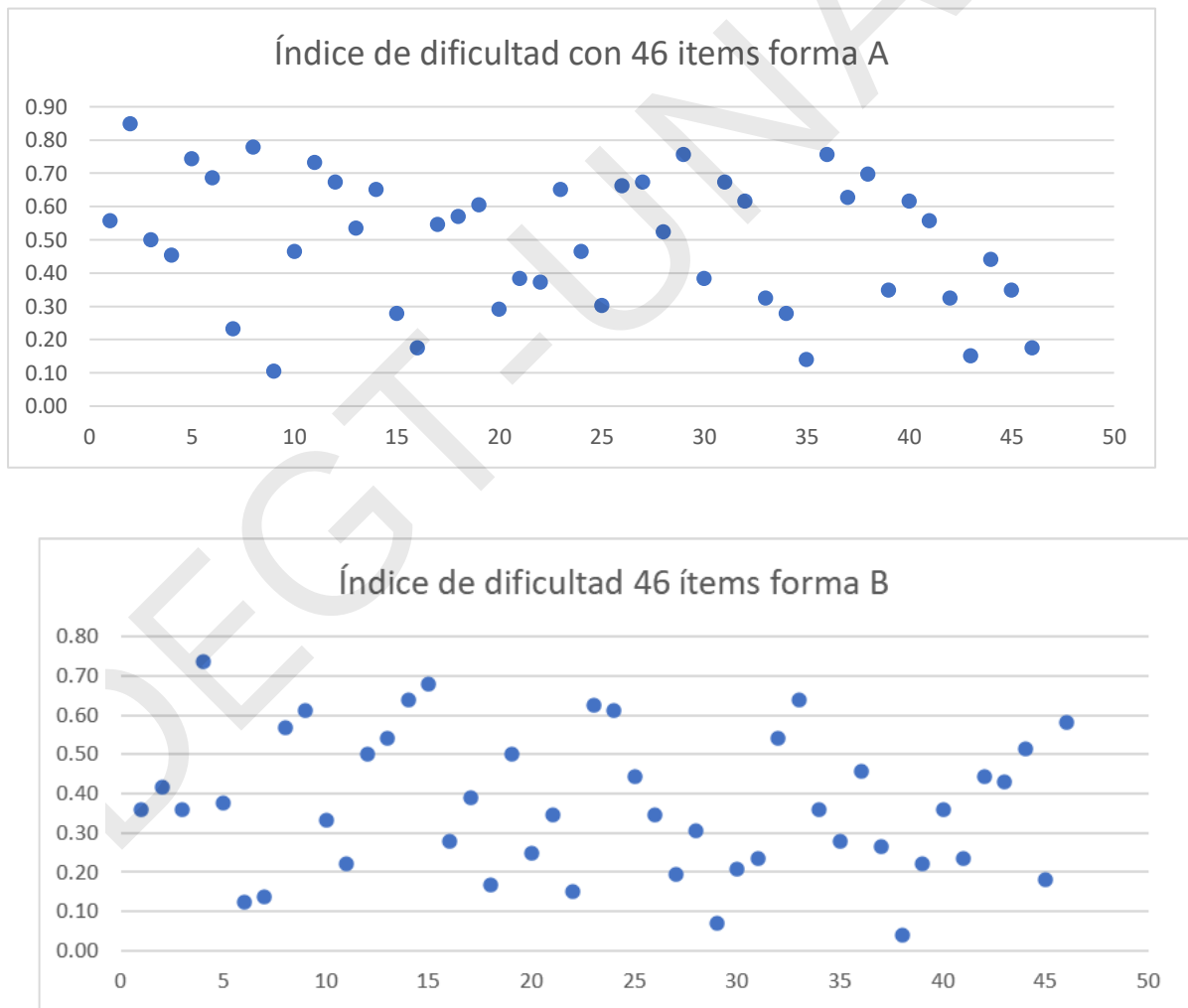
Figura 16: Verificación datos atípico de las pruebas pilotos

Fuente: Elaboración propia

En la figura 16 muestra que ambas pruebas pilotadas no poseen datos atípicos, sin embargo, es importante mencionar que en ambas cuenta con un estadístico donde la puntuación mínima fue de 1 y la máxima fue de 4, pero con respecto a la media se observa que en la forma A es equivalente a 3.02 y en la forma B la media es de 2.10, con respecto a la mediana en la forma A es de 3 y la forma B es de 2, tal y como se puede observar en la caja de bigotes, en el centro del diagrama indica los puntajes que están por debajo de la mediana y la otra mitades de los puntajes se encuentra arriba de la mediana.

En el siguiente gráfico se puede observar más a profundidad la dificultad de ambas pruebas piloto:

Gráfico 1: índice de dificultad forma A y B



Fuente: Elaboración propia

En el gráfico 1 la forma B, fue una de las pruebas con mayor dificultad para ser contestada por los estudiantes, para conformar la prueba operativa se tomaron en cuenta reactivos de ambas pruebas.

5.1.2 Digitalización de los resultados, obtenidos en aplicación de la prueba operativa.

Dándole el seguimiento al objetivo número tres de esta investigación: Determinar las propiedades psicométricas de la prueba que evalúe el desempeño académico de los estudiantes de noveno grado en el área de Ciencias Sociales. La prueba se aplicó en los centros educativos, Héctor Pineda Ugarte con una muestra representativa del 27%, en el Mixto Hibuera con un 45% de muestra representación y el Técnico Abelardo Fortín con el 28%.

Se verificó la confiabilidad de la prueba operativa obteniendo como resultado lo siguiente:
Tabla 20: Confiabilidad de la prueba operativa

Estadísticas de fiabilidad	
Alfa de Cronbach	N de elementos
0.798	50

Fuente: Elaboración propia

Con una cantidad de 50 reactivos contenidos en la prueba operativa, su confiabilidad es de 0.798, para que la confiabilidad aumentara se suprimieron 5 reactivos obtenidos como resultado el siguiente dato:

Tabla 21: Confiabilidad con 45 reactivo

Estadísticas de fiabilidad		
Alfa de Cronbach	Alfa de Cronbach basada en elementos estandarizados	N de elementos
0.817	0.818	45

Fuente: Elaboración propia

En esta prueba operativa 5 reactivos fueron suprimidos, el alfa de Cronbach subió a 0.817, haciendo que la prueba operativa fuera más confiable.

Tabla 22: Parámetro estadísticos de 45 reactivos de la prueba operativa

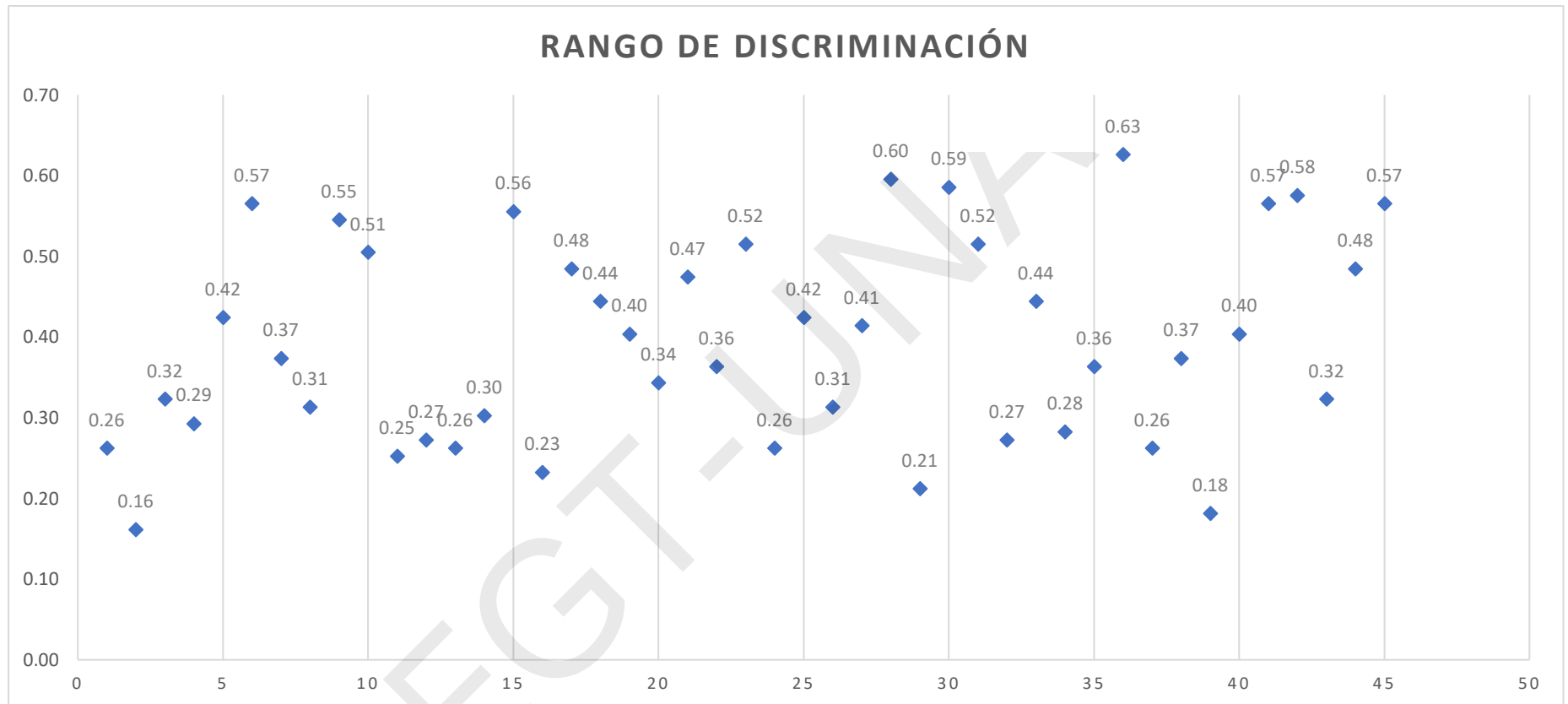
Ítem	Discriminación	Dificultad	Correlación total de elementos corregida
CCSS18N001F1O	0.26	0.67	0.15
CCSS18N002F1O	0.16	0.90	0.15
CCSS18N003F1O	0.32	0.49	0.17
CCSS18N004F1A	0.29	0.35	0.17
CCSS18N006F1O	0.42	0.66	0.26
CCSS18N010F1O	0.57	0.33	0.35
CCSS18N011F1O	0.37	0.74	0.27
CCSS18N012F1O	0.31	0.78	0.21
CCSS18N013F1O	0.55	0.64	0.33
CCSS18N014F1O	0.51	0.73	0.34
CCSS18N015F1O	0.25	0.34	0.18
CCSS18N016F1A	0.27	0.17	0.26
CCSS18N017F1A	0.26	0.49	0.10
CCSS18N018F1A	0.30	0.51	0.16
CCSS18N021F1A	0.56	0.37	0.40
CCSS18N022F1A	0.23	0.35	0.14
CCSS18N023F1O	0.48	0.76	0.40
CCSS18N024F1A	0.44	0.42	0.27
CCSS18N025F1A	0.40	0.31	0.28
CCSS18N026F1A	0.34	0.73	0.22
CCSS18N027F1O	0.47	0.72	0.38
CCSS18N028F1O	0.36	0.53	0.24
CCSS18N029F1O	0.52	0.79	0.41
CCSS18N031F1A	0.26	0.82	0.23
CCSS18N032F1O	0.42	0.60	0.23
CCSS18N034F1A	0.31	0.28	0.21
CCSS18N036F1O	0.41	0.83	0.39
CCSS18N037F1O	0.60	0.58	0.39
CCSS18N038F1O	0.21	0.72	0.17
CCSS18N040F1O	0.59	0.69	0.44
CCSS18N041F1A	0.52	0.61	0.32
CCSS18N002F2O	0.27	0.46	0.23
CCSS18N009F2O	0.44	0.65	0.30
CCSS18N010F2O	0.28	0.36	0.21
CCSS18N012F2O	0.36	0.43	0.25
CCSS18N013F2O	0.63	0.41	0.44
CCSS18N014F2O	0.26	0.81	0.25
CCSS18N015F2O	0.37	0.63	0.25

CCSS18N017F2O	0.18	0.25	0.12
CCSS18N008F1A	0.40	0.73	0.31
CCSS18N032F2O	0.57	0.54	0.37
CCSS18N036F2O	0.58	0.54	0.38
CCSS18N043F2O	0.32	0.40	0.18
CCSS18N044F2O	0.48	0.41	0.34
CCSS18N046F2O	0.57	0.56	0.43

Fuente: Elaboración propia

En la tabla 20 se puede observar la discriminación y la dificultad de cada reactivo obtenidos en la prueba operativa. Teniendo resultados positivos después de eliminar los ítems que estadísticamente fueran negativos o con una puntuación baja.

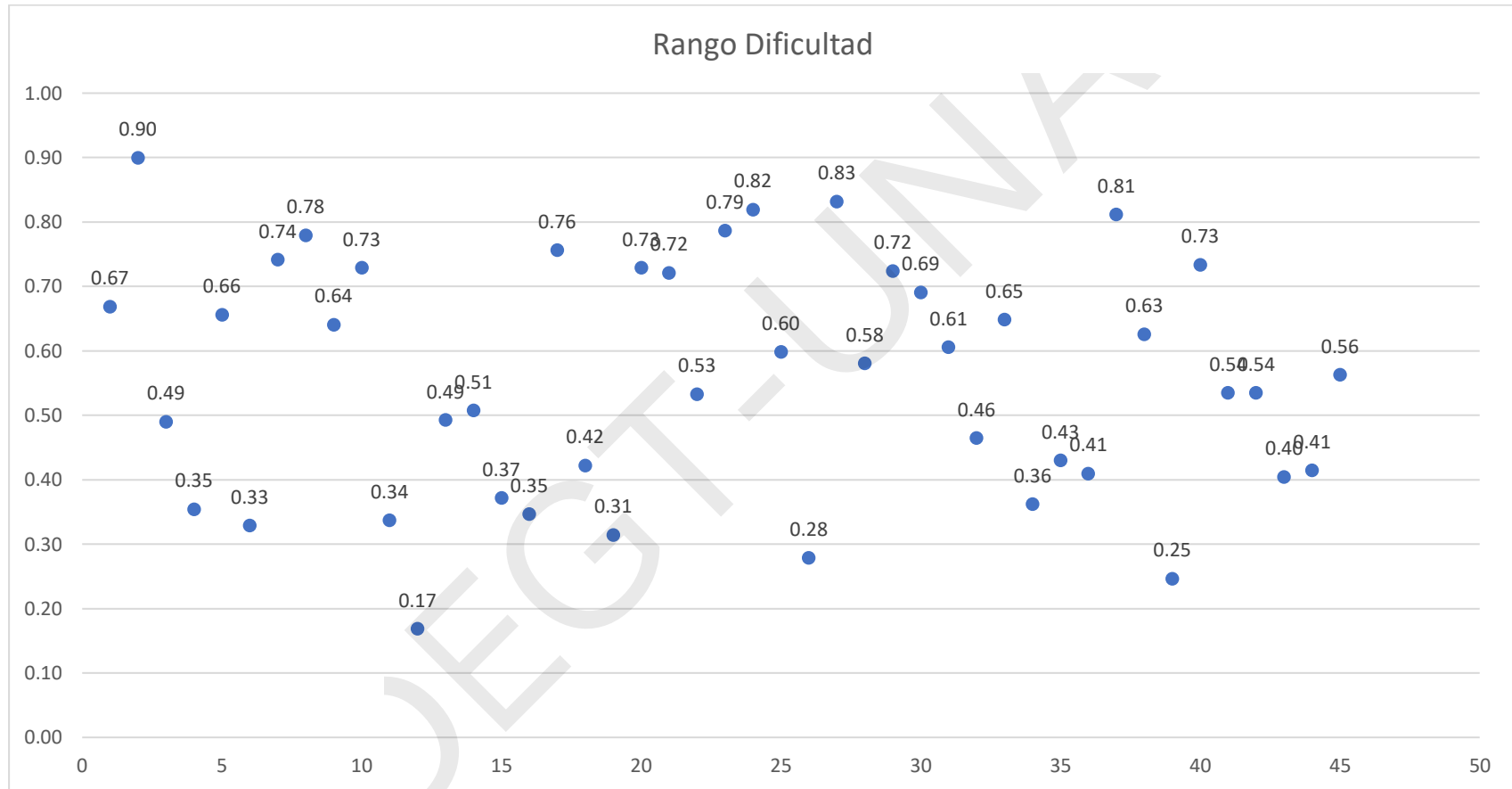
Gráfico 2:Rango de discriminación prueba operativa



Fuente: Elaboración propia

En el gráfico 2 se puede evidenciar cada uno de los valores de discriminación de cada ítem, siempre basándose en los valores de Küder Richardson, el reactivo con el código CCSS18N002F11O cuenta a con una discriminación deficiente por lo tanto tiene que ser revisado; al contrario de los demás que se encuentran en el rango de un índice razonable a un buen índice de discriminación

Gráfico 3:Rango de dificultad prueba operativa

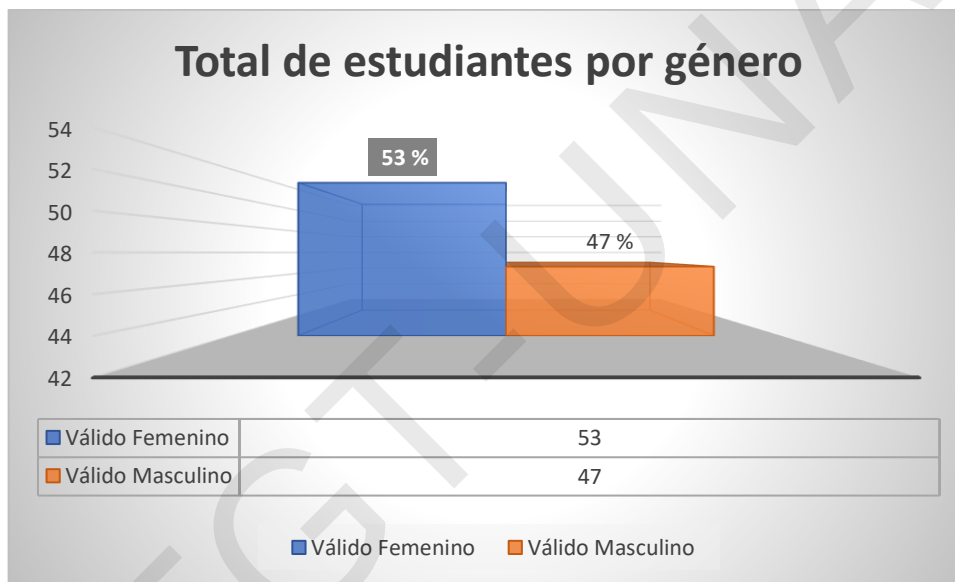


Fuente: Elaboración propia

En el gráfico 3 se puede evidenciar la dificultad de los ítems en la prueba operativa siempre basándose en los valores de Küder Richardson, se puede observar que el ítem difícil de contestar por los participantes es el reactivo con el código CCSS18N016F1A con el 0.17. Con respecto a los demás reactivos se encuentra en los rangos de relativamente difícil hasta muy fácil.

La prueba estandariza de fin de grado para evaluar el desempeño académico de los estudiantes de noveno grado en el área curricular Ciencia Sociales se aplicó en tres centros educativos descritos anteriormente; Con respecto al género en los tres centros educativos se obtuvo los siguientes resultados:

Gráfico 4: Total de estudiantes por género

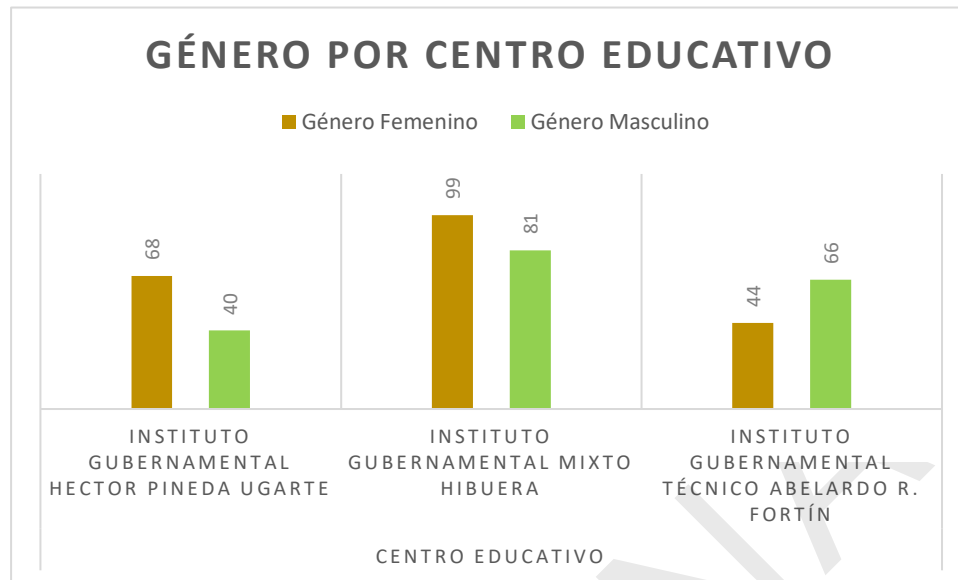


Fuente: Elaboración propia.

El gráfico 4 describe el género de estudiantes investigados, 53% estudiantes pertenecen al género femenino y 47% al masculino. Se puede concluir que, de los participantes, el género femenino es mayoritario. Los datos obtenidos en este gráfico coinciden con la tesis doctoral de Valladares (2018), dicha tesis fue realizada en un grado mayor (Décimo grado)

El estadístico por centro educativo se representa de la siguiente manera:

Gráfico 5: Género por centro educativo



Fuente: Elaboración propia.

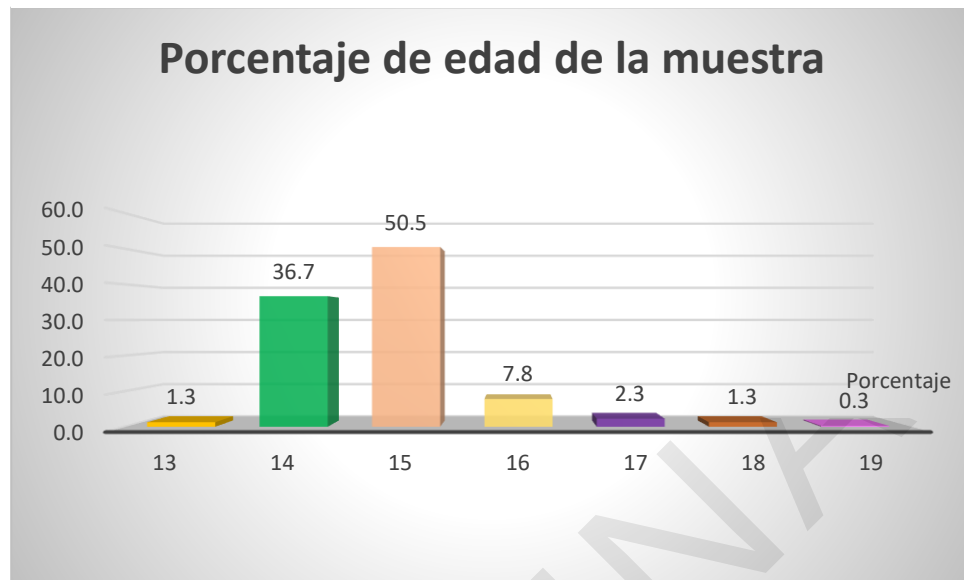
La mayor representación por centro educativo se encontró el mayor porcentaje de género femenino se encuentra en el instituto Mixto Hibueras con el 99 % seguido de Héctor Pineda Ugarte con el 68% y con solo un 44% en Técnico Abelardo Fortín, con el género masculino sigue liderando el Mixto Hibueras con el 81%, seguido de Técnico Abelardo Fortín el 66% y con el 40% el Héctor Pineda Ugarte.

La Secretaría de Educación (2019) expresa que:

el sistema educativo hondureño ha tenido un ligero avance en el período 2015-2018 en la Educación Pre-Básica con una mejoría cercana a 5% hasta tener 44.3% de cobertura en el año 2018, y en la atención de los jóvenes en la educación media con una mejoría de 3% (45.7% en el año 2018). (Secretaría de Educación , 2019, pág. 19)

En el gráfico 7 se evidencia la edad de la muestra a la cual se le administró la prueba de Ciencia Sociales de fin de grado.

Gráfico 6: Porcentaje de edad de la muestra



Fuente: Elaboración propia

Los estudiantes que respondieron la prueba de fin de grado del área de Ciencia Sociales se encuentran en la edad establecida por el CNB; entre doce a 15 años, observando que el mayor porcentaje de estudiantes se encuentran en la edad de 15 años con el 50.5%, seguido de la edad de 14 años con el 36.7% y con 1.3% de 13 años, sin embargo, hay jóvenes que están rezagados en noveno grado hay estudiantes de 16 a 19 años de edad, es decir que sobre pasan la edad establecida por el CNB.

Es valioso rescatar la información proyectada por el Instituto Nacional de Estadísticas (INE), donde proyecta que para el 2016-2020 crecerá la población entres estas edades que corresponde a séptimo y noveno grado, asimismo se debe realizar esfuerzos para mejorar la cobertura educativa y la calidad de la misma (Secretaria de Educación, 2017)

Por otro lado, se puede mencionar que la investigación realizada por Valladares (2018) demuestran que aún en grados superiores se tiene estudiantes de sobre edad “Los alumnos de 14 años, 18 y 19 años, estarían como valores atípicos para estar en décimo grado, pero el sistema educativo hondureño presenta problemas de sobreedad.” (Valladares, 2018, pág. 188)

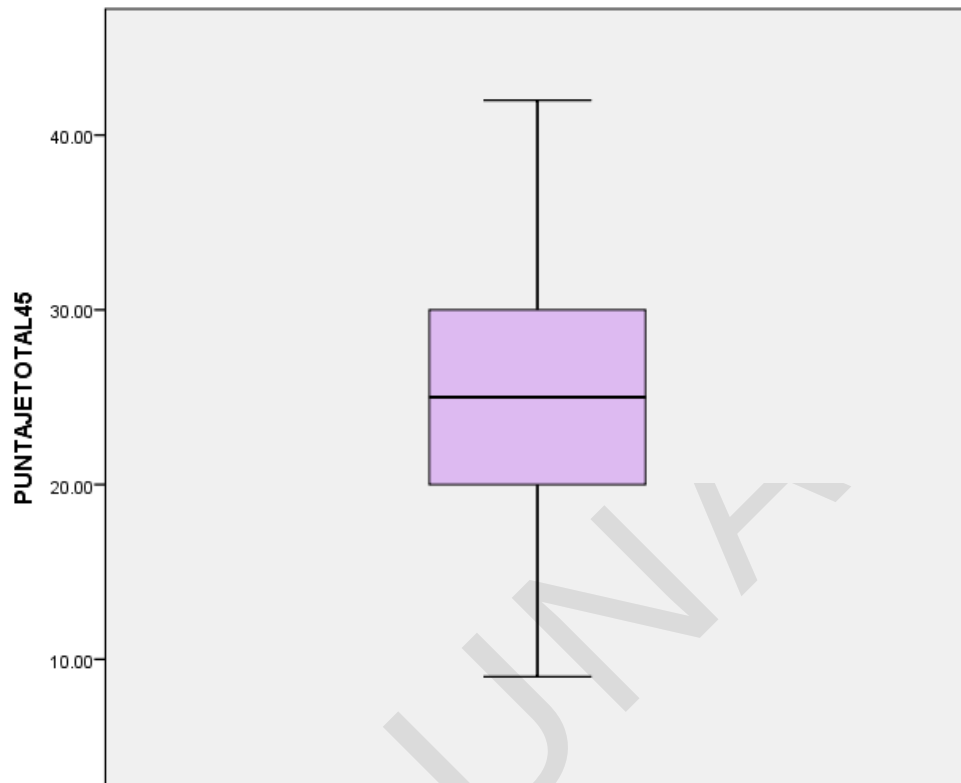
Tabla 23:Análisis estadístico descriptivo por bloques del área curricular de Ciencias Sociales

Estadísticos descriptivos							
	N	Rango	Mínimo	Máximo	Media	Desviación estándar	Varianza
BLOQUE1La_Persona_y su ser social Rec	398	.83	.17	1.00	.6340	.18894	.036
BLOQUE2La_Sociedad_y los espacios geográficos Rec	398	.86	.10	.95	.5072	.17537	.031
BLOQUE3Las_sociedades_y el tiempo social Rec	398	1.00	.00	1.00	.5075	.23216	.054
BLOQUE4La_sociedad_y las actividades humanas Rec	398	1.00	.00	1.00	.5921	.23546	.055
N válido (por lista)	398						

Fuente: Elaboración propia.

En la tabla anterior se evidencia los resultados estadísticos descriptivos de los cuatro bloques que conforman el área de Ciencias Sociales en noveno grado. Según este estadístico el bloque que obtuvo un mayor promedio fue la persona y su ser social con el 0.63, es decir que la mayoría de los estudiantes respondieron correctamente los reactivos de este bloque, y con una desviación del 0.19; con respecto al bloque 2 y al bloque tres son equivalente con respecto a la dificultad ya que ambas obtuvieron un promedio de 0.51 a diferencia de la desviación estándar ya que en el bloque 2 la sociedad y los espacios geográficos fue de 0.18 y en el bloque 3 las sociedades y el tiempo social obtuvo una desviación de 0.23.

Gráfico 7: Diagrama de caja del puntaje total de 45 reactivos de la prueba de noveno grado CCSS



Fuente: Elaboración propia

Con 45 reactivos se puede observar que no hay dato atípico es la prueba es decir que existe una distribución normal en la prueba. Los puntajes obtenidos son del 9 como mínimo y 40 como máximo y la media de 25 y se encuentra situada en medio del diagrama de la caja de bigotes, lo que muestra que la mitad de los puntajes totales están por debajo de la media y la otra mitad de los puntajes totales están sobre la media.

Dándole seguimiento al tercer objetivo de la investigación, *describir el desempeño académico de los estudiantes de noveno grado en el área de Ciencias Sociales*, para darle respuesta a este objetivo primero se tuvo que ordenar los reactivos para conocer la escala donde se ubicaría cada uno de los ítems; lo cual quedó de la siguiente manera:

Tabla 24: Escala de puntuación de desempeño académico

PUNTAJETOTAL45						
		Frecuencia	Frecuencia Acumulada	Percentiles	Puntaje Zi Media: 25.0829 Desviación: 6.93	Escala de puntuación Estandarizada = (50* Zi) +300
Válido	9.00	2	2	1	-2.3	184
	10.00	1	3	0	-2.2	191
	11.00	3	6	1	-2.0	198
	12.00	6	12	2	-1.9	206
	13.00	5	17	1	-1.7	213
	14.00	11	28	3	-1.6	220
	15.00	10	38	3	-1.5	227
	16.00	13	51	3	-1.3	234
	17.00	11	62	3	-1.2	242
	18.00	11	73	3	-1.0	249
	19.00	19	92	5	-.9	256
	20.00	19	111	5	-.7	263
	21.00	24	135	6	-.6	271
	22.00	17	152	4	-.4	278
	23.00	13	165	3	-.3	285
	24.00	17	182	4	-.2	292
	25.00	28	210	7	.0	299
	26.00	16	226	4	.1	307
	27.00	16	242	4	.3	314
	28.00	16	258	4	.4	321
	29.00	22	280	6	.6	328
	30.00	24	304	6	.7	335
	31.00	16	320	4	.9	343
	32.00	13	333	3	1.0	350
	33.00	15	348	4	1.1	357
	34.00	15	363	4	1.3	364
	35.00	8	371	2	1.4	372
	36.00	9	380	2	1.6	379
	37.00	9	389	2	1.7	386
	38.00	4	393	1	1.9	393
	39.00	2	395	1	2.0	400
	41.00	2	397	1	2.3	415
	42.00	1	398	0	2.4	422
	Total	398		100.0		

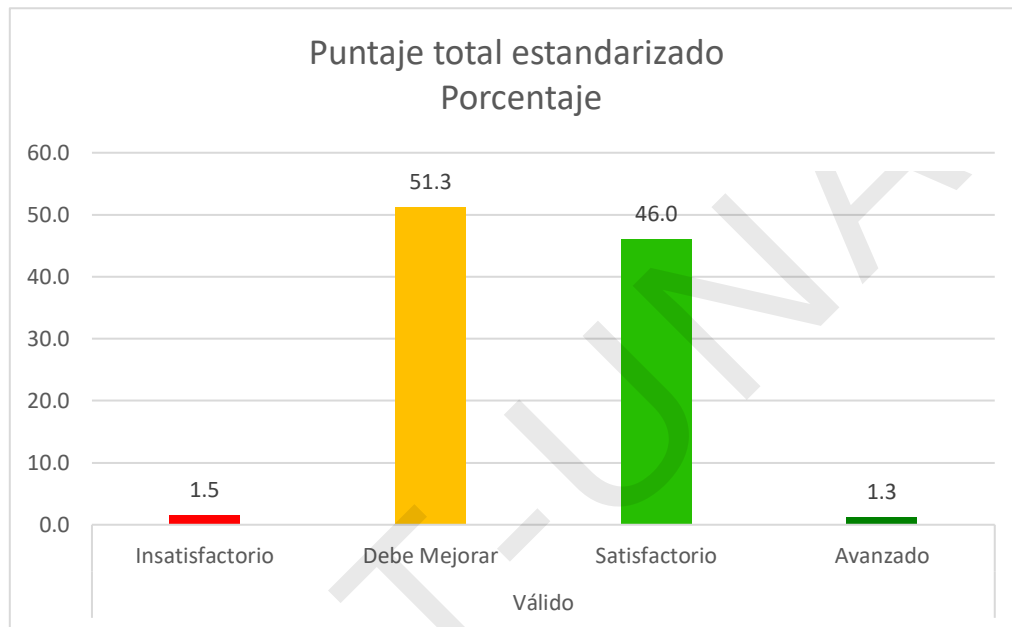
Fuente: Elaboración propia

En la tabla se puede observar en la escala de puntuación estandarizada que se utilizó para poder conocer el desempeño académico de los participantes. (Ver tabla 13). A continuación, se le presenta una tabla por puntaje por cada centro educativo.

CAPITULO VI: Discusión de los resultados

En la prueba de Ciencias Sociales aplicada a los estudiantes de noveno grado se han obtenido los siguientes resultados:

Gráfico 8: Porcentaje de estudiante por nivel de desempeño



Fuente: Elaboración propia

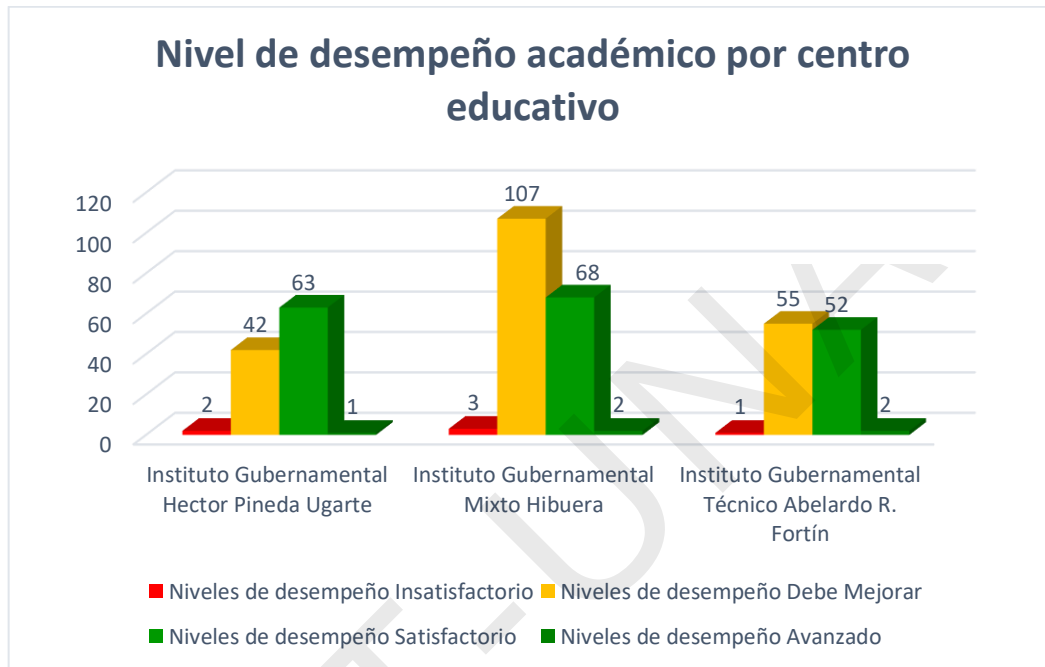
En noveno grado en el área de Ciencias Sociales, la mayoría de los estudiantes se encuentra localizados en el nivel de Debe Mejorar con el 51.3% el cual queda evidenciado que los estudiantes no están logrando por completo las competencias establecidas por el CNB, por otro lado se observa que son pocos los estudiantes que presentan conocimientos mínimos para ser promovidos para el siguiente grado, en cambio un dato preocupante para el sistema educativo nacional es el nivel avanzado ya que solo el 1.3% se ubican en este nivel, y en comparación en el nivel insatisfactoria la diferencia es mínima.

Según la Secretaría de Educación en el 2019 son pocos los estudiantes que se encuentran en el nivel avanzado esto con respecto áreas curriculares de español y matemáticas, sin embargo en el área de Ciencias Sociales no deja de ser la excepción (tan solo el 1.3% se encuentra en nivel avanzado) aún hay una cantidad de estudiantes considerables de estudiantes que no poseen las habilidades, destrezas, conocimientos y estandares de

contenidos para avanzar al siguiente grado y se encuentran en el nivel de desempeño debe mejorar (con el 51.3%). (Secretaría de Educación , 2019, pág. 31)

Una vez obtenido el porcentaje de los niveles de desempeño académico de los estudiantes se procedió a sacar un gráfico por sientio educativo dando como resultado lo siguientes

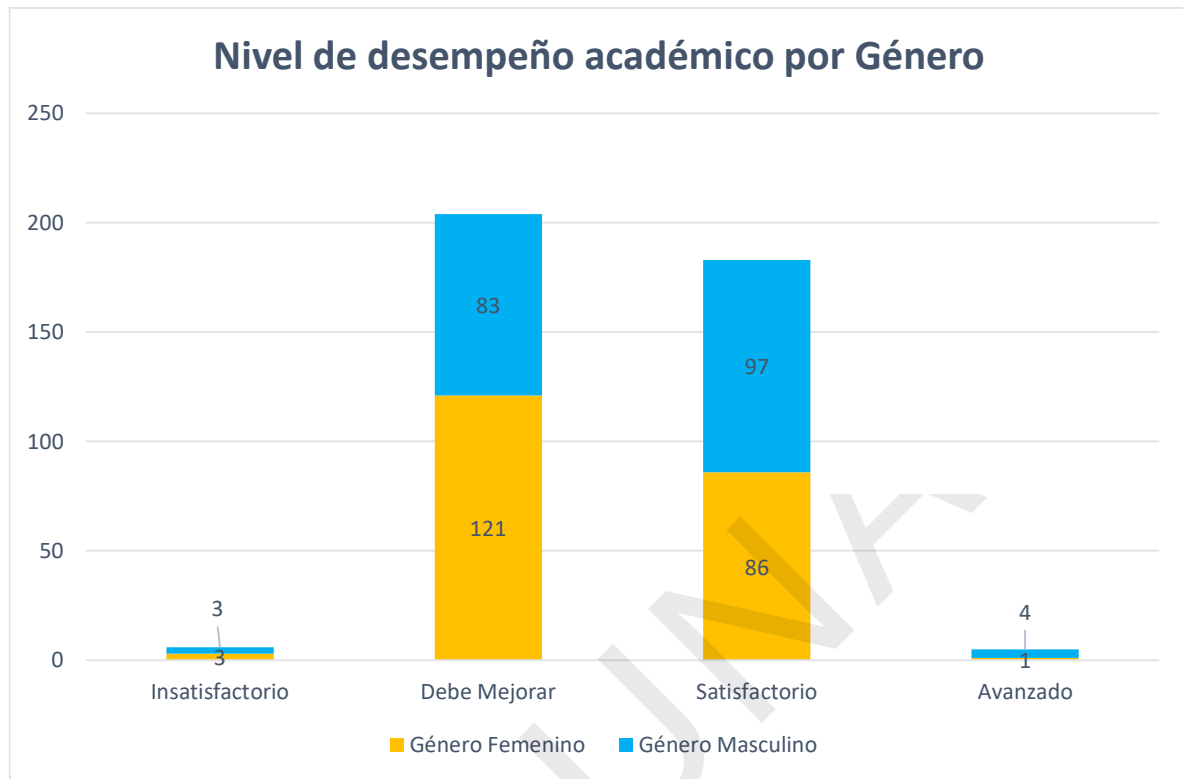
Gráfico 9: Nivel de desempeño académico por centro educativo



Fuente Elaboración propia

Con respecto al desempeño de los estudiantes por centro educativos, se puede observar que es en el centro educativo Mixto Hibueras donde se encuentra la mayor cantidad de estudiantes el nivel de debe mejorar, seguido del Técnico Abelardo Fortín y por consiguiente el Héctor Pineda Ugarte, con respecto al nivel satisfactorio lo lidera Mixto Hibueras con 68 estudiantes, seguido del Héctor Pineda Ugarte con 42 estudiantes y con 52 estudiantes se encuentra el Técnico Abelardo Fortín, referente al nivel insatisfactorio y el avanzado ambos están representado por 2 en el centro educativo Héctor Pineda Ugarte y el Técnico Abelardo Fortín, lo contrario de Mixto Hibueras que lo representan 3 estudiantes.

Gráfico 10:Nivel de desempeño académico por género



Elaboración propia

En el gráfico 11 se puede observar que con respecto al nivel de desempeño por género en el nivel insatisfactorio tanto el género femenino como el masculino es de 3: sin embargo en el nivel de debe mejorar predomina más el género femenino con el 121 de las participantes mientras el 83 es de género masculino, por otro lado en el nivel satisfactorio se encuentra liderado por el género masculino con el 97 mientras que el género femenino el 86 en el nivel avanzado se encuentra el género masculino con un 4 y solo el 1 es del género femenino.

Estos datos coinciden con los estudios realizado por la Secretaría de Educación donde en el área de matemáticas predominan más el género masculino “En Matemáticas, en la escala de puntuación estandarizada, se encontró que los varones de 9no grado están levemente arriba que las jóvenes mujeres” (Secretaría de Educación , 2019;p.39)

Conclusiones

1. El desarrollo de una prueba de fin de grado que se encargue de evaluar el desempeño académico de los estudiantes de noveno grado en Ciencias Sociales es un reto para la secretaria de educación, porque hay que concientizar a los docentes la importancia que son las pruebas estandarizadas y los beneficios que estas pueden traer en el sistema educativo

2. Para el diseño de los reactivos se capacitaron a docentes del área de CCSS, para que los reactivos contaran con la confiabilidad y la validez que se requiere para que los ítems obtengan la calidad pertinente, dichos docentes tenían que estar laborando en una institución como docente de noveno grado y con experiencia en la misma

3. Para verificar si los reactivos cumplían su función se realizaron los estadísticos de alfa de Cronbach, mediante el cual se obtuvieron resultados positivos, aun así, se eliminaron algunos reactivos que no cumplieron su función y de esta manera se obtuvo una mejor confiabilidad, índice de discriminación para reconocer que estudiantes cuentan con los conocimientos y que estudiantes no. Asimismo, se calculó la dificultad de los reactivos para saber que cual de los ítems era más fácil de contestar y cual no lo era por parte de los participantes.

4. Con respecto al desempeño académico en el área curricular de CCSS, se concluyó que de la muestra evaluada en la investigación el 46%, están alcanzando los estándares establecidos en el CNB, el 1.3. % poseen, más conocimientos de lo que demandas el CNB Y el 51.3% de los estudiantes evaluados en este estudio no están logrando las habilidad y destrezas estipuladas por el sistema educativo de Honduras.

5. Las investigaciones realizadas a nivel nacional han realizado análisis comparativo entre el nivel de desempeño por género en las áreas de español y matemáticas; en el cual se evidencia que en el área de español el género femenino tiene un mejor desempeño al contrario en el área de matemáticas que lo lidera el género masculino. Realizando la comparación de estas dos áreas en CCSS quienes obtuvieron mejores resultados fueron el género masculino.

Recomendaciones

Las pruebas estandarizadas son instrumentos de evaluación de suma importancia dentro de un sistema educativo, dado que nos proporciona un panorama general de las fortalezas y debilidades del mismo sistema, en el caso de los estudiantes; en el cual se logran detectar que grupo de los participantes o de la población con necesidades de mejoras educativas, asimismo permite identificar qué factores están impactando en el desempeño de los estudiantes.

Al momento de elaborar un instrumento de evaluación que se aplicará a nivel nacional es recomendable asegurarse que todos los docentes que participarán en el proceso de elaboración de reactivos, conozcan los estándares que establece el CNB y de esta manera asegurar que la información a obtener será valiosa, asimismo se deben tener en cuenta la confiabilidad y la validez de dicha evaluación también debe estar acorde a las especificaciones técnicas que permita analizar la información proporcionada por los participantes. Es útil informar a las instituciones educativas sobre el desempeño académico de sus educandos y tomar decisiones pertinentes para que se realicen un plan de mejora.

Con los resultados obtenidos en la prueba de Ciencias Sociales se le sugiere a toda la comunidad educativa que se inicie a realizar una evaluación estandarizada en las 4 áreas curriculares establecidas en el CNB para formular políticas educativas pertinentes para mejorar el desempeño académico de la población estudiantil; a la vez se recomienda que se capacite a los docentes en los tópicos que más debilidades poseen y de esta manera mejorar el proceso de enseñanza-aprendizaje en esta área curricular. A la vez se recomienda indagar sobre los estudiantes que se encuentra cursando noveno grado en una sobreedad y buscar que su aprendizaje se más significativo.

Con la información proporcionada en esta prueba se le sugiere al docente readecuar sus estrategias de enseñanza-aprendizajes según el desempeño académico obtenido las estudiantes en esta evaluación, ya que son los educadores que más protagonismo tienen en el aula de clases y son los principales agentes para ayudar a las autoridades a transformar los resultados negativos a positivos

Referencias bibliografía

- Martón, J. L. (2008). *Calentamiento Global: al borde del límite*. España: Córdoba, España,.
- Abad, F., Garrido, J., Olea, J., & Ponsoda, V. (2006). *Introducción a la Psicometría Teoría Clásica de los Tests y Teoría de la Respuesta al Ítem*. Madrid.
- Aiken, L. R. (2003). Diseño y elaboración de tests. En L. R. Aiken, & L. G. Figueroa (Ed.), *Test psicológicos y evaluación* (M. E. Vega, Trad., págs. 18-42). México: Pearson Educación .
- Alas, M., & Moncada, G. (2011). Problemas de Equidad en el Sistema Educativo Hondureño. *INIEES*, 234-235.
- Alonso, R. S. (2016). Teoría de la Educación: Conocimiento de la Educación, Investigación, Disciplina Académica. *EDITORIAL Revista Virtual Redipe*, VIII(8), 19-38.
- Aragón, B. L. (2015). *Evaluación psicológica: historia, fundamentos teórico-conceptuales y psicometría*. México: El Manual Moderno.
- Arancibia, J. (2015). La estandarización de la evaluación. Las pruebas nacionales e internacionales ¿medición o evaluación? *Internacional de la Educación para América Latina*, 21.
- Ardila, R. (2011). Inteligencia. ¿Qué sabemos y qué nos falta por investigar? *Revista Academia Colombiana de Ciencias Exactas, Físicas y Naturales*, XXXV(134), 97-103.
- Arredondo, S. C., & Diago, J. C. (2006). *Evaluación Educativa y Promoción Escolar*. Madrid: Pearson Educación, S.A.
- Arregui, P., Valverde, G., Ravela, P., Wolfe, R., Ferrer, G., Rizo, F. M., . . . Wolff, L. (2008). *Las Evaluaciones Educativas que América Latina Necesita*. Santiago de Chile: PREAL.
- Attorresi, H. F., Lozzia, G. S., Abal, F. J., Galibert, M. S., & Aguerri, M. E. (2009). Teoría de Respuesta al Ítem. Conceptos básicos y aplicaciones para la medición de constructos psicológicos. *Revista Argentina de Clínica Psicológica*, XVIII,(2), 179-188.
- Ayala, R. A. (2006). *Herramientas de medición y evaluación*. Universidad de Puerto Rico, Recinto de Humacao.
- Badii, M. A., Rodríguez, O. L., Aguilar, J., & Acuña, M. (2015). Pérdida de Biodiversidad: Causas y Efectos. *Daena: International Journal of Good Conscience.*, 156-174.
- Barajas, Y. E. (2011). Una reseña sobre la validez de constructo de pruebas referidas a criterio. *Perfiles Educativos*, XXXIII(131), pp. 131-154.

- Barrenechea, I. (2010). Evaluaciones Estandarizadas: Seis Reflexiones Críticas. *Archivos Analíticos de Políticas Educativas*, 1-27.
- Barros, R. E. (2012). El Debate en torno a la Ciencia y las Ciencias Sociales. *Estrategias pedagógicas para la solución de conflictos escolares*, 10(2), 57-61.
- Cabrero, B. G. (2010). Modelos teóricos e indicadores de evaluación educativa. *Revista Electrónica Sinéctica*(35), 1-21.
- Camilloni, A. R. (2009). Estándares, evaluación y currículo. *Archivos de Ciencias de la Educación*, III(3), 55-68.
- Castillo, M. A. (2009). La evaluación de los aprendizajes basada en el desempeño por competencias. *Revista Electrónica "Actualidades Investigativas en Educación*, 9(2), 1-25.
- Central Test. (2013). *El libro blanco de la evaluación psicométrica*. Obtenido de <https://www.centraltest.es/>
- Cerón, M. S., & Cruz, F. M. (2013). Las evaluaciones estandarizadas: sus efectos en tres países latinoamericanos. *Revista Latinoamericana de Estudios Educativos*, XLIII(1), 97-124.
- Consejo Internacional de Ciencias Sociales/UNESCO. (2015). *Informe Mundial sobre Ciencias Sociales 2013; Cambios ambientales globales*. Paris: UNESCO.
- Córdova, K. E., & Lozano, D. E. (2015). Las prácticas de evaluación del aprendizaje en relación con los estándares internacionales: un estudio exploratorio. *Innovación Educativa*, 15(68), 117-133.
- Cortés, M. E., & León, M. I. (2004). *Generalidades sobre Metodología de la Investigación*. Universidad Autónoma del Carmen, Campeche, México.
- Costa, P. A. (2017). La responsabilidad y los desafíos de las ciencias sociales. *Psicoperspectiva*, 16(1), 1-6.
- Covacevich, C. (2014). Cómo seleccionar un instrumento para evaluar aprendizajes. *Banco Interamericano de Desarrollo (BID)*, 40.
- Cueto, S. (2009). La evaluación estandarizada de la educación ciudadana en América Latina: Estado de la cuestión y algunas propuestas. *Programa de Promoción de la Reforma Educativa en América Latina y el Caribe*, 40.
- Delgado, J. R., & Palos, P. A. (2003). Desempeño académico y conductas de riesgo en adolescentes. *Revista de Educación y Desarrollo*, 5-16. Obtenido de http://www.cucs.udg.mx/revistas/edu_desarrollo/anteriores/7/007_Palacios.pdf
- Delgado, V. M., & Rivilla, A. M. (2014). Características de la enseñanza y la evaluación de los aprendizajes. *Revista Papeles*, 6(11), 20-35.

- Dirección Nacional de Información y Evaluación de la Calidad Educativa / Ministerio de Educación. (2009). *Hacia una cultura de evaluación*. Informativo, Buenos Aires, Argentina.
- EACEA. (2010). *Pruebas nacionales de evaluación del alumnado en Europa: objetivos, organización y utilización de los resultados*. España. doi:10.2797/3471
- Escorcía, J. (2008). ¿Que es evaluación educativa? *Instituto de estudios en educación* (7), 9.
- Escorza, Y. H., & Aradillas, A. L. (2013). *Teorías del Aprendizaje en el contexto Educativo*. México, México: Editorial Digital Tecnológico de Monterrey.
- FAREMA. (2017). *Educación: Una Deuda Pendiente. Informe de progreso Educativo de Honduras*.
- Fast, M., & Vukmirovic, Z. (2016). *Psicometría I*. Universidad Nacional Autónoma de Honduras, Ciencias Social, Escuelas Psicológicas. Tegucigalpa, Honduras: Maestría en Psicometría y evaluación Educativa.
- Ferrer, G. (2007). *Estudio comparado internacional sobre procesos de elaboración e implementación de estándares de curriculum en América Latina*. El Salvador: Grupo de Trabajo sobre Estándares y Evaluación PREAL.
- Gallegos, M. (2008). La nueva fase de políticas de la calidad educativa en Latinoamérica y el Caribe: medición y evaluación del rendimiento académico. *Revista Latinoamericana de Estudios Educativos*, XXXVIII(1-2,), 9-34.
- Gil, R. (2007). Teoría Andragógico-Integradora para la transformación universitaria. *FERMENTUM*, 17(48), 210-233.
- González, F. M. (2007). *Instrumento de Evaluación Psicológica*. La Habana: Ciencias Médicas.
- Guadrón, L. V., Matos, W. G., Castro, L. M., & Mendi, L. F. (2013). La evaluación de los aprendizajes en la educación universitaria en el marco del enfoque Andragógico. *Negotium*, 9(26), 86-115.
- Guevara, D., Hernández, K., & Cabrera, L. (2014). *Evaluación Institucional en educación Media Educativa*, Departamento de Administración y Gestión de la Educación, Tegucigalpa, Honduras.
- Guevara, R. S. (Mayo de 2017). Índice de la calidad de los instrumentos de evaluación. *Boletín Virtual*, VI, 1-8.
- Guzmán, J. C. (2010). La evaluación de los aprendizajes vista desde los profesores efectivos que enseñan psicología. *Revista Electrónica Sinéctica*(34), 1-15.
- Guzmán, M. P. (2012). *Modelos predictivos y explicativos del rendimiento académico universitario: caso de una institución privada en México*. Tesis doctoral, Universidad Complutense de Madrid, Departamento de Métodos de Investigación y Diagnóstico en Educación, Madrid.

- Hackmann, H., & Moser, S. (2013). *Informe mundial sobre ciencias sociales. Cambios ambientales globales*. doi:ISBN 978-92-3-104254-6
- Hernández, R. (2018). Mejora de la capacidad técnica nacional para alcanzar las Metas EFA bajo liderazgo fortalecido de la Secretaría de Educación . *MIDEH un proyecto que deja huella en el sistema educativo nacional* (pág. 11). Tegucigalpa, Honduras: Pueblo de lo Estado Unidos de América .
- Hidalgo-Montesinos, M. D., & French, B. F. (2016). Una introducción didáctica a la Teoría de Respuesta al Ítem para comprender la construcción de escalas. *Revista de Psicología Clínica con Niños y Adolescentes*, III(2), 13-21.
- Islas, A. M. (2010). Evaluación de la educación. *Congreso Iberoamericano de educación: Un congreso para que pensemos entre todos la educación que queremos* , (págs. 1-23). Buenos Aires, República de Argentina.
- Lamas, H. A. (2015). Sobre el rendimiento escolar. *Propósitos y Representaciones*, III(1), 313-386.
- Lansdale, J., Vukmirovic, Z., & Estrada, M. (2012). Estableciendo estándares educativos de desempeño en Honduras: El Primer paso en crear una cultura de evaluación y rendición de cuentas en educación. *Innovare*, 151-171.
- López, E. A. (2011). *Política fiscal y estrategia como factor de desarrollo de la mediana empresa comercial Sinaloense: Un estudio de caso*. Tesis doctoral, Culiacán, Sinaloa, México.
- López, J. E. (2008). *La reforma educativa en Honduras, el currículo nacional básico y la actitud de los docentes del primer ciclo de la educación básica desde la perspectiva de la Teoría de la acción razonada*. . Tesis de maestria , Tegucigalpa.
- López, K. M., & Lever, C. O. (2016). Las evaluaciones estandarizadas del aprendizaje y la mejora de la calidad educativa. *Temas de educación*, 22(I), 131-146.
- Maldonado, C. E. (2009). Complejidad de los Sistemas Sociales: Un reto para las ciencias sociales. *Cinta Moebio*(36), 146-157.
- Martin, R., Moore, J., & Schindler, S. (2016). Definiendo la desigualdad. *ARQ*(93), 30-41.
- Martínez, B. H. (2013). Causas del bajo rendimiento escolar origina un alto nivel de deserción escolar y habilidades para estudiar ayudan a mejorar el rendimiento escolar. *Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 1-17.
- Medina, A. M., García, M. A., Martínez, M. G., & Abundez, G. M. (2011). *Evaluación de los aprendizajes en el aula. Opiniones y prácticas de docentes de primaria en México*. informativo, México .

- Mejía, B. E. (2008). *Impacto de las reformas educativas en el movimiento magisterial hondureño*. Tesis de maestría, FLACSO. Programa Centroamericano de Postgrado, Facultad Latinoamericana de Ciencias Sociales, Guatemala.
- Meliá, J. M. (2017). Evaluación Estandarizada. *Revista Iberoamericana de Evaluación Educativa*, 10(1), 5-8. Obtenido de <https://revistas.uam.es/riee>
- Méndez, A. (2013). *Tutorías solidarias de niños, a niños, en las escuelas multigrados de la zona rural, como una estrategia metodológica para mejorar el rendimiento académico en el área de matemática, en el municipio de Yamaranguila, departamento de Intibucá*.
- Ministerio de Educación de España. (2010). *Pruebas nacionales de evaluación del alumnado en Europa: objetivos, organización y utilización de los resultados*. Eurydice, España.
- Ministerio de Educación de Guatemala/USAID. (2007). *Estándares Educativos para Guatemala/ Programa estándares e investigación educativa*. Guatemala.
- Ministerio de Educación de Nicaragua. (1999). *Estándares de contenido y desempeño educación primaria*. Nicaragua. Obtenido de <http://www.oei.es/estandares/nicaragua.htm>
- Molinare, S. C., Meriño, A. B., & Barro, M. R. (2014). Factores pedagógicos relacionados con el rendimiento académico en estudiantes de cinco instituciones educativas del distrito de Santa Marta, Colombia. *Revista Intercontinental de Psicología y Educación*, 16(2), 151-169.
- Moncada, G., & Abuslim, K. (2016). Uso de los datos de las evaluaciones de rendimiento académico en los centros educativos, desde la perspectiva de King & Peachman. En MIDEH, *Investigación educativa y política pública* (pág. 304). Tegucigalpa, Honduras.
- Morales, F. S., & Etchegaray, M. V. (2016). Funcionamiento Diferencial del Ítem en una Evaluación Estandarizada según Necesidades Educativas Especiales Transitorias. *Revista Iberoamericana de Evaluación Educativa*, 145-160.
- Moscoso, S. C., & Pérez-Gil, J. A. (2008). *Diseño y medición de programas de intervención neuropsicológica: aspectos fundamentales*. Tesis de magister, Universidad de Sevilla, España.
- Muñiz, J. (2010). Las teorías de los tests: Teoría clásica y teoría de respuesta a los ítems. *Papeles del Psicólogo*, 31(1), 57-66. Obtenido de <http://www.cop.es/papeles>
- Muñiz, J., & Fonseca, E. (2008). Construcción de instrumento de medida para la evaluación universitaria. *Revista de Investigación en Educación*(5), 13-25.
- Murillo, F. J., & Román, M. (2008). La evaluación educativa como derecho humano. *Revista Iberoamericana de Evaluación Educativa* 2008, 1(1), 5.
- Nieto, N. (2016). Los campos asesinos de la desigualdad. *Política y Cultura*(45), 273-277.

- Otzen , T., & Manterola , C. (2017). Técnicas de muestreo sobre una población a estudio. *Int. J. Morphol.*,35(1):227-232, 2017., 230.
- Paz, C. L. (2014). *Competencias Docentes para la Atención a la Diversidad: Investigación-acción en la Universidad Pedagógica Nacional Francisco Morazán Honduras*. Tesis doctoral , Universidad de Alicante , España .
- Pereira, J. (2015). Evaluación, medición o verificación de los aprendizajes en el aula: Un estudio de caso en el Colegio Humanístico Costarricense de Heredia. *Revista Electrónica Educare*, 19(2), 405-428.
- Prieto, G., & Delgado, A. R. (2003). Análisis de un test mediante el modelo de Rasch. *Psicothema*, 15(1), 94-100.
- Prieto, G., & Delgado, A. R. (2010). Fiabilidad y Validez . *Papeles del Psicólogo*, 31((1)), 67-74.
- Programa de Promoción de la Reforma Educativa en América Latina y el Caribe. (2002). *Informe de Progreso Educativo El Salvador*. El Salvador.
- Programa Estado de la Nación. (2013). *Estado de la educación Costarricense*. San José Costa Rica : Editorama S.A.
- Ramírez, D. B. (2010). El Conductismo: Orígenes, Trayectoria y Significado. *ResearchGate*, 73. Obtenido de <https://www.researchgate.net/publication/266261125>
- Rápalo, R. (2010). Estado de la Educación en Honduras. *UPNFM*, 24-25.
- Ravela , P., Arregui, P., & Wolfe, R. (2008). Las Evaluaciones Educativas que America Latina necesita. *Revista Iberoamericana de Evaluacion Educativa*, 46-63.
- Real Academia Española. (2019). *DLE*. Obtenido de DLE: <https://dle.rae.es/?id=ZVMWXKy>
- Reidl-Martínez., L. M. (2013). Confiabilidad en la medición. *Investigación en educación médica* , 11(6), 108-111.
- República de Honduras/Sector Educativo. (2004). *Articulación de las Intervenciones en Educación en el marco del Plan de Gobierno 2002 – 2006 y de la Estrategia para la Reducción de la Pobreza*. Tegucigalpa.
- Reyes, R. (2016). Expresión Escrita: Desempeño de los estudiantes con relación a la metodología utilizada por los docentes. En MIDEH, *Proyecto Mejorando el impacto al desempeño estudiantil de Honduras (MIDEH)* (págs. 43-73).
- Rivera, G. P. (2007). La evaluación de los aprendizajes. *Reencuentro*, 48, 20-26.
- Sampieri, R. H., Collado, C. F., & Lucio, M. d. (2014). *Metodología de la Investigación*. México: McGraw-Hill/Interamericana editores,S.A. de C.V.

- Sánchez, M. R. (2004). *Introducción a la Teoría de respuesta al Item, una herramienta para el análisis de variables latentes: aplicada a la medición de la calidad de vida de la infancia*. Universidad de Extremadura, Departamento de Economía Aplicada y Organización de Empresas.
- Schunk, D. H. (2012). *Teorías del aprendizaje. Una perspectiva educativa*. Naucalpan de Juárez, Estado de México, México: Pearson Educación de México, S.A. de C.V.
- SE/MIDEH. (2016). *Informe Nacional de Desempeño Académico*. Tegucigalpa: Secretaría de Educación.
- SE/MIDEH. (2016). *Informe Nacional de Desempeño Académico 2016 Español y Matemáticas. 1ro a 9no grado*. Tegucigalpa, Honduras : Secretaría de Educación .
- SE/USINIEH. (2015). *SISTEMA EDUCATIVO HONDUREÑO EN CIFRAS PERIODO ACADNÉMICO 2015*. Tegucigalpa.
- Secretaría de Educación . (2019). *Informe Nacional de Desempeño Académico 2018 Español (Lectura) y Matemáticas. 3ro, 6to y 9no grado*. Tegucigalpa, Honduras .
- Secretaría de Educación. (2003). *Currículo Nacional Básico*. Tegucigalpa, Honduras.
- Secretaría de Educación. (2004). *Currículo Nacional Básico sintetizado*. Tegucigalpa, Honduras.
- Secretaría de Educación. (2009). *Informe nacional de evaluación de los aprendizajes educación básica: Primero y Segundo ciclo*. Tegucigalpa, Honduras.
- Secretaría de Educación. (2011). *Estándares educativos nacionales*. Tegucigalpa, Honduras.
- Secretaria de Educación. (2016). *Sistema Educativo Hondureño en Cifras Período Académico 2015*. Francisco Morazán , Tegucigalpa, Honduras. Recuperado el 5 de Abril de 2017
- Secretaría de Educación. (2016). *Sistema Educativo Hondureño en Cifras, Período Académico 2015*. Estadístico.
- Secretaria de Educación. (2017). *Informe sistema educativo Hondureño en cifras periodo académico 2014-2016. PERIODO ACADÉMICO* , Tegucigalpa.
- Secretaría de Educación y Proyecto Mejorando el Impacto al Desempeño Estudiantil de Honduras. (2017). *Informe nacional de desempeño académico 2016 Español y Matemática 1ro a 9no grado*. Tegucigalpa. Honduras.
- Secretaría del Convenio sobre la Diversidad Biológica. (2010). *La Perspectiva Mundial sobre la Diversidad Biológica 3*. Montreal. Obtenido de <https://www.cbd.int/doc/publications/gbo/gbo3-final-es.pdf>

- SE-MIDEH. (2017). *Informe nacional de desempeño académico 2016 Español y Matemática 1ro a 9no grado*. Tegucigalpa. Honduras.
- SE-MIDEH. (2017). *Informe Nacional de Desempeño Académico 2017 Español y Matemáticas. 1ro a 9no grado*. Tegucigalpa: Secretaría de Educación.
- Sistema Nacional de Evaluación de la Calidad de los Aprendizajes. (2007). *Construcción de pruebas basado en estándares*. Panamá.
- Terrón, A. M. (2010). Introducción al análisis de la Teoría de Respuesta al Ítem . *Aidesoc.net*, 17.
- Toranzos, L. V. (2014). Evaluación educativa: hacia la construcción de un espacio de aprendizaje. *Propuesta Educativa*, 1-13.
- UNESCO. (2015). *Educación para la ciudadanía mundial, temas y objetivos de aprendizaje*. Francia.
- UNESCO y Foro Consultivo. (2011). *Informe sobre las ciencias sociales en el mundo: Las brechas del conocimiento*. Expositivo, México.
- UNICEF. (2011). *La Infancia en Honduras, Análisis de Situación 2010*. Tegucigalpa, Honduras. Obtenido de [http:// www.unicef.org/honduras](http://www.unicef.org/honduras)
- Valcárcel, N. M., Cabellos, A. d., & Avilés, R. H. (2012). La evaluación del aprendizaje en entornos virtuales de enseñanza aprendizaje: notas para una reflexión . *Revista Iberoamericana de Educación / Revista Ibero-americana de Educação*, 1-16.
- Valladares, D. F. (2018). *Calidad y Política Educativa para el Desarrollo Humano en Honduras*. Tesis Doctoral , Tegucigalpa, Honduras .
- Vergara, C. F., & Quiles, J. O. (2016). *Diseño y construcción de pruebas*. Oficina de Evaluación del Aprendizaje Estudiantil.
- Zapata-Ros, M. (23 de Marzo de 2015). Teorías y modelos sobre el aprendizaje en entornos conectados y ubicuos. Bases para un nuevo modelo teórico a partir de una visión crítica del “conectivismo”. *Education in the Knowledge Society*, 16(1), 69-102.

Declaración de Ética de la Investigación

La presente investigación titulada “Diseño de una prueba de desempeño académico en el área de Ciencia Sociales para los estudiantes de noveno grado en Honduras 2017-2018” ha sido diseñada, planificada y aplicada a estudiantes que cursaban el noveno grado de los centros educativos correspondientes en el Distrito Central de la Ciudad de Tegucigalpa.

De acuerdo a lo anterior, yo, Dulce Cristina Guevara Castillo, Licenciada en Administración y Gestión de la Educación y estudiantes de la Maestría en Psicometría y Evaluación Educativa, con número de cuenta **PSE100113**, declaro que he desarrollado esta investigación siguiendo las instrucciones proporcionada por la coordinación de la maestría y mi asesor de tesis.

De tal manera, la información presentada en este documento es el resultado de trabajo personal, teniendo en cuenta la legislación sobre propiedad intelectual, sin haber incurrido en la falsificación de la información o a cualquier fraude

Dulce Cristina Guevara Castillo

Nombre y Firma



Anexos

Anexo 1: Tabla de especificaciones para el CCSS

Bloque/ Área clave	Componente	Contenido a evaluar		Niveles			Total	A	B	C	Total
		Estándar/	Habilidad o información precisa que se quiere evaluar	Conocimiento	Comprensión	Aplicación					
1. La Persona y su ser social.	Convivencia armónica		Conocen los antecedentes que determinan la razón de ser de los Derechos Humanos, identificando el carácter utilitario de sus regulaciones.	2	2	2	6	10	10	10	30
2. La Sociedad y los espacios geográficos	Espacio Geográfico		Reconocen la importancia actual del estudio de la Geografía Económica y Política, analizando la condición internacional de Honduras en el orden económico, político, diplomático y social.	2	2	2	6	10	10	10	30
			Conocen y manejan la división política de cada continente, relacionando las características físico- geográfico y la distribución de la población a nivel mundial.	2	2	2	6	10	10	10	30
3. Las sociedades y el tiempo social	Tiempo histórico		Establecen relaciones entre las diferentes formas de periodizar la Historia Mundial, identificando los aportes de las civilizaciones en el desarrollo de la humanidad.	1	2	2	5	5	10	10	25
			Comprenden las grandes revoluciones sociales, identificando causa y consecuencias en el mundo actual.	2	2	2	6	10	10	10	30
			Comprenden e identifican los diferentes conflictos mundiales, por medio de conocimiento de sus origen, antecedentes y lecciones aprendidas.	1	1	2	4	5	5	10	20
4. La sociedad y las actividades humanas	Acción social del ser humano		Comprenden las diferentes acciones y actividades económicas que realiza la población mundial, por medio de la identificación y caracterización de los sectores de la economía.	1	2	2	5	5	10	10	25
			Comprenden los indicadores de Desarrollo Humano (IDH), analizando comparativamente la situación en Honduras y estableciendo relaciones con las políticas y sociales actuales.	2	3	2	7	10	15	10	35

Fuente: Elaboración propia (2018) adaptado del Currículo Nacional Básico (2004)



UNAH
UNIVERSIDAD NACIONAL
AUTÓNOMA DE HONDURAS

UNIVERSIDAD NACIONAL AUTONOMA DE HONDURAS
FACULTAD DE CIENCIAS SOCIALES
ESCUELA DE CIENCIAS PSICOLÓGICAS
MAESTRIA EN PSICOMETRIA Y EVALUACION EDUCATIVA
SECRETARIA DE EDUCACION

Anexo 2: Guión metodológico para el taller de evaluación educativa y redacción de ítems

GUIÓN METODOLÓGICO

Nombre del Taller	Taller de Evaluación Educativa y Redacción de Ítems	Duración de la actividad	3 días	Número de Participantes
				101
Objetivos Generales	1. Validar a través de la revisión de especialistas las tablas de especificaciones de pruebas en las áreas del conocimiento de español, matemáticas, ciencias naturales y sociales. 2. Contribuir en la formación de un banco de reactivos de opción múltiple para la construcción de pruebas estandarizadas del nivel básico y medio.			

Día	Objetivo	Contenido	Actividad	Recursos	Responsables	Tiempo		
						Hora de inicio	Hora final	Duración
Miércoles 2/05/18	Explicar la importancia que tiene la evaluación como herramienta en el proceso de enseñanza-aprendizaje	La evaluación para el aprendizaje: una herramienta para la mejora educativa.	Exponer sobre la evaluación	Data show, computador Gafete Carpeta	Nanci Castro Dulce Guevara	08:30 am	9:30 am	1 hora
	Conocer el Marco Normativo para la evaluación de los aprendizajes en Honduras Acuerdo 1796 –SE-2017.	Marco Normativo para la evaluación de los aprendizajes en Honduras Acuerdo 1796 –SE-2017.	Explicar el Acuerdo 1796 – SE-2017 Espacio de preguntas y respuestas	Data show, computador Acuerdo 1796 –SE-2017.	MSc. José Luis Cabrera Director General de Currículo y Evaluación	09:30am	10:20 am	40 minutos 10 minutos
	Receso					10:20 am	10:40 am	20 minutos
	Describir los pasos para la construcción de una prueba que mida los conocimientos en las áreas curriculares de español, Matemáticas, Ciencias Naturales y Ciencias Sociales. A través del marco conceptual y operativo.	Diseño de una prueba de rendimiento nacional: Marco conceptual y operativo para la construcción de una prueba.	Presentar los marcos conceptual y operativo para la construcción de una prueba.	Data show, computador Papel bond Lápiz, borrador	Darwin Dávila Carlos Aguilar Isabel Hernández	10:40 am	12:20 am	1 hora, 40 minutos
	Almuerzo					12:20 pm	01:20p m	1 hora
	Explicar la tabla de especificaciones de las áreas de español, Matemáticas, Ciencias	Tabla de especificacion es Por área.	Revisión en equipos de la Tabla de Especificaciones por área.	Data show, computador Tabla de especificacio	Integrantes Línea Investigación por área	01:20pm	02:30p m	1 hora, 10 minutos

Naturales y Ciencias Sociales.	nes por área.							
Validar la tabla de especificaciones de los estándares educativos a través de juicios de expertos en las áreas del conocimiento de español, Matemáticas, Ciencias Naturales y Ciencias Sociales.	Tabla de especificaciones.	Discutir la tabla de especificaciones con los especialistas en las materias para la aprobación.	Data show, computador	Integrantes Línea Investigación por área	02:30pm	03:30 pm	1 hora	
Retroalimentación las actividades desarrolladas durante el proceso para consolidar criterios.	Conclusiones, acuerdos y cierre de la jornada				03:30pm	4:00 pm	30 minutos	

Día	Objetivo	Contenido	Actividad	Recursos	Responsables	Tiempo		
						Hora	Hora	Duración
						inicio	finalización	
	Establecer los criterios de redacción de ítems para la construcción de una prueba	Criterios para la redacción de ítems	Exponer los criterios de redacción de ítems para el diseño de las pruebas.	Data show, computador Manual de redacción de ítems	Cindy Arauz Kesly García	08:00 am	9:20 am	1 hora, 20 minutos
	Elaborar ítems según los criterios establecidos para la construcción de la prueba	Practica 1 y Discusión Redacción de ítems (Elaboración de 3 ítems)	Redactar y discusión ítems.	Data show, computador Plantillas de ítems Libros de texto Tabla de especificaciones	Integrantes Línea Investigación por área y docentes.	9:20 am	10:00am	40 minutos

Jueves 3/05/18	Receso				10:00 am	10:20 am	20 min.	
	Explicar la metodología para la redacción de ítems en las diferentes áreas.	Metodología para la redacción de ítems	Guiar el procedimiento para la redacción de ítems	Data show, computador Plantillas de ítems	Nanci Castro Damary Sosa	10:20 am	10:40 am	20 minutos
	Crear 10 ítems de selección única para cada área curricular.	Construcción de 10 ítems, por área curricular.	Formular 10 ítems en pares por área y retroalimentar en grupo	Data show, computador Plantillas de ítems Libros de texto Tabla de especificaciones	Docentes por área curricular	10:40am	11:40 am	1 hora
	Revisar los ítems para la unificación criterios entre los expertos de cada área.	Desarrollo de ítems	Discutir que ítems se acepta, se rechaza o se mejora	Data show, computador Plantillas de ítems Libros de texto Tabla de especificaciones	Docentes por área curricular	11:40 am	12:20 am	40 min.
	Almuerzo				12:20	1:20	1 hora	
	Construir 10 ítems de selección única para cada área curricular.	Redacción de 10 ítems por área curricular	Formular 10 ítems en pares por área y retroalimentar en grupo	Data show, computador Plantillas de ítems Libros de texto Tabla de especificaciones	Docentes por área curricular e integrantes de la línea de investigación	01:20 pm	02:30 pm	1 hora
Receso				02:30 pm	02:50 pm	20 minutos		

	Construir 10 ítems de selección única por área curricular.	Redacción de 10 ítems individual y retroalimentación	Formular 10 ítems individualmente por área y dar retroalimentación	Data show, computador Plantillas de ítems Libros de texto Tabla de especificaciones	Docentes por área curricular e integrantes de la línea de investigación	02:50 pm	04:00 pm	1 hora 10 minutos
Día	Objetivo	Contenido	Actividad	Recursos	Responsables	Tiempo		
						Hora inicio	Hora finalización	Duración
		Retroalimentación			Darwin Dávila	08.00 am	08:20 am	20 min.
	Construir 10 ítems de selección única por área curricular.	Redacción de 10 ítems individual y retroalimentación	Formular 10 ítems individualmente por área y dar retroalimentación	Data show, computador Plantillas de ítems Libros de texto Tabla de especificaciones	Docentes por área curricular e integrantes de la línea de investigación	08:20 am	9:20 am	1 hora
Viernes			Receso			10:20 am	10:40 am	20 min.
4/05/18	Construir 10 ítems de selección única por área curricular.	Redacción de 10 ítems en pares y retroalimentación	Formular 10 ítems individualmente por área y dar retroalimentación	Data show, computador Plantillas de ítems Libros de texto Tabla de especificaciones	Docentes por área curricular e integrantes de la línea de investigación	10:40 am	12:20 pm	1 hora 40 minutos
			Almuerzo			12:20 pm	01:20 pm	1 hora

Revisar la actividad de creación de ítems para la retroalimentación de del taller.	Revisión y retroalimentación final en grupo	Monitorear el proceso de redacción de ítems	Data show, computado Plantillas de ítems Libros de texto Tabla de especificaciones	Docentes por área curricular e integrantes de la línea de investigación	01:20 pm	03:30 pm	1 hora y 10 minutos
Agradecer a los expertos por área curricular y a los invitados especiales por su valiosa participación en el taller de redacción de ítems para la evaluación educativa	Cierre del Evento Palabras de Agradecimiento Entrega de Diplomas	Brindar palabras de Agradecimiento y Entrega de Diplomas	Data show, computador Plantillas de ítems Libros de texto Tabla de especificaciones		03:30 pm	04:00 pm	30 minutos

Fuente: Adaptado del “Taller de Evaluación Educativa y Redacción de Ítems” (2018).



Anexo 3: Consentimiento informado para el director

HOJA DE CONSENTIMIENTO INFORMADO PERSONA ENCARGADA DEL/LA JOVEN

Los(as) estudiantes del centro educativo que usted acertadamente dirige han sido seleccionado(a) para participar en un estudio de investigación dirigido por la Lic. Dulce Cristina Guevara Castillo, de la Universidad Nacional Autónoma de Honduras (UNAH), y la Maestría en Psicometría y Evaluación Educativa.

1. ACLARACIÓN

La participación de los estudiantes es de carácter **VOLUNTARIO**, si no está de acuerdo con los términos que a continuación se le presentan, si se siente incómodo(a) con los procedimientos que se le proponen en el estudio o si se siente incapaz de contribuir a esta investigación no está obligado(a) a participar. En todo momento usted o los estudiantes están en la facultad de retirarse de la investigación si así lo desearan.

2. PROPÓSITOS DEL ESTUDIO

El propósito de este estudio es investigar el desempeño académico con que cuentan los estudiantes de noveno grado en el área curricular de Ciencias Sociales en Honduras, por medio de la evaluación con una evaluación de cuestionario educativo que incluye 50 preguntas de el área mencionada

3. PROCEDIMIENTOS

Al aceptar participar en este estudio, le pediremos a los estudiantes que colabore realizando las siguientes actividades.

- a. Escuchar las instrucciones brindadas por el aplicador de la prueba. b) Responder los 50 ítems de Ciencias Sociales, c) Durante la aplicación guardar silencio.
- b. La prueba de desempeño académico será realizada por un educador capacitado para esta tarea quien explicará de forma detallada y precisa cada actividad a realizar.
- c. Este estudio requiere de un tiempo de 2 horas y consiste en realizar una prueba con el objetivo de evaluar las habilidades cognitivas de los estudiantes.
- d. Los resultados de esta prueba y el diagnóstico serán guardados y a éstos sólo tendrán acceso la persona responsable del estudio en cuestión.

4. RIESGOS POTENCIALES

Las evaluaciones implicadas pueden generar cierta frustración por la dificultad de alcanzar un desempeño esperado por los estudiantes. Sin embargo, el investigador está capacitado para reducir mediante técnicas de relajación la ansiedad que se le pueda causar.



5. BENEFICIOS PREVISTOS

a. Para los participantes

Los estudiantes NO obtendrán beneficios monetarios ni pago alguno por participar en este estudio.

b. Para la sociedad

Los beneficios que promete este estudio van encaminados a mejorar la calidad de la educación en el área de Ciencias Sociales de los niños y adolescentes de Honduras. Los resultados serán la base para propuestas de intervención y podrán ser utilizados por los tomadores de decisiones para proponer políticas educativas y nuevas líneas de investigación.

6. RESPONSABILIDADES DEL INVESTIGADOR(A)

a. Obligaciones financieras

Los procedimientos para este estudio, será hecho sin costos por los estudiantes. Esto quiere decir que todo gasto que se relacione con la realización de esta investigación corre por cuenta del investigador.

b. Identificación de los investigadores

En caso de necesitar información sobre el estudio, por favor comunicarse al correo guevaradulcecastillo@gmail.com

7. PRIVACIDAD Y CONFIDENCIALIDAD

Las únicas personas que sabrán que los estudiantes son participantes de este estudio son las autoridades de la maestría y del investigador. Ninguna información acerca de los estudiantes proporcionada por usted durante la investigación, será revelada a otros sin su permiso escrito.

Cuando los resultados de la investigación sean publicados o discutidos en conferencias, la información que se proporcione no revelará la identidad de los estudiantes. Si se utilizan fotografías, vídeos o grabaciones de su caso, las mismas serán usados para propósitos de educación y su identidad será protegida.

8. DERECHOS DE LOS SUJETOS DE INVESTIGACIÓN

a. Retiro de la investigación

Los estudiantes pueden retirar su consentimiento en cualquier momento y suspender su participación en esta investigación sin sanción alguna. Si tiene alguna pregunta acerca de sus derechos como sujeto de investigación, puede contactarse por medio del correo guevaradulcecastillo@gmail.com

b. Información sobre su muestra

Al final de este documento se le preguntará a los estudiantes si desea recibir información general acerca de lo que se encontró en este estudio (o las conclusiones del estudio), tanto él como Ud. puede decidir si desea información sobre los resultados obtenidos por esta investigación.



UNAH
UNIVERSIDAD NACIONAL
AUTÓNOMA DE HONDURAS



Procesamiento Técnico Documental Digital.UDI-DEGT-UNAH
**MAESTRÍA EN
PSICOMETRÍA Y
EVALUACIÓN EDUCATIVA**

cienciassociales.unah.edu.hn
mpse@unah.edu.hn
Tel: (504) 2216-6100 Ext. 100470

FIRMA DE LOS PADRES O REPRESENTANTE LEGAL

He leído (o alguien me ha leído) la información anteriormente mencionada. Se me ha dado la oportunidad de hacer todas mis preguntas y a todas ellas se les ha dado respuesta y las he comprendido. Me comprometo a ser puntual y cumplir con todas las indicaciones que me sean dadas en este estudio. Se me ha dado una copia de este consentimiento.

FIRMANDO ESTE CONSENTIMIENTO INFORMADO, **ACEPTO** TODAS LAS CONDICIONES Y RESPONSABILIDADES QUE CONLLEVA PARTICIPAR VOLUNTARIAMENTE EN ESTA INVESTIGACION.

Nombre completo: _____

Firma del representante legal (director del Centro Educativo): _____

Fecha ____/____/____

FIRMA DEL INVESTIGADOR

He dado explicación al representante legal o director del centro educativo y he resuelto todas sus dudas. Doy constancia que estos han entendido la información descrita en este documento y aceptan participar de forma voluntaria.

Nombre del investigador: _____

Firma del investigador: _____

Fecha: ____/____/____ (debe ser la misma que la del participante).

Anexo 4: Conocimiento informado para participantes en la investigación

CONSENTIMIENTO INFORMADO PARA PARTICIPANTES EN LA INVESTIGACIÓN:

Diseño de una prueba de desempeño académico en el área de Ciencia Sociales para los estudiantes de noveno grado en Honduras 2017-2018

El propósito de este documento llamado **consentimiento informado** es explicar claramente a los participantes de esta investigación acerca de cuál será su papel en este proceso, advertir en qué consiste el estudio y obtener el consentimiento para incluirlo como participante.

La presente investigación está dirigida por la Licenciada en Administración y Gestión en educación Dulce Cristina Guevara Castillo, de la Universidad Nacional Autónoma de Honduras, UNAH. El propósito del estudio es diseñar una prueba de rendimiento académico, que permita medir las habilidades cognitivas con que cuentan los estudiantes de Educación Básica en el área de ciencias sociales.

Lo(a) hemos seleccionado para participar en este estudio. Si está de acuerdo en ser parte de esta investigación será tratado por especialista en psicología y en educación y se le pedirá participar en las siguientes actividades: a) Escuchar las instrucciones brindadas por el aplicador de la prueba. b) Responder los 50 ítems de Ciencias Sociales, c) Durante la aplicación guardar silencio, d) la Prueba tiene una duración de 2 horas.

Lo que usted manifieste durante esta actividad académica podría ser fotografiado, grabado en audio o video, con el único objetivo de respaldar posteriormente la información, pero una vez hecho el estudio, si hubiere grabaciones se destruirán

Su participación es totalmente voluntaria y la información recolectada será confidencial, no se usará para ningún otro fin que no sea esta investigación, sus respuestas serán agrupadas usando una codificación, de modo que lo que usted conteste, será completamente anónimo.

Esta investigación no representa ningún riesgo ético, moral, emocional o físico, sin embargo, debe considerar que se puedes sentir un poco ansioso(a), así que es posible que sientas algo de fatiga al finalizar. Sin embargo, el investigador está capacitado para reducir mediante técnicas de relajación la ansiedad que se te pueda causar.

Los participantes no recibirán ningún pago económico o material, ni ningún beneficio adicional. Si tiene alguna duda puede hacer las consultas pertinentes en cualquier momento durante su participación en el proceso.

Gracias por participar.

Anexo 5: Asentimiento informado

ASENTIMIENTO INFORMADO

Manifiesto estar de acuerdo en participar en la investigación llamada: **Diseño de una prueba de desempeño académico en el área de Ciencias Sociales para los estudiantes de noveno grado en Honduras 2017-2018** dirigida por la Licenciada en Administración y Gestión en educación Dulce Cristina Guevara Castillo, y cuyo propósito del estudio es diseñar una prueba de rendimiento académico, que permita medir las habilidades cognitivas con que cuentan los estudiantes de Educación Básica (Media) en el grado de noveno en el área curricular de Ciencias Sociales.

Indico que se me ha informado que mi participación en este proceso consiste en responder a un cuestionario que contiene 50 ítems de Ciencias Sociales y ello tendrá una duración aproximada de 2:00 hora. Todo cuanto haga en esta participación no conlleva ningún riesgo ético, moral, emocional o físico para mi persona.

Entiendo que la información que aporte para este estudio es confidencial, que no se utilizará para ningún otro propósito. Se me ha informado de mi derecho de hacer las consultas pertinentes, así como de retirarme o de no responder si considero que lo que se me plantea puede perjudicarme en alguna forma.

He sido informado de que no recibiré retribución económica o material, ni ningún beneficio adicional por esta participación.

Se me ha informado que una copia de este documento que yo firme me será entregada y que al finalizar el estudio puedo preguntar sobre los resultados del mismo. Ante cualquier consulta, comprendo que puedo contactar a la Licenciada Guevara por el siguiente medio [guevaradulcecastillo @g mail.com](mailto:guevaradulcecastillo@gmail.com)

Nombre del participante

Firma del participante

Lugar y Fecha

Anexo 6: Descriptores para el área de ciencias sociales en noveno grado

Bloques	Niveles de desempeño		
	Debe mejorar	Satisfactorio	Avanzado
	Los estudiantes en este nivel evidencian dominio mínimo de los contenidos y capacidades cognitivas establecidas en el CNB; requieren constante supervisión, muestran poca habilidad de comunicación tanto de manera explícita como implícita, con poca capacidad de interpretar información para solucionar problemas elementales	Los estudiantes en este nivel evidencian el logro de los aprendizajes demostrando dominio suficiente de los contenidos y capacidades cognitivas, establecidas en el CNB; trabajan de forma independiente con poca supervisión, tienen habilidades para comunicar ideas de forma clara y estructurada y sobre contenido no solamente explícito sino también implícito; interpretando y analizando información para solucionar problemas cognitivamente complejos.	Los estudiantes en este nivel, demuestran un desempeño destacado de los contenidos y capacidades cognitivas, establecidas en el CNB; trabajan de forma independiente sin requerir supervisión y poseen altas habilidades para comunicar ideas claras; interpretando y analizando información para generar soluciones y toma de decisiones ante situaciones complejas
1- La persona y su ser social	Los estudiantes en este nivel muestran poca capacidad para explicar los antecedentes que determinan la razón de ser de los derechos humanos y su regulación, se les dificulta comparar e identificar las normativas y las organizaciones que aplican los derechos humanos en el país y a nivel mundial.	Los estudiantes en este nivel manifiestan la capacidad de explicar los antecedentes que determinan la razón de ser de los derechos humanos y sus regulaciones, además puede comparar e identificar las normativas y las organizaciones que aplican los derechos humanos en el país y a nivel mundial.	Los estudiantes en este nivel poseen una capacidad superior para explicar los antecedentes que determinan la razón de ser de los derechos humanos y sus regulaciones, comparan e identifican las normativas y las organizaciones que aplican los derechos humanos en el país y a nivel mundial.

2- La sociedad y los espacios geográfico	Los estudiantes en este nivel identifican y describen con poca capacidad la importancia de la ubicación geográfica, las características de la geografía, las vías de comunicación y transporte, así como los recursos económicos de Honduras y en el mundo, además se les dificulta establecer relaciones entre las características geográficas de los continentes y sus recursos naturales no poseen la capacidad de comparar y analizar las formas de prevención de riesgo y desastres en otros continentes.	Los estudiantes en este nivel son capaces de identificar y describir la importancia de la ubicación geográfica, las características de la geografía, las vías de comunicación y transporte, así como los recursos económicos de Honduras y en el mundo, además tienen la capacidad de establecer y relacionar las características geográficas de los continentes y sus recursos naturales, así como las diferentes formas de prevenir los riesgos y desastres en otros continentes	Los estudiantes en este nivel manifiestan una capacidad superior de identificar y describir la importancia de la ubicación geográfica, las características de la geografía, las vías de comunicación y transporte, así como los recursos económicos de Honduras y en el mundo, consiguen establecer y relacionar las características geográficas de los continentes y sus recursos naturales, así como las diferentes formas de prevenir los riesgos y desastres en otros continentes
3. Las sociedades y el tiempo social	Los estudiantes en este nivel muestran poca capacidad para comprender las diferentes formas de periodizar la Historia, así como las causas y la consecuencia de las grandes revoluciones teniendo en cuenta los errores y lecciones que se obtuvieron de estas mismas.	Los estudiantes en este nivel poseen la capacidad de analizar las diferentes formas de periodizar la Historia, así como las causas y la consecuencia de las grandes revoluciones teniendo en cuenta los errores y lecciones que se obtuvieron de estas revoluciones.	Los estudiantes en este nivel demuestran una capacidad superior de analizar las diferentes formas de periodizar la Historia, así como las causas y la consecuencia de las grandes revoluciones teniendo en cuenta los errores y lecciones que se obtuvieron de estas revoluciones.

4- La sociedad y las actividades humanas	Los estudiantes en este nivel muestran poca capacidad para analizar las diferentes organizaciones sociales, políticas y económicas del mundo, a la vez se les dificulta reconocer el trabajo del Índice de Desarrollo Humano comparándolo con la situación de nuestro país.	Los estudiantes en este nivel cuentan con la capacidad de analizar, comprender y comparar las diferentes organizaciones sociales, políticas y económicas del mundo, así como reconocer el trabajo del Índice de Desarrollo Humano comparándolo con la situación de nuestro país	Los estudiantes en este nivel muestran una capacidad superior para de analizar, comprender y comparar las diferentes organizaciones sociales, políticas y económicas del mundo, así como reconocer el trabajo del Índice de Desarrollo Humano comparándolo con la situación de nuestro país.
---	---	---	--

Fuente. Elaboración propia

Anexo 7: Instrumentos de recolección de información


UNAH UNIVERSIDAD NACIONAL AUTÓNOMA DE HONDURAS
 MAESTRÍA EN PSICOMETRÍA Y EVALUACIÓN EDUCATIVA
 SECRETARÍA DE EDUCACIÓN

PRUEBA DE CIENCIAS SOCIALES

PILOTAJE




2018

Nombre del Centro Educativo: _____
 Nombre del estudiante: _____
 Edad: _____ Género: F ☐ M ☐
 Número de lista: _____ Sección _____
 Fecha _____

CODIGO
 FORMA A

9^{TO}
GRADO


UNAH UNIVERSIDAD NACIONAL AUTÓNOMA DE HONDURAS
 MAESTRÍA EN PSICOMETRÍA Y EVALUACIÓN EDUCATIVA
 SECRETARÍA DE EDUCACIÓN

PRUEBA DE CIENCIAS SOCIALES

PILOTAJE




2018

Nombre del Centro Educativo: _____
 Nombre del estudiante: _____
 Edad: _____ Género: F ☐ M ☐
 Número de lista: _____ Sección _____
 Fecha _____

CODIGO
 FORMA B

9^{TO}
GRADO

Anexo 8: Prueba operativa

UNAH
UNIVERSIDAD NACIONAL AUTÓNOMA DE HONDURAS
MAESTRÍA EN PSICOMETRÍA Y EVALUACIÓN EDUCATIVA

PRUEBA DE CIENCIAS SOCIALES

Nombre del Centro Educativo: _____
 Nombre del estudiante: _____
 Edad: _____ Género: F ☐ M ☐
 Número de lista: _____ Sección: _____ Jornada: _____
 Fecha: _____

PRUEBA OPERATIVA

9^{NO}
GRADO

CÓDIGO

Anexo 9: Manual de instrucciones para la aplicación de la prueba piloto y operativa



Universidad Nacional Autónoma de Honduras
Escuela de Ciencias Psicológicas
Maestría en Psicometría y Evaluación Educativa



**MANUAL DE INSTRUCCIONES PARA LA APLICACIÓN
DE LA PRUEBA DE CIENCIAS SOCIALES 9no GRADO
PILOTAJE**

2018



Universidad Nacional Autónoma de Honduras
Escuela de Ciencias Psicológicas
Maestría en Psicometría y Evaluación Educativa



**MANUAL DE INSTRUCCIONES PARA LA APLICACIÓN
DE LA PRUEBA DE CIENCIAS SOCIALES 9no GRADO
PRUEBA OPOERATIVA**



2018

Anexo 10: Escala de puntuación por centro educativo

Ítems	Centro Educativo	Puntaje directo	Escala
398	Instituto Gubernamental Héctor Pineda Ugarte	36	AVANZADO
397	Instituto Gubernamental Héctor Pineda Ugarte	34	AVANZADO
396	Instituto Gubernamental Héctor Pineda Ugarte	30	SASTIFACTORIO
395	Instituto Gubernamental Héctor Pineda Ugarte	28	SASTIFACTORIO
394	Instituto Gubernamental Héctor Pineda Ugarte	38	AVANZADO
393	Instituto Gubernamental Héctor Pineda Ugarte	31	SASTIFACTORIO
392	Instituto Gubernamental Héctor Pineda Ugarte	35	AVANZADO
391	Instituto Gubernamental Héctor Pineda Ugarte	25	SASTIFACTORIO
390	Instituto Gubernamental Héctor Pineda Ugarte	27	SASTIFACTORIO
389	Instituto Gubernamental Héctor Pineda Ugarte	14	DEBE MEJORAR
388	Instituto Gubernamental Héctor Pineda Ugarte	35	AVANZADO
387	Instituto Gubernamental Héctor Pineda Ugarte	12	DEBE MEJORAR
386	Instituto Gubernamental Héctor Pineda Ugarte	35	AVANZADO
385	Instituto Gubernamental Héctor Pineda Ugarte	21	DEBE MEJORAR
384	Instituto Gubernamental Héctor Pineda Ugarte	28	SASTIFACTORIO
383	Instituto Gubernamental Héctor Pineda Ugarte	34	AVANZADO
382	Instituto Gubernamental Héctor Pineda Ugarte	37	AVANZADO
381	Instituto Gubernamental Héctor Pineda Ugarte	34	AVANZADO
380	Instituto Gubernamental Héctor Pineda Ugarte	19	DEBE MEJORAR
379	Instituto Gubernamental Héctor Pineda Ugarte	17	DEBE MEJORAR
378	Instituto Gubernamental Héctor Pineda Ugarte	29	SASTIFACTORIO
377	Instituto Gubernamental Héctor Pineda Ugarte	20	DEBE MEJORAR
376	Instituto Gubernamental Héctor Pineda Ugarte	37	AVANZADO
375	Instituto Gubernamental Héctor Pineda Ugarte	32	SASTIFACTORIO
374	Instituto Gubernamental Héctor Pineda Ugarte	31	SASTIFACTORIO
373	Instituto Gubernamental Héctor Pineda Ugarte	36	AVANZADO
372	Instituto Gubernamental Héctor Pineda Ugarte	21	DEBE MEJORAR
371	Instituto Gubernamental Héctor Pineda Ugarte	25	SASTIFACTORIO
370	Instituto Gubernamental Héctor Pineda Ugarte	18	DEBE MEJORAR
369	Instituto Gubernamental Héctor Pineda Ugarte	22	DEBE MEJORAR
368	Instituto Gubernamental Héctor Pineda Ugarte	30	SASTIFACTORIO
367	Instituto Gubernamental Héctor Pineda Ugarte	30	SASTIFACTORIO
366	Instituto Gubernamental Héctor Pineda Ugarte	36	AVANZADO
365	Instituto Gubernamental Héctor Pineda Ugarte	31	SASTIFACTORIO
364	Instituto Gubernamental Héctor Pineda Ugarte	15	DEBE MEJORAR
363	Instituto Gubernamental Héctor Pineda Ugarte	24	SASTIFACTORIO

362	Instituto Gubernamental Héctor Pineda Ugarte	9	INSASTIFACTORIO
361	Instituto Gubernamental Héctor Pineda Ugarte	26	SASTIFACTORIO
360	Instituto Gubernamental Héctor Pineda Ugarte	26	SASTIFACTORIO
359	Instituto Gubernamental Héctor Pineda Ugarte	26	SASTIFACTORIO
358	Instituto Gubernamental Héctor Pineda Ugarte	32	SASTIFACTORIO
357	Instituto Gubernamental Héctor Pineda Ugarte	26	SASTIFACTORIO
356	Instituto Gubernamental Héctor Pineda Ugarte	24	SASTIFACTORIO
355	Instituto Gubernamental Héctor Pineda Ugarte	16	DEBE MEJORAR
354	Instituto Gubernamental Héctor Pineda Ugarte	25	SASTIFACTORIO
353	Instituto Gubernamental Héctor Pineda Ugarte	29	SASTIFACTORIO
352	Instituto Gubernamental Héctor Pineda Ugarte	17	DEBE MEJORAR
351	Instituto Gubernamental Héctor Pineda Ugarte	19	DEBE MEJORAR
350	Instituto Gubernamental Héctor Pineda Ugarte	16	DEBE MEJORAR
349	Instituto Gubernamental Héctor Pineda Ugarte	29	SASTIFACTORIO
348	Instituto Gubernamental Héctor Pineda Ugarte	35	AVANZADO
347	Instituto Gubernamental Héctor Pineda Ugarte	14	DEBE MEJORAR
346	Instituto Gubernamental Héctor Pineda Ugarte	15	DEBE MEJORAR
345	Instituto Gubernamental Héctor Pineda Ugarte	33	SASTIFACTORIO
344	Instituto Gubernamental Héctor Pineda Ugarte	32	SASTIFACTORIO
343	Instituto Gubernamental Héctor Pineda Ugarte	27	SASTIFACTORIO
342	Instituto Gubernamental Héctor Pineda Ugarte	23	SASTIFACTORIO
341	Instituto Gubernamental Héctor Pineda Ugarte	25	SASTIFACTORIO
340	Instituto Gubernamental Héctor Pineda Ugarte	11	INSASTIFACTORIO
339	Instituto Gubernamental Héctor Pineda Ugarte	23	SASTIFACTORIO
338	Instituto Gubernamental Héctor Pineda Ugarte	31	SASTIFACTORIO
337	Instituto Gubernamental Héctor Pineda Ugarte	19	DEBE MEJORAR
336	Instituto Gubernamental Héctor Pineda Ugarte	32	SASTIFACTORIO
335	Instituto Gubernamental Héctor Pineda Ugarte	25	SASTIFACTORIO
334	Instituto Gubernamental Héctor Pineda Ugarte	19	DEBE MEJORAR
333	Instituto Gubernamental Héctor Pineda Ugarte	28	SASTIFACTORIO
332	Instituto Gubernamental Héctor Pineda Ugarte	21	DEBE MEJORAR
331	Instituto Gubernamental Héctor Pineda Ugarte	19	DEBE MEJORAR
330	Instituto Gubernamental Héctor Pineda Ugarte	19	DEBE MEJORAR
329	Instituto Gubernamental Héctor Pineda Ugarte	36	AVANZADO
328	Instituto Gubernamental Héctor Pineda Ugarte	24	SASTIFACTORIO
327	Instituto Gubernamental Héctor Pineda Ugarte	22	DEBE MEJORAR
326	Instituto Gubernamental Héctor Pineda Ugarte	18	DEBE MEJORAR
325	Instituto Gubernamental Héctor Pineda Ugarte	31	SASTIFACTORIO
324	Instituto Gubernamental Héctor Pineda Ugarte	25	SASTIFACTORIO
323	Instituto Gubernamental Héctor Pineda Ugarte	14	DEBE MEJORAR

322	Instituto Gubernamental Héctor Pineda Ugarte	35	AVANZADO
321	Instituto Gubernamental Héctor Pineda Ugarte	36	AVANZADO
320	Instituto Gubernamental Héctor Pineda Ugarte	20	DEBE MEJORAR
319	Instituto Gubernamental Héctor Pineda Ugarte	29	SASTIFACTORIO
318	Instituto Gubernamental Héctor Pineda Ugarte	34	AVANZADO
317	Instituto Gubernamental Héctor Pineda Ugarte	42	AVANZADO
316	Instituto Gubernamental Héctor Pineda Ugarte	30	SASTIFACTORIO
315	Instituto Gubernamental Héctor Pineda Ugarte	36	AVANZADO
314	Instituto Gubernamental Héctor Pineda Ugarte	30	SASTIFACTORIO
313	Instituto Gubernamental Héctor Pineda Ugarte	25	SASTIFACTORIO
312	Instituto Gubernamental Héctor Pineda Ugarte	30	SASTIFACTORIO
311	Instituto Gubernamental Héctor Pineda Ugarte	20	DEBE MEJORAR
310	Instituto Gubernamental Héctor Pineda Ugarte	28	SASTIFACTORIO
309	Instituto Gubernamental Héctor Pineda Ugarte	33	SASTIFACTORIO
308	Instituto Gubernamental Héctor Pineda Ugarte	33	SASTIFACTORIO
307	Instituto Gubernamental Héctor Pineda Ugarte	33	SASTIFACTORIO
306	Instituto Gubernamental Héctor Pineda Ugarte	26	SASTIFACTORIO
305	Instituto Gubernamental Héctor Pineda Ugarte	21	DEBE MEJORAR
304	Instituto Gubernamental Héctor Pineda Ugarte	30	SASTIFACTORIO
303	Instituto Gubernamental Héctor Pineda Ugarte	23	SASTIFACTORIO
302	Instituto Gubernamental Héctor Pineda Ugarte	33	SASTIFACTORIO
301	Instituto Gubernamental Héctor Pineda Ugarte	34	AVANZADO
300	Instituto Gubernamental Héctor Pineda Ugarte	13	DEBE MEJORAR
299	Instituto Gubernamental Héctor Pineda Ugarte	30	SASTIFACTORIO
298	Instituto Gubernamental Héctor Pineda Ugarte	28	SASTIFACTORIO
297	Instituto Gubernamental Héctor Pineda Ugarte	18	DEBE MEJORAR
296	Instituto Gubernamental Héctor Pineda Ugarte	28	SASTIFACTORIO
295	Instituto Gubernamental Héctor Pineda Ugarte	29	SASTIFACTORIO
294	Instituto Gubernamental Héctor Pineda Ugarte	38	AVANZADO
293	Instituto Gubernamental Héctor Pineda Ugarte	31	SASTIFACTORIO
292	Instituto Gubernamental Héctor Pineda Ugarte	30	SASTIFACTORIO
291	Instituto Gubernamental Héctor Pineda Ugarte	34	AVANZADO
290	Instituto Gubernamental Mixto Hibuera	23	SASTIFACTORIO
289	Instituto Gubernamental Mixto Hibuera	14	DEBE MEJORAR
288	Instituto Gubernamental Mixto Hibuera	32	SASTIFACTORIO
287	Instituto Gubernamental Mixto Hibuera	27	SASTIFACTORIO
286	Instituto Gubernamental Mixto Hibuera	11	INSASTIFACTORIO
285	Instituto Gubernamental Mixto Hibuera	37	AVANZADO
284	Instituto Gubernamental Mixto Hibuera	29	SASTIFACTORIO
283	Instituto Gubernamental Mixto Hibuera	12	DEBE MEJORAR

282	Instituto Gubernamental Mixto Hibuera	23	SASTIFACTORIO
281	Instituto Gubernamental Mixto Hibuera	25	SASTIFACTORIO
280	Instituto Gubernamental Mixto Hibuera	27	SASTIFACTORIO
279	Instituto Gubernamental Mixto Hibuera	32	SASTIFACTORIO
278	Instituto Gubernamental Mixto Hibuera	18	DEBE MEJORAR
277	Instituto Gubernamental Mixto Hibuera	21	DEBE MEJORAR
276	Instituto Gubernamental Mixto Hibuera	37	AVANZADO
275	Instituto Gubernamental Mixto Hibuera	32	SASTIFACTORIO
274	Instituto Gubernamental Mixto Hibuera	26	SASTIFACTORIO
273	Instituto Gubernamental Mixto Hibuera	14	DEBE MEJORAR
272	Instituto Gubernamental Mixto Hibuera	18	DEBE MEJORAR
271	Instituto Gubernamental Mixto Hibuera	22	DEBE MEJORAR
270	Instituto Gubernamental Mixto Hibuera	28	SASTIFACTORIO
269	Instituto Gubernamental Mixto Hibuera	30	SASTIFACTORIO
268	Instituto Gubernamental Mixto Hibuera	29	SASTIFACTORIO
267	Instituto Gubernamental Mixto Hibuera	21	DEBE MEJORAR
266	Instituto Gubernamental Mixto Hibuera	36	AVANZADO
265	Instituto Gubernamental Mixto Hibuera	31	SASTIFACTORIO
264	Instituto Gubernamental Mixto Hibuera	21	DEBE MEJORAR
263	Instituto Gubernamental Mixto Hibuera	16	DEBE MEJORAR
262	Instituto Gubernamental Mixto Hibuera	35	AVANZADO
261	Instituto Gubernamental Mixto Hibuera	15	DEBE MEJORAR
260	Instituto Gubernamental Mixto Hibuera	27	SASTIFACTORIO
259	Instituto Gubernamental Mixto Hibuera	15	DEBE MEJORAR
258	Instituto Gubernamental Mixto Hibuera	24	SASTIFACTORIO
257	Instituto Gubernamental Mixto Hibuera	25	SASTIFACTORIO
256	Instituto Gubernamental Mixto Hibuera	23	SASTIFACTORIO
255	Instituto Gubernamental Mixto Hibuera	14	DEBE MEJORAR
254	Instituto Gubernamental Mixto Hibuera	21	DEBE MEJORAR
253	Instituto Gubernamental Mixto Hibuera	24	SASTIFACTORIO
252	Instituto Gubernamental Mixto Hibuera	17	DEBE MEJORAR
251	Instituto Gubernamental Mixto Hibuera	21	DEBE MEJORAR
250	Instituto Gubernamental Mixto Hibuera	22	DEBE MEJORAR
249	Instituto Gubernamental Mixto Hibuera	17	DEBE MEJORAR
248	Instituto Gubernamental Mixto Hibuera	21	DEBE MEJORAR
247	Instituto Gubernamental Mixto Hibuera	20	DEBE MEJORAR
246	Instituto Gubernamental Mixto Hibuera	20	DEBE MEJORAR
245	Instituto Gubernamental Mixto Hibuera	19	DEBE MEJORAR
244	Instituto Gubernamental Mixto Hibuera	27	SASTIFACTORIO
243	Instituto Gubernamental Mixto Hibuera	15	DEBE MEJORAR

242	Instituto Gubernamental Mixto Hibuera	28	SASTIFACTORIO
241	Instituto Gubernamental Mixto Hibuera	16	DEBE MEJORAR
240	Instituto Gubernamental Mixto Hibuera	20	DEBE MEJORAR
239	Instituto Gubernamental Mixto Hibuera	12	DEBE MEJORAR
238	Instituto Gubernamental Mixto Hibuera	13	DEBE MEJORAR
237	Instituto Gubernamental Mixto Hibuera	24	SASTIFACTORIO
236	Instituto Gubernamental Mixto Hibuera	38	AVANZADO
235	Instituto Gubernamental Mixto Hibuera	18	DEBE MEJORAR
234	Instituto Gubernamental Mixto Hibuera	25	SASTIFACTORIO
233	Instituto Gubernamental Mixto Hibuera	24	SASTIFACTORIO
232	Instituto Gubernamental Mixto Hibuera	18	DEBE MEJORAR
231	Instituto Gubernamental Mixto Hibuera	20	DEBE MEJORAR
230	Instituto Gubernamental Mixto Hibuera	23	SASTIFACTORIO
229	Instituto Gubernamental Mixto Hibuera	15	DEBE MEJORAR
228	Instituto Gubernamental Mixto Hibuera	28	SASTIFACTORIO
227	Instituto Gubernamental Mixto Hibuera	26	SASTIFACTORIO
226	Instituto Gubernamental Mixto Hibuera	23	SASTIFACTORIO
225	Instituto Gubernamental Mixto Hibuera	25	SASTIFACTORIO
224	Instituto Gubernamental Mixto Hibuera	24	SASTIFACTORIO
223	Instituto Gubernamental Mixto Hibuera	30	SASTIFACTORIO
222	Instituto Gubernamental Mixto Hibuera	34	AVANZADO
221	Instituto Gubernamental Mixto Hibuera	41	AVANZADO
220	Instituto Gubernamental Mixto Hibuera	21	DEBE MEJORAR
219	Instituto Gubernamental Mixto Hibuera	26	SASTIFACTORIO
218	Instituto Gubernamental Mixto Hibuera	22	DEBE MEJORAR
217	Instituto Gubernamental Mixto Hibuera	29	SASTIFACTORIO
216	Instituto Gubernamental Mixto Hibuera	20	DEBE MEJORAR
215	Instituto Gubernamental Mixto Hibuera	31	SASTIFACTORIO
214	Instituto Gubernamental Mixto Hibuera	38	AVANZADO
213	Instituto Gubernamental Mixto Hibuera	25	SASTIFACTORIO
212	Instituto Gubernamental Mixto Hibuera	30	SASTIFACTORIO
211	Instituto Gubernamental Mixto Hibuera	28	SASTIFACTORIO
210	Instituto Gubernamental Mixto Hibuera	24	SASTIFACTORIO
209	Instituto Gubernamental Mixto Hibuera	24	SASTIFACTORIO
208	Instituto Gubernamental Mixto Hibuera	23	SASTIFACTORIO
207	Instituto Gubernamental Mixto Hibuera	37	AVANZADO
206	Instituto Gubernamental Mixto Hibuera	31	SASTIFACTORIO
205	Instituto Gubernamental Mixto Hibuera	29	SASTIFACTORIO
204	Instituto Gubernamental Mixto Hibuera	30	SASTIFACTORIO
203	Instituto Gubernamental Mixto Hibuera	17	DEBE MEJORAR

202	Instituto Gubernamental Mixto Hibuera	30	SASTIFACTORIO
201	Instituto Gubernamental Mixto Hibuera	35	AVANZADO
200	Instituto Gubernamental Mixto Hibuera	33	SASTIFACTORIO
199	Instituto Gubernamental Mixto Hibuera	16	DEBE MEJORAR
198	Instituto Gubernamental Mixto Hibuera	17	DEBE MEJORAR
197	Instituto Gubernamental Mixto Hibuera	29	SASTIFACTORIO
196	Instituto Gubernamental Mixto Hibuera	20	DEBE MEJORAR
195	Instituto Gubernamental Mixto Hibuera	29	SASTIFACTORIO
194	Instituto Gubernamental Mixto Hibuera	25	SASTIFACTORIO
193	Instituto Gubernamental Mixto Hibuera	27	SASTIFACTORIO
192	Instituto Gubernamental Mixto Hibuera	23	SASTIFACTORIO
191	Instituto Gubernamental Mixto Hibuera	25	SASTIFACTORIO
190	Instituto Gubernamental Mixto Hibuera	25	SASTIFACTORIO
189	Instituto Gubernamental Mixto Hibuera	26	SASTIFACTORIO
188	Instituto Gubernamental Mixto Hibuera	19	DEBE MEJORAR
187	Instituto Gubernamental Mixto Hibuera	20	DEBE MEJORAR
186	Instituto Gubernamental Mixto Hibuera	25	SASTIFACTORIO
185	Instituto Gubernamental Mixto Hibuera	33	SASTIFACTORIO
184	Instituto Gubernamental Mixto Hibuera	21	DEBE MEJORAR
183	Instituto Gubernamental Mixto Hibuera	19	DEBE MEJORAR
182	Instituto Gubernamental Mixto Hibuera	33	SASTIFACTORIO
181	Instituto Gubernamental Mixto Hibuera	17	DEBE MEJORAR
180	Instituto Gubernamental Mixto Hibuera	29	SASTIFACTORIO
179	Instituto Gubernamental Mixto Hibuera	16	DEBE MEJORAR
178	Instituto Gubernamental Mixto Hibuera	25	SASTIFACTORIO
177	Instituto Gubernamental Mixto Hibuera	26	SASTIFACTORIO
176	Instituto Gubernamental Mixto Hibuera	17	DEBE MEJORAR
175	Instituto Gubernamental Mixto Hibuera	31	SASTIFACTORIO
174	Instituto Gubernamental Mixto Hibuera	21	DEBE MEJORAR
173	Instituto Gubernamental Mixto Hibuera	13	DEBE MEJORAR
172	Instituto Gubernamental Mixto Hibuera	20	DEBE MEJORAR
171	Instituto Gubernamental Mixto Hibuera	16	DEBE MEJORAR
170	Instituto Gubernamental Mixto Hibuera	18	DEBE MEJORAR
169	Instituto Gubernamental Mixto Hibuera	20	DEBE MEJORAR
168	Instituto Gubernamental Mixto Hibuera	21	DEBE MEJORAR
167	Instituto Gubernamental Mixto Hibuera	27	SASTIFACTORIO
166	Instituto Gubernamental Mixto Hibuera	29	SASTIFACTORIO
165	Instituto Gubernamental Mixto Hibuera	20	DEBE MEJORAR
164	Instituto Gubernamental Mixto Hibuera	27	SASTIFACTORIO
163	Instituto Gubernamental Mixto Hibuera	27	SASTIFACTORIO

162	Instituto Gubernamental Mixto Hibuera	21	DEBE MEJORAR
161	Instituto Gubernamental Mixto Hibuera	21	DEBE MEJORAR
160	Instituto Gubernamental Mixto Hibuera	15	DEBE MEJORAR
159	Instituto Gubernamental Mixto Hibuera	33	SASTIFACTORIO
158	Instituto Gubernamental Mixto Hibuera	26	SASTIFACTORIO
157	Instituto Gubernamental Mixto Hibuera	24	SASTIFACTORIO
156	Instituto Gubernamental Mixto Hibuera	20	DEBE MEJORAR
155	Instituto Gubernamental Mixto Hibuera	16	DEBE MEJORAR
154	Instituto Gubernamental Mixto Hibuera	14	DEBE MEJORAR
153	Instituto Gubernamental Mixto Hibuera	25	SASTIFACTORIO
152	Instituto Gubernamental Mixto Hibuera	19	DEBE MEJORAR
151	Instituto Gubernamental Mixto Hibuera	19	DEBE MEJORAR
150	Instituto Gubernamental Mixto Hibuera	21	DEBE MEJORAR
149	Instituto Gubernamental Mixto Hibuera	25	SASTIFACTORIO
148	Instituto Gubernamental Mixto Hibuera	20	DEBE MEJORAR
147	Instituto Gubernamental Mixto Hibuera	20	DEBE MEJORAR
146	Instituto Gubernamental Mixto Hibuera	22	DEBE MEJORAR
145	Instituto Gubernamental Mixto Hibuera	16	DEBE MEJORAR
144	Instituto Gubernamental Mixto Hibuera	12	DEBE MEJORAR
143	Instituto Gubernamental Mixto Hibuera	14	DEBE MEJORAR
142	Instituto Gubernamental Mixto Hibuera	16	DEBE MEJORAR
141	Instituto Gubernamental Mixto Hibuera	13	DEBE MEJORAR
140	Instituto Gubernamental Mixto Hibuera	11	INSASTIFACTORIO
139	Instituto Gubernamental Mixto Hibuera	10	INSASTIFACTORIO
138	Instituto Gubernamental Mixto Hibuera	34	AVANZADO
137	Instituto Gubernamental Mixto Hibuera	33	SASTIFACTORIO
136	Instituto Gubernamental Mixto Hibuera	34	AVANZADO
135	Instituto Gubernamental Mixto Hibuera	20	DEBE MEJORAR
134	Instituto Gubernamental Mixto Hibuera	28	SASTIFACTORIO
133	Instituto Gubernamental Mixto Hibuera	24	SASTIFACTORIO
132	Instituto Gubernamental Mixto Hibuera	29	SASTIFACTORIO
131	Instituto Gubernamental Mixto Hibuera	33	SASTIFACTORIO
130	Instituto Gubernamental Mixto Hibuera	23	SASTIFACTORIO
129	Instituto Gubernamental Mixto Hibuera	27	SASTIFACTORIO
128	Instituto Gubernamental Mixto Hibuera	25	SASTIFACTORIO
127	Instituto Gubernamental Mixto Hibuera	32	SASTIFACTORIO
126	Instituto Gubernamental Mixto Hibuera	28	SASTIFACTORIO
125	Instituto Gubernamental Mixto Hibuera	33	SASTIFACTORIO
124	Instituto Gubernamental Mixto Hibuera	34	AVANZADO
123	Instituto Gubernamental Mixto Hibuera	24	SASTIFACTORIO

122	Instituto Gubernamental Mixto Hibuera	34	AVANZADO
121	Instituto Gubernamental Mixto Hibuera	25	SASTIFACTORIO
120	Instituto Gubernamental Mixto Hibuera	39	AVANZADO
119	Instituto Gubernamental Mixto Hibuera	24	SASTIFACTORIO
118	Instituto Gubernamental Mixto Hibuera	32	SASTIFACTORIO
117	Instituto Gubernamental Mixto Hibuera	30	SASTIFACTORIO
116	Instituto Gubernamental Mixto Hibuera	21	DEBE MEJORAR
115	Instituto Gubernamental Mixto Hibuera	35	AVANZADO
114	Instituto Gubernamental Mixto Hibuera	36	AVANZADO
113	Instituto Gubernamental Mixto Hibuera	22	DEBE MEJORAR
112	Instituto Gubernamental Mixto Hibuera	21	DEBE MEJORAR
111	Instituto Gubernamental Mixto Hibuera	29	SASTIFACTORIO
110	Instituto Gubernamental Técnico Abelardo R. Fortín	32	SASTIFACTORIO
109	Instituto Gubernamental Técnico Abelardo R. Fortín	29	SASTIFACTORIO
108	Instituto Gubernamental Técnico Abelardo R. Fortín	34	AVANZADO
107	Instituto Gubernamental Técnico Abelardo R. Fortín	29	SASTIFACTORIO
106	Instituto Gubernamental Técnico Abelardo R. Fortín	21	DEBE MEJORAR
105	Instituto Gubernamental Técnico Abelardo R. Fortín	31	SASTIFACTORIO
104	Instituto Gubernamental Técnico Abelardo R. Fortín	27	SASTIFACTORIO
103	Instituto Gubernamental Técnico Abelardo R. Fortín	27	SASTIFACTORIO
102	Instituto Gubernamental Técnico Abelardo R. Fortín	19	DEBE MEJORAR
101	Instituto Gubernamental Técnico Abelardo R. Fortín	29	SASTIFACTORIO
100	Instituto Gubernamental Técnico Abelardo R. Fortín	22	DEBE MEJORAR
99	Instituto Gubernamental Técnico Abelardo R. Fortín	37	AVANZADO
98	Instituto Gubernamental Técnico Abelardo R. Fortín	16	DEBE MEJORAR
97	Instituto Gubernamental Técnico Abelardo R. Fortín	24	SASTIFACTORIO
96	Instituto Gubernamental Técnico Abelardo R. Fortín	16	DEBE MEJORAR

95	Instituto Gubernamental Técnico Abelardo R. Fortín	26	SASTIFACTORIO
94	Instituto Gubernamental Técnico Abelardo R. Fortín	13	DEBE MEJORAR
93	Instituto Gubernamental Técnico Abelardo R. Fortín	31	SASTIFACTORIO
92	Instituto Gubernamental Técnico Abelardo R. Fortín	32	SASTIFACTORIO
91	Instituto Gubernamental Técnico Abelardo R. Fortín	33	SASTIFACTORIO
90	Instituto Gubernamental Técnico Abelardo R. Fortín	30	SASTIFACTORIO
89	Instituto Gubernamental Técnico Abelardo R. Fortín	26	SASTIFACTORIO
88	Instituto Gubernamental Técnico Abelardo R. Fortín	34	AVANZADO
87	Instituto Gubernamental Técnico Abelardo R. Fortín	15	DEBE MEJORAR
86	Instituto Gubernamental Técnico Abelardo R. Fortín	25	SASTIFACTORIO
85	Instituto Gubernamental Técnico Abelardo R. Fortín	19	DEBE MEJORAR
84	Instituto Gubernamental Técnico Abelardo R. Fortín	25	SASTIFACTORIO
83	Instituto Gubernamental Técnico Abelardo R. Fortín	26	SASTIFACTORIO
82	Instituto Gubernamental Técnico Abelardo R. Fortín	22	DEBE MEJORAR
81	Instituto Gubernamental Técnico Abelardo R. Fortín	29	SASTIFACTORIO
80	Instituto Gubernamental Técnico Abelardo R. Fortín	22	DEBE MEJORAR
79	Instituto Gubernamental Técnico Abelardo R. Fortín	28	SASTIFACTORIO
78	Instituto Gubernamental Técnico Abelardo R. Fortín	22	DEBE MEJORAR
77	Instituto Gubernamental Técnico Abelardo R. Fortín	31	SASTIFACTORIO
76	Instituto Gubernamental Técnico Abelardo R. Fortín	32	SASTIFACTORIO
75	Instituto Gubernamental Técnico Abelardo R. Fortín	29	SASTIFACTORIO
74	Instituto Gubernamental Técnico Abelardo R. Fortín	24	SASTIFACTORIO

73	Instituto Gubernamental Técnico Abelardo R. Fortín	22	DEBE MEJORAR
72	Instituto Gubernamental Técnico Abelardo R. Fortín	17	DEBE MEJORAR
71	Instituto Gubernamental Técnico Abelardo R. Fortín	12	DEBE MEJORAR
70	Instituto Gubernamental Técnico Abelardo R. Fortín	22	DEBE MEJORAR
69	Instituto Gubernamental Técnico Abelardo R. Fortín	15	DEBE MEJORAR
68	Instituto Gubernamental Técnico Abelardo R. Fortín	30	SASTIFACTORIO
67	Instituto Gubernamental Técnico Abelardo R. Fortín	25	SASTIFACTORIO
66	Instituto Gubernamental Técnico Abelardo R. Fortín	30	SASTIFACTORIO
65	Instituto Gubernamental Técnico Abelardo R. Fortín	18	DEBE MEJORAR
64	Instituto Gubernamental Técnico Abelardo R. Fortín	39	AVANZADO
63	Instituto Gubernamental Técnico Abelardo R. Fortín	24	SASTIFACTORIO
62	Instituto Gubernamental Técnico Abelardo R. Fortín	22	DEBE MEJORAR
61	Instituto Gubernamental Técnico Abelardo R. Fortín	25	SASTIFACTORIO
60	Instituto Gubernamental Técnico Abelardo R. Fortín	31	SASTIFACTORIO
59	Instituto Gubernamental Técnico Abelardo R. Fortín	26	SASTIFACTORIO
58	Instituto Gubernamental Técnico Abelardo R. Fortín	31	SASTIFACTORIO
57	Instituto Gubernamental Técnico Abelardo R. Fortín	19	DEBE MEJORAR
56	Instituto Gubernamental Técnico Abelardo R. Fortín	34	AVANZADO
55	Instituto Gubernamental Técnico Abelardo R. Fortín	25	SASTIFACTORIO
54	Instituto Gubernamental Técnico Abelardo R. Fortín	41	AVANZADO
53	Instituto Gubernamental Técnico Abelardo R. Fortín	19	DEBE MEJORAR
52	Instituto Gubernamental Técnico Abelardo R. Fortín	37	AVANZADO

51	Instituto Gubernamental Técnico Abelardo R. Fortín	30	SASTIFACTORIO
50	Instituto Gubernamental Técnico Abelardo R. Fortín	18	DEBE MEJORAR
49	Instituto Gubernamental Técnico Abelardo R. Fortín	30	SASTIFACTORIO
48	Instituto Gubernamental Técnico Abelardo R. Fortín	36	AVANZADO
47	Instituto Gubernamental Técnico Abelardo R. Fortín	21	DEBE MEJORAR
46	Instituto Gubernamental Técnico Abelardo R. Fortín	19	DEBE MEJORAR
45	Instituto Gubernamental Técnico Abelardo R. Fortín	27	SASTIFACTORIO
44	Instituto Gubernamental Técnico Abelardo R. Fortín	22	DEBE MEJORAR
43	Instituto Gubernamental Técnico Abelardo R. Fortín	28	SASTIFACTORIO
42	Instituto Gubernamental Técnico Abelardo R. Fortín	25	SASTIFACTORIO
41	Instituto Gubernamental Técnico Abelardo R. Fortín	37	AVANZADO
40	Instituto Gubernamental Técnico Abelardo R. Fortín	37	AVANZADO
39	Instituto Gubernamental Técnico Abelardo R. Fortín	27	SASTIFACTORIO
38	Instituto Gubernamental Técnico Abelardo R. Fortín	27	SASTIFACTORIO
37	Instituto Gubernamental Técnico Abelardo R. Fortín	32	SASTIFACTORIO
36	Instituto Gubernamental Técnico Abelardo R. Fortín	30	SASTIFACTORIO
35	Instituto Gubernamental Técnico Abelardo R. Fortín	21	DEBE MEJORAR
34	Instituto Gubernamental Técnico Abelardo R. Fortín	30	SASTIFACTORIO
33	Instituto Gubernamental Técnico Abelardo R. Fortín	33	SASTIFACTORIO
32	Instituto Gubernamental Técnico Abelardo R. Fortín	34	AVANZADO
31	Instituto Gubernamental Técnico Abelardo R. Fortín	26	SASTIFACTORIO
30	Instituto Gubernamental Técnico Abelardo R. Fortín	23	SASTIFACTORIO

29	Instituto Gubernamental Técnico Abelardo R. Fortín	29	SASTIFACTORIO
28	Instituto Gubernamental Técnico Abelardo R. Fortín	22	DEBE MEJORAR
27	Instituto Gubernamental Técnico Abelardo R. Fortín	14	DEBE MEJORAR
26	Instituto Gubernamental Técnico Abelardo R. Fortín	23	SASTIFACTORIO
25	Instituto Gubernamental Técnico Abelardo R. Fortín	17	DEBE MEJORAR
24	Instituto Gubernamental Técnico Abelardo R. Fortín	29	SASTIFACTORIO
23	Instituto Gubernamental Técnico Abelardo R. Fortín	12	DEBE MEJORAR
22	Instituto Gubernamental Técnico Abelardo R. Fortín	14	DEBE MEJORAR
21	Instituto Gubernamental Técnico Abelardo R. Fortín	19	DEBE MEJORAR
20	Instituto Gubernamental Técnico Abelardo R. Fortín	30	SASTIFACTORIO
19	Instituto Gubernamental Técnico Abelardo R. Fortín	30	SASTIFACTORIO
18	Instituto Gubernamental Técnico Abelardo R. Fortín	25	SASTIFACTORIO
17	Instituto Gubernamental Técnico Abelardo R. Fortín	9	INSASTIFACTORIO
16	Instituto Gubernamental Técnico Abelardo R. Fortín	28	SASTIFACTORIO
15	Instituto Gubernamental Técnico Abelardo R. Fortín	20	DEBE MEJORAR
14	Instituto Gubernamental Técnico Abelardo R. Fortín	28	SASTIFACTORIO
13	Instituto Gubernamental Técnico Abelardo R. Fortín	20	DEBE MEJORAR
12	Instituto Gubernamental Técnico Abelardo R. Fortín	15	DEBE MEJORAR
11	Instituto Gubernamental Técnico Abelardo R. Fortín	17	DEBE MEJORAR
10	Instituto Gubernamental Técnico Abelardo R. Fortín	21	DEBE MEJORAR
9	Instituto Gubernamental Técnico Abelardo R. Fortín	19	DEBE MEJORAR
8	Instituto Gubernamental Técnico Abelardo R. Fortín	21	DEBE MEJORAR

7	Instituto Gubernamental Técnico Abelardo R. Fortín	31	SASTIFACTORIO
6	Instituto Gubernamental Técnico Abelardo R. Fortín	19	DEBE MEJORAR
5	Instituto Gubernamental Técnico Abelardo R. Fortín	16	DEBE MEJORAR
4	Instituto Gubernamental Técnico Abelardo R. Fortín	18	DEBE MEJORAR
3	Instituto Gubernamental Técnico Abelardo R. Fortín	22	DEBE MEJORAR
2	Instituto Gubernamental Técnico Abelardo R. Fortín	14	DEBE MEJORAR
1	Instituto Gubernamental Técnico Abelardo R. Fortín	33	SASTIFACTORIO

Fuente: Elaboración propia

En la tabla 23 se observa a los centros educativos que obtuvieron un buen desempeño. El mayor puntaje lo obtuvo el Instituto Héctor Pineda Ugarte con el 422 y con el puntaje más bajo se encuentra entre el Instituto Técnico Abelardo R. Fortín y el Instituto Héctor Pineda Ugarte con el 90.

Anexo 11:Taller de evaluación y redacción de ítems











Anexo 12: Docentes del área de Ciencias Sociales validando el instrumento





