

327054



**Universidad Virtual**

**Escuela de Graduados en Educación**

**Prácticas que permiten mejorar la calidad en el proceso de instrumentación para pruebas que miden rendimiento escolar.**

**Tesis que para obtener el grado de:**

**Maestría en Administración de Instituciones Educativas**

presenta:

**Arely Guadalupe Morales Blancas**

Asesor tutor:

**Mtra. Georgina González Ávila**

Asesor titular:

**Dra. Katherina E. Gallardo Córdova**

**Texcoco, Estado de México, México**

**Enero, 2011**

## Resumen

El problema que en esta investigación se presenta es de vital importancia para el ámbito educativo ya que da cuenta de la forma en la que se deben de realizar instrumentos de evaluación. La investigación que se presenta se ubica en un enfoque cuantitativo, en donde se destaca la pregunta general de investigación que refiere: ¿En qué medida llevar a cabo prácticas de planeación, construcción y análisis de datos de pruebas diagnósticas de rendimiento académico, favorece el mejoramiento en el diseño de este tipo de instrumentos?. Posibilidad que da en ello la Taxonomía de Marzano y Kendall como propuesta para su realización ya que permite en su proceso de pensamiento y sistema de conocimiento, verificar el desempeño académico de los alumnos. Esta propuesta de investigación permitió se confrontará la pregunta general de investigación con los paradigmas teórico-instrumentales para realizar pruebas objetivas con lo que se favorece el mejoramiento en el diseño de este tipo de instrumentos como aporte de investigación y en el campo de aplicación para otros espacios académicos. Los resultados, dan panorama del rendimiento académico de los alumnos de 4°,5° y 6° de primaria y de 1° y 3° de secundaria en una de las zonas de oriente del Estado de México. Asimismo se hace un recuento del estado de las pruebas objetivas en su construcción y análisis de datos, determinando que la mayoría son pruebas homogéneas excepto la de 3°, y que en su generalidad, son consideradas pruebas difíciles ya que tienen un índice de dificultad alto en su habilidad de adecuación de medición. Dicha evaluación permitió emitir inferencias sobre los dominios de conocimiento y niveles de pensamiento a partir de la Taxonomía de Marzano y presentar conclusiones a partir de los hallazgos encontrados.

## Índice

Introducción.....	8
Capítulo 1. Planteamiento del problema.....	12
Antecedentes.....	12
Definición del problema.....	20
Preguntas de investigación.....	22
Objetivos de investigación.....	23
Justificación.....	23
Limitaciones.....	24
Capítulo 2. Revisión de la literatura.....	26
Evaluar con calidad.....	27
Proceso histórico de las pruebas.....	28
Evaluación Educativa.....	31
Medición Educativa.....	32
Bases teóricas y técnicas para la construcción de instrumentos de medición: Pruebas objetivas.....	33
La nueva taxonomía de Marzano y Kendall.....	33
Bases conceptuales: Pruebas o exámenes.....	34
Pruebas objetivas.....	35
Pruebas de aprovechamiento.....	36
Pruebas de diagnóstico.....	37
Proceso de planeación para el diseño de pruebas objetivas.....	38

Objetivos de aprendizaje.....	39
Uso de la taxonomía de Bloom en la Educación.....	39
Uso de la nueva taxonomía de Marzano y Kendall.....	40
Dominios de conocimiento.....	42
Sistemas de pensamiento.....	43
Contenido que debe abarcarse.....	50
Tablas de especificaciones.....	53
Diseño de pruebas de reactivos de opción múltiple.....	54
Objetivos de aprendizaje como punto de partida de la planeación de la evaluación.....	54
Objetivos de aprendizaje y su relación con el nivel taxonómico.....	55
Bases teórico-prácticas para construcción de reactivos de opción múltiple...	56
Constitución de la base de la pregunta.....	57
Opciones de respuesta.....	58
Distractores.....	59
Justificación de distractores.....	60
Retroalimentación.....	61
Otros tipos de reactivos de opción múltiple.....	62
Reactivos de elección múltiple complejo.....	62
Reactivos apareados o de respuesta por pares.....	63
Reactivos multi-ítem de base común.....	63
Ventajas y desventajas del diseño de reactivos de opción múltiple.....	65

Propiedades de las pruebas de reactivos de opción múltiple: Psicometría.....	66
Capítulo 3. Enfoque metodológico.....	67
Descripción de enfoque.....	67
Diseño de investigación.....	71
Método.....	73
Participantes.....	73
Instrumentos.....	78
Procedimientos.....	83
Estrategias de análisis de datos.....	84
Capítulo 4. Análisis de resultados.....	88
Proceso de validación.....	88
Descripción de trabajo realizado por los expertos.....	94
Resultados para la toma de decisiones en relación al instrumento y su conformación final.....	96
Resultados de la aplicación de las pruebas.....	108
Análisis de los resultados desde la perspectiva de Marzano y Kendall.....	113
Capítulo 5. Conclusiones.....	116
Discusión.....	116
Conclusiones.....	120
Recomendaciones.....	120
Referencias.....	124
Apéndice A. Forma de Consentimiento.....	132

Apéndice B. Lista de cotejo.....	133
Apéndice C. Prueba preliminar de 3° de primaria en la asignatura de Historia.....	140
Apéndice D. Sabanas de datos de la prueba de Historia, que se aplicó a 4° de primaria.....	144
Apéndice E. Currículum Vitae.....	148

## **Introducción**

La evaluación de los aprendizajes, es una posibilidad de transformar los ámbitos educativos ya que permite generar una cultura de rendición de cuentas. En el aula el proceso de evaluación se torna un poco más complejo, ya que es el docente quien verifica el aprendizaje del alumno, por medio de instrumentos que dan cuenta de su desempeño académico. Existen varios mecanismos de evaluación internos como son las pruebas realizadas por los docentes y externos como son las pruebas objetivas realizadas por organismos como PISA (Program for International Student Assessment) e INEE (Instituto Nacional para la Evaluación de la Educación), que cumplen la función de evaluar los objetivos de las principales áreas de conocimiento del currículo nacional.

Los fines de esta investigación promovieron un proceso de evaluación, centrado en la verificación del desempeño del alumno por medio de la planeación, construcción y análisis de pruebas diagnósticas, como instrumentos de medición de procesos de pensamiento y sistemas de conocimiento, desde la propuesta de Marzano y Kendall. Este instrumento que fue aplicado en la educación básica (primaria y secundaria) en grados en donde se encuentra la especialidad de Historia como parte de las asignaturas curriculares. Asimismo, estos instrumentos permitieron, tomar decisiones en la construcción y en los procesos que evaluaron conductas específicas.

Bajo esta idea cabe mencionar, que la investigación que se presenta se ubica en un enfoque cuantitativo. En el documento se expone el proceso y resultados de la investigación educativa en los siguientes capítulos: Planteamiento de problema, Revisión de la literatura, Enfoque metodológico, Análisis de Resultados y Conclusiones.

En el primer capítulo se presenta el planteamiento de problema; el cual expone los procesos de evaluación y sus diseños instrumentales derivados de las políticas internacionales y nacionales con el fin de contextualizar la construcción de instrumentos. La definición de una realidad permite al investigador plantearse cuestionamientos precisos sobre lo que se quiere investigar de tal manera que las preguntas y objetivos fueron la guía para la generación de nuevos conocimientos. La justificación dio muestra de la relevancia de la problemática detectada así, como de los alcances y limitantes de la investigación.

En el segundo capítulo, se definen los conceptos y teorías que se encuentran alrededor de la evaluación en el campo educativo, razón por la que se realizó una revisión exhaustiva sobre las mediciones que en este momento se realizan en educación básica, centrados en las diferentes enfoques que se han trabajado y que permean en la actual escuela. En este tenor las perspectivas que se toman en cuenta es el proceso histórico de la construcción de pruebas en el campo psicológico y educativo son referidas de autores como: Aiken (2003), Escudero (2003) y Cummings (1990).

Por lo que la evaluación y medición cobraron sentido en relación a las bases teóricas e instrumentales que promueven las aportaciones de: Marzano y Kendall (2007), Airasiam (2002), Gronlund (1983), Guirtz y Palamidess (2000), Karmel (1974), López e Hinojosa (2001), Lafourcade (1997), Thorndike (1989) y Bloom (1971), en la construcción de pruebas capaces de medir objetivos educativos con calidad, para direccionar procesos de calidad en los modelos de aprendizaje. Estas aportaciones son el marco de referencia que dio, cuenta para responder la pregunta de investigación.



En el tercer capítulo se incluye el enfoque metodológico que da cuenta del proceso seguido la selección del enfoque de la investigación cuantitativo y su alcance exploratorio y descriptivo. Dentro de esta investigación se encuentran dos diseños que fueron seleccionados para llevarla a cabo la investigación: el diseño transeccional exploratorio y diseño transeccional descriptivo, detallado con mayor énfasis en esta sección. Para la interpretación de variables, el estudio descriptivo, como estrategia para la aplicación y decodificación de los datos arrojados. En la sección de participantes se muestra como los profesores de educación básica validaron los instrumentos realizados y los alumnos a quienes se les aplicó el instrumento que validaron los profesores.

En la segunda sección llamada los instrumentos; describe los exámenes preliminares, que se elaboraron por especialistas del Tecnológico de Monterrey, basándose en los planes y programas de estudio de la asignatura de Historia vigentes en la educación básica (primaria 3°, 4°, 5° y 6° y secundaria 2°) y que se aplicaron a los alumnos de diferentes niveles como muestra representativa.

Por último la recolección de datos permitió confrontar las variables y validar el instrumento con el fin de promover una prueba objetiva que se utilice y aplique en estudios posteriores por medio de medidas de tendencia central para verificar la confiabilidad por medio del procedimiento alfa Cronbach y la validez de criterio.

En el cuarto capítulo. Los resultados del análisis dieron respuesta las preguntas planteadas en la investigación con lo que se confronta la pregunta general de investigación con los paradigmas teórico- instrumentales. En la primera parte se presenta el proceso de validación mediante la descripción de los instrumentos así como

los niveles, cantidad de preguntas y niveles de procesamiento y dominios de conocimiento de cada prueba.

Posteriormente se presentan los resultados que emitieron los jueces a partir de los criterios de evaluación emitidos en la lista de cotejo partir de cada nivel educativo por lo que se muestra un ejemplo de las diferencias y de las similitudes encontradas en los reactivos por los expertos con el fin de tomar una decisión en el ajuste y de cómo quedaron constituidos los reactivos a partir de esta validación.

En la segunda parte se muestran los resultados más destacados a partir de verificar las puntuaciones de los alumnos más altos y los más bajos con el fin de determinar la complejidad de la prueba a través de la postura de Cohen (2000) y Carey (2002) en donde se muestra la confiabilidad de las pruebas y su homogeneidad.

Por último se hace una correlación entre los reactivos y la Nueva Taxonomía de Marzano y Kendall con el fin de mostrar los niveles en los que se realizó la prueba y el acercamiento a los objetivos planificados de la prueba.

En el capítulo cinco. Conclusiones, en ellas se destaca la importancia de la aplicación y verificación de instrumentos que miden rendimiento académico los hallazgos encontrados a partir de la construcción de este tipo de pruebas bajo la propuesta de Marzano y Kendall son la posibilidad de innovar pedagógicamente en el espacio áulico ya que permiten la verificación de rendimiento académico. Las recomendaciones que se hacen es continuar con en el diseño de este tipo de instrumentos como aporte para futuras investigaciones y campos de aplicación en otros espacios académico.

## **Capítulo 1. Planteamiento del problema**

El planteamiento del problema, para esta investigación permitió visualizar la importancia que tiene la evaluación en nuestros tiempos, por lo que este capítulo muestra el contexto en el que la investigación se desarrollo. Asimismo se presenta la descripción y objeto de investigación, se describen las directrices mediadas por los objetivos de la investigación así como el estado que guarda en relación a las limitaciones y delimitaciones del proyecto .Por lo que se presenta en este segmento el cómo las prácticas de evaluación en educación han permitido la toma de decisiones en relación a los instrumentos que miden aprendizajes. Asimismo se pretende que el lector identifique la problemática, respecto a las prácticas que permiten mejorar la calidad de instrumentos de medición del desempeño académico en relación a su planeación, construcción y análisis de datos de pruebas diagnósticas de rendimiento escolar a través de la concreción de un objetivo general que guiará la investigación y permitirá justificar, por qué es importante este tema en el contexto actual.

### *Antecedentes*

Para entender la direccionalidad que tiene la educación en México en cuanto al proceso de calidad en la evaluación educativa, es necesario ubicarnos temporalmente en los procesos vividos desde los años setenta a los noventa y su participación en las pruebas de rendimiento académico. Durante esta época se destacan proyectos investigativos de evaluación tales como: TIMSS (Tercer estudio Internacional sobre Matemáticas y Ciencias): LLECE (Laboratorio Latinoamericano de Evaluación de la Calidad Educativa de la OREALC) de acuerdo a su incorporación de México en 1994, la OCDE

(Organización para la Cooperación y Desarrollo Económicas), a través de su proyecto de evaluación llamado PISA (Program for International Student Assessment), diseñado este último para incidir en políticas educativas de los países incorporados a esta organización y para medir aptitudes o competencias de los alumnos de 15 y 16 años de manera integrada, teniendo como áreas o evaluación de dominio de contenidos a la lectura, Matemáticas y Ciencias, organizadas en tres dimensiones: “procesos, contenido y contexto o situación” (PISA, 2005, p.16).

Durante el 2002 surge en México el INEE (Instituto Nacional para la Evaluación de la Educación). Este organismo tiene como función ofrecer a las autoridades de los sectores educativos de educación básica (preescolar, primaria y secundaria) y media superior, el desarrollo de pruebas de rendimiento de alta calidad con características de tipo criterial y matricial. En este sentido cada año se utiliza una muestra representativa en donde se valora el “grado en el que los alumnos alcanzan los objetivos de las principales áreas de los planes de estudio centrado en el currículo nacional” (INEE, 2006, p.13).

Bajo estas necesidades, el INEE conforma un plan general de evaluación del aprendizaje (PGEA), que tiene como estrategia de evaluación las pruebas Excale (Exámenes de la calidad y logro educativos) las cuales destacan:

“la complejidad e interacción de los factores que intervienen y afectan en el logro educativo, oportunidades de aprendizaje de los alumnos, coleccionar información de las diferentes esferas educativas y reconocer que el contenido de los aprendizajes se manifiestan en tres niveles de currículo: formal, implementado y logrado” (Excale, 2005, p.2).

La característica principal de éstas es evaluar cuatrianualmente a un grado de modo que de 2005 a 2016 se tengan resultados desde educación básica hasta bachillerato en todas las áreas temáticas (Español, Matemáticas, Ciencias Naturales, Ciencias Sociales).

Derivado de las políticas educacionales surge ENLACE (Examen Nacional de Logro Académico en Centros Escolares), prueba objetiva y estandarizada que mide los conocimientos y habilidades definidos en los planes y programas oficiales de estudio de educación básica (primaria y secundaria) y media superior.

Otro documento en donde se especifica qué es lo que se evalúa y las estrategias normativas para evaluar la calidad desde la política pública federal, es el que se incluye en el Programa Sectorial de Educación 2007-2012, que manifiesta la promoción de la evaluación en todos los procesos y sectores de la educación. Asimismo, contempla la evaluación como un instrumento que en su diseño pone en práctica políticas públicas por medio de indicadores pertinentes que promueven la mejora y la calidad (SEP, 2007).

En marzo de 2007 se llevó a cabo en Buenos Aires Argentina, la Segunda Reunión de Ministros de Educación para América Latina y el Caribe (PRELAC). El tema central fue la educación de calidad para todos, entendida “como un bien público y un derecho humano fundamental que los estados tienen la obligación de respetar, promover y proteger, con el fin de asegurar la igualdad de oportunidades en el acceso al conocimiento de toda la población” (UNESCO, 2007, p.5).

Lo antes mencionado da cuenta de una preocupación que existe no sólo en Latinoamérica, sino a nivel mundial, ya que promueve posicionarse en estándares de calidad de los servicios que se ofrecen en la educación, donde la medición está

determinada por niveles de competitividad social con características de gratuidad como base legal, la que se manifiesta el derecho de acceso a una educación de relevancia, pertinencia, equidad, eficiencia y eficacia; ejemplo de ello en México es el programa Alianza por la Calidad Educativa (ALIANZA, 2006).

Para el campo de la educación básica existen reformas a los planes y programas de estudio que contemplan en la evaluación atender y dar cuenta de logros en el aprendizaje de manera integral (SEP, 2006). Shmelkes (1997) opina que la calidad de la evaluación busca, en educación básica, ser orientada al desarrollo de las competencias cognitivas fundamentales de los alumnos, con el objetivo de promover la autonomía, y que las experiencias de aprendizaje sean apropiadas por el sujeto.

Es necesario resaltar que el acuerdo 384 emitido por el Diario Oficial de la Federación (2006), explica que la educación básica pretende articular la continuidad y progresión de contenidos, asegurando la congruencia de los propósitos de éstos en los diferentes niveles. Asimismo, que los planes y programas de estudio sean congruentes con los requerimientos de la sociedad. De tal manera que se busca instrumentar un proceso de evaluación con metodologías que midan la confiabilidad en cuanto al desempeño del sistema educativo, con parámetros sobre resultados de aprendizaje, recursos, procesos y contexto social (SEP, 2007).

Los resultados de PISA son otro referente de medición en la educación básica, ya que en el informe de la Organización de Estados Iberoamericanos (OEI) muestran que el Sistema Educativo Mexicano debe enfrentar dos retos importantes: el primero es que los jóvenes estén preparados para enfrentar los desafíos del contexto actual; y el segundo,

que los alumnos que tienen resultados altos sean competentes para desarrollar puestos de alto nivel (OEI, 2006).

El instrumento de evaluación PISA se aplica en los grados obligatorios de educación básica y bachillerato, en consecuencia se caracteriza por la evaluación de conocimientos y habilidades, su metodología se apoya de las pruebas matriciales y cuenta con una base de 167 reactivos de opción múltiple en donde se incluyen cuatro respuestas (opciones de falso y verdadero, si/no o de respuesta breve o corta), integrados en 85 unidades que son agrupados en diferentes versiones, por lo que cada ítem es puesto en diferentes cuadernillos con la finalidad de ser aplicado a una muestra representativa, su resolución aproximada es de 2 horas (PISA, 2005).

Otro referente es la prueba ENLACE, que consta de 50 a 70 reactivos de opción múltiple, con sólo una respuesta correcta. Las asignaturas curriculares de los programa de estudios oficial de educación básica que se evalúan son: Matemáticas, Español y Formación Cívica y Ética, en 3°, 4°, 5° y 6° de primaria, y en secundaria de 1° a 3°. Es necesario destacar que a partir del 2010, se introduce en las pruebas la asignatura de Historia, considerándose para cada grado la estructura siguiente: 3°, 35 reactivos con 1 tema y 7 subtemas del programa de estudios; en 4°, 40 reactivos 8 temas y 31 subtemas; 5°, 40 reactivos, 7 temas y 23 subtemas; en 6° 40 reactivos, 5 temas y 30 subtemas. Para el caso de secundaria: en 2° 50 reactivos contemplándose 5 bloques y 7 temas, mientras que en 3° son 50 reactivos ubicados en 5 unidades y 11 temas. Este instrumento se ha considerando como una prueba alineada al currículo nacional.

Este instrumento evaluó a 15, 766,608 alumnos en 2009 (9, 604,980 de primaria y 6, 161,628 de secundaria) situados en 133,497 escuelas del país, obteniendo los

resultados siguientes: de 2006 a 2009: existe un crecimiento continuo de porcentajes del alumnado en niveles de insuficiente y elemental (indicadores con los que son medidos estos exámenes), disminuyendo en 11.6 puntos, lo que significa que 1.1 millones de alumnos entre 3° y 6° de primaria pasaron a los niveles de bueno y excelente en Español, mientras que en Matemáticas, los niveles de insuficiente y elemental disminuyeron en 13.4 puntos porcentuales, es decir, 1.3 millones de alumnos se promovieron del nivel bueno al excelente.

Para 2010, en la asignatura de Historia se consideró una muestra de 8, 322,078 alumnos en primaria. Los resultados muestran que 62.9% manifiesta un dominio de contenidos elemental, mientras que en secundaria de 3, 431,282 de los alumnos que presentaron esta prueba 62.5% se encuentra en un nivel elemental. Estos resultados muestran el nivel de desempeño del alumno en relación a los contenidos curriculares. Se identifica que en el nivel de secundaria en las asignaturas Español y Matemáticas más de 99 mil alumnos de 3° se promovieron a los niveles de bueno y excelente (ENLACE, 2008).

Los resultados antes mencionados originan una política de evaluación del desempeño de los alumnos congruente con indicadores de calidad que se establecen en la OCDE, a través de PISA.

Se reconoce, por tanto, que el diseño de los instrumentos de evaluación que se promueven en nuestro país por medio del INEE, cumple con políticas de evaluación de calidad, teniendo como característica la medición de habilidades cognitivas focales sobre algún conocimiento particular. Su rigurosidad en cuanto a la administración y calificación e interpretación de resultados se deriva de la propuesta del National Center



for Reserch on Evaluations, Standards and Student Testing (CRESST), quien aplica criterios para revisar su calidad técnica bajo los siguientes criterios: alineamiento con el currículo nacional, calidad del contenido, complejidad cognitiva, evaluación significativa, propiedad del lenguaje, transferencia y generalización. Por tanto, la visión es realizar mediciones de desempeño, que permitan posicionarnos en estándares de calidad de la enseñanza, muestra de ello son las pruebas de ENLACE (INEE, 2005).

De esta manera, el propósito de la investigación es revisar los procesos de evaluación de organismos nacionales, con la intención de dar respuesta a una de las líneas de investigación de un proyecto operado por la Universidad Virtual del Tecnológico de Monterrey denominado. “La evaluación externa como sistema de evaluación educativa a escuelas de bajo rendimiento”. Dicho proyecto centra su propuesta en la evaluación educativa, por lo que está diseñado para escuelas públicas de educación básica (primaria y secundaria), de bajo logro académico. Este sistema permite mejorar la calidad en el servicio educativo.

Bajo este contexto, la investigación se ubicó en la primera etapa del diseño de sistemas de evaluación, de la que se destacan las siguientes fases: diseño de instrumentos con base en los criterios que realizan los expertos y selección de la muestra en donde se promovieron tanto las cartas de consentimiento como los permisos correspondientes a las escuelas para aplicar el instrumento preliminar y llevar a cabo el primer pilotaje.

De acuerdo con la Secretaría de Educación Pública, en el año escolar 1993-1994 entra en vigor la reforma de los planes y programas de estudio de la educación primaria, por lo que en sus planes de estudios vigentes, la asignatura de Historia se aborda durante los seis años, dando prioridad en el primer año a crear en el alumno la reflexión de

nociones de cambio a través del tiempo, así como de la idea de la identificación de secuencias históricas del calendario cívico; en 3º, los alumnos conocen los rasgos generales de la historia, geografía y la organización social y política de la entidad federativa en donde viven; en 4º, se tiene como propósito que el alumno identifique características principales y la secuencia de las etapas de la historia en donde se apliquen nociones de tiempo y cambio a periodos prolongados; en 5º y 6º se estudian temas de Historia Universal que van desde el origen del hombre hasta la época actual, en donde el eje es la Historia de México, por lo que el estudio es de lo inductivo a lo deductivo (SEP, 2000). Para el caso de secundaria, la formación se ubica en 2º y 3º, cuando el alumno sea capaz de analizar la realidad y actúe con una perspectiva histórica bajo la comprensión de hechos y procesos, desde lo general a lo particular, por lo que se hace una revisión de Historia Universal y en el siguiente año escolar Historia de México (SEP, 2006).

La razón por la que se retoma para esta investigación la aplicación de los exámenes perfilados y revisados por los expertos de 4º a 6º de primaria, es que en estos grados el alumno es capaz de desarrollar su capacidad para comprender procesos históricos e identificarlos, organizar e interpretar información. Por ello, sólo se aplica el instrumento a 1º y 3º de secundaria.

La aplicación de los instrumentos se realizó a los alumnos de 4º, 5º y 6º de primaria de la Escuela Primaria “Miguel Hidalgo” ubicada en el municipio de Papalotla; en 1º y 3º de secundaria en la Escuela Secundaria Oficial No. 0151, “Carmen Serdán” ubicada en la comunidad de San Miguel Tlaixpan en el municipio de Texcoco. Cada grupo que se consideró tiene entre 37 y 40 alumnos como máximo. Estas escuelas se

encuentran ubicadas en la zona oriente del Estado de México y pertenecen a la zona VII del la región de Texcoco.

### *Definición del problema*

Uno de los problemas que presenta la educación se refiere al proceso de planeación e instrumentación de escenarios que sirvan de referente de evaluación de aprendizajes del alumno con características de medición objetiva, en donde la calidad de la evaluación se encuentre centrada en la “renovación de pensamiento sobre la educación y el proceso de enseñanza, se convierta en una relación pedagógica en la que se estimule el aprendizaje significativo” (Sacristán y Pérez, 1999, p.339).

La evaluación es una tarea compartida ante todos los actores que intervienen en el ámbito educativo. Por esta razón en la educación básica tiene como regla principal una evaluación formativa, en la que se promueve la verificación del desarrollo de habilidades y competencias de la estructura curricular.

En la actualidad, todos los maestros elaboran instrumentos que pueden medir el desempeño de los alumnos, por lo que se identifica bajo este paradigma que muchas veces no son lo suficientemente claros en lo que desean medir, ya que se centran en el tema de una secuencia didáctica parcializando los conocimientos de los alumnos y, por tanto, su rendimiento (Díaz- Barriga, 2002).

Las pruebas objetivas como instrumentos objetivos y estandarizados, permiten la medición de conductas (Anastasi, 1982), de tal manera que en la búsqueda por diseñar instrumentos que permitiesen al docente emitir inferencias sobre el procesos de

pensamiento y dominios de conocimiento, se reconoce la taxonomía de Marzano y Kendall como medio para verificar, por medio de la construcción de pruebas objetivas, el logro del proceso de enseñanza y aprendizaje escolar.

Otra problemática detectada es el desconocimiento de los docentes para diseñar pruebas de diagnóstico, bajo una metodología que permita, en el proceso de construcción, tener conocimientos de qué y para qué sirve evaluar el aprendizaje. Muestra de ello es la escases de investigaciones realizadas en este campo por los profesores de educación básica. De la misma manera hay una ausencia de propuestas didácticas innovadoras que rompan con las vertientes tradicionalistas de evaluación sumativa, en donde se proporcionen elementos de aspectos formativos que reflejen el cúmulo de habilidades y aprendizajes de los alumnos y no sólo la información de una o varias unidades sin sentido (Airasian, 2002).

Una de las debilidades que se tienen en educación básica es, la existencia de profesores que tienen entre 20 y 25 años de egreso de una Escuela Normal, con paradigmas de transmitir conocimientos, mientras que las políticas actuales son el ser facilitadores de situaciones de enseñanza y aprendizaje con características de construir y verificar procesos de evaluación (Chadwick y Rivera, 1997).

Con ello, se pretende centrar la atención en la construcción de instrumentos que sean capaces de medir el desempeño académico del alumno, por medio de instrumentos objetivos como las pruebas educativas para dar testimonio del ejercicio y aplicación, tomando como referentes teóricos las bases conceptuales de Marzano y Kendall en su modelo de conducta ante el aprendizaje, y las instrumentales de Thorndike (1989) y

Lafourcade (1997), ya que estos teóricos promueven en su propuesta cómo construir pruebas objetivas con innovación evaluativa.

### *Preguntas de investigación*

Bajo este tenor se propone la siguiente pregunta general de investigación:

¿En qué medida llevar a cabo prácticas de planeación, construcción y análisis de datos de pruebas diagnósticas de rendimiento académico, favorece el mejoramiento en el diseño de este tipo de instrumentos?

Y se desprenden las siguientes preguntas específicas:

1) ¿En qué medida el uso de la taxonomía de Marzano y Kendall en el diseño de instrumentos de medición del desempeño académico permitirá emitir inferencias sobre los procesos de pensamiento y dominios de conocimiento que ha logrado alcanzar un alumno?

2) ¿En qué medida el uso de pruebas estadísticas para el análisis de resultados y niveles de dificultad favorece la toma de decisiones en cuanto a cómo debe quedar finalmente constituida una prueba diagnóstica de rendimiento académico?

3) ¿En qué medida el diseño de preguntas de opción múltiple constituye una prueba objetiva que permite medir el rendimiento académico?

### *Objetivos de investigación*

Una vez establecidas las preguntas de investigación, se plantean los siguientes objetivos:

Objetivo general: Identificar y describir los elementos que intervienen en el proceso de construcción y mejoramiento de pruebas diagnósticas de rendimiento académico, tomando en cuenta ciertos estándares de calidad.

Objetivos específicos:

1) Utilizar para el diseño de instrumentos de medición del desempeño académico la nueva taxonomía de Marzano y Kendall.

2) Describir el proceso de diseño de reactivos de opción múltiple que conforman una prueba de tipo objetiva que permite medir el rendimiento académico.

3) Analizar e interpretar los resultados y niveles de dificultad que arrojó la validación de pruebas objetivas de reactivos de opción múltiple que favorecen la eficaz toma de decisiones que impactan en el proceso evaluativo.

### *Justificación*

El desarrollo de un diseño de instrumentos de medición de desempeño académico que permitan valorar la calidad de un proceso de instrumentación con prácticas de planeación, construcción y análisis de pruebas objetivas encuentra su justificación en los siguientes puntos:

1. Mejorar instrumentos de evaluación que midan el desempeño académico a través de pruebas objetivas. Con ello se garantiza que construir una prueba

con características centradas en el aprendizaje, constituye un factor para elevar la calidad de la enseñanza como parte de un sistema de evaluación de calidad.

2. Construir instrumentos de evaluación utilizando la nueva taxonomía de Marzano y Kendall que permitan emitir inferencias sobre los procesos de pensamiento y dominio de conocimiento que ha logrado el alumno, como parte de las actividades de enseñanza de los docentes y centrados en promover una alternativa más de evaluación formal.
3. Análisis de uso de pruebas estadísticas, por lo que para dar sentido a la sistematización de la información y valorar el nivel de dificultad de la prueba, las decisiones que se realizarán son en relación a la constitución de una prueba de rendimiento académico, capaz de medir una asignatura o varias, ya que la verificación de cada componente en el diseño dará cuenta del grado de factibilidad y validez de los instrumentos de medición aplicados.
4. El diseño de pruebas objetivas con preguntas de opción múltiple, como instrumento que mida eficazmente cualquier proceso que se presente en el aula, y que garantice en su construcción la evaluación de una conducta.

### *Limitaciones*

Algunas de las limitaciones que se tuvieron en el estudio fueron: tiempos acortados para contactar las escuelas en donde se aplicaron los instrumentos,

burocratización imperante en las supervisiones que, aun cuando posibilita una oportunidad para verificar el desempeño de los alumnos, lo observan como un persecuimiento a su labor docente.

En cuanto a los docentes a quienes se les invitó como expertos para validar el instrumento preliminar, manifestaron tener tiempos limitados para la revisión del instrumento. Sin embargo, existió la disposición para la revisión antecediendo, para el caso de primaria, no ser expertos de la asignatura de Historia, ya que en su formación inicial las especificidades por asignatura no son vistas de manera puntual.



## Capítulo 2. Revisión de la Literatura

El presente capítulo tiene como finalidad explicitar las prácticas de evaluación a la construcción de pruebas objetivas como marco de referencia para la construcción de reactivos de opción múltiple, en donde para su planeación, construcción y evaluación, favorecen el rendimiento académico. En este sentido, el referente teórico que sustenta este proceso son las bases teóricas y técnicas de Lafourcade (1997), Thorndike (1987) y la nueva taxonomía de Marzano y Kendall (2007).

Es necesario hacer hincapié que el trabajo se dividió en cuatro apartados. En el primero se exponen los antecedentes que demarcan el proceso histórico de las pruebas como proceso de evaluación, hasta llegar a lo que se conoce como pruebas objetivas y su utilización en el campo educativo. Además se toca la evaluación educativa como punto de partida para entender su actual postura en la educación, así como los principios que la definen. Como parte de los antecedentes se incluye la medición como referente instrumental en la concreción de juicios valorativos y de atributo que nos permitirán medir los objetivos que se pretenden en la educación.

En el segundo apartado se revisan las bases teóricas y técnicas para la elaboración de instrumentos de medición de pruebas objetivas. Aquí se promueve los referentes teóricos de la nueva taxonomía de Marzano y Kendall como modelo de revisión de conductas de aprendizaje. Asimismo se manifiestan las bases conceptuales expuestas en las pruebas y los tipos que se utilizan en la educación para verificar el desempeño de los alumnos.

El tercer apartado trata sobre el proceso de planeación para el diseño de pruebas objetivas, por lo que se definen los objetivos de aprendizaje desde la taxonomía de Bloom y la nueva de Marzano y Kendall como parte de la construcción de objetivos más apegados al conocimiento y a las situaciones de aprendizaje en el aula. También contiene el tipo de contenido que debe abarcarse en la prueba y su concreción en una tabla de especificación del contenido.

Se finaliza con el diseño de pruebas de reactivos de opción múltiple, teniendo como base la planeación, relación de objetivos con nivel taxonómico, los elementos que se requieren para su construcción, ventajas y desventajas de diseño y las propiedades psicométricas que lo constituyen.

### *Evaluar con calidad*

El INEE desde 2003 creó un sistema de indicadores en México para evaluar la calidad de la educación, con la finalidad de dar respuesta a los objetivos de la educación a partir de la pertinencia, relevancia, eficacia interna y externa, suficiencia, eficiencia, equidad e impacto con fines de adecuarse al contexto social. De la misma manera, se exige comprender la función de la escuela y el impacto que ésta tiene en el servicio que ofrece a través de la rendición de cuentas y transparencia en los procesos (INEE, 2006b). Por lo que el objeto de la evaluación de calidad es obtener un juicio de atributo valorativo en relación al grado de cumplimiento por medio de estándares.

### *Proceso histórico de las pruebas*

En el contexto de las pruebas de carácter escrito, se encuentra el compendio de un sistema de evaluación que ha dado sentido y significado a la verificación de la medida a través de un proceso histórico. Bajo estas condicionantes se identificaron los conceptos y teorías que atañen a la elaboración de una prueba objetiva en el contexto actual. Siendo necesario ubicar teóricamente a ciertos autores (Aiken 2003; Cummings 1990; Escudero 2003; Thorndike, 1982) para contextualizar la implicación de la evaluación en cuanto a los logros educativos formales, que tiene sus inicios en los modelos; oriental, occidental y americano.

Dichos autores atribuyen a los chinos específicamente a la Dinastía Sui (289-618 d.c.), el modelo oriental, considerado como el primer sistema de evaluación, en donde el alumno debía memorizar los documentos básicos para legitimarse como futuro gobernante de un sistema aristocrático, garantizando con ello una táctica de control y selección de funcionarios.

En este siglo XXI, destaca el modelo japonés como parte de un sistema oriental que, pese haber iniciado sus procesos de evaluación en 1872, actualmente promueve el “uso de exámenes para estimular el aprendizaje”, garantizando con ello tener a los alumnos más destacados para el nivel que cursa con la caución de no deserción por causas tales como el desconocimiento de los programas por parte del docente o la escuela (Cummings *et al.*, 1990).

Otro modelo que sobresale es el occidental. Nace en Prusia en el siglo XVII como una propuesta de construcción de exámenes frente a la necesidad de seleccionar a sus

funcionarios para sus cargos gubernamentales. Alemania y Francia, en los siglos XVIII y XIX, utilizan como eje creativo para la realización de evaluaciones la elaboración de exámenes de educación que se especializaban en el rendimiento académico de sus alumnos.

En Inglaterra el proceso es tardío, nace con la necesidad y en beneficio de una clase alta. Sin embargo, durante el siglo XVIII se promueve en la Universidad de Cambridge un examen, que tenía la consigna de reconocer a sus estudiantes más destacados. Posteriormente, la Universidad de Oxford adopta el mismo procedimiento con mayor apertura dejando un legado para las escuelas públicas anglosajonas.

Por último, se revisó el modelo americano que tuvo influencia del Reino Unido. En el siglo XVIII se construyeron las primeras pruebas de selección de ingreso a estudiantes universitarios. En este siglo también destacan las pruebas educativas centradas en el examen oral, pero con poca efectividad como instrumento. De tal manera que se le atribuye a George Fisher, en 1864, la creación de la primera prueba objetiva de aprovechamiento, teniendo como característica su escala de medición confiable, misma que hasta nuestros días persiste como paradigma en la construcción de pruebas.

Escudero *et al.*, (2003) destacan la aparición de los *tests* psicométricos durante el siglo XIX los cuales, por su carácter de medición de las conductas humanas eran, reconocidos como una actividad evaluativa que tenían como características la medición como término universal en la evaluación. Los objetivos, tenían carácter de atribución de la norma y los test de rendimiento como indicador de evaluación educativa. Thorndike *et*

*al.*, (1982) enfatiza que la aparición de la medición psicológica y educativa en las escalas de inteligencia de Binet.

En 1926 se sustituyen las pruebas de ensayo por las de opción múltiple. Paralelamente, se crea maquinaria automatizada para la calificación, promoviéndose una masificación del uso de las pruebas estandarizadas para el rendimiento académico de los alumnos (Aiken *et al.*, 2003).

Después de la segunda guerra mundial crean SAT (*Scholastic Aptitud Test*) “prueba de aptitudes que se aplica por primera vez en 1947, que en su diseño se identificaba la opción múltiple” (Cummings *et al.*, 1990, p.94)

En el siglo XXI en México se crea el INEE como organismo independiente de la Secretaría de Educación Pública (SEP) que tiene como misión establecer políticas para evaluar aprendizajes a través de su Plan de Desarrollo de manera externa al Sistema Educativo Nacional. Su intención es obtener ventaja de los modelos sobre teoría curricular y aprendizaje escolar, para “proveer información del logro académico de estudiantes a nivel nacional y estatal, contribuyendo a conocer la calidad de los servicios que se ofrecen en los niveles de educación básica y media superior” (INEE, 2005a, p. 5).

Cabe recalcar que este sistema de evaluación se planteó con el fin de contar con elementos teóricos y técnicos concentrados en las pruebas nacionales denominadas Excale, que son de carácter criterial alineadas al currículo y caracterizadas por su diseño matricial (INEE, 2005a).

Por otra parte, la prueba ENLACE, ha sido una política pública implementada para proporcionar información diagnóstica del grado en el que los alumnos adquieren los

temas y contenidos de las asignaturas. Estos instrumentos de evaluación se caracterizan en educación básica por ser pruebas objetivas, estandarizadas, alineadas al currículo, centradas en el conocimiento de planes y programas de estudio, oficiales y vigentes (SEP, 2008).

### *Evaluación Educativa*

La evaluación educativa, para Lafourcade (1997), es una actividad crítica de aprendizaje, ya que por medio de ella se adquiere un conocimiento del estado real de una Institución en relación al aprovechamiento de sus alumnos. Escudero *et al.*, (2003) y Stufflebeam y Shinkfield (2005) refieren que para Tyler la evaluación no sólo se quedaba en el maestro y en su propuesta. Manifestaba que debían elaborarse taxonomías formulando objetivos en donde debían incluirse fuentes de información como son los tests, exámenes y otros criterios como revisión de juicios de valoración unificados.

La evaluación educativa como sistema de evaluación, en Tyler (1973), permite emitir juicios racionales acerca de áreas programáticas, en donde se miden los objetivos alcanzados del currículo a través del qué y el cómo de la enseñanza, es decir, el saber oficializado que impacta en los resultados que se obtienen a través de la comprobación de los objetivos alcanzados.

Por su parte, Valenzuela (2008) la define como un proceso cuyo producto, estima el grado en el que un proceso educativo favorece el logro de las metas para las que fue creado. En este tenor la revisión de un proceso, permite direccionar acciones a partir de

la identificación de un propósito en donde el objeto de evaluación estará determinado por la particularización de lo que se desea cuantificar.

Casanova (1998) y Guirtz y Palamidessi (2000) tienen como supuestos que en la direccionalidad de los planes y programas de estudio, el currículum como artificio promueve la atención a la diversidad de alumnos y a su formación, garantizando con ella una calidad que contribuye a un estado óptimo del alumno en lo cognitivo, afectivo y social.

Por tanto, la evaluación educativa en la actualidad se caracteriza por ser un proceso integral, que determina la sistematización y acumulación permanente, al atribuir y valorar los objetivos alcanzados por los alumnos, por lo que dentro de sus funciones aprueban evaluar un sistema en conjunto, sus elementos, conocimientos y habilidades de los alumnos, ya que promueve la toma de decisiones en el proceso y permite la verificación de juicios de valor por medio de la medición (INEE, 2005a).

### *Medición Educativa*

Algunos autores (Airasiam, 2002; Gronlund, 1983; Guirtz y Palamidessi, 2000; Karmel, 1974; López e Hinojosa 2001; Thorndike *et al.*, 1989) definen que la medición es un proceso que determina el grado de cuantificación y asignación de un valor ya sea numérico o en forma de letra; es un atributo que define la cualidad a medir de manera perceptible.

En este caso, la medición permite a través de sus mecanismos proporcionarnos información de tal forma que se debe seleccionar el atributo y definirlo claramente,

realizando las operaciones necesarias para ponerlo de manifiesto con los instrumentos (test o pruebas) que permiten descifrar un dato de forma estandarizada. En este sentido la medición se ocupa de hacer la valoración numérica de cómo descifrar todas aquellas posibilidades tanto cualitativas como cuantitativas del atributo de un dato generado en el juicio de valor (Lafourcade, 1997).

*Bases teóricas y técnicas para la construcción de instrumentos de medición: Pruebas  
Objetivas.*

Las bases teóricas para la construcción de instrumentos dan cuenta del estado del arte de un constructo, que permite medir con efectividad un proceso educativo que se destaca en su instrumentación por una relación objetiva entre los contenidos de currículo y la evaluación que subraye las normas y técnicas de planeación, en donde la calidad que se requiere o exigen las actuales políticas educativas centren su logro en la medición del desempeño bajo estándares de mejora en la calidad educativa por medio de la Nueva Taxonomía de Marzano y Kendall como referente para su cimentación.

*La nueva taxonomía de Marzano y Kendall*

La nueva taxonomía de Marzano y Kendall tiene como característica la facultad de ejecutar un proceso mental basado en dos factores: la complejidad inherente a los procesos en término, de control centrados en reestructurar el curriculum, la instrucción y la evaluación como se presenta en la Figura 1.



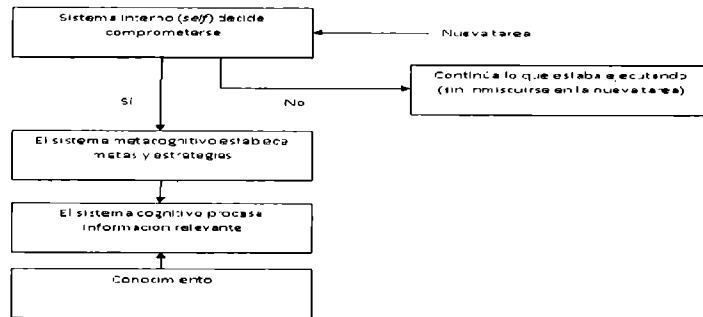


Figura 1. Modelo de conducta ante el aprendizaje (Tomado de Marzano y Kendall, 2007).

En la Figura 1 se describe el proceso humano de decidir o no centrarse en una nueva tarea en algún punto del tiempo. El modelo presenta tres sistemas mentales: el interno (self), el metacognitivo y cognitivo (Gallardo, 2009).

En los tres sistemas mentales se realiza un trabajo de interrelación, creencias y metas que se utilizan para hacer juicios en relación a una nueva tarea, sirve de motivación con posibilidades de éxito para que el nuevo saber se lleve a cabo, por lo que en estos niveles la definición de las tareas va a permitir la baja probabilidad de éxito o el impacto en ellas. “Estas nuevas tareas de la Taxonomía de Marzano permiten en el curriculum la conformación de nuevos diseños curriculares” (Marzano y Kendall, 2007, p. 15). Bajo este modelo se pretende evaluar el aprendizaje por medio de un instrumento que permita en su objetividad dar cuenta clara de qué se mide y cómo se mide el aprendizaje.

#### *Bases conceptuales: pruebas o exámenes*

Valenzuela (2008) comenta que un test o prueba es un diseño sistémico que permite medir conductas del alumno en donde el evaluador obtiene un valor numérico

del grado de conocimiento de un tema. Para Airasiam et al. (2002) es un procedimiento formal, sistemático generalmente de papel y lápiz mediante el cual se recaba información sobre el desempeño de los alumnos. Mientras que para Díaz-Barriga (2002) y Kaplan (2006) son instrumentos de control, donde se intenta la verificación del grado de rendimiento o aprendizaje logrado por los alumnos.

Las pruebas contienen propiedades psicométricas como característica metodológica. Según Valdés (2010) deben contener validez ya que permiten medir lo que se diseñó, identificándose que puede medirse por su contenido (en donde el rango mide en la prueba una conducta o habilidad), criterio (esta forma de validez de contenido es de forma criterial y se realiza por medio del juicio de expertos de la asignatura) y predicción (refiere el grado de precisión en el que las calificaciones de la prueba se correlacionan y su criterio y desempeño futuro). En cuanto a la confiabilidad, Cohen y Sherdlik (2001) manifiestan que su atributo de consistencia interna permite verificar su homogeneidad por medio del método de alfa Cronbach (por lo que deben ser reales, prudentes, representativos, precisos y consistentes).

### *Pruebas objetivas*

Dentro de las forma de evaluar el aprendizaje tenemos las pruebas objetivas que se caracterizan por ser instrumentos de medición que permiten señalar por medio de sus objetivos el logro alcanzado de los alumnos en relación al contenido.

Por lo que las pruebas objetivas en su concepto son una “medida objetiva y estandarizada de una muestra de conducta” (Anastasi, 1982, p. 21). Para Aiken *et al.* (2003) son capaces de medir objetivos de la instrucción de orden superior.

Las ventajas de la utilización de pruebas objetivas para la evaluación del desempeño escolar tiene una serie de características sobre las pautas tradicionales de evaluación por lo que cuando se centra en el rendimiento establece vínculos de reconocimiento entre la enseñanza y el aprendizaje como parte del proceso formativo en donde se busca el desarrollo de habilidades (Hargreaves, Earl y Ryan, 2000). Como ejemplo tenemos las que se aplican en ENLACE, Excale, PISA, CENEVAL, entre otros, que se caracterizan por la evaluación no sólo de contenidos, sino por su amplitud de verificación de habilidades obtenidas por el alumno y por su capacidad de respuesta a los contextos que se les pregunta.

### *Pruebas de aprovechamiento*

Lo característico de todas las pruebas en sentido estricto es la verificación del rendimiento en donde la objetividad, validez y confiabilidad son determinantes para su construcción (Thorndike *et al.*, 1989). Este tipo de instrumentos en su planificación se utilizan para conocer cuánto es que sabe un sujeto en relación a un cierto número de temas o del desempeño de habilidades. En sus propósitos destacan: el diagnóstico, la evaluación de una competencia adquirida, la asignación de una calificación, la promoción de grado o certificación, la evaluación del currículo y programa de estudios, por lo que

proporciona información confiable para la toma de decisiones en las políticas educativas (Aiken *et al.*, 2003). Otra característica que se les atribuye en la medición es que en sus ítems se estiman conocimientos y destrezas de diferentes niveles por ello pueden ser de criterio y de norma (Thorndike *et al.*, 1989).

### *Pruebas de diagnóstico*

Este tipo de pruebas tienen la función de identificar áreas de oportunidad en el aprendizaje de una asignatura. Por lo anterior, para elaborar una prueba de esta índole, se debe analizar el tipo de contenido que se incluye, ya que se realizan por grupos de reactivos para su medición. Por lo que en su calificación son capaces de medir varias subhabilidades de forma confiable (Vincenzi y Angelis, 2008).

Thorndike *et al.* (1987) afirman que la objetividad que se busca en este tipo de pruebas se refiere a la perspectiva de la evaluación con relación a la norma y la vinculación con referencia a un dominio, esto es, pruebas que respondan a determinados valores estadísticos como índice de dificultad (permite el evaluar la calidad del instrumento bajo normas o criterios, siendo la proporción de estudiantes que responden correctamente un reactivo) y poder de discriminación (permite el estimar si el reactivo lo aciertan con más frecuencia los alumnos de un grupo superior que los alumnos considerados como malos en grado inferior) (SEP, 1997).

### *Proceso de planeación para el diseño de pruebas objetivas*

En el proceso para elaborar una prueba se caracteriza una visión previa a este proceso llamada planeación, con lo que se pretende verificar el aprendizaje de los alumnos. En este sentido, el grado de apropiación de un conocimiento se medirá en relación a la dimensión pedagógica y a la discriminación de saberes por medio de una prueba objetiva con el fin de tomar decisiones que fortalezcan las acciones en el campo educativo (Valenzuela, 2008).

En el proceso de evaluación de un aprendizaje a través de una prueba objetiva se debe considerar lo siguiente, según Thorndike et al. (1989):

1. Objetivos que planifica el maestro en función de la enseñanza y que dan muestra de su desempeño académico del alumno.
2. El abordaje de los tipos de contenido y su relación con lo que se espera que el alumno aprenda en el proceso de instrucción basado en una taxonomía (habilidades o conocimientos).
3. Construcción de tabla de especificaciones.
4. Diseño conceptual y métrico de instrumento, elaboración, redacción, prueba y revisión de reactivos.
5. Análisis del instrumento en su dificultad y discriminación.
6. Análisis técnico de la prueba (confiabilidad y validez)

En este apartado se denota el proceso que se ha de seguir para el diseño de una prueba objetiva con el fin de entender la progresión de un aprendizaje y posteriormente asumir una propuesta que permita mejorar la calidad del proceso de instrumentación de una prueba de carácter objetivo.

### *Objetivos de aprendizaje*

Los objetivos de aprendizaje permiten en la evaluación determinar qué se va a evaluar, y manifiestan una clara visión de estar en función de conductas observables, medibles y unitarias por lo que deben de representar el proceso de la enseñanza y el aprendizaje así como el logro y alcance (Thorndike *et al.*, 1989). Para Lafourcade (1997) representan la base de lo que un alumno puede evidenciar a través de comportamientos específicos, por lo que propone la clasificación de estos con base en la taxonomía de Bloom.

Lo que caracteriza a los objetivos de aprendizaje según diversos autores (Huerta, 1978; Lafourcade, 1997; Marzano y Kendall, 2007; Thorndike *et al.*, 1989), es que:

1. Los objetivos deben enunciar la relación entre la conducta de los alumnos y los propósitos educativos.
2. Deben de iniciar con un verbo activo que revele su conducta y deje ver la apropiación del contenido observado y medido.
3. Deben ser precisos con significados uniformes que refieran un solo proceso de enseñanza y aprendizaje.

*Uso de la taxonomía de Bloom en la educación.* La taxonomía propuesta por Benjamin Bloom, en 1964, intenta ordenar por nivel de complejidad las conductas que se promueven en la escuela teniendo como propósito la clasificación del conocimiento organizando por categorías de orden jerárquico. Esta segregación permite para las

pruebas objetivas potenciar la razón mediante la verificación del grado de dificultad de un aprendizaje.

De tal manera que una taxonomía puede ser útil para ordenar términos, por lo que Bloom (1971) la dividió en tres campos o dominios de aprendizaje: cognoscitivo, afectivo y psicomotor. Esta clasificación ha sido el referente teórico-instrumental por excelencia que pretende en su medición una complejidad cognitiva buscando con ello evaluar diversos grados de habilidades como son la comprensión de conceptos y procedimientos, la solución de problemas y el pensamiento crítico (INEE, 2005a).

Según Lafourcade *et al.* (1997), la taxonomía de Bloom utiliza el campo cognoscitivo para evaluar habilidades y conocimientos dentro de las pruebas objetivas clasificándose en las siguientes jerarquías: conocimiento, comprensión, aplicación, análisis, síntesis y evaluación. Por lo que las áreas de oportunidad que se detectan en esta taxonomía son que, por su naturaleza, el aprendizaje se relaciona de forma simple, única, basada en un modelo conductista más que en uno constructivista, en donde los procesos son de carácter más coordinado y subordinados el uno del otro en los procesos cognitivos (Marzano y Kendall *et al.*, 2007). Aun cuando la taxonomía de Bloom realiza una expansión de la concepción de aprendizaje, la mayoría de los especialistas en construcción de pruebas coinciden en que es sólo una herramienta para el diseño de ítems (INEE, 2005a).

*Uso de la nueva taxonomía de Marzano y Kendall.* La nueva taxonomía de Marzano y Kendall es una propuesta alternativa que cumple con los lineamientos para

planear las pruebas objetivas, ya que toma en cuenta el proceso de aprendizaje desde cinco dimensiones (Marzano y Pickering, 2005): actitud de percepción, integración del conocimiento, extensión y refinación, uso significativo y hábitos mentales, que en la evaluación realiza una mediación entre el proceso de construcción de la planeación del aprendizaje y la medición del proceso mental que se desea medir por medio del objetivo de aprendizaje, medir un proceso mental en términos de control del alumno (Gallardo, 2009).

La nueva taxonomía de Marzano y Kendall tiene sus bases teóricas en la concreción de objetivos de aprendizaje de acuerdo al tipo de conocimiento con que se trabaja, está conformada por dos dimensiones: niveles de procesamiento conformado por tres sistemas: interno o self, metacognitivo y cognitivo y los dominios de conocimiento: información, procedimientos mentales y procedimientos psicomotores (Gallardo, 2009) (Figura 2).

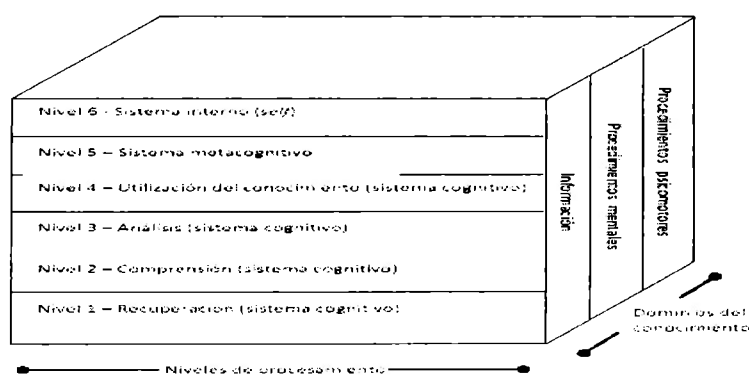


Figura 2. La nueva taxonomía (tomado de Marzano y Kendall *et al.*, 2007).

Bajo la clasificación de sus dos dimensiones que se aprecia en la Figura 2, se promueve la ubicación de objetivos de aprendizaje por niveles de especificación que



pueden utilizarse en el desarrollo curricular centrado en el desarrollo de habilidades de pensamiento, como una forma de evaluar aprendizajes a partir de la construcción de pruebas objetivas (Gallardo, 2009). Por lo que se describen cada uno de estos dominios de conocimiento y de niveles de procesamiento con el fin de darle sentido a esta propuesta teórica.

*Dominios de conocimiento.* Los dominios de conocimiento en la nueva taxonomía de Marzano permiten el clasificar el proceso de aprendizaje bajo tres componentes: Información: en ella al agregar contenidos se forma la construcción de sus esquemas imaginarios que cobran importancia para entender de qué manera se conciben los procesos mediante los procedimientos, por lo que el docente debe de proponer una práctica educativa diferente en los cuales se alude a la construcción de significados por medio del conocimiento declarativo Marzano (2000). Para organizar sus redes proposicionales dando respuesta de manera jerárquica a niveles de; vocabulario, hechos, secuencias temporales, generalizaciones, principios de causa y efecto y los principios de correlación (Gallardo, 2009).

El segundo son los procedimientos mentales: conocidos también como conocimientos procedimentales considerando el cómo, en donde el cerebro humano es capaz de construir “estructuras de producción o de tipo si-entonces” (Anderson, 1983; en Marzano y Kendall, 2007, p.16). Se busca ligar los conocimientos que se conocen o la habilidad que se está tratando de aprender en donde el alumno es capaz de resignificar (Marzano, 2000).

El aprendizaje que se implica en este tipo de procedimientos contiene tres fases; cognitiva, asociación y autónoma. Los procedimientos mentales también se jerarquizan consistiendo en interiorizar procesos que van desde macro procedimientos ubicados en reconocer modelos o principios que en los niveles descendentes construyen las habilidades que consisten en entender el proceso principal de la habilidad a desarrollar conformadas por; reglas simples, algoritmos y tácticas Marzano y Kendall et al. (2007).

Por último los procedimientos psicomotores: En este procedimiento debe almacenarse el conocimiento con el fin de utilizarlo, se practica para desarrollar la habilidad con facilidad y en el que se va instituyendo la construcción de significado considerados así por dos razones; son almacenados en la memoria como cualquier otro tipo de procedimiento mental, con base en la relación si-entonces y el segundo ubicado en donde primero se aprende o se adquieren los conocimientos para después llevarlo a la práctica y finalmente se maneja de manera automática (Gallardo, 2009).

Su jerarquía se caracteriza por procedimientos complejos tales como; fuerza estática, equilibrio corporal total, velocidad de reflejos, velocidad en movimientos de muñeca, precisión en movimientos de dedos, precisión en movimiento manual, firmeza de mano-brazo y control y precisión. Cabe destacar que en este procedimiento se considera la aportación de (Marzano y Kendall *et al.*, 2007).

*Sistemas de pensamiento.* Los sistemas de pensamiento coadyuvan, al igual que los dominios de conocimiento, a entender los procesos mentales por medio de explicar el aprendizaje a través de dos elementos: procesos mentales y clasificación de los sistemas y

subsistemas de pensamiento. Esta taxonomía retoma la postura de Anderson (1995), en Marzano y Kendall *et al.*, (2007) que manifiesta que existe una sola memoria con diferentes funciones: sensorial, permanente y de trabajo el conjunto de las tres permite en el alumno tener la conciencia de lo que se procesa en la memoria de trabajo (Gallardo, 2009).

Los sistemas de pensamiento y los tipos de conocimiento permiten concretar el objetivo de aprendizaje que será evaluado, siendo quien determine el nivel que cada reactivo debe alcanzar en una prueba objetiva.

Los sistemas de pensamiento reflejan la jerarquización de los tres sistemas; cognitivo, metacognitivo y sistema interno (self). Como se muestra en la siguiente Figura 3.

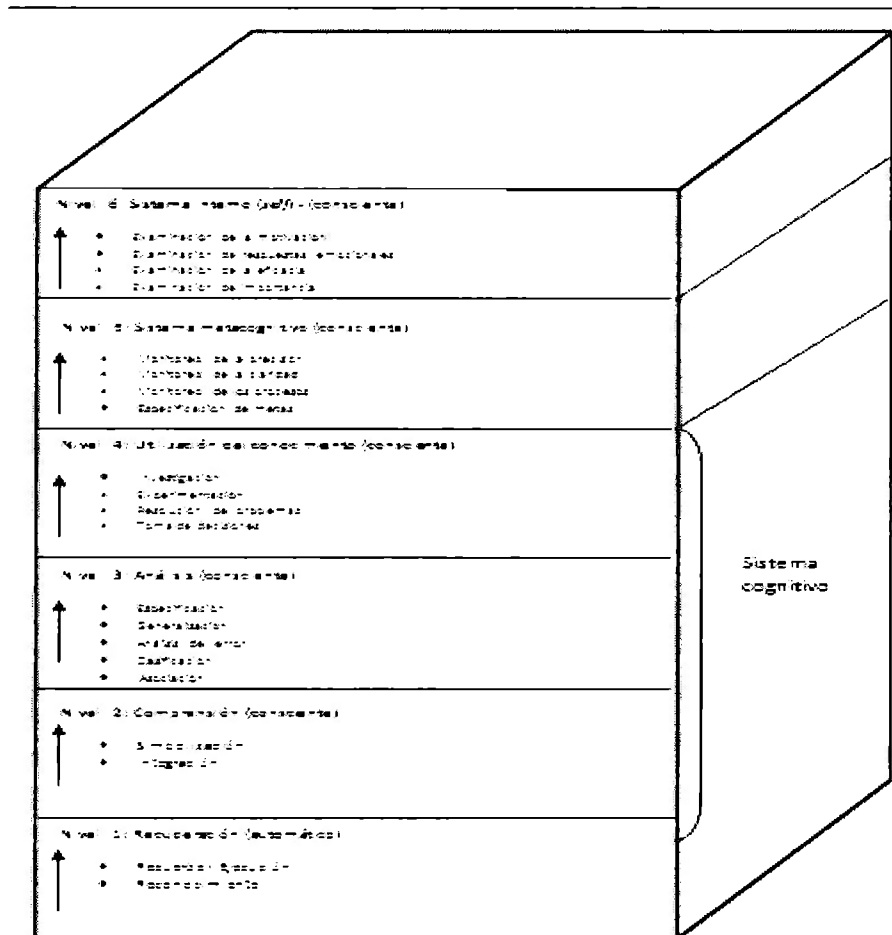


Figura 3. Operaciones mentales del sistema de pensamiento en la nueva taxonomía. Manual nueva taxonomía de Marzano (Gallardo, 2009).

En esta representación el primer nivel es el de *recuperación*, el cual en su estructuración mental del alumno pertenece al sistema cognitivo y tiene como característica activar y transferir el conocimiento de la memoria permanente al de la memoria de trabajo conscientemente procesada (Marzano y Kendall *et al.*, 2007). Bruner (1997), en relación a este proceso de recuperación, comenta que “el rasgo más importante de la memoria a largo plazo para el aprendizaje y la instrucción no es su capacidad, si no su organización interna” (37, p.).

En una prueba objetiva, este nivel de procesamiento es básico ya que en los contenidos de la disciplina de Historia se recupera información del alumno que éste sea capaz de reconocer eventos o sucesos que acontecieron e identificar personajes relevantes de forma automática.

La *comprensión* es el segundo nivel dentro del sistema cognitivo, se encarga de traducir el conocimiento de forma adecuada para su almacenaje en la memoria permanente, es decir toma, el formato para que la información clave se preserve. Las asociaciones en la memoria procesual forman reglas, proporcionando al conocimiento declarativo por medio de principios, hechos, etc. en donde la memoria logre realizar asociaciones para percibirlo, interpretarlo y recordarlo.

Según la taxonomía de Marzano y Kendall *et al.* (2007), las operaciones mentales que coadyuvan a la comprensión son: la integración y simbolización pues permiten organizar enseñanzas situadas con el fin de obtener en sus organizadores los mejores logros de desempeño en el alumno.

Si bien no compete hablar de la forma en la que se organizan los contenidos, es importante recalcar que el nivel de procesamiento sigue patrones en donde estos ordenadores gráficos le permiten al alumno, de manera eficaz, jerarquizar la información. Por ello en una evaluación por medio de una prueba objetiva permite al profesor ir construyendo los símbolos y el lenguaje centrados en un objetivo de aprendizaje. Si el alumno se enfrenta con este tipo de instrumentos de medición en donde deba de completar o identificar características, automáticamente procesa la información recibida y

realiza presuposiciones, deducciones lógicas e inferir sucesos o hechos, de tal manera que su respuesta se acerca más a lo que pretende medir en el proceso.

El *análisis* pertenece al tercer nivel de las operaciones mentales del sistema de pensamiento, en donde se manifiesta una extensión razonada del conocimiento, por lo que Marzano y Kendall *et al.* (2007) refieren que para resignificar y profundizar el conocimiento, es necesario contener procesos de; asociación, clasificar, análisis del error, generalización y especificación. Asimismo, crean un nuevo concepto denominado retroducción, que se define como el proceso para generar una idea a partir de varios casos y definir así la generalización.

La *utilización del conocimiento*, ubicada en el nivel 4, busca la satisfacción de la persona con relación a tareas específicas. Marzano y Kendall *et al.* (2007) refieren en su taxonomía que su categorización se conforma de: toma de decisiones, resolución de problemas, experimentación e investigación. Este tipo de objetivos en la planificación de un prueba objetiva son de alto impacto, ya que en ellos se pretende que el alumno sea capaz de evaluar todas las alternativas posibles con el fin de aplicar sus conocimientos, determinar fines o corroborar hipótesis.

La *metacognición*, que se encuentra en el nivel 5, pretende la verificación del proceso de control del pensamiento. Asimismo, se encarga de la ejecución en sus funciones por lo que son consideradas para la nueva taxonomía como: especificación de metas, monitoreo de los procesos, de la claridad y de la precisión. Esta fase permite la toma de decisiones en relación a los objetivos planificados determinados en lo que se quiere medir y verificar su efectividad.

Self es el último nivel de los sistemas de pensamiento por lo que existe una interrelación en el sistema interno de pensamiento en donde la toma de decisión del aprendizaje se ve intervenida por la motivación y la atención, se interponen en el proceso de aprendizaje las actitudes, creencias y emociones. Existen cuatro tipos de pensamiento que lo conforman: examinación de la importancia, de la eficacia, de las respuestas emocionales y de la motivación de acuerdo con Marzano y Kendall *et al.* (2007). Cabe destacar que las operaciones mentales presentadas en este segmento y los conocimientos conforman la interacción existente entre ellos.

Se encuentra como beneficio trabajar con la taxonomía de Marzano en la construcción de objetivos de aprendizaje para el diseño de pruebas objetivas como un proceso de evaluar aprendizaje. Por lo que se presentan los verbos que se pueden utilizar en cada una de las jerarquizaciones para los sistemas de pensamiento, de tal manera que el lector identifique el uso de la nueva taxonomía de Marzano y Kendall para su construcción que se apeguen a evaluar el desempeño académico de los alumnos (ver Tabla 1).

Tabla 1.  
*Niveles de operación mental del sistema de pensamiento y verbos que se sugieren.*

Niveles de Operación mental del sistema de pensamiento.	Subniveles.	Verbos sugeridos para redactar objetivos.
Nivel 1 Recuperación	Reconocimiento	Identificar, reconocer.
	Recuerdo	Recordar, ejemplificar, nombrar, enlistar, etiquetar, establecer, describir.
Nivel 2 Comprensión	Integración	Describir, explicar, hacer conexiones, parafrasear, resumir.
	Simbolización	Representar gráficamente, ilustrar, dibujar, mostrar, utilizar modelos, diagramar, esquematizar.
	Asociación	Categorizar, diferenciar, discriminar, distinguir, ordenar, crear analogías o metáforas.
Nivel 3 Análisis	Clasificación	Organizar, ordenar, clasificar, identificar.
	Análisis del error	Identificar, evaluar, criticar, diagnosticar, editar y revisar.
	Generalización	Concluir, inferir, proponer principios o reglas, trazar rutas.
Nivel 4 Utilización del conocimiento	Especificación	Defender, Predecir, juzgar, deducir, argumentar.
	Toma de decisiones	Decidir, seleccionar, elegir.
	Resolución de problemas	Adaptar, sobrellevar, proponer, desarrollar estrategias, resolver, alcanzar metas.
	Experimentación	Probar, examinar.
Nivel 5 Sistema Metacognitivo	Investigación	Investigar, indagar, tomar postura, explicar.
	Especificar metas	Trazar, lograr, conseguir lo que se quiere.
	Monitoreo de los procesos	Evaluar, determinar que tan bien se ha hecho algo, determinar que tan efectivo es.
	Monitoreo de la claridad	Estar claros, mejor entendimiento.
Nivel 6 Sistema interno (self)	Monitoreo de la precisión	No hay verbos sugeridos.
	Examinación de la Importancia	No hay verbos sugeridos.
	Examinación de la eficacia	Mejorar, calificar el propio desempeño.
	Examinación de la respuesta emocional	Identificar, expresar emociones y sentimientos.
	Examinación de la motivación	No hay verbos sugeridos.

En la Tabla 1 se muestran los verbos que pueden ser utilizados para ubicar el nivel taxonómico para cada uno de los objetivos, con la intención de conformar y construir



reactivos que midan dominios de conocimiento y sistemas de pensamiento en una prueba de desempeño estandarizada.

### *Contenido que debe abarcarse en las pruebas objetivas*

El contenido según Thorndike *et al.* (1989) es “un vehículo por medio del cual se logran los objetivos procesados” (p. 203). Pretendiendo con ello concretar una eficiencia de lo que se medirá del currículo escolar. Nigro (1995) y Sacristán (1995) promueven que los componentes que definen a este proceso son; el *input* o entrada del sistema y los recursos que colaboran en el proceso, las limitaciones, las estrategias en donde especifica que en ellas se garantizará la utilización del recurso considerando las limitaciones; el *feedback*, que es el conjunto de afirmaciones que el sistema proporciona en su funcionamiento, sobre todo a partir del producto que va logrando; y por último el *output*, que es el logro del objetivo planificado en una secuencia de enseñanza de carácter objetivo.

Asimismo en la actualidad el currículo escolar está permeado en el contenido a abarcarse por tres áreas de conocimiento, como refiere Díaz-Barriga (2002):

1. Contenido declarativo (saber qué): Este tipo de saber es imprescindible ya que constituye el entramado fundamental sobre el que se estructuran las pruebas objetivas, en él se ubica el conocimiento de datos, hechos, principios y conceptos. En éste, se realiza una distinción taxonómica determinada en el *conocimiento factual* (en él se refieren datos y hechos que se deben de aprender de forma literal), el *conocimiento conceptual* (en él se destaca que el

alumno debe de aprender principios, conceptos y explicaciones de manera que sea capaz de abstraer sus características que lo definen). Por tanto se deben contar con cuatro elementos que validan la emisión de juicios para rechazar o aceptar un nuevo conocimiento; bases, garantías, respaldo y calificadores (Gallardo, 2009).

2. Contenidos procedimentales (saber hacer o saber): en él se ejecutan los procedimientos, estrategias, técnicas, habilidades, destrezas, métodos por mencionar algunos. Con ello se persigue el ordenamiento de acciones dirigidas a una meta. En esta fase se promueve el uso de la metacognición.
3. Contenidos actitudinal-valora (saber ser): este proceso se manifiesta de forma gradual y lenta pues en él se manifiestan una serie de situaciones externas de contexto sociocultural, por lo que se requiere de aprendizajes significativos que incidan realmente en el comportamiento del alumno.

Con todo ello, se hace una manifestación de que el contenido implícito en un programa de estudios o en un curso, da muestra de la pretensión que se tiene en la formación del alumno que, para el caso de esta investigación, se ubica en la revisión de contenidos de educación básica (Historia) como guía para la planificación de una prueba objetiva que dé cuenta de su rendimiento académico por medio de los contenidos que se enseñan en los currículos.

Dentro de las condiciones que se utilizan para que la prueba abarque los contenidos necesarios, según Nigro (1995) y Thorndike *et al.* (1989), se destacan las siguientes: realizar un bosquejo de los contenidos y de los objetivos, abarcar por cada

combinación de contenido y objetivo del proceso la redacción de cuestionamientos que se adecuen a la aplicación de ellos de acuerdo a los tiempos escolares, asignar porcentajes a cada área de contenido y a cada objetivo hasta totalizar 100%, el número de reactivos debe estar en relación a las conductas y objetivos planificados, por último aprobar la dificultad de la prueba objetiva a través de verificar el porcentaje de examinados que contestaron bien las preguntas, con el fin de verificar la calidad del instrumento y su grado de dispersión en relación al aprovechamiento de los alumnos.

Carey (2001) manifiesta que los criterios y las normas del contenido que debe abarcarse en la prueba, se identifican por medio de variables, que influyen en el índice de dificultad y normas para determinar la calidad del producto. Lafourcade *et al.* (1997) manifiestan que el nivel de dificultad del contenido de una prueba objetiva otorga la posibilidad de discriminación entre los alumnos excelentes y los deficientes, midiendo el atributo en la elaboración de un instrumento confiable. Para Cohen (2001), la consistencia interna de una prueba que mide contenido es el grado de correlación entre todos los reactivos en una escala, con el fin de evaluar su homogeneidad (se dice que éstas son homogéneas cuando son capaces de medir un solo rasgo) por lo que entre más homogénea sea una prueba es más consistente entre sus reactivos. La heterogeneidad radica en la descripción del grado en que una prueba mide factores diferentes. Para esta investigación, las medidas de tendencia central y el método de coeficiente de alfa desarrollado por Cronbach en 1951, es considerado para estimar la confiabilidad y consistencia interna de la prueba.

Estas condicionantes que se incluyen en una prueba de papel y lápiz permitirán verificar el uso de la nueva taxonomía de Marzano y Kendall, como estrategia en la elaboración de borradores o tablas con la finalidad de que en sus casillas se sitúe el objetivo de aprendizaje el contenido que debe abarcarse, e identificar los dominios de conocimiento y sistemas de pensamiento para concretar una prueba objetiva que permita medir desempeño académico.

### *Tablas de especificaciones*

Las tablas de especificaciones permiten, según Lafourcade *et al.* (1997), contener las ideas-eje que desarrolla la unidad, bloque o modulo asimismo las conductas que se intentaron promover, por lo que advierte el equilibrio mantenido en el número de ítems. Para Thorndike *et al.* (1989), es un esbozo bidimensional que permite guiar la proporción temática contenida en un programa de estudios en relación a cada objetivo de aprendizaje, con la finalidad de determinar el número de ítems equilibrado, que en la construcción de la prueba sea capaz de medir lo que se requiere.

Cabe destacar que para otros autores (Airasian *et al.* 2002; Hills, 1982; Valenzuela, 2008), la tabla de especificaciones es una matriz que permite planificar una prueba de acuerdo al número de reactivos. Esta representación gráfica da cuenta del contenido que se abarcó en la construcción de un instrumento previo a la prueba y su diseño de reactivos acorde a la especificación del contenido y a la conducta que se desea

medir tal y como se especifica en la Figura 4, que muestra un ejemplo de contenidos de la Asignatura de Historia en 5° año de primaria.

CONTENIDOS DE LA ASIGNATURA DE HISTORIA QUINTO AÑO								
Dominios del conocimiento			Niveles de procesamiento					
a Información	b Procedimientos Mentales		Nivel 1 Recuperación	Nivel 2 Comprensión	Nivel 3 Análisis	Nivel 4 Utiliza el conocimiento		
A.1 Vocabulario	B.1 Reglas simples		1.1 Reconocimiento	2.1 Integración	3.1 Asociación	4.1 Toma de decisiones		
A.2 Hechos	B.2 Algoritmos		1.2. Recordar/ejecución	2.2. Simbolización	3.2. Clasificación	4.2 Resolución de problemas		
A.3. Sec. De eventos	B.3. Tácticas				3.3. Análisis del error	4.3. Experimentación		
Objetivos	Temas						Peso	Cantidad
Conocer los primeros años de nuestro país y los cambios que le ha sufrido con el tiempo	3. Los primeros años del México independiente	A.3		2.1			8%	13
Saber acerca de la Revolución de la Independencia	1. La Revolución de Independencia	A.3		2.1			8%	8
Conocer acerca de las reformas que tuvo nuestro país	4. La Reforma	A.3		2.1			5%	8

Figura 4. Tabla de especificaciones de contenidos de la disciplina de Historia (tomada de Gallardo, 2009).

### *Diseño de pruebas de reactivos de opción múltiple.*

Según Marzano y Kendall *et al.* (2007), el diseño de pruebas de reactivos de opción múltiple contempla un proceso que requiere considerar el bosquejo de preguntas, los principios de su esbozo y las herramientas de diseño. Bajo este tenor, los principios que se abordan en este segmento contemplarán todos esos componentes para su creación, uso y diseño de pruebas de reactivos de opción múltiple.

### *Objetivos de aprendizaje como punto de partida de la planeación de la evaluación*

El diseño del proceso de evaluación en todo ejercicio previo exige una planificación, con el fin de incidir en la toma de decisiones y mejora. En la calidad del instrumento. El primer paso es la detección de los objetivos de aprendizaje, ya que ellos permiten asegurar la direccionalidad del qué y hasta dónde llegar en el proceso de

enseñanza y aprendizaje. Thorndike *et al.* (1989) manifiestan que las características que deben tener los objetivos de aprendizaje son: enunciar términos de conducta medibles en la enseñanza y no en propósitos educativos, su formulación debe medir y ser observada en relación al verbo que connota y ser plausibles en relación al tiempo áulico en la enseñanza y el aprendizaje.

El docente trata de medir los procesos de su unidad o modulo. Sin embargo cabe la posibilidad de que si estos objetivos no tienen el sentido exacto de la conducta a medir, se puede desviar el propósito. De esta manera se destacan cinco imprecisiones que a considerarse para su planificación según Lafourcade *et al.* (1997): ser comunes a un grupo o clase de sujetos, captarse por los sentidos, ser definibles en su claridad y expresión, que sean variables en relación con los resultados y que promuevan reacciones entre los evaluadores. Si se logra evitar estas imprecisiones, la medición de la conducta dará como resultado un objetivo claro.

#### *Objetivos de aprendizaje y su relación con el nivel taxonómico de Marzano y Kendall.*

Ahora bien, cabe destacar la utilización de la nueva taxonomía de Marzano y Kendall basada en la propuesta anterior de Bloom, con el fin de construir reactivos a partir de los dominios de conocimiento y los sistemas de pensamiento, por lo que esta alianza se verá reflejada en el principio básico denominado objetivo, el cual determina en cada reactivo la complejidad que debe alcanzar su base (Gallardo, 2009). También se

contempla determinar el porcentaje a los temas que integran, con la finalidad de identificar cuantos reactivos se deben de incluir en la prueba (Tabla 2).

Tabla 2  
*Ejemplos de objetivos de aprendizaje de la disciplina de sexto año de primaria de la disciplina de Historia y su ubicación en la nueva taxonomía.*

Objetivos	Dimensiones de la nueva Taxonomía de Marzano y Kendall			
	Tipo de dominio de Conocimiento	Nivel en sistema de pensamiento	Peso	Cantidad
Al final el alumno será capaz de:				
Conocer los primeros años de nuestro país y los cambios que ha sufrido con el tiempo.	Secuencia de eventos	Integración	28%	13
Conocer el México actual.	Secuencia de eventos	Asociación	10%	5
Identificar las causas de la Revolución Mexicana.	Hechos	Recuerdo	7%	4

La Tabla 2 muestra la utilización de los objetivos de aprendizaje y su relación directa con el nivel taxonómico que permite identificar la ponderación y determinación porcentual, proporcional a los temas importantes del programa de estudios de Historia. De tal manera que se aprecia el equilibrio que se le debe otorgar a cada objetivo y la cantidad de reactivos que deben realizarse.

*Bases teórico-prácticas para construcción de reactivos de opción múltiple.*

Kaplan (2006) define los reactivos como “preguntas o problemas específicos que componen a la prueba” (p. 6), mientras que para el CENEVAL, (2010) es un “planteamiento (estímulo) que demanda cierta tarea del individuo. Su propósito es evidenciar la presencia o ausencia de un conocimiento, habilidad o destreza” (p. 4).

Existen varios tipos de reactivos conforme al proceso de planeación de la prueba que se haya adaptado, para esta investigación se utiliza el que se centra en pruebas objetivas, específicamente en el de reactivos de opción múltiple.

Los reactivos de opción múltiple (ROM) se definen como cuestionamientos o enunciados que tienen varias opciones de respuesta estructurada, de las cuales sólo una es correcta, la cual el alumno será capaz de seleccionar según su criterio. Los ROM se caracterizan por conocer el grado de aprendizaje de los alumnos en su nivel, naturaleza e índole (CENEVAL, 2010). Para Lafourcade *et al.* (1997), el reactivo consta de dos partes: en la primera se manifiesta el problema o proposición y en la segunda se destacan las soluciones por medio de opciones de respuesta, en donde mínimamente se deben presentar cuatro o cinco alternativas, destacándose que una de ellas es la correcta.

Bajo esta idea se plantea que una prueba objetiva de reactivos de opción múltiple permite con su diseño, apreciar el aprovechamiento de los alumnos a través de la planificación de reactivos estructurados, que en el ejercicio de su redacción permite la valoración de uno o varios evaluadores con el fin de que sea lo más objetiva posible para los propósitos educativos (Thorndike *et al.*, 1989).

*Constitución de la base de la pregunta.* El primer paso previo para la construcción de los ROM, constituir la base que es considerada como la oración, que se expresa de forma directa o en una oración incompleta como lo afirman Airasian *et al.* (2002) y Lafourcade *et al.* (1997). Para el CENEVAL (2010) se constituye por una pregunta, afirmación gráfica o enunciado que describe una situación o problema de forma explícita.



Por ello, la base en su representación emana el contenido que se pretende evaluar, porque en su manifestación el contenido es relevante por su nivel taxonómico, en ellos se destaca la utilización de un lenguaje sencillo y se evita la cita textual en su contenido por lo que no debe de emitir nada que debilite o confunda la mejor respuesta (Lafourcade *et al.*, 1997).

A continuación se presenta la siguiente información con la posibilidad de detectar tres diferentes formas de realización de las bases de reactivos como ejemplo ilustrativo de la constitución de bases de reactivos, de tal manera que permite retomar cualquiera de las opciones para su realización (Tabla 3).

Tabla 3

*Constitución de diferentes bases para la realización de reactivos de opción múltiple*

Base directa a través de una pregunta	Base expresada de manera incompleta:	Base de oración incompleta
1. ¿Quién fue el primer emperador de México?	1. De los emperadores que hubo en México quien es el que más se destacó por su patriotismo y amor a la Nueva España...	1. El análisis de la productividad parcial muestra: A) la baja producción por unidad de factor B) el grado óptimo que se puede lograr C) el máximo de producción que se puede obtener D) la alternativa más adecuada

Los ejemplos anteriores dan cuenta de cómo se conforma una base. Esta idea permite construir una visión de cuál es la mejor opción apegada al tipo de contenido del cual se requiere verificar y al objetivo de aprendizaje que dé cuenta de la conducta que se pretende medir.

*Opciones de respuesta.* El punto que se destaca a continuación es sobre las opciones de respuesta y su construcción Lafourcade *et al.* (1997) sugieren que se deben

seleccionar y expresar las respuestas de tal manera que sean veraces y pertinentes, y que contengan un grado de discriminación en relación a una operación mental, dando cuenta con ello de los aprendizajes adquiridos por los alumnos, de tal manera que se procura que pertenezcan al mismo campo semántico por lo que en su generalidad tienen el mismo nivel, en la respuesta correcta sólo una contiene la solución a la base (Tablas 4, 5 y 6).

*Distractores.* Gallardo (2009) y Lafourcade *et al.* (1997) explicitan que los distractores son todas aquellas opciones de respuesta que no son correctas, con ellos se puede verificar la validación y precisión en su estructura. Cada base debe contener opciones de respuesta en donde los distractores son situaciones que en su conformación “son plausibles” (Thorndike, *et al.* 1989, p. 243), creíbles y se relacionan con el tema. En ellos se incluyen los errores más comunes del sustentante, por lo que se debe de omitir en su construcción los que pueden confundir o ser parcialmente correctos. Con el propósito de clarificar las recomendaciones para su elaboración, a continuación se presenta un ejemplo (Tablas 4, 5 y 6).

Tabla 4  
*Ejemplos de elaboración de distractores en un reactivo de opción múltiple (Gallardo, 2009).*

Reactivo #1	ccss_2sec_r_001	
	Base	Según los especialistas, se ha considerado al indigenismo como:
a.	Opción correcta	Un conjunto de ideas y actividades que realizan los estados a favor de las poblaciones indígenas.
b.	Distractor 1	Un conjunto de ideas y actividades enfocadas a desprestigiar a los indígenas.
c.	Distractor 2	Un conjunto de ideas y actividades que se filman para promover el indigenismo.
d.	Distractor 3	Una corriente de pensamiento opuesta a las prácticas de los indígenas.

Como se denota en el ejemplo anterior, los distractores cumplen la función de ser excluidos por el contenido erróneo de respuesta, también realizan la función de ser la posibilidad de discriminación por dominio de contenido y sistema de pensamiento por parte del alumno. Asimismo le da énfasis a determinar por qué la opción correcta cumple con los lineamientos y sustento de aprobación en su contenido y estructura.

*Justificación de distractores.* En el proceso, los distractores conforman un grupo de opciones en las que se suma una base y una respuesta correcta, por tanto, quien los diseña debe tener conocimiento amplio en la disciplina a evaluar, ya que garantiza poseer distractores confiables, en donde su sólida elección impactará en el poder de predicción o de discriminación del reactivo (Tablas 5 y 6).

Tabla 5  
*Ejemplos de elaboración de justificaciones en un reactivo de opción múltiple (Gallardo, 2009).*

Reactivo #9	ccss_6prim_r_009	
	Base	¿Cuál era la situación de México hacia 1845 poco antes de la guerra con Estados Unidos?
a.	Opción correcta	México no podía cubrir los gastos y las deudas, había varios pleitos internos, poca gente pagaba los impuestos, había mucha inseguridad e insalubridad.
b.	Distractor 1	Se promovieron varias nuevas leyes que mejoraron la situación de México. Por fin se restableció la paz en México y fue un país consolidado.
c.	Distractor 2	México se declaró un país independiente. El comercio marítimo se hallaba suspendido. Las fronteras no estaban bien señaladas, había muchas muertes.
d.	Distractor 3	Se realizaron obras importantes, apareció el ferrocarril, y los medios de comunicación como el correo y el telégrafo.
	Justificación 1	Esto no corresponde con lo que se vivió en aquella época.
	Justificación 2	Ésta fue una época anterior a la que estamos hablando.
	Justificación 3	Ésta fue una época posterior a la que se habla.

Las justificaciones que muestra la Tabla 5 permiten aprobar la respuesta correcta por medio de la identificación de los contenidos implícitos, validando el contenido en los distractores, de tal manera que las respuestas erróneas justificadas promueven en el

alumno el descartar por medio del dominio de conocimiento. La no correspondencia del contenido con lo que pide la base permite medir en esta relación la confiabilidad de la prueba (Airasian *et al.*, 2002).

*Retroalimentación.* La retroalimentación, y las reflexiones que se generan de ésta, es un proceso en donde se dan relaciones para valorar lo aprendido, se basa en criterios claros y específicos, por lo que es de carácter compartido, y puntualiza la falta cometida durante el proceso de evaluación. Gallardo (2009) retoma de Wiggings (1998) la retroalimentación de forma clara y precisa; en ella se debe incluir los errores cometidos, revelar la precisión donde se presentan las fallas con el fin de comprender su naturaleza y mejorarlo, ser breves en la temporalidad para la retroalimentación, por lo que entre más se acorte el tiempo de ésta, permitirá mejores aprendizajes (Tabla 6).

Tabla 6  
*Ejemplos de elaboración de retroalimentación en un reactivo de opción múltiple (Gallardo, 2009)*

Reactivo #4		ccss_3prim_r_004
	Base	Son comunidades que se caracterizan por tener actividades basadas en la agricultura, la ganadería, la pesca y en la elaboración de artesanías:
a.	Opción correcta	Rurales.
b.	Distractor 1	Urbanas.
c.	Distractor 2	Internacionales.
d.	Distractor 3	Estatales.
	Justificación 1	No es la respuesta correcta ya que las comunidades urbanas son las que cuentan medios de comunicación rápidos, servicios públicos, médicos y escuelas de nivel superior.
	Justificación 2	No es la respuesta correcta ya que las comunidades internacionales se refieren a aquellas compuestas por gente de diverso origen.
	Justificación 3	No es la respuesta correcta ya que no hay comunidades estatales.
	Retroalimentación	(IN) Vocabulario- reconocimiento

En la Tabla 6 se muestra la retroalimentación, resultado de los objetivos de acuerdo con los dominios de conocimiento y sistemas de pensamiento. Las

especificaciones de está permiten ubicar la intención del proceso, que se observa en el planteamiento de la base del reactivo.

#### *Otros tipos de reactivos de opción múltiple*

Existe una diversidad de reactivos de opción múltiple por lo que en su característica existe una manifestación tanto en la estructura de la base como en la cantidad de información que se vierte en cada una de las respuestas para lo que se especifican algunos de los tipos que no se incluirán para esta investigación sino que se tiene el conocimiento que son empleados en pruebas objetivas actuales, realizadas por PISA y ENLACE.

#### *Reactivos de elección múltiple complejo*

Los reactivos de opción múltiple complejo son denominados así por el tipo de formato y las variaciones de combinaciones de opciones que se emplean como elecciones de respuesta. Su efectividad radica en estimar el conocimiento que se tenga de la causa, efectos, funciones y usos múltiples, así como la capacidad para aplicarlos e interpretarlos (Thorndike *et al.*, 1989). Sus bases son amplias ya que debe identificar el alumno relaciones entre las instrucciones que se le piden y cada una de las vertientes del reactivo. En él se destacan las distinciones que se hacen entre el contenido de cada respuesta. Por lo regular, son amplios en las respuestas, ya que se pretende medir más de un conocimiento. Sus discriminaciones implican una relación entre más de una respuesta.

### *Reactivos apareados o de respuesta por pares*

Se consideran reactivos apareados o de respuesta porque su eficacia permite medir el juicio de cantidades relativas y sus efectos en la cronología relativa. La relación entre columnas es una especificación que los caracteriza en ellos, se manifiesta la asociación o relación entre columnas en función del tipo de base que se haya determinado. Este formato debe emplearse para contenidos en donde se incluyen habilidades entre hechos o conceptos.

Ciertos autores (De Vizenci y De Angelis, 2008; Lafourcade *et al.* 1997; Thorndike *et al.* 1989) proponen que para la construcción de reactivos apareados o de respuesta eficiente se contemplen especificaciones: en las instrucciones, explicar si las repuestas pueden repetirse, las proposiciones que se construyen sean lo suficientemente homogéneas por lo que deben reflejar un solo ejercicio de apareamiento, limitar el conjunto de reactivos, guardar un equilibrio y consistencia entre columnas en cuanto a su extensión, usar para cada columna un encabezado que describa adecuadamente su contenido, procurar que el número de opciones de respuestas sea mayor que el número de enunciados que haya que responder y las opciones de respuesta deben tener un orden lógico.

### *Reactivos multi-ítem de base común*

Los reactivos multi-ítem de base común se destacan por desprenderse de una base esquematizada de información estos, son llamados de apareamiento ya que de una

pregunta se derivan varias respuestas de forma hilada, representado en un estudio de caso, un gráfico o problema a resolver (Tabla 7).

Tabla 7  
*Ejemplos de diferentes reactivos de opción múltiple (tomado de CENEVAL, 2005)*

	Otros tipos de reactivos
Reactivo multi-item	
Caso seriado (tres reactivos)	
Un paciente de 38 años de edad acude a consulta por sufrir astenia, anorexia y náusea de dos meses de evolución; hace un mes se agrega dolor abdominal, impotencia sexual e ictericia leve. El enfermo refiere que ha perdido 4 kg de peso en ese lapso. A la exploración física la tensión arterial es de 110/80 mm Hg, la frecuencia cardiaca de 82 por minuto, peso de 74 kg y talla de 1.69 m. Se observa ictericia conjuntival leve, caries múltiples, hipertrofia parótida, tórax sin datos anormales. En el abdomen se aprecia una hernia umbilical pequeña y hay hemorroides. La biometría hemática muestra hemoglobina de 11g/dl.	
	1. ¿Es muy probable que el paciente tenga antecedentes de? A) ingestión crónica de alcohol. B) exposición a plaguicidas. C) Intoxicación por plomo. D) Intoxicación por anilinas.
	Los datos principales de laboratorio son glucemia de 106 mg/dl; creatinina 1.2 mg/dl, leucocitos 4,500mm <sup>3</sup> , tiempo de protrombina alargado, hierro sérico normal, transaminasa glutámico oxalacética de 85 UI/L y albúmina de suero 2.6 g/dl.
	2. Con estos datos, el procedimiento más adecuado para el diagnóstico es: A) determinar TGP y TGO B) gammagrama de hígado. C) ultrasonido abdominal. D) biopsia hepática.
Reactivo de opción múltiple complejo	
Son dos características metodológicas del Diario del Profesor.	
I. Permite recordar los hechos más problemáticos del día.	
II. Sirve de guía para la reflexión sobre la práctica.	
III. Propicia el desarrollo de los niveles descriptivo, analítico-explicativo y valorativo del proceso.	
IV. Favorece que los alumnos del grupo escriban sus emociones y experiencias de aprendizaje.	
	A) I, 3. B) 4, 2. C) 1, 4. D) 2, 3.

Los ejemplos de la Tabla 7, dan cuenta de otros tipos de reactivos utilizados en pruebas objetivas y su objetivo es la medición de contenidos que abarcan una transdisciplinariedad o simplemente miden habilidades en algunos casos.

#### *Ventajas y desventajas del diseño de reactivos de opción múltiple*

Las ventajas en el diseño de reactivos de opción múltiple según Valenzuela *et al.* (2008) son: fáciles de calificar, por lo que cubren grandes cantidades de materiales curriculares, pueden estar escritos en cualquier nivel del dominio cognitivo, están diseñados de tal manera que el alumno tenga la capacidad de diferenciar la mejor opción entre varias opciones correctas en diversos grados y se pueden mejorar si se usan técnicas de análisis de reactivos.

Se determina que las desventajas son: su elaboración requiere de ciertas habilidades por parte del evaluador, por el contenido a evaluar. Se considera complicada en su instrumentación, tanto para la creación de la base como para las opciones de respuesta. Estos exámenes permiten al alumno adivinar la respuesta, aunque hay sistemas (nueva taxonomía de Marzano) de penalización por respuestas erróneas, para evitar que esto suceda.

Todo lo anterior permite hacer una valoración de qué es lo que se pretende medir en la evaluación de aprendizajes, aun cuando se considera que son pruebas objetivas que requieren de mucha elaboración y dominio de lo que se pretende evaluar, son confiables en su puntuación pues no le permiten al alumno la dispersión de la respuesta y se especifica en el contenido un objetivo de aprendizaje.



### *Propiedades de las pruebas de reactivos de opción múltiple: Psicometría*

La psicometría es un conjunto de modelos formales que posibilitan la medición de variables psicológicas y establecer bases para efectuarlo adecuadamente (CENEVAL, 2010). Casanova (2004) refiere que, el enfoque psicométrico estudia rasgos aislados de la personalidad (inteligencia, fluidez verbal, estilo cognitivo etc.) cuya valoración separada se ofrece mediante una puntuación.

La prueba de opción múltiple en sus propiedades “representa, una posibilidad de evaluar a las personas por sus aptitudes, habilidades y conocimientos” (Castrillón, 2003, p.102). Debe cumplir con ser: *objetiva*, su puntuación se produce al margen del juicio personal del evaluador e incluye una muestra de los contenidos; *confiable*, por la consistencia del atributo interna y externa entre reactivos determinado en la homogeneidad de la prueba, al medir un solo rasgo por medio del método de coeficiente alfa Cronbach, (Cohen, 2001).

Por último, debe contener *validez* ya que su “juicio es basado en las inferencias extraídas de la puntuación de la prueba” (Cohen, 2001, p.184) y de la valoración de expertos, esta validez es medida en tres categorías; validez de contenido (describe un juicio que concierne a la unidad de análisis y a la muestra en donde se aplicó dicha prueba con relación al contenido que se abarcó en un curso de una disciplina), validez relacionada con un criterio (juicio mediante criterios o normas en donde es valorada la prueba o la puntuación de la prueba) y validez de constructo (es el juicio de lo apropiado de las inferencias para explicar o describir el comportamiento) (Cohen, 2001).

### **Capítulo 3. Enfoque metodológico**

En este apartado se hace una descripción general del enfoque metodológico utilizado en la investigación así, como la forma en la que se recolectaron los datos a partir de los instrumentos aplicados y su decodificación. La metodología utilizada es de carácter cuantitativo, por lo que dentro de ella se especifica que se hace una aproximación al campo de la investigación aplicada (Giroux y Tremblay, 2004). Determinado por un instrumento denominado prueba objetiva de reactivos de opción múltiple que en su diseño hace uso de la nueva taxonomía de Marzano y Kendall. El propósito es dar respuesta a la pregunta general de investigación, con la finalidad de proponer instrumentos de valoración que midan rendimiento académico.

#### *Descripción de enfoque*

El enfoque que se contemplo por las cualidades del problema de investigación es cuantitativo porque se adhiere a principios de ciencia en donde se estudian fenómenos que obedecen leyes de funcionamiento y razonamiento lógico para descubrir el orden de las cosas por medio de la medición y el análisis de datos (Giroux y Tremblay, 2004) pues su conformación, como lo plantean Hernández, Fernández y Baptista (2008), contiene elementos que se encuentran “relacionados entre sí: los objetivos que persigue la investigación, las preguntas de investigación, la justificación de la investigación, la viabilidad del estudio y la valuación de las deficiencias en el conocimiento del problema” (p. 47).

Posteriormente, se realizó una búsqueda de información que permitió ubicar el estado del arte y dar respuesta a la pregunta de investigación general; ¿En qué medida llevar a cabo prácticas de planeación, construcción y análisis de datos de pruebas diagnósticas de rendimiento académico, favorece el mejoramiento en el diseño de este tipo de instrumentos?, Por lo que derivado de la revisión de la literatura se utilizó el método de mapeo determinado por el índice general como vertebra para la conformación de un marco teórico con fuentes primarias directas que dio cuenta del fenómeno visto desde los antecedentes. La investigación persiguió, en este sentido, explicar la naturaleza de los datos con el fin de que su recopilación respuesta a fenómenos estudiados (Giroux y Tremblay, 2004).

Una vez identificado el bagaje teórico, se realizó un análisis conceptual de tal manera que se planteó el objetivo general y los específicos con el fin de orientar la investigación y dar sentido a la recolección de datos.

De igual manera esta revisión se orientó, a identificar otras perspectivas teóricas, históricas y conceptuales acerca del tema, como pauta para determinar el alcance del estudio identificándose variables para agrupar medidas mediante índices, ya que como propone Giroux y Tremblay (2004), permiten aprehender una realidad en conjunto, en su combinación describe el efecto de los factores siguientes: rendimiento escolar, prácticas de planeación, construcción y análisis de datos de pruebas diagnósticas de rendimiento académico, procesos de pensamiento y dominio de conocimiento que alcanzó el alumno desde la nueva taxonomía de Marzano y Kendall, índices de dificultad de reactivos,

índices de discriminación, análisis de distractores y estándares de calidad ubicados en cada uno de los objetivos de la investigación.

Las escalas de medición utilizadas son de razón, ya que se atribuyeron valores numéricos al fenómeno estudiado. La justificación para omitir el uso de una hipótesis es porque no la investigación se ajusta a la afirmación anticipada a la pregunta de investigación (Giroux y Tremblay, 2004).

Desde este punto de vista, el alcance inicial del estudio es exploratorio (permite el familiarizarse con un tópico desconocido o poco estudiado) y descriptivo (ya que se identificó lo que se medirá variable y se recolectaron datos sobre personas-profesores y grupos-alumnos) (Hernández *et al.*, 2008).

En esta investigación se utilizó, para la interpretación de variables, el estudio descriptivo, como estrategia para la aplicación y decodificación de los datos arrojados, un instrumento diseñado por un grupo de expertos en la asignatura de Historia del Tecnológico de Monterrey, validado por profesores de educación básica y aplicado a los alumnos de diferentes niveles como muestra representativa. La recolección de datos permitió confrontar las variables y validar el instrumento con el fin de promover una prueba objetiva que se utilice y aplique en estudios posteriores.

Para la recolección de datos, se elaboró un plan en donde se consideraron variables y cualidades de los participantes con el propósito de obtener de las fuentes fidedignas, confiables y objetivas la información ubicada en la muestra. Los instrumentos de medición utilizados fueron la entrevista a profesores expertos en la asignatura de Historia, examen o prueba preliminar validada por los expertos antes mencionados, lista

de cotejo para la validación y pruebas aplicadas a los alumnos de educación básica. Estos instrumentos de medición representan el registro de los datos contenidos en las variables, decodificadas por medio de medidas de tendencia central (Hernández *et al.*, 2008).

Se destaca que el estudio no se queda en el plano de la descripción, sino que además identifica, utiliza, analiza e interpreta algunas correlaciones entre las prácticas que permiten mejorar la calidad e instrumentos capaces de medir el desempeño académico, específicamente, las pruebas objetivas de reactivos de opción múltiple utilizando la nueva taxonomía de Marzano y Kendall como parte de una práctica en la mejora de la calidad de un instrumento de evaluación en la educación básica.

Se justifica para esta investigación el uso de los principios de Marzano y Kendall como una teoría de vanguardia, que permite la construcción de reactivos de opción múltiple en pruebas objetivas, en las que se miden dominios de conocimiento y sistemas de pensamiento que dan cuenta del desempeño ya sea de un alumno o del conjunto de alumnos de una escuela o varias (Figura 5).

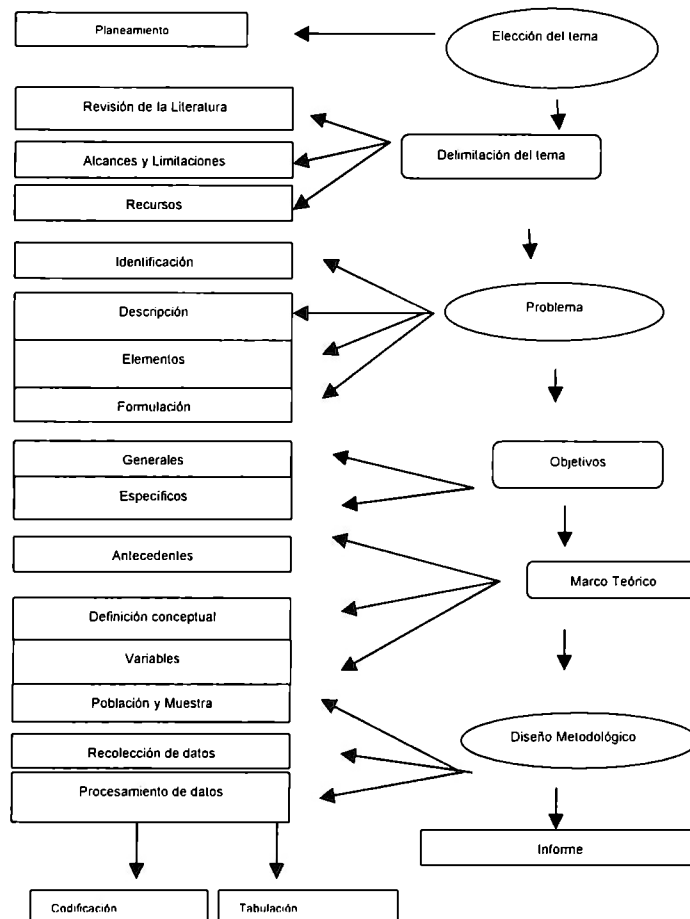


Figura 5. Esquema del modelo de investigación cuantitativa a seguir (Giroux y Tremblay, 2004).

### *Diseño de investigación*

El diseño es, según Wenger (2001), una colonización sistémica, planificada y reflexiva del tiempo y del espacio al servicio de una tarea, mientras que para Hernández *et al.* (2008) es un plan o estrategia concebida para obtener la información que se desea. Ante esta visión de lo que implica el diseño de una investigación, es conveniente precisar que entre más objetiva sea una investigación tiene la posibilidad de aportar conocimientos a la comunidad educativa.

Dentro de esta investigación se encuentran dos diseños que fueron seleccionados para llevarla a cabo: el diseño transeccional exploratorio (permite el conocer una variable o varias a partir de recopilar datos en un momento único con el fin analizar su incidencia e interrelación) y diseño transeccional descriptivo (tienen como primordialidad indagar la incidencia de las variables en una población) (Hernández *et al.*, 2008).

La combinación de estos diseños permitió, en el primero, hacer una exploración por medio de la revisión de la literatura, identificándose que en el contexto educativo se construyen pruebas de diagnóstico con reactivos de opción múltiple (PISA, ENLACE y Excale), por lo que los puntajes obtenidos en dichas pruebas de papel y lápiz cumplen con ubicar tanto a las escuelas como a los alumnos bajo rankings y estándares de calidad. Los profesores sólo se basan en este tipo de pruebas para especificar las habilidades complejas que tienen los alumnos dejando de lado el nivel de apropiación de los contenidos apegados al currículum.

El uso del segundo diseño que se eligió cuenta con características de “verificar la incidencia de las modalidades o niveles de una o más variables en una población” (Hernández *et al.* 2008, p. 210). Bajo esta lógica, la descripción de estas variables permitió identificar el tipo de participantes y la correlación con éstas.

Para aplicar los instrumentos se utilizó como unidad de análisis el Estado de México como la posibilidad de llevar a cabo para la aplicación de la muestra y recolección de los datos, de acuerdo a las características poblacionales y por el planteamiento del problema de investigación.

## *Método*

En este apartado se describe la forma en la que se utilizó la investigación científica. Giroux y Tremblay (2004) y Hernández *et al.* (2008) destacan que es una forma de adquirir conocimientos bajo ciertos paradigmas de forma sistémica, empírica y crítica. Este principio nos sirve para darle sentido de pertenencia al problema de estudio que aquí se plantea, reconociendo el uso de un conjunto de posturas incluidas en el método de investigación determinando y especificando el tipo de población, así como los instrumentos de aplicación en las unidades de análisis.

### *Participantes*

Este proyecto define como unidades de análisis a las escuelas públicas de la región escolar VII del Estado de México, a los profesores expertos que evaluaron la calidad del instrumento preliminar y a los alumnos a los que se les aplicaron los instrumentos reformulados a partir de los resultados del juicio de los expertos.

Así, la población comprende a los profesores expertos en educación básica (primaria y secundaria) y alumnos de entre 8 y 14 años de edad de las escuelas primarias y secundarias de la Región Escolar VII del Estado de México, que tiene una tasa de cobertura en 2008 del nivel educativo primaria general de 101.8 hombres y 102.9 de mujeres, lo que equivale a 6,430 de alumnos en primaria general de 10,998 de otros subsistemas atendido por 1669 docentes; en Secundaria general es de 82.1 hombres y 85.8 mujeres, que equivale a 864,365 alumnos en secundaria general de 1, 407,928 de



otros subsistemas atendido por 1332 docentes (INEE, 2007). Por lo que es necesario destacar que esta cobertura pertenece a 8 municipios.

Los profesores expertos que se seleccionaron de entre 1669 docentes tienen como característica ser profesionales de la educación en su formación inicial: son egresados de las escuelas Normales de Texcoco, Los Reyes y Superior de Toluca; de los 15 docentes seleccionados, 9 (7 mujeres y 2 hombres) ejercen la docencia en nivel primaria 3°, 4°, 5° y 6°, destacándose sólo un docente con estudios de maestría en Ciencias de la Educación por la Universidad del Valle de México.

De secundaria se seleccionaron a seis docentes (cuatro mujeres y dos hombres) de ellos, dos han realizado estudios de posgrado en Ciencias de la Educación por la Universidad del Valle de México y La Normal Superior del Estado de México. De la población especificada se destaca su amplia experiencia en los niveles educativos que oscila entre 13 y 37 años de servicio en sus planteles. En el nivel de secundaria los docentes son egresados de la especialidad de Ciencias Sociales e Historia y actualmente son titulares de esta última asignatura.

Por otra parte el total de los alumnos a los que se les aplicó el instrumento reformulado por los expertos es derivado de una población escolar de 80 alumnos en 4° de primaria, 80 en 5° y en 6° de primaria de 73 alumnos. En secundaria, de 185 alumnos en 1° y 187 de alumnos 3°. Las características no experimentales presentadas por Hernández *et al.* (2008) sugieren seleccionar muestras probabilísticas, con ello se descarta la aplicación y realización de pruebas objetivas en 1°, 2° de primaria y 2° de secundaria. Por lo que se reduce la muestra a sólo un segmento de la población, con la

idea de que todos los elementos tienen la posibilidad de ser elegidos. Sin embargo así se disminuye el error estándar para los fines de aplicación de la prueba objetiva.

Ahora bien, el Estado de México se divide en regiones escolares. Los municipios de Papalotla y Texcoco pertenecen a la región VII, esta información contextual cobra importancia para esta investigación ya, que en ahí se contempla la recolección de datos y el muestreo probabilístico.

Es necesario destacar el grado de marginación que existe en estos municipios ya que da cuenta del contexto socioeconómico, identificándose que Papalotla y Texcoco se encuentran en niveles bajo y muy bajo. El primero se destaca porque atiende en primaria un total de 724 alumnos en 2 planteles de educación pública atendidos por 21 docentes, y en secundaria, a 1,102 alumnos en 2 planteles atendidos por 53 docentes, también de educación pública; mientras que el segundo atiende a 31,136 estudiantes de primaria en 107 planteles con una capacidad docente de 1,038 profesores y a 15,282 estudiantes de secundaria en 58 planteles con características de inclusión de escuelas públicas y privadas atendido por 849 docentes (CONAPO, 2005).

La primaria “Miguel Hidalgo” con clave del centro de trabajo 15EPR0508L, se encuentra situada en el municipio de Papalotla. Esta institución de carácter público tiene dos grupos de estudiantes de 1º a 6º con un aproximado de 36-40 alumnos por salón. En la región es considerada como una escuela incluida en programas de calidad educativa, por lo que los resultados de su desempeño en la prueba ENLACE muestran, lo siguiente (Tabla 8).

Tabla 8

*Puntajes promedio de los alumnos de primaria, 2009 (tomado de resultados ENLACE, 2008).*

Grado Escolar Primaria	Puntaje promedio en 2009, obtenido en la Escuela Primaria "Miguel Hidalgo"	Puntaje promedio en 2009, obtenido a nivel Estatal	Puntaje promedio en 2009, obtenido a nivel Nacional
3°	549	552	529
4°	502	503	516
5°	543	508	516
6°	557	515	510

Como se observa en la Tabla 8, los promedios de puntaje obtenidos están por encima de los nacionales exceptuando a 4° grado, infiriéndose que tienen un buen desempeño en la prueba de ENLACE.

La secundaria oficial No.0151 "Carmen Serdán", con clave del centro de trabajo 15EES0405I, se ubica en el municipio de Texcoco, en una de sus comunidades denominada San Miguel Tlaixpan, y se caracteriza por ser una escuela de alto desempeño, en ella se contempla la existencia de 5 grupos por cada nivel con un aproximado de 35-40 alumnos por salón a quienes se les aplicó la prueba ENLACE obteniendo los siguientes resultados (Tabla 9).

Tabla 9

*Puntajes promedio de los alumnos de secundaria, 2009 (tomada de Resultados ENLACE, 2008).*

Grado Escolar Secundaria	Puntaje promedio en 2009, obtenido en la Escuela Secundaria Ofic.No.0151 "Miguel Hidalgo"	Puntaje promedio en 2009, obtenido a nivel Estatal	Puntaje promedio en 2009, obtenido a nivel Nacional
1°	514	504	497
2°	529	504	497
3°	534	522	513

En la Tabla 9 se percibe que todos los grados se encuentran por encima del puntaje promedio nacional, infiriéndose que se tiene buen desempeño en la prueba de ENLACE.

La información antes mencionada contempla el por qué se utilizaron como muestra las escuelas antes mencionadas, pues tienen un parámetro de medición en el que no se tiene ingerencia pero se utiliza como referente de desempeño del alumnado.

Con ello se justifica el uso de una muestreo probabilístico por racimos, ya que como especifica Hernández *et al.* (2008), implica diferenciar entre las unidades de análisis (quienes son medidos) y la unidad muestral (acceso a la unidad de análisis), determinándose que las primeras son las escuelas seleccionadas, profesores expertos y los alumnos a quienes se les aplica el instrumento preliminar, y la segunda es la validación del contenido de los exámenes y la aplicación de los instrumentos reformulados en las escuelas seleccionadas.

Para el cálculo de esta muestra se utilizó la siguiente fórmula:  $n' = s^2 / V^2$ .

Así, la selección aleatoria por número de estrato determinado por cada racimo se estableció tal y como se muestra en la Tabla 10.

Tabla 10

*Selección aleatoria por número de estrato de cada nivel educativo en educación básica (primaria y secundaria).*

Estrato	Número de alumnos en Población de nivel Primaria	Número de alumnos a los que se les aplicó el instrumento preliminar	Número de expertos que evaluaron la calidad del instrumento
4°	80	40	3
5°	80	40	3
6°	69	35	3
Total	N=229	n=115	n=9
Estrato	Número de alumnos en Población de nivel Secundaria	Número de alumnos a los que se les aplicó el instrumento preliminar	Número de expertos que evaluaron la calidad del instrumento
1°	185	38	3
3°	187	37	3
Total	N=372	n=75	n=9

Los criterios que se utilizaron para la selección de la muestra estuvieron determinados por aquellas escuelas que se destacarán por sus puntajes en la prueba ENLACE. En los expertos que evaluaron, la calidad del instrumento fue tener como mínimo 5 años de ejercer en la docencia, conocer los planes de estudios vigentes, así como el enfoque de la asignatura de Historia y estar frente a grupo ejecutándola. En los alumnos a los que se les aplicó la prueba, tener entre 8 y 15 años de edad que es el promedio en la que se sitúan a los niveles especificados en la investigación y pertenecer al grupo piloto.

### *Instrumentos*

Los instrumentos de medición utilizados para esta investigación son exámenes preliminares, que se elaboraron por especialistas del Tecnológico de Monterrey,

basándose en los planes y programas de estudio de la asignatura de Historia vigentes en la educación básica (primaria 3º, 4º, 5º y 6º y secundaria 2º). En su estructura son pruebas objetivas con reactivos de opción múltiple, que se caracterizan por contener una base y cuatro respuestas. En su diseño, midieron dominios de conocimiento y sistemas de pensamiento destacándose para ello la utilización de la nueva taxonomía de Marzano y Kendall.

Cada uno de los exámenes cuenta con el siguiente número de reactivos en la asignatura de Historia:

1. 3º de primaria 24 reactivos y se aplicó a 40 alumnos del grado posterior, es decir, 4º año.
2. 4º de primaria 39 reactivos y se aplicó a 40 alumnos del grado posterior, es decir, 5º año.
3. 5º de primaria 66 reactivos y se aplicó a 35 alumnos del grado posterior, es decir, 6º año.
4. 6º de primaria 50 reactivos y se aplicó a 38 alumnos del grado posterior, es decir, 1º año de secundaria.
5. 2º de secundaria 54 reactivos y se aplicó a 37 alumnos del grado posterior, es decir, 3º año de secundaria.

Antes de aplicar este instrumento fue verificado por expertos, los cuales evaluaron su calidad por medio de una lista de cotejo que midió la validez, entendida según Cohen (2001), como el atributo que se le asigna por medio de términos como aceptable o débil por medio de un juicio de utilidad.

Asimismo, las categorías se basaron en validez de relación, con un criterio en donde se destaca la calidad de las preguntas que conforman el examen preliminar. En tanto, se evaluó la calidad de cada reactivo en cuanto a su forma (planteamiento) y contenido (temas abordados, respuesta correcta, y distractores así como justificación de las mismas). Estos instrumentos después de haber sido evaluados por los docentes y haber hecho las modificaciones sugeridas por los mismos, se aplicaron en el contexto, por lo que la duración de las pruebas fue de 1 hora con 45 minutos con variación de entre 2 y 3 minutos en el caso de 3° y 4° de primaria.

Una vez aplicadas las pruebas preliminares, se realizó el primer vaciado de la información en hojas de Excel especificándose la siguiente información: nombre del profesor experto, fecha de evaluación de la prueba, nombre del entrevistador, nombre de la escuela en que labora, clave de la escuela, grupo al que le da clase, en caso de secundaria materia que imparte, fecha de nacimiento, género, formación profesional, institución en donde realizó estudios de licenciatura, formación profesional de la maestría e institución en donde la estudió.

Posteriormente, en ese mismo archivo de Excel se insertaron datos de los resultados de la evaluación de la prueba de rendimiento de forma general por cada grado y por la valoración negativa o positiva de los expertos hacia cada reactivo codificándolo con 0 o 1. Asimismo, en un vaciado específico, se le asignaron valores tales como 1 = deficiente, 2 = regular y 3 = mala de acuerdo a los criterios establecidos para evaluar cada pregunta.

La información que se derivó de este concentrado permitió verificar la validez interna de la prueba y permitió tomar decisiones, como lo manifiesta Cohen (2001), bajo la teoría de las decisiones y utilidad de las pruebas de Cronbach, que permiten asignar un término como el de índice de aciertos, índice de fallas, falso positivo y falso negativo especificado bajo el coeficiente de validez. Asimismo, se pretende la validez de constructo que dé cuenta de la homogeneidad de la prueba.

En el segundo vaciado se anotaron los datos que arrojaron los exámenes en cuanto a la respuesta correcta y no correcta, asignándole valores de 1 y 0 con el fin de obtener una sabana de datos que tuviera la suma de los índices de dificultad por reactivo, así como la calificación por alumno. Estos estadísticos permitieron determinar, 30% de lo alumnos con mayor rendimiento y del menor rendimiento, por cada uno de las pruebas realizadas totalizando cinco instrumentos valorados de esta forma, obteniendo con ello el índice de dificultad de la prueba.

Según Hernández *et al.* (2008), el enfoque cuantitativo de la investigación tiene como estrategia la recolección de los datos para explicitar deductivamente las variables con base en la medición numérica, por lo que para el análisis de la confiabilidad se utilizaron las medidas de tendencia central derivadas de los intervalos de razón (puntos en una distribución) que a continuación se describen:

1. Moda (puntuación que ocurre con mayor frecuencia)
2. Mediana (mitad de casos o posición intermedia de la distribución)

Su fórmula es:  $N+1/2$



3. Media (promedio aritmético de una distribución se simboliza X)

$$\text{Su fórmula es: } \bar{X} = \frac{X_1 + X_2 + X_3 + X_4}{N}$$

Estas tres formulas fueron utilizadas para sacar las medidas de dispersión apoyadas en la teoría de Aiken (2003):

1. Rango (diferencia entre la puntuación mayor y la menor siendo el valor de la calificación más alta y la más baja). Su fórmula es:  $X_M - X_m$
2. Desviación estándar (promedio de desviación de las puntuaciones con respecto a la media por lo que representa la variabilidad entre los reactivos de la prueba)
3. Varianza (mide las variabilidades de forma aleatoria)

Lo que se obtuvo fueron las calificaciones, los máximos y mínimos valores de la prueba, con lo que se verificó su simetría. Asimismo, se verificó la confiabilidad de las pruebas mediante el método de consistencia interna del coeficiente de alfa Cronbach, para medir en la calidad, según Carey (2001), la dificultad que tuvo la prueba para el grupo, la capacidad de discriminación de alumnos que tienen mayor conocimiento del dominio de habilidad y la plausibilidad de grupo.

Las medidas antes mencionadas permitieron en la investigación obtener el valor mínimo y máximo de los alumnos en la prueba. Asimismo, permitió sacar los índices de discriminación en donde se ve reflejado el desempeño de cada alumno por reactivo y en

general. En cuanto al índice de dificultad, se determinó mediante el porcentaje de los alumnos que contestaron correctamente a cada pregunta.

Por último, el método de coeficiente alfa Cronbach, al trabajar con variables de razón, midió la consistencia interna de la prueba que oscilaron entre 0 (nula confiabilidad) y 1 (confiabilidad total), determinándose en cada prueba aplicada la confiabilidad (Cohen, 2001), por lo que se utilizó la siguiente fórmula estadística para sacar el coeficiente:

31 Mediante la varianza de los ítems
--------------------------------------

$$\alpha = \frac{K}{K-1} \left[ 1 - \frac{\sum I_i^2}{T^2} \right]$$

### *Procedimientos*

En esa intención, se tiene como estrategia del proyecto de investigación para la recolección idónea de datos, la contemplación de las siguientes etapas en el proceso: 1) Planeación de ubicación del contexto para la validación de los instrumentos (permiso a supervisiones y cartas de consentimiento de los expertos para evaluar la calidad del instrumento); 2) Proceso de planeación temporal para verificación por parte de los expertos que evaluaron la calidad del instrumento (validación del contenido por medio de lista de cotejo para evaluar la calidad de los instrumentos); 3) Modificaciones y reformulaciones realizadas por los expertos y que se contemplaron en la realización del instrumento preliminar; 4) Muestreo poblacional (ubicación del contexto para aplicación de instrumento preliminar); 5) Pilotaje de Instrumento preliminar; 6) Captura de resultados; 7) Análisis e interpretación de resultados del juicio de los expertos y

resultados de aplicación de los exámenes, verificando los niveles de dificultad que arrojo la validación y la consistencia interna vista en la confiabilidad del instrumento (por medio de la estadística descriptiva, específicamente por medidas de tendencia central y coeficiente alfa Cronbach); 8) Validación de instrumentos y toma de decisiones.

### *Estrategias de análisis de datos*

Esta fase permitió la elaboración de un informe técnico sobre los resultados de la aplicación de las pruebas objetivas de opción múltiple que miden rendimiento académico desde la nueva taxonomía de Marzano y Kendall; elementos que fueron el sustento de la investigación. Las estrategias que se utilizaron para el análisis de datos son la utilización de las unidades de análisis, que se dividió en dos: la primera parte fue la de exploración, en donde se identificó a los especialistas (15 de 4 escuelas diferentes), quienes validaron el instrumento o prueba. Lo que se pretende con esa información es verificar si el profesor con base en las listas de cotejo emitió la información correcta con relación al conocimiento que tiene sobre los planes y programas de estudio vigentes. Asimismo, si los objetivos de aprendizaje ubicados en cada reactivo corresponden al nivel taxonómico de Marzano y Kendall. Se analizó si las pruebas reflejan las habilidades los profesores detectan los problemas previos a la aplicación del instrumento en donde se incluyen la tres características del instrumento para que este sea de calidad: dificultad para el grupo, capacidad de discriminación entre los estudiantes que poseen más conocimientos y habilidades en un dominio, así como la plausibilidad del elemento para el grupo por medio de las listas de cotejo (Carey, 2001).

La segunda parte fue la aplicación de los instrumentos reformulados. A partir del juicio de los expertos se realizó la verificación de las variables operacionales propuestas en esta investigación por medio de: vaciado de datos en hojas de cálculo estadístico del programa computacional de Excel, en donde se sistematizó la información y se sacaron medidas de tendencia central (moda, mediana y media). Posterior a ello, se clasificaron los alumnos que obtuvieron el mayor puntaje y el menor. Asumiendo las siguientes variaciones en la información (Tabla 11).

Tabla 11.  
*Resultados de las medidas de tendencia central*

Medidas de Tendencia Central				
Grupos/Alumnos	Número de Reactivos	Media	Mediana	Moda
4º/40	24	45.20	41.6	41.6
5º/40	39	49.61	48.71	38.46
6º/35	66	53.29	54.55	59.09
1º/38	50	47.58	49	50
3º/37	54	47.05	48.15	53.70

De la información vertida en la Tabla 11, se identificó en la moda la calificación de mayor frecuencia, en la mediana, el valor de 50% que se encuentra por debajo de las calificaciones, y la media representa la sumatoria de todas las calificaciones dividida entre el número de calificaciones. Esta estrategia permitió identificar los niveles de logro de los estudiantes en su desempeño académico y la estandarización de cada una de las pruebas.

La primera medida que se evaluó de los datos es la calidad de la prueba a través de la dificultad que tuvo para el grupo, utilizando el índice de dificultad por medio del elemento  $p$  (proporción o porcentaje de alumnos que respondieron correctamente a la

prueba que van desde 0.00 a 1,00). Otro mecanismo son los criterios y las normas establecidas en la lista de cotejo y en las variables (complejidad de la habilidad, características del logro del grupo, la comparación de los índices de dificultad que se midieron en cada habilidad y la comparación de los índices de dificultad de medir habilidades jerárquicamente.), con el fin de verificar la calidad de la prueba y evaluar la calidad de la enseñanza.

El segundo análisis, da a los índices de discriminación de cada prueba por medio de la utilización de los rangos en los que se establece el número máximo y el mínimo de acuerdo a su puntuación, midiendo con ellos la capacidad de cada elemento de la prueba. Los criterios que se establecieron dependieron de factores inherentes a la habilidad medida en la nueva taxonomía de Marzano y Kendall y las características de los estudiantes en su desempeño académico. El análisis de estos datos reflejó si la prueba fue homogénea o heterogénea por medio de las medidas de dispersión en donde el acercamiento o concentración de puntuaciones promovieron la consistencia (desviación estándar).

La consistencia interna de la prueba estima la frecuencia entre reactivos dentro de una escala, como afirma Cohen (2001), se determinó con ello la homogeneidad por el método de coeficiente de confiabilidad alfa Cronbach, ya que con ello se obtuvo un juicio de valor de la eficacia de la prueba y del índice de error de la construcción de la prueba en su interpretación por lo que se trató que la medición se mantuviera libre de errores (Valenzuela, 2008) (Tabla 12).

Tabla 12  
*Resultados de las medidas de dispersión.*

Grado escolar al que se aplico prueba	Medidas de Dispersión				
	Valor mínimo obtenido	Valor máximo obtenido	Desviación estándar	Varianza	alfa Cronbach
4° primaria	16.67	83.33	12.60	154.99	0.1671
5° primaria	25.64	82.05	12.67	156.6	0.9308
6° primaria	15.15	72.73	12.12	142.79	0.9889
1° secundaria	24	72	13.44	175.82	1.0062
3° secundaria	20.4	79.6	11.27	123.54	0.9993

En la Tabla 12 se muestra el vaciado de las medidas de dispersión con la finalidad de revisar la validez y confiabilidad del contenido en donde se midieron: procesos de pensamiento y dominios de conocimiento que ha logrado alcanzar el alumno y realizar inferencias de logro de aprendizaje de los alumnos, identificar niveles de dificultad de la prueba y el diseño de preguntas dentro de la misma.

La validez de criterio, midió la estandarización de la prueba que, en palabras de Hernández *et al*, (2008), son las puntuaciones obtenidas por los alumnos en la prueba, en donde se realiza una correlación de los puntajes más altos y de los más bajos para identificar el rango en el que se encuentran. La validez de constructo ubicó de forma jerárquica, si el alumno contestó el reactivo correctamente e incorrectamente diagnosticando el funcionamiento de la prueba.

## Capítulo 4. Análisis de resultados

El presente capítulo presenta el análisis de resultados en tres secciones, en la primera se trata: el proceso de validación de los expertos con relación al banco de reactivos, por lo que se especifican cada una de las acciones que realizaron los validadores y se dan ejemplos de las modificaciones que se realizaron a las pruebas en donde se destaca la toma de decisiones para su reestructura.

En el segundo apartado se hace una correlación entre los resultados emitidos en cada una de las pruebas, con relación a los índices de dificultad de la prueba y de discriminación, con el fin de verificar los resultados emitidos en los puntajes de los alumnos y su desempeño académico, también se muestra la homogeneidad y heterogeneidad de cada una de las pruebas mediante la confiabilidad y validez expuesta en el método de alfa Cronbach.

En el tercer apartado se concluye con las inferencias en donde se especifica la correlación existente entre las pruebas y la nueva taxonomía de Marzano y Kendall.

### *Proceso de validación*

En el proceso de validación se destaca el análisis de los resultados a partir de dos secciones: en la primera se incluye la forma en la que el dato numérico representa las cualidades de cada uno de los participantes que evaluaron el instrumento preliminar (Tabla 13), en el que se incluye la validez del contenido de cada uno de los reactivos, destacándose la medición de cada habilidad o conocimiento en cuanto a su esencialidad, utilidad y necesidad de los contenidos que abarcan (Cohen, 2001). Mientras que en la

segunda se presentan los resultados obtenidos del instrumento lista de cotejo (Apéndice B), que permitieron en la evaluación una toma de decisión en relación a como quedaron conformadas las pruebas de 3º, 4º, 5º, 6º de primaria y 2º de secundaria tomando ciertos estándares de calidad.

De la aplicación de los exámenes perfilados determinado como último instrumento 003 (Tabla 13), una vez que se ha recurrido a un diagnóstico de validez predictiva y confiabilidad mediante el método de consistencia interna por medio del coeficiente alfa Cronbach, se verifica la toma de decisiones en relación a cómo debe quedar conformada finalmente una prueba objetiva. Asimismo, se determina su homogeneidad o heterogeneidad.

Por último, en este proceso se emiten inferencias con respecto al uso de la taxonomía de Marzano y Kendall con relación a los procesos de pensamiento y dominios de conocimiento que alcanzaron los alumnos.

En este tenor, se busca que el lector comprenda la correlación entre las preguntas específicas derivadas de la pregunta general de investigación y la utilización de cada instrumento, para comprender la forma en la que se analizaron los resultados arrojados mediante la siguiente matriz (Tabla 13).



Tabla 13

*Matriz de datos del proceso de validación de los instrumentos y su aplicación con las preguntas de investigación y variables.*

Preguntas específicas	Objetivos específicos	Variables	Datos: Docentes expertos		Datos: alumnos
			Instrumento 001	Instrumento 002	Instrumento 003
¿En qué medida el uso de la Taxonomía de Marzano y Kendall en el diseño de instrumentos de medición del desempeño académico permitirá emitir inferencias sobre los procesos de pensamiento y dominios de conocimiento que ha logrado alcanzar un alumno?	Utilizar para el diseño de instrumentos de medición del desempeño académico la Nueva Taxonomía de Marzano y Kendall.	Prácticas de planeación, construcción y análisis de datos de pruebas diagnósticas. Rendimiento escolar. Procesos de pensamiento y dominio de conocimiento desde la Nueva Taxonomía de Marzano y Kendall. Índice de dificultad. Índice de discriminación. Estándares de calidad	Exámenes preliminares de Primaria y Secundaria	Lista de cotejo con criterios para evaluación de exámenes	Exámenes perfilados a partir del juicio de los expertos de Primaria y Secundaria
¿En qué medida el uso de pruebas estadísticas para el análisis de resultados y niveles de dificultad favorece la toma de decisiones en cuanto a cómo debe quedar finalmente constituida una prueba diagnóstica de rendimiento académico?	Describir el proceso de diseño de reactivos de opción múltiple que conforman una prueba de tipo objetiva que permite medir el rendimiento académico				

En la Tabla 13 se confirma que el uso de la sistematización del proceso de investigación es coherente con las acciones posteriores, de los instrumentos y el resultado derivado de cada uno de ellos, en contraste con el planteamiento inicial.

Una vez que se realizó la revisión de la literatura y se concretó el marco teórico que guió la investigación, se recibieron vía electrónica por parte de la tutora titular del Tecnológico de Monterrey, el banco de datos que debía suministrarse a los participantes. Que en este primer acercamiento de validación, fueron los profesores expertos que revisaron el instrumento preliminar. Posteriormente, se realizaron las gestiones necesarias en la supervisión escolar del Municipio de Texcoco (parte de las unidades de análisis) en donde se presentó el proyecto de investigación para que se proporcionarán las facilidades de acceso a las bases de los docentes que cubrían el perfil para validar el banco de reactivos elaborado por el Tecnológico de Monterrey.

Por lo que se determinó realizarles una entrevista a los docentes, quienes validaron los instrumentos, en sus escuelas de origen para hacerles la invitación y firmaran las cartas de consentimiento (Apéndice A) que explicitaba el compromiso adquirido, así como la confidencialidad y anonimato de sus posturas. Una vez firmadas las cartas se realizó una reunión con los directivos de las escuelas y los docentes expertos, en la que se presentó el proyecto y entregaron los instrumentos. Cabe destacar que los 15 expertos mostraron una actitud de respeto e interés en la participación del trabajo.

Después se entregaron las pruebas preliminares, de acuerdo al grado y las listas de evaluación de reactivos de opción múltiple (listas de cotejo) (Apéndice B) para que evaluaran el instrumento, indicándose que el dominio de su especialidad era de gran importancia para la revisión de las pruebas, ya que éstas comprendían temas del programa de estudios de la asignatura de Historia. Los criterios fueron revisar cada reactivo de opción múltiple, sus opciones de respuestas, distractores y justificaciones.

Esta valoración permite, según Lafourcade *et al.* (1997), conocer la calidad de un instrumento en sus elementos teóricos y técnicos. En esa reunión se acordó el tiempo de entrega, disponiéndose de 12 días para la validación. Cabe mencionar que todos los profesores cumplieron con lo establecido.

Ahora bien, las características que contienen las pruebas de reactivos de opción múltiple de los exámenes preliminares de 3°, 4°, 5° y 6° de primaria y 2° de secundaria que se les proporcionaron a los expertos validadores, son las siguientes:

1. Todos los reactivos de las pruebas pertenecen a los diferentes programas de la asignatura de Historia, por lo que sus temas fueron escogidos de acuerdo a la importancia del enfoque del plan de estudios.
2. Cada uno de los reactivos se realizó conforme a los propósitos de la asignatura en donde se objetivo del tema para medir los dominios de conocimiento y niveles de procesamiento de acuerdo a la nueva taxonomía de Marzano y Kendall.
3. Por cada prueba se realizó una tabla de especificaciones que da cuenta del equilibrio de temas de la asignatura y muestra la cantidad de reactivos utilizados (Figura 4).
4. Cada reactivo contiene una base y cuatro opciones de respuesta, de las cuales una es la correcta y las otras tres son distractores, justificándose para cada una de las incorrectas el por qué no eran la opción correcta. Asimismo, se incluye la retroalimentación en donde se concreta el resultado de los aprendizajes de

los alumnos conforme a la nueva taxonomía de Marzano y Kendall (Tablas 6 y 14).

En cuanto al instrumento de evaluación de reactivos de opción múltiple como parte del proceso de validación de contenido (lista de cotejo), sus características son que: presenta las instrucciones en donde se explica el proceso que se sigue para validar las pruebas preliminares, contienen información general sobre el experto (datos personales, formación profesional, género, experiencia laboral y datos de la institución en la que trabaja), incluyen los criterios para evaluar cada uno de los reactivos a partir de cuatro indicadores.

Tabla 14  
*Niveles de pensamiento y dominios de conocimiento de las pruebas de Historia.*

Grado Escolar	Número de reactivos	Niveles de sistema de pensamiento							Dominios de conocimiento	Porcentaje
		Recuperación	Comprensión	Análisis	Utilización del conocimiento	Sistema cognitivo	Sistema interno	Información		
3° primaria	24	*						*		100%
4° primaria	39	*						*		100%
5° primaria	66	*						*		100%
6° primaria	50	*	*					*	*	99%
2° secundaria	54	*						*		100%

Los bancos de reactivos se caracterizan en 3°, 4°, 5° de primaria y 2° de secundaria. El nivel de sistema de pensamiento que permea en los reactivos se encuentra ubicado en el de recuperación (se ubican en subniveles de recuerdo y reconocimiento) y el dominio de conocimiento es el de información (en los subniveles de vocabulario, hechos y generalizaciones) de un 100%.

Con respecto a 6° de primaria 98% de los reactivos tienen como característica, estar realizados bajo el nivel taxonómico en el sistema de pensamiento de recuperación con subniveles de recuerdo y reconocimiento. En su dominio de conocimiento en la información, el otro 2% se ubica en el siguiente nivel, el de comprensión y en dominio de conocimiento el de procedimientos mentales de asociación.

En la Tabla 14 se observa que todos los reactivos concretaron sus objetivos de aprendizaje en el campo del conocimiento declarativo implicado en el sistema de pensamiento uno, que tiene que ver con el sistema cognitivo básico, en donde la relación que se tiene con los estímulos crea representaciones que producen cierta información en la memoria para recordar o reconocer (Marzano *et al.*, 2007).

#### *Descripción de trabajo realizado por los expertos*

En el proceso de evaluación que desempeñaron cada uno de los expertos se destacan los resultados siguientes en cuanto a la valoración realizada por ellos de tal manera que se observan el número de reactivos revisados, el porcentaje de comentarios y los porcentajes de las observaciones en redacción y ortografía (Tabla 15).

Tabla 15

*Resultados de la valoración de los expertos para validar los instrumentos.*

Grado escolar	No. reactivos	Comentarios en relación a la Claridad en el lenguaje (%)				Respuesta correcta (%)	Ajustado al nivel (%)	Observaciones en Redacción y Ortografía (%)
		Deficiente	Regular	Mala	Total			
3° primaria	24	4.16	25	8.3	37.5	100	100	20.8
4° primaria	39	17.9	2.5	2.5	23	100	100	17.9
5° primaria	66	3	3	0	6	100	1.5	4.5
6° primaria	50	0	6	2	8	100	8	8
2° secundaria	54	3.7	25.9	0	30	100	3.7	22.2

En la Tabla 15 se manifiestan gráficamente los datos de los 3 expertos por cada uno de los niveles de educación básica, destacándose que los comentarios en relación al porcentaje de claridad en el lenguaje, el total representa los comentarios del 100%.

En 3° de primaria se consideran los expertos que la prueba tiene 25% de claridad en el lenguaje y 20.8 % de incongruencias de redacción y ortografía. Para 4° de primaria se considera que 17.9% del lenguaje en la prueba no es claro asimismo en redacción y ortografía; para 5°, consideran los expertos que es deficiente y regular que 1.5% no tiene el nivel taxonómico para el cual fue pensada la prueba y tiene un 4.5% de redacción y ortografía. En 6° de primaria la prueba considera que su claridad en lenguaje es regular, mientras en su nivel taxonómico que nos indica que un 8% se propone realizar ajustes o escribir el nivel con el que fue valorado, además de realizar ajustes de ortografía y redacción.

Por último, en 2° de secundaria los expertos consideran que la prueba, en 25.9%, tiene un lenguaje regular, mientras que en 3.7% no tiene el nivel taxonómico en su diseño y en 22.2% tienen observaciones en redacción y ortografía.

*Resultados para la toma de decisiones en relación al instrumento y su conformación final*

Ahora corresponde explicar cómo los expertos realizaron la validación por cada uno de los grados, en donde se destacan las diferencias y semejanzas de los criterios aplicados a la validación, por lo que sólo se presentan ejemplos de algunos reactivos más representativos en donde se realizaron ajustes o en donde se toma la decisión de dejar el reactivo tal cual por falta de argumento consensado por los expertos para su cambio.

Tabla 16

*Resultados del número y porcentaje de reactivos a los que se realizaron comentarios por parte de los docentes expertos validadores.*

Grado Escolar	Total de número de reactivos de cada prueba	Expertos que validaron instrumento	Reactivos que tuvieron comentarios.	Porcentaje total de comentarios realizados por los expertos.
3° primaria	24	001,002 y 003	3, 7, 11, 16, 18, 20, 22, 23 y 24	3 3.33%
4° primaria	39	004,005 y 006	5, 7, 12, 17, 20, 21, 22, 32 y 34	23.07%
5° primaria	66	007,008 y 009	5, 37, 38 y 40	6.06%
6° primaria	50	010,011 y 012	5, 8, 9 y 12	8%
2° secundaria	54	013,014 y 015	2, 3, 4, 7, 8, 9, 16, 17, 25, 29, 32, 37, 40, 46, 49 y 52	29.62%

En los resultados de la Tabla 16 se pueden observar que: la prueba de 3° tiene 33.33% (8 reactivos) de observaciones con relación a la claridad en el lenguaje o a la redacción u ortografía (Apéndice C). En 4°, 23.07% (9 reactivos) presenta observaciones en el mismo rubro anterior, mientras que en 5°, 6.06% (4 reactivos) presenta observaciones en claridad en el lenguaje, nivel taxonómico, redacción y ortografía con lo que se observa que este es el examen que presenta menor porcentaje de observaciones, al igual que 6° de primaria, con 8% de (4 reactivos), en donde la diferencia es que presenta comentarios en relación a la claridad del lenguaje y la redacción y ortografía.

Por último, la prueba preliminar de 2° de secundaria (Tabla 16) especifica que en 29.62% (16 reactivos) de la prueba existe un mayor porcentaje de comentarios por parte de los expertos en cuanto a la claridad en el lenguaje, nivel taxonómico y redacción y ortografía. Con ello, ahora corresponde mostrar algunas ejemplificaciones de la toma de decisiones que se realizó en relación a los comentarios de los expertos.

En tercer grado los expertos comentan que el reactivo No. 7:

#7	ccss_3prim_r_007
Base	Son dos de las fuentes de información que necesita un historiador para elaborar la historia de un pueblo:
a. Opción correcta	Objetos y documentos.
b. Distractor 1	Comidas y cultura.
c. Distractor 2	Materiales y documentos.
d. Distractor 3	Objetos y comercio.
Justificación 1	No es la respuesta correcta ya que las comidas y materiales no son fuentes de información para un historiador.
Justificación 2	No es la respuesta correcta ya que la comida y la cultura por sí solas no son fuentes de información para un historiador.
Justificación 3	No es la respuesta correcta ya que el comercio es una actividad y no una fuente de información.
Retroalimentación	(IN) Generalización- reconocimiento



El experto 001 comenta del reactivo lo siguiente:

No puede ser un distractor ya que los “materiales” como concepto pueden ser para los niños en la asignatura de Historia los documentos (papeles) por lo que se sugiere que se realice el cambio (Entrevista docente experto-001).

El experto 002 de este mismo grado comenta:

En el reactivo la redacción de la base no es la más acorde ya que da indicios de respuesta, es muy obvia se debe cambiar por un cuestionamiento (Entrevista docente experto-002).

Por lo que se determina que el reactivo debía quedar conformado en su estado original. Con ello se puede observar que existen diferencias de juicio valorativo con relación a la estructura del reactivo ya que para el experto 001 se debe cambiar un distractor mientras que para el experto 002 tiene que cambiar la base del reactivo, determinándose por el investigador dejar el reactivo de acuerdo a como aparece en la banco de reactivos, la toma de decisión para no cambiarlo es que no se encuentra un argumento en los expertos que tenga mayor precisión, pues se formuló en términos de medir el reconocimiento de una fuente (Tenbrink, 2002).

Ejemplo de semejanzas:

En el reactivo número 9 perteneciente a este 3° convergen los expertos 001, 002 y 003 en que se repiten las justificaciones por lo que la propuesta para el cambio es la siguiente en su opinión. El reactivo estaba conformado de la siguiente forma:

#9		ccss_3prim_r_009
	Base	Los olmecas se establecieron en La Venta lo que actualmente es el municipio de:
a.	Opción correcta	Huimanguillo.
b.	Distractor 1	Nacajuca
c.	Distractor 2	Centla
d.	Distractor 3	Paraíso
	Justificación 1	No es la correcta ya que no se establecieron en este municipio.
	Justificación 2	No es la correcta ya que no se establecieron en este municipio.
	Justificación 3	No es la correcta ya que no se establecieron en este municipio.
	Retroalimentación	(IN) Hechos- recomendaciones

Los comentarios de los expertos fueron:

... La justificación dos como se repite yo pondría que “no es la correcta por qué no se establecieron en esta región” y en la tres que “no es la correcta por qué no se establecieron en esa localidad”. (Entrevista docente experto-001.)

...La justificación dos del reactivo debe quedar” no es la correcta ya que no se establecieron en esa comunidad”; en la tercera justificación” no es la correcta ya que no se establecieron en esta región”. (Entrevista docente experto-002.)

...No estoy de acuerdo que se tengan en la justificación 2 y 3 las mismas justificaciones por lo que cambiaria en lugar de municipio en la 2 por comunidad y 3 por estado. (Entrevista docente experto-003.)

Derivado de las valoraciones de los tres expertos, se toma la decisión de modificar la justificación 2 y 3, en donde bajo la revisión del programa de estudios permitió identificar, que la palabra de municipio no abarca un concepto de espacialidad y agregándose las de comunidad y región. Por lo que el reactivo quedó conformado de la siguiente forma:

#9		ccss_3prim_r_009
	Base	Los olmecas se establecieron en La Venta lo que actualmente es el municipio de:
a.	Opción correcta	Huimanguillo.
b.	Distractor 1	Nacajuca
c.	Distractor 2	Centla
d.	Distractor 3	Paraíso
	Justificación 1	No es la correcta ya que no se establecieron en este municipio.
	Justificación 2	No es la correcta ya que no se establecieron en esta <i>comunidad</i> .
	Justificación 3	No es la correcta ya que no se establecieron en esta <i>región</i> .
	Retroalimentación	(IN) Hechos- recomendaciones

A continuación se presentan las semejanzas en la prueba de 4° de primaria en donde se destacan los siguientes ejemplos.

Ejemplo de semejanzas que se detectaron. El reactivo en su estado preliminar se conforma de la siguiente manera:

#7		ccss_4prim_r_007
	Base	En el Valle de México se levantó un pueblo que se convertiría en la ciudad más grande de Mesoamérica.
a.	Opción correcta	Teotihuacán.
b.	Distractor 1	Xochicalco.
c.	Distractor 2	Bonampak.
d.	Distractor 3	Mitla.
	Justificación 1	No es la respuesta correcta ya que Xochicalco no era más grande que Teotihuacán.
	Justificación 2	No es la respuesta correcta ya que Bonampak no era más grande que Teotihuacán.
	Justificación 3	No es la respuesta correcta ya que Mitla no era más grande que Teotihuacán.
	Retroalimentación	(IN) Hechos- reconocimiento

....El reactivo 7 tienen en su base que cambiarse la redacción ya que no se entiende de lo que se está hablando. (Entrevista docente experto-005.)

....La base del reactivo en su estructura, considero que se tiene que cambiar en su redacción para que el lenguaje sea mejor identificado por el alumno.  
(Entrevista docente experto-006.)

Posteriormente, a partir del conocimiento y argumento, se verifica el reactivo y se toma la decisión de cambiar su estructura de la redacción de la base, ya que dos expertos argumentan que no es claro el lenguaje utilizado quedando el reactivo de la siguiente manera:

#7	Número	ccss_4prim_r_007
	Base	En el Valle de México se levantó un pueblo que se convertiría en la ciudad de más <i>concentración urbana</i> : Mesoamérica.
a.	Opción correcta	Teotihuacán.
b.	Distractor 1	Xochicalco.
c.	Distractor 2	Bonampak.
d.	Distractor 3	Mitla.
	Justificación 1	No es la respuesta correcta ya que Xochicalco no era más grande que Teotihuacán.
	Justificación 2	No es la respuesta correcta ya que Bonampak no era más grande que Teotihuacán.
	Justificación 3	No es la respuesta correcta ya que Mitla no era más grande que Teotihuacán.
	Retroalimentación	(IN) Hechos- reconocimiento

En esta prueba se destaca que los expertos participantes no tienen discrepancias en las aportaciones que hacen, más bien se manifiestan semejanzas en todos los casos.

Ejemplo de semejanzas. El reactivo se conforma de la siguiente manera en su estructura preliminar:

#5		ccss_4prim_r_005
	Base	La venta es uno de los centros de la Cultura Madre mesoamericana, estamos hablando de los:
a.	Opción correcta	Olmecas.
b.	Distractor 1	Zapotecas.
c.	Distractor 2	Aztecas.
d.	Distractor 3	Toltecas.
	Justificación 1	No es la respuesta correcta ya que la cultura zapoteca no fue la primera en Mesoamérica.
	Justificación 2	No es la respuesta correcta ya que la cultura azteca no fue la primera en Mesoamérica.
	Justificación 3	No es la respuesta correcta ya que la cultura tolteca no fue la primera en Mesoamérica.
	Retroalimentación	(IN) hechos- reconocimiento

El reactivo número 5 de la prueba en su base no es clara la redacción por lo que considero que se debe de cambiar. (Entrevista docente experto-004.)

En la pregunta 5 la base no es clara se debe de cambiar la redacción. (Entrevista docente experto-005.)

Después de la valoración queda estructurado de la siguiente forma:

#5		ccss_4prim_r_005
	Base	Uno de los centros ceremoniales de la Cultura Mesoamericana que se ubica en el municipio de la venta está representado por:
a.	Opción correcta	Olmecas.
b.	Distractor 1	Zapotecas.
c.	Distractor 2	Aztecas.
d.	Distractor 3	Toltecas.
	Justificación 1	No es la respuesta correcta ya que la cultura zapoteca no fue la primera en Mesoamérica.
	Justificación 2	No es la respuesta correcta ya que la cultura azteca no fue la primera en Mesoamérica.
	Justificación 3	No es la respuesta correcta ya que la cultura tolteca no fue la primera en Mesoamérica.
	Retroalimentación	(IN) hechos- reconocimiento

Ahora bien, los reactivos de 5° de primaria que se revisaron por los expertos 007, 008 y 009, nos explican lo que encontraron al validar la prueba.

Ejemplo de diferencias. Se presenta la forma en la que estaba estructurado el reactivo previo a que se tomara la decisión de cambiarlo. Aun cuando sólo un experto el que realiza el juicio valorativo se realiza el cambio sin afectar el objetivo ni la retroalimentación. Se toma en cuenta este ejemplo ya que ningún especialista sugiere realizar cambios en las justificaciones.

#37		ccss_5prim_r_037
	Base	Fue el gran impulsor del desarrollo en esta época:
a.	Opción correcta	El comercio entre ciudades
b.	Distractor 1	La religión entre pueblos
c.	Distractor 2	La plática entre gobernantes
d.	Distractor 3	El casamiento entre reyes
	Justificación 1	No es precisa
	Justificación 2	No es precisa
	Justificación 3	No es precisa
	Retroalimentación	(IN) Hechos- reconocimiento

Según comentario del experto:

En la base del reactivo 37 la redacción y el lenguaje no permiten que el alumno comprenda lo que se pregunta. (Entrevista docente experto-007.)

Posteriormente, cuando se reestructuró el reactivo, quedó de la siguiente manera:

#37		ccss_5prim_r_037
	Base	<i>¿Qué fue lo que impulsó el desarrollo durante el Renacimiento?</i>
a.	Opción correcta	El comercio entre ciudades
b.	Distractor 1	La religión entre pueblos
c.	Distractor 2	La plática entre gobernantes
d.	Distractor 3	El casamiento entre reyes
	Justificación 1	No es precisa
	Justificación 2	No es precisa
	Justificación 3	No es precisa
	Retroalimentación	(IN) Hechos- reconocimiento

Continuando con esta valoración, corresponde verificar los datos de 6° de primaria en donde los expertos comentaron lo siguiente.

Ejemplo de diferencias. Se toma la decisión de realizar los ajustes en la redacción del reactivo que se encontraba establecido de la manera siguiente:

#5		ccss_6prim_r_005
	Base	<i>¿Cómo se llamaba el partido que formo Valentín Gómez Farías y que pretendía?</i>
a.	Opción correcta	Partido reformista, deseaba la igualdad entre todos los mexicanos ante la ley.
b.	Distractor 1	Partido de los conservadores, querían un gobierno central fuerte.
c.	Distractor 2	Rito yorkino, pretendía apoyar el sistema federal.
d.	Distractor 3	Rito escocés, pretendía defender los derechos de los criollos.
	Justificación 1	El partido de los conservadores era el oponente del partido reformista, estaba formado por el clero y el ejército.
	Justificación 2	El rito yorkino fue una logia formada por el embajador Poinsett.
	Justificación 3	El rito escoces fue otra logia formada por los criollos del país.
	Retroalimentación	IN: hechos, recuperación, recuerdo.

Los expertos opinan:

El reactivo número 5 en su base es complicado ya que no se entiende la redacción. (Entrevista docente experto-010.)

La base presenta incongruencias de redacción lo que sugiero se modifique además que en el distractor 3 existen faltas de ortografía. (Entrevista docente experto-0012.)

Con el ajuste que se realizó al reactivo, sugerido por los expertos, quedó conformado de la siguiente forma:

#5	Número	ccss_6prim_r_005
	Base	<i>¿Cuáles fueron las pretensiones del partido que conformó Valentín Gómez Farías y qué pretendía?</i>
a.	Opción correcta	Partido reformista, deseaba la igualdad entre todos los mexicanos ante la ley.
b.	Distractor 1	Partido de los conservadores, querían un gobierno central fuerte.
c.	Distractor 2	Rito yorkino, pretendía apoyar el sistema federal.
d.	Distractor 3	Rito escocés, pretendía defender los derechos de los criollos.
	Justificación 1	El partido de los conservadores era el oponente del partido reformista, estaba formado por el clero y el ejército.
	Justificación 2	El rito yorkino fue una logia formada por el embajador Poinsett.
	Justificación 3	El rito <i>escocés</i> fue otra logia formada por los criollos del país.
	Retroalimentación	IN: hechos, recuperación, recuerdo.

Es necesario recalcar que en este reactivo sólo uno de los expertos se percató de las faltas ortográficas existentes, por lo que se determina que es una diferencia entre los expertos de saberes gramaticales.



En el siguiente ejemplo de semejanzas, el reactivo se encontraba conformado de la siguiente manera:

#12		ccss_6prim_r_012
	Base	¿Cómo fue que se puso fin a la ocupación de Estados Unidos?
a.	Opción correcta	Se firmó el tratado de Guadalupe Hidalgo, en el cual México cedía su territorio y a cambio recibía 15 millones de pesos.
b.	Distractor 1	El ejército Mexicano gano la batalla y expulso a los soldados Estadounidenses del país.
c.	Distractor 2	México regalo algunos territorios del país a Estados Unidos
d.	Distractor 3	Dejaron que Texas y todos los estados que quisieran se unieran voluntariamente a Estados Unidos.
	Justificación 1	El ejército Mexicano no gano nunca la batalla contra Estados Unidos
	Justificación 2	No los regalaron en realidad.
	Justificación 3	Esto no fue lo que paso.
	Retroalimentación	IN: secuencia de eventos, comprensión, integración.

Los expertos argumentaron que:

En la base del reactivo existe una implicación de lenguaje muy elevado que los alumnos no comprenderían a la hora de contestar, por lo que se sugiere cambiar la redacción y las palabras que sean más acordes al nivel cognitivo del niño. (Entrevista docente experto-010.)

El reactivo es complejo a la lectura y muy técnico para el nivel.

(Entrevista docente experto-0012.)

Después de valorar el juicio de los expertos se toma la decisión de no cambiar el reactivo, ya que en ninguno de los dos casos se presenta un argumento correcto para ajustar la base, porque es acorde al nivel cognitivo en el que se encuentra el reactivo.

En los reactivos relacionados con 2° de secundaria, se inicia con el reactivo #2 en donde se tiene un ejemplo de diferencias.

#2	ccss_2sec_r_002
Base	El Síndrome de Inmunodeficiencia Adquirida (SIDA) es un padecimiento que ataca:
a. Opción correcta	Las defensas naturales del cuerpo humano.
b. Distractor 1	A hombres en edades tempranas: niñez y adolescencia.
c. Distractor 2	Las células cancerígenas que podríamos tener en el cuerpo.
d. Distractor 3	A la población del continente africano.
Justificación 1	Esta razón no es válida, ya que el sida puede presentarse en cualquier etapa de la vida del ser humano.
Justificación 2	Esta razón no es válida ya que el sida no está directamente ligado con el cáncer.
Justificación 3	Esta razón no es válida ya que es un padecimiento se da en los cinco continentes.
Retroalimentación	(IN) Hechos- reconocimiento

El experto comentó que:

El reactivo número 2 es un reactivo que no tiene que ver con el programa de estudios de Historia, por lo que es conveniente eliminarlo. (Entrevista docente experto-014.)

Con relación a este reactivo, se destaca que sólo uno de los expertos propuso su anulación. Los otros expertos omiten cualquier comentario, considerándose que la validez para eliminar el reactivo no es esencial quedando el reactivo tal y como se planteó desde un inicio.

Con los ejemplos antes descritos, se concluye esta sección que se realizó para la validez de contenido, quedando como resultado el contribuir a juicios de cómo queda construida la prueba preliminar posterior a realizarle los ajustes necesarios, destacándose

en general el ser pruebas objetivas, que cubren en su totalidad con los programas contemplados, tanto para primaria como para secundaria, de tal manera que son instrumentos alineados al curriculum.

### *Resultados de aplicación de las pruebas*

Los resultados de la aplicación de las pruebas objetivas representó, por medio del uso de medidas de tendencia central (media, mediana y moda), el desempeño académico de los alumnos a quienes se les aplicaron los diferentes instrumentos. Se destaca el índice de dificultad con porcentajes de 100%, por cada prueba, y las medidas de dispersión (valores mínimos y máximos obtenidos en la prueba así como la desviación estándar y varianza), de donde se obtiene el índice de discriminación por cada examen, bajo un rango de 0.3, considerándose como muy bajo, y al 0.75 como muy alto (Tabla 17).

Tabla 17

*Resultados de medidas de tendencia central, medidas de dispersión, índices de dificultad, índices de discriminación y alfa Cronbach.*

Grado/ Alumnos	# Reactivos	Moda	Mediana	$\Sigma$	Índice de correctos (Índice de dificultad)	% del rango de valores promedio de	Rango índice de discriminación arrojado por la	Índice de discriminación	alfa Cronbach
4°/40	24	41.6	41.6	45.2	55-31 (prueba difícil)		-0.3 a 0.8	0.5 alto	0.16710
5°/40	39	38.4	48.7	49.6	55-31 (prueba difícil)		-0.1 a 0.6	0.3 muy bajo	0.93084
6°/35	66	59.0	54.5	53.2	55-31 (prueba difícil)		-0.3 a 0.6	0.2 muy bajo	0.98898
1°/38	50	50	49	47.5	55-31 (prueba difícil)		-0.3 a 0.6	0.55 alto	1.00629
3°/37	54	53.7	48.1	47.0	55-31 (prueba difícil)		-0.5 a 0.6	0.4 bajo	0.99935

En la Tabla 17 se representan los resultados emitidos posteriores a la aplicación del instrumento 003 (pruebas o exámenes perfilados a partir del proceso de juicio de

expertos) en cada uno de los grados (4°, 5°, 6° de primaria y 1° y 3° de secundaria) en donde se subraya por cada uno de los grupos lo siguiente:

En 4° se aplicó una prueba a 40 alumnos, dando como resultado en la moda 41.6, que indica el número de veces que se repite la puntuación de la prueba. Asimismo la mediana es de 41.6 reflejando la posición intermedia de la distribución entre los valores de correspondencia de los reactivos en los que se acertó por debajo de los que fallaron en la respuesta correcta.

El índice de dificultad se encuentra ubicado entre 5-20%, (Tabla 18), por lo que se considera dentro del rango, como una prueba difícil. Su proporción es la representación de los estudiantes que contestaron correctamente, con lo que la variable nos permite visualizar que la habilidad que se desea medir se encuentra en el nivel adecuado, mientras que la capacidad del logro de índice de discriminación de 0.5 indica que es una prueba heterogénea, ya que los reactivos muestran inconsistencia en las habilidades que se miden pues se encuentra por debajo del nivel ( 0.1671007) de confiabilidad mínimo, que es de 0.7 determinado por el alfa Cronbach (Carey 2001).

En 5° se aplicó la prueba objetiva a 40 alumnos y los resultados que arrojaron son: la moda es de 38.46 siendo la puntuación biserial de la prueba. La mediana de 48.71 proyecta que la mitad de los alumnos tiene un nivel de desempeño adecuado en relación a la otra mitad que contestó incorrectamente sus reactivos. La prueba se encuentra con un índice de dificultad, ubicado entre 31-55%, considerándose como una prueba difícil ya que está dentro del rango establecido, además, es apropiada en el nivel taxonómico para el que fue diseñada. En consecuencia, se contempla que la prueba es homogénea ya que

su confiabilidad (0.9308445) se encuentra por encima del valor igual o mayor a 0.7, de acuerdo con el método del alfa Cronbach, por lo que es considerada como ideal en la confiabilidad, por su cercanía al valor máximo (1) (Cohen, 2001).

En cuanto al índice de discriminación, se encuentra en un rango de puntuación bajo, representado por un valor de 0.3, lo que indica que el desempeño de los alumnos en la prueba es muy bajo. Respaldándose esta variable con los puntajes, en donde el valor máximo obtenido fue de 82.05 y el mínimo de 25.64 de 100, contemplado en 39 reactivos de la prueba.

Para 6° los resultados muestran que los alumnos tienen un buen desempeño académico, representado por el 59.09 que contestó los reactivos de forma correcta. Aun cuando se vislumbra que de acuerdo al índice de dificultad la prueba se contempla como difícil (aun cuando el reactivo el No. 6 tiene un porcentaje muy alto de puntuación, 69), su confiabilidad se centra en (0.9889895), por lo que se caracteriza por ser una prueba de carácter homogéneo y confiable, ya que se acerca al valor máximo del método del alfa Cronbach. El valor máximo de la prueba es de 72.73 mientras que el mínimo es 15.15, de un porcentaje máximo de 100% del índice de dificultad (Lafourcade *et al.*, 1987). En cuanto a la discriminación, se encuentra en un nivel muy bajo reflejado por 0.2.

En 1° de secundaria los resultados muestran que la moda se encuentra en una relación directa con la mediana, lo que representa que 50% de los alumnos tuvieron un desempeño adecuado al contestar correctamente las preguntas, ubicándose la otra mitad por debajo de esta valor. Los alumnos se encuentran en la media de 47.58, considerada como favorable en relación al puntaje el rango de dificultad de la prueba es de 55-31% en

sus valores correctos, por lo que es considerada una prueba difícil que cumple con los objetivos establecidos para realizar la medición de las habilidades de acuerdo con la taxonomía de Marzano y Kendall.

Es necesario recalcar que en esta prueba se destaca el valor máximo de confiabilidad representado por el valor de 1.0062985 de acuerdo con el método de alfa Cronbach, lo que determina que tiene la consistencia interna entre reactivos, ya que cumple con los lineamientos de tener el error mínimo de estandarización, por lo que se considera ser una prueba homogénea (Carey, 2001).

Por último, en 3° de secundaria la prueba arrojó dentro de sus resultados que la moda representa 53.70 de dificultad para resolver la prueba, por lo que el resto se encuentra por debajo del desempeño, recalándose que los alumnos se encuentran en una media de 47.05, ubicada como favorable en cuanto al índice de dificultad, afirmándose que es una prueba difícil, ya que se encuentra en el rango de 55-31%, por lo que la complejidad de la habilidad que se midió de acuerdo a las características del grupo representó los valores obtenidos en la prueba máximo de 79.6 y mínimo de 20.4. En cuanto al índice de discriminación se proyecta un valor de 0.4, que es considerado como bajo, lo que representa que la prueba midió la capacidad que el alumno tuvo para contestar acertadamente considerándose que los objetivos planificados midieron las habilidades. El método de alfa Cronbach proyectó que la prueba tiene el valor máximo implicado en 0.9993576 para determinar que es confiable, de consistencia interna y homogénea (Tabla 18 y Apéndice D).

Tabla 18

*Resultado de Índices de dificultad que presentan las pruebas.*

Grado Escolar	# reactivos	Reactivos de bajo rendimiento	Valor mínimo obtenido	Valor máximo obtenido	Desviación estándar	Índice de dificultad medido %	Rango 0-30%
4°	24	9, 10, 17 y 18	16.67	83.33	12.60	5-20%	Muy difícil
5°	39	1, 12, 16 y 38	25.64	82.05	12.67	20-28%	Muy difícil
6°	66	5, 6, 11, 34, 55, 58 y 60	15.15	72.73	12.12	6-69%	Muy difícil
1°	50	2, 13 y 45	24.0	72.0	13.44	21-29%	Muy difícil
3°	54	1, 6, 12, 15, 16, 36, 46 y 48	20.4	79.6	11.27	19-30%	Muy difícil

La capacidad de discriminación de los ítems presentados en los resultados de las pruebas estimulan las respuestas de medición de conductas (Lafourcade, 1987), por lo que se considera que las pruebas en general, exceptuando la de 4°, cumplen con los lineamientos de estandarización en donde la correlación de los ítems miden la conducta deseada y la habilidad para la que fue creada. Asimismo, se destaca que en general los exámenes tienen la consistencia interna, por lo que se considera que las pruebas, exceptuando la que se aplicó a 4°, son homogéneas.

El uso de las pruebas estadísticas a las que se sometió la validez y confiabilidad de los exámenes, permitieron conformar pruebas de las que su diagnóstico consintió en reconocer el rendimiento académico de los alumnos, así como los niveles de dificultad. El favorecimiento de esta instrumentación permitió la conformación de pruebas objetivas con preguntas de opción múltiple acordes a los requerimientos de educación básica.

### *Análisis de los resultados desde la perspectiva de Marzano y Kendall*

Los resultados que se vierten en este análisis permiten reconocer el uso de la taxonomía de Marzano y Kendall como una alternativa de medición del desempeño académico. Su sustento teórico permitió, a través del uso práctico de éste, establecer los parámetros de habilidad para cada uno de los reactivos propuestos en los exámenes preliminares y en los exámenes perfilados, con lo que se demostró que el nivel de dominio de conocimiento y nivel de pensamiento, planteado en cada prueba, proporcionó elementos en la conformación de tablas de especificaciones mostrándose el rendimiento académico de los alumnos.

Se evidenció la presencia de los conocimientos declarativos o procedimentales, que dieron cuenta de habilidades mentales y procesamiento de razonamiento complejo, lo que denotó el análisis y aplicación en relación a los contenidos de la asignatura (Marzano, 2005).

De tal manera que los resultados identificados en la prueba ENLACE coinciden con los aplicados con esta prueba objetiva bajo la taxonomía de Marzano y Kendall. La inferencia que se realiza es, que cuando se plantean estructuras de planeación en relación a mediciones objetivas de aprendizaje a través de pruebas objetivas, las prácticas de evaluación identifican el desempeño real de los alumnos en relación a los conocimientos adquiridos, además que se verifica del diseño curricular y enfoque de la asignatura de Historia la alineación con el curriculum.



De esta explicación se infiere que; en los casos de 4°, 5° de primaria y 3° de secundaria, se debe atender en relación a los dominios de conocimiento, el primer nivel denominado de información a través de la construcción de significados por medio del conocimiento declarativo en donde se incluyen niveles de vocabulario, hechos y generalizaciones con respecto al niveles de pensamiento. Por lo que se deben realizar refuerzos en las áreas del primer sistema cognitivo denominado recuperación, estableciendo procesos de reconocimiento y de recuerdo.

La siguiente inferencia corresponde, a las pruebas objetivas a 6° de primaria y 1° de secundaria. Los reactivos muestran en su planificación objetiva los dominios de conocimiento donde tienen dificultad los alumnos, por lo que se detecta que es el conocimiento declarativo un área de dificultad para la ubicación del sujeto-verbo-sujeto, de tal forma que el sujeto no identifica la acción que ejerce del sujeto al objeto, no identificando hechos, jerarquías de secuencias temporales, generalizaciones. En los sistemas de pensamiento del nivel 1 que incluye la recuperación, se requieren activar y transferir conocimientos por medio del reconocimiento y el recuerdo. Asimismo, hace falta reforzar la comprensión por medio de la promoción del conocimiento declarativo (Tabla 19).

Tabla 19  
*Resultados de dominio de conocimiento y niveles de pensamiento de la taxonomía de Marzano y Kendall.*

Grado Escolar	Número de reactivos	Reactivos de bajo rendimiento	Nivel de dominio de conocimiento y niveles de pensamiento de la taxonomía de Marzano y Kendall
4 ° primaria	24	9, 10, 17 y 18	Principio- Reconocimiento Vocabulario-Reconocimiento Hechos- Reconocimiento Generalización
5° primaria	39	1, 12, 16 y 38	Hechos- Reconocimiento Vocabulario-Reconocimiento Hechos- Recuerdo Generalización- Reconocimiento
6° primaria	66	5, 6, 11, 34, 55, 58 y 60	Hechos- Reconocimiento Hechos- Recuerdo/Ejecución Hechos- Recuerdo Generalización- Hechos Generalización Generalización- Recuerdo Comprensión Secuencia de eventos- Comprensión/Integración
1° secundaria	50	2, 13 y 45	Hechos- Recuperación, Recuerdo Hechos-Comprensión, Simbolización Hechos-Análisis
3°secundaria	54	1, 6, 12, 15, 16, 37, 46 y 48	Vocabulario- Reconocimiento Hechos- Reconocimiento Generalización- Reconocimiento Hechos- Recuerdo

Se concluye en la secuencia de los resultados que las prácticas de evaluación que se efectuaron para la planeación, construcción y análisis de datos permitieron consolidar pruebas diagnósticas que mejoraron los instrumentos de medición, que en su diseño favorecieron y verificaron del rendimiento académico de los alumnos. Identificándose que este proceso permitió, bajo estándares de calidad, medir la construcción de instrumentos viables, confiables y objetivos, aptos para aplicarse en otros espacios educativos.

## Capítulo 5. Conclusiones

En este capítulo se explican los hallazgos encontrados en la investigación y se promueven algunas recomendaciones para la mejora de los procesos de evaluación a través del diseño de pruebas objetivas con reactivos de opción múltiple, que en su conformación permiten verificar el desempeño académico de los alumnos. Esto con la intención de mejorar no sólo la evaluación, sino objetivizar la planeación de enseñanzas congruentes con las necesidades actuales de aprendizaje de los alumnos, en donde se denoten la medición de conductas a través de la taxonomía de Marzano y Kendall.

### *Discusión*

El objetivo de la investigación fue identificar y describir elementos en la construcción de pruebas objetivas de reactivos de opción múltiple bajo el referente de la taxonomía de Marzano y Kendall. Lo anterior con la intención de tomar decisiones con respecto a cómo quedó finalmente constituida una prueba diagnóstica, que en su intención permitió medir rendimiento académico por medio de ciertos estándares de calidad, teniendo como finalidad el contribuir en el fortalecimiento de este tipo de procesos de evaluación.

En este análisis se establecieron tres preguntas y objetivos específicos que dieron pauta a su esclarecimiento, a partir de la revisión de la literatura y en la aplicación de instrumentos (Tabla 13), dando con ello respuesta a su explicación en este capítulo. Siendo estos cuestionamientos los siguientes:

1. ¿En qué medida el uso de la taxonomía de Marzano y Kendall en el diseño de instrumentos de medición del desempeño académico permitirá emitir inferencias sobre los procesos de pensamiento y dominios de conocimiento que ha logrado alcanzar un alumno?

El uso de la taxonomía de Marzano y Kendall proporcionó elementos teórico-instrumentales para la identificación de contenidos declarativos, factuales, procedimentales y actitudinales que deben de incluirse en una prueba objetiva, en donde se ubicó en cada objetivo de aprendizaje los conocimientos que se querían medir.

Dio como resultado el diseño y aplicación de instrumentos de medición de habilidades que dan cuenta de los dominios de conocimientos (clasificados en procesos de información y procedimientos mentales) y sistemas de pensamiento, que permitieron explicar el aprendizaje adquirido por los alumnos en sus procesos formativos de la asignatura de Historia (proceso mental, clasificación de sistema y subsistema de pensamiento) (Marzano, 2007) de los alumnos de 4°, 5° y 6° de primaria y de 1° y 3° de secundaria en donde se midió su habilidad para la resolución de la prueba. Asimismo, se probó el pilotaje de instrumentos de evaluación que midieron su desempeño académico.

Se identificó la posibilidad de emisión de inferencias de dominio de contenidos, del proceso de enseñanza y aprendizaje, y la complejidad de medir la enseñanza de forma compleja y objetivizada, resultado de los sistemas de pensamiento cognitivo, metacognitivo y sistema interno (self). El juicio valorativo de las inferencias permitió realizar deducciones para determinar cómo quedaron conformados los instrumentos de

medición de desempeño académico, por lo que se infiere dominar el campo de acción para emitir juicios acertados.

De la valoración que se hace del desempeño de los alumnos en las pruebas, se considera que se encuentran en un estado de aprendizaje que compete al primer nivel de dominio de conocimientos de procedimientos mentales y en los sistemas de pensamiento cognitivo de comprensión. De ésta, las inferencias realizadas nos permiten concluir que los reactivos se centran en los primeros niveles de cognición por lo que este juicio permitió tomar decisiones respecto al estado del aprendizaje evaluado y a la conformación de las pruebas.

2. ¿En qué medida el uso de pruebas estadísticas para el análisis de resultados y niveles de dificultad favorece la toma de decisiones en cuanto a cómo debe quedar finalmente constituida una prueba diagnóstica de rendimiento académico?

El uso de las medidas de tendencia central y de dispersión permitió validar el proceso de los datos duros y cuantificables de las pruebas aplicadas. Asimismo, determinó las cualificaciones de las pruebas realizadas por los expertos, su homogeneidad, el grado de objetividad, validez y confiabilidad. La consistencia interna de una prueba medida mediante el método de alfa Cronbach dio credibilidad al proceso y a la consistencia interna de las pruebas.

Los indicadores que se encuentran dentro de estas interrogaciones son: evaluación de calidad, evaluación educativa, bases teóricas-técnicas y teórico-prácticas para la construcción de pruebas objetivas con reactivos de opción múltiple, bajo el referente de la

taxonomía de Marzano y Kendall, desempeño académico, índices de dificultad de reactivos, índices de discriminación, análisis de distractores y estándares de calidad.

Dentro de los procesos de recolección de datos, los instrumentos que se utilizaron son: los exámenes preliminares, listas de cotejo y los exámenes ya aprobados. Se destaca la participación de los docentes, quienes validaron la prueba preliminar y la de los alumnos a quienes se les aplicó el instrumento.

3. ¿En qué medida el diseño de preguntas de opción múltiple constituye una prueba objetiva que permite medir el rendimiento académico?

Someter las pruebas de reactivos de opción múltiple a estos procesos de factibilidad concede evaluar, bajo estándares de calidad, el desempeño académico de los alumnos. Por su sistematicidad no sólo evalúan la cuantificación, sino la cualificación. Con lo que se prevé que el uso de estas valoraciones estadísticas permite estandarizar una prueba de manera confiable con resultados lo más acercarnos a la realidad del fenómeno educativo, llamado evaluación de calidad.

Esta información permitió aprobar un proceso de construcción de pruebas objetivas mediante la triangulación de validez en su constructo, contenido y criterio, con la finalidad de refutar y asignar el atributo de la consistencia de medición, por medio de la factibilidad de las teorías expuestas para la construcción de los instrumentos. Promover procesos de evaluación del desempeño bajo este tipo de instrumentos metodológicos o pruebas objetivas de reactivos de opción múltiple, busca destacar la estandarización y verificar el aprovechamiento real de los alumnos.

### *Conclusiones*

Los procesos de evaluación en la actualidad se consideran un obstáculo. Sin embargo, representan la posibilidad de medir la calidad de los procesos de aprendizaje de los alumnos. Bajo este paradigma se emiten las siguientes conclusiones:

Se reconoce la falta de investigaciones en relación al desempeño académico identificándose que los estudios existentes sólo devienen de instituciones públicas en donde sus resultados son ubicados en el desempeño de los alumnos pero bajo referentes teóricos de construcción instrumental tradicional.

Se identifica que el experto, quien es el sabedor de estos procesos de enseñanza y aprendizaje, ve en la construcción de pruebas objetivas de opción múltiple la posibilidad nula de realizar propuestas de evaluación de aprendizajes, ya que desconoce los procedimientos y la gama de posibilidades de medición de procesos mentales bajo la taxonomía de Marzano y Kendall para su construcción.

El diseño de pruebas objetivas de reactivos de opción múltiple bajo la taxonomía de Marzano y Kendall, permite como sistema de evaluación la posibilidad de medir rendimiento académico reflejado en los procesos de aprendizaje.

### *Recomendaciones*

Durante la evaluación de calidad medir los procesos de enseñanza y aprendizaje bajo metodologías centradas en el desempeño de los alumnos, debe considerarse como

innovación pedagógica en el campo educativo. Por lo que realizar investigaciones en este rubro permitirá al docente evidenciar la calidad y efectividad del diseño de pruebas objetivas con reactivos de opción múltiple bajo el referente de la taxonomía de Marzano y Kendall con la finalidad de medir conductas y habilidades cercanas a las competencias de los alumnos.

Es importante recalcar que para el diseño de instrumentos de desempeño bajo la taxonomía de Marzano y Kendall, se debe establecer en la planeación una evaluación que sea congruente con lo que se quiere medir, por lo que se deben ponderar el plan y los propósitos de los programas de estudios vigentes, así como objetivizar la enseñanza. El uso de las tablas de especificaciones para la construcción de los exámenes es primordial, ya que es el mapa que da el equilibrio de las habilidades de pensamiento y dominios de conocimiento que se desean medir.

Por otra parte, el equilibrio de una prueba debe someterse a sistemas de validación. Se sugiere trabajar bajo academias de grado regionales donde se utilicen como indicadores las listas de cotejo en las que los expertos creen juicios y tomen decisiones de cómo debe de quedar conformada la prueba que mida el desempeño académico de los alumnos, permitiendo con ello verificar el impacto del aprendizaje de los alumnos de forma estandarizada y no esperar las pruebas de ENLACE.

Un recurso fundamental de una prueba objetiva de reactivos de opción múltiple, es contar con los elementos que la constituyen, por lo que la omisión de cualquiera de sus componentes limita las expectativas de construcción para la que fue creada. Es necesario en la construcción y diseño de pruebas objetivas, bajo la taxonomía de Marzano y



Kendall exponer los resultados a la comunidad educativa con el fin de promover este tipo de prácticas de innovación pedagógica que tienen validez, confiabilidad y objetividad, ya que estas propuestas distan de los exámenes tradicionales, pues permiten medir la comprensión de lo aprendido y refieren la manifestación de una habilidad.

El sentido didáctico de la planeación permite evaluar el desempeño de los alumnos por lo que se mide la competencia real dándole sentido a las exigencias sociales con respecto a los estándares de aprendizaje de los alumnos. El rendimiento académico es visto en la evaluación educativa como un sistema complejo donde se perfila que el alumno sea congruente con los requerimientos específicos de su formación, de tal manera que el diseño de pruebas objetivas permite en la evaluación el control de calidad de la educación.

Otra recomendación es crear centros de evaluación nacionales con características académicas de asociación civil, con la intención de promover este tipo de instrumentos a nivel nacional. En este sentido los resultados que se muestren impulsaran en las escuelas prácticas congruentes con el sistema de calidad de evaluación y de concientización de docentes.

Bajo este tenor, las preguntas de investigación abren la posibilidad de aplicación de este sistema de evaluación en todos los campos educativos, por lo que se sugiere como punto final aplicar esta propuesta en educación media superior con el fin de tener una visión más amplia del impacto de formación de los alumnos en una o varias disciplinas.

Queda como compromiso para el investigador el explorar otros campos de aplicación, como son bachillerato y licenciatura, con el fin de seguir poniendo en práctica estas metodologías teórico-instrumentales como estrategia de evaluación de calidad en el desempeño académico del alumnado.

El planteamiento de esta propuesta abre el panorama para impulsar otros instrumentos de evaluación bajo la taxonomía de Marzano y Kendall, que en su diseño permitan evaluar de igual manera el desempeño del alumno de forma objetiva, confiable y viable.

## Referencias

- Aiken. L. R., Montes de Oca. G. y Domínguez, W. (2003). *Test psicológicos*. México: Thompson.
- Airasiam, P. W. (2002). *La evaluación en el salón de clases*. Biblioteca Pública para la actualización del maestro. Distrito Federal, México: SEP.
- Anastasi, A. A. (1982). *Test Psicológicos*. España, Madrid: Aguilar.
- Bloom, B. C. (1971). *Taxonomía de los objetivos de la educación, la clasificación de las metas educacionales*. Buenos Aires: Ateneo.
- Brunner, J. T. (1997). *Una ciencia del aprendizaje en el aula, en Escuelas para pensar*. Biblioteca Normalista. Ed. Primera, Cooperación Española. México: SEP.
- Carey, L. M. (2001). *Analyzing test, items, tasks, and tests*. Boston, EE.UU: Allyn and Bacon.
- Casanova, M. A. (1998). *La Evaluación Educativa*. Biblioteca Normalista. Ed. Primera, Cooperación Española. Distrito Federal, México: SEP.
- Casanova, M. A. (2004). *Evaluación y calidad de los centros educativos*. Madrid, España: Muralla.
- Castrillón, M. D. A. (2003). *Psicometría: historia y conceptos básicos para la estandarización de pruebas objetivas*. Informes Psicológicos ISSN: 0124-4906, vol: 5 fasc: págs: 101 – 110. Medellin, Colombia: UCC.
- Centro Nacional de Evaluación para la Educación Superior A.C. (2005).

- Centro Nacional de Evaluación para la Educación Superior A.C. (2010). *Manual para la elaboración de reactivos*. Distrito Federal, México: CENEVAL.
- Chadwick, C. B. y Rivera, N. I. (1997). *Evaluación formativa para el docente*, Distrito Federal, México: Paidós Educador.
- Cohen, R. J. y Sherdllik, M. E. (2001). *Pruebas y evaluación psicológicas: Introducción a las pruebas y a la medición*. Capítulos 6 y 7, Distrito Federal, México: Mc Graw-Hill.
- Consejo Nacional de Población y Vivienda. En cifras. Proyecciones de la Población 2005-2030 en el conteo 2005. Consultado el día 04 de Abril del 2010. En <http://www.conapo.gob.mx/index.México> .
- Cummings, W. K., Thomas, R. M. (1990). *International comparative education: Practices, issues, & prospects*. Inglaterra: Butterworth-Heinemann.
- De Vizenci, A. y De Angelis, P. (2008). La evaluación de los aprendizajes de los alumnos: Orientaciones para el diseño de instrumentos de evaluación. *Revista de Educación y Desarrollo*, Abril- junio.
- Díaz-Barriga, A. F. (2002). *Estrategias para una aprendizaje significativo una interpretación constructivista*. México: Mc Graw Hill.
- Díaz-Barriga, A. F. (2006). *Enseñanza situada: vínculo entre la escuela y la vida*. Distrito Federal, México: Mc Graw Hill.

Escudero, E.T. (2003). *Desde los test hasta la investigación evaluativa actual*, Un siglo, el XX, de intenso desarrollo de la evaluación en educación. España: RELIEVE: v.9, n.1, p.11-43.

Evaluación Nacional de Logro Académico en Centros Escolares (2008). *ENLACE*  
Consultado el día 04 de Abril del 2010. En  
[http://www.edomexico.gob.mx/evaluacioneducativa/anexos/resultados\\_ENLACE\\_Basica\\_2008.pdf](http://www.edomexico.gob.mx/evaluacioneducativa/anexos/resultados_ENLACE_Basica_2008.pdf).

Evaluación Nacional del Logro Académico en Centros Escolares. (2008). *ENLACE, Educación Básica*. Consultado en Abril 05 del 2010. En  
<http://enlace.sep.gob.mx/ba/?p=caracteristicas>

Fernández, B. R. (1995). *Evaluación de programas una guía práctica en ámbitos sociales educativos y de salud*. Madrid: Síntesis.

Gallardo, K. (2009). *Manual Nueva Taxonomía Marzano y Kendall*. Distrito Federal, México: ITESM.

Giroux, S. y Tremblay, G. (2004). *Metodología de las Ciencias Humanas*. México: FCE.

Gronlund, N. E. (1983). *Medición y Evaluación en la enseñanza*. Distrito Federal, México: Pax- México.

Guirtz, S. y Palamidessi. (2000). *El ABC de la tarea Docente: Currículum y Enseñanza*. Argentina: AIQUE.

- Hargraves, A. Earl, L. y Ryan, J. (2000). Reinventar la educación de los adolescentes, *en Una educación para el cambio*. Biblioteca Normalista. OCTAEDRO. Distrito Federal, México: SEP.
- Hernández, Fernández y Baptista (2008). *Metodología de la Investigación*. 4ta.ed. México. Mc Graw-Hill.
- Hills, J. R. (1982). *Planificar la prueba*, en evaluación y medición en la escuela. Distrito Federal, México: Kapelusz.
- Huerta, J. (1978). *La clasificación de los objetivos de aprendizaje su función y utilidad*. Distrito Federal, México: Trillas.
- Instituto Nacional de Estadística y Geografía (INEGI). Población: Segundo conteo de población y vivienda. Consultado el día 04 de abril del 2010. En <http://www.inegi.org.mx/inegi/default.aspx?s=est&c=124>.
- Instituto Nacional para la Evaluación Educativa (2005a). *Exámenes de la calidad y Logro Educativos*, Proceso de construcción y características básicas. Los temas de evaluación. Col. 8. Distrito Federal, México, 1-14: INEE.
- Instituto Nacional para la Evaluación Educativa (2005b). *PISA para docentes*. La evaluación como oportunidad. Distrito Federal, México: INEE.
- Instituto Nacional para la Evaluación Educativa (2006). *La calidad de la Educación Básica en México*, 2006. Informe Anual. Distrito Federal, México: INEE.
- Instituto Nacional para la Evaluación Educativa (2006b). *Plataforma para el Desarrollo del Sistema de Indicadores educativos de México*. Carpeta con

materiales de trabajo de la 1º Reunión Plenaria, junio 27, 2006. Distrito Federal, México: INEE.

Instituto Nacional para la Evaluación Educativa (2007). *Estimaciones a partir del sistema de estadísticas continuas de la DGPP/SEP*. Inicio del ciclo escolar 2006-2007 y de las proyecciones de la población de México, 200-2050, CONAPO, en INEE Panorama educativo 2007. Distrito Federal, México: INEE.

Instituto Nacional para la Evaluación Educativa (2008). *Estudio comparativo del aprendizaje en sexto de primaria en México 2005-2007: Español y Matemáticas*. Distrito Federal, México: INEE.

Kaplan, M. R. y Saccuzzo, D. P. (2006). *Pruebas Psicológicas/Psychological Testing: Principios, aplicaciones y temas*. 6ta. Ed., Distrito Federal, México: Thompson.

Karmel, L. J. (1974). *Medición y Evaluación escolar*. Distrito Federal, México: Trillas.

Lafourcade, P. D. (1997). *Evaluación de los aprendizajes*. Buenos Aires, Argentina: Kapelusz.

López, F. S. y Hinojosa, K. M. (2001). *Evaluación del aprendizaje*. Distrito Federal, México: Trillas.

Marzano, R. J. y Kendall, J. S. (2007). *The new taxonomy of educational objectives*. California, EE.UU. : Corwin Press.

- Marzano, R. J. (2000). *Designing a new taxonomy of educational objectives*.  
Thousand Oaks, CA: Corwin Press.
- Marzano, R. J. (2000). *Dimensiones del aprendizaje*. Distrito Federal, México:  
ITESO.
- Marzano, R.J. y Pickering, D. (2005). *Dimensiones del aprendizaje*. Manual para el  
Maestro, 2da. Ed., Distrito Federal, México: ITESO.
- Nigro, R. C. (1995). *Un modelo de prueba escrita que revela capacidades relacionadas  
con el proceso de aprendizaje*. Brasil: Enseñanza de las Ciencias, 13(3), 347-361.
- Organización de Estados Iberoamericanos (5, Diciembre, 2007). PISA 2006 en  
México. Consultado el 22 de octubre de 2009. En  
<http://www.oei.es/evaluacioneducativa/pisa2006-w.pdf>
- Programa Internacional de Evaluación de Estudiantes. (2008). *Guía de Orientación  
Prueba Piloto*. Consultado el 22 de octubre de 2009. En  
[http://hydra.icfes.gov.co/pisa/Documentos/Guia\\_Orientacion\\_PISA2009\\_Final  
.pdf](http://hydra.icfes.gov.co/pisa/Documentos/Guia_Orientacion_PISA2009_Final.pdf).
- Sacristán, J. G. (1995). *La pedagogía de los objetivos: obsesión por la eficiencia*.  
España: Morata.
- Sacristán J. G. y Pérez, G. A. I. (1999). *Comprender, Enseñar y Transformar la  
Enseñanza*. España: Morata.



- Schmelkes, S. (1997). *La Calidad en la educación primaria, un estudio de caso*. Distrito Federal, México: SEP/FCE.
- Secretaría de Educación Pública (1997). *Prácticas de evaluación en el aula*. Distrito Federal, México: SEP.
- Secretaría de Educación Pública (2005). *PISA para Docentes: La evaluación como oportunidad de aprendizaje*. Distrito Federal, México: INEE.
- Secretaría de Educación Pública (2006). *Prioridades y retos de la educación básica*. Alianza por la Calidad de la Educación. Distrito Federal, México SENTE.
- Secretaría de Educación Pública (2007). *Educación Básica. Secundaria. Plan de Estudios 2006 de Secundaria*. Distrito Federal, México: SEP.
- Secretaría de Educación Pública (2007). *Objetivos del Programa Sectorial de Educación 2007-2012*. 14. Indicadores. Consultado el 20 de octubre de 2009. En [www.dgest.gob.mx/web/archivos/.../objetivos\\_noticia\\_31.pdf](http://www.dgest.gob.mx/web/archivos/.../objetivos_noticia_31.pdf).
- Secretaría de Educación Pública. Boletín (16, Mayo, 2008). *Alianza por la Calidad de la Educación*. Consultado el 22 Octubre de 2009. En <http://www.sep.gob.mx/wb/sep1/alianzaporlacalidaddelaeducacion.pdf>.
- Tenbrink, T. D. (2002). *Evaluación guía práctica para profesores*. (6a. ed.). España, Madrid: NARCEA.
- Thorndike, R. L. y Hagen, E. P. (1989). *Medición y evaluación en psicología y educación* (2a. ed.). Distrito Federal, México: Trillas.

Tyler, R. W. (1973). *Principios básicos del currículo*. Buenos Aires: Troquel.

UNESCO. (2007). *Educación de calidad para todos un asunto de derechos humanos*.

Documento de discusión sobre políticas educativas en el marco de la II Reunión Intergubernamental del Proyecto Regional de Educación para América Latina y el Caribe (EPT/PRELAC) 29 y 30 de marzo de 2007; Buenos Aires, Argentina. Santiago de Chile: UNESCO para América Latina y el Caribe.

Valenzuela, J. (2008). *Evaluación de instituciones educativas*. Distrito Federal, México: ITESM.

Valdés, A. A. (2010). *Propiedades psicométricas de los instrumentos de evaluación*

[Recurso Educativo Abierto]. Disponible en el sitio web:

[http://reacudi.blogspot.com/2010/07/rea-instituto-tecnologico-de-sonora\\_07.html](http://reacudi.blogspot.com/2010/07/rea-instituto-tecnologico-de-sonora_07.html)

Wenger, E. (2001). *Comunidades de práctica. Aprendizaje significado e identidad*.

España: Paidós.

# Apéndice A

## Forma de Consentimiento

### Forma de Consentimiento

#### Calidad en la instrumentación para la evaluación del aprendizaje

Por medio de la presente quiero invitarte a participar en un estudio que estoy realizando sobre los procesos de validez de contenido y confiabilidad para pruebas objetivas que evalúan el nivel de desempeño académico. Yo soy alumno (a) de la Maestría en Educación del Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM). Este estudio está siendo realizado por mí como parte de mi tesis de maestría, con el respaldo de las autoridades del ITESM y la asesoría de profesores de dicha institución. Se espera que en este estudio participe una muestra de aproximadamente 3 profesores y alumnos de un aula que estén cursando el cuarto, quinto, sexto de primaria y primero y tercero de secundaria.

Si tú decides aceptar esta invitación, tu participación consistirá básicamente en alguna de estas tareas (yo te diré cuál):

1.- Fungir como juez experto para evaluar los reactivos de opción múltiple que conforman un banco de reactivos para las materias de español, matemáticas, ciencias naturales e historia;

2.- Permitirme evaluar a tus alumnos (por lo menos 40 de ellos) con estos instrumentos para coleccionar datos sobre la manera en que el instrumento está prediciendo el desempeño escolar.

Toda información obtenida en este estudio será estrictamente confidencial. Ni los profesores ni las autoridades del ITESM tendrán acceso a la información que cada uno de ustedes proporcione en lo particular. Todos los cuestionarios que ustedes hayan respondido serán llevados a mi casa y guardados bajo llave. Los nombres serán codificados, de tal forma que ningún nombre aparecerá en la base de datos que será analizada por mis asesores y por un servidor. Si los resultados de este estudio son publicados, los resultados contendrán únicamente información global del conjunto de alumnos participantes.

Tu participación en este estudio es voluntaria y de ninguna forma afectará en tu entorno laboral. Si tú decides participar ahora, pero más tarde deseas cancelar tu participación, lo puedes hacer cuando así lo deseas sin que exista problema alguno. Si tú tienes alguna pregunta, por favor hazla. Si tú tienes alguna pregunta que quieras hacer más tarde, yo te responderé gustosamente. En este último caso, copia mi número telefónico [**Arely Guadalupe Morales Blancas—Tel.:0445959574751**] o el de la titular de la investigación Dra. Katherina Gallardo Córdova 01-81-83-58-2000 ext. 6488 para que nos llames. Si deseas conservar una copia de esta carta, solicítamela y te la daré.

Si tú decides participar en este estudio, por favor anota tu nombre, firma y fecha en la parte inferior de esta carta, como una forma de manifestar tu aceptación y consentimiento a lo aquí estipulado. Recuerda que tú podrás cancelar tu participación en este estudio en cualquier momento que lo desees, aun cuando hayas firmado esta carta.

Nombre del participante

Firma

Fecha

Nombre del investigador

Firma

Fecha

## Apéndice B

### Lista de cotejo

#### Instrucciones

#### Evaluación de reactivos de opción múltiple como parte del proceso de validación de contenido

***Estructura del instrumento:***

Este instrumento consta de dos partes. La parte "A" la conforman preguntas de opción múltiple diseñadas a partir del plan de estudios de primaria o secundaria que ha emitido la Secretaría de Educación Pública a nivel federal y Estatal en casos específicos (como la materia de historia para el Estado de Tabasco, nivel tercer grado de primaria y primer año de secundaria).

La parte "B" la conforman una serie de criterios que se han seleccionado para evaluar la calidad de preguntas que conforman la parte "A", diseñadas para evaluar cada reactivo en cuanto a su forma (planteamiento) y contenido (temas abordados, respuesta correcta y distractores así como justificación de las mismas).

***Tiempo estimado para su aplicación:***

Máximo 2 horas.

***Instrucciones para los expertos evaluadores:***

1. Leer detenidamente **cada una de las preguntas, la opción correcta, distractores y justificaciones** que conforman los reactivos de opción múltiple.
  2. Calificar o dar su opinión sobre el nivel de calidad alcanzado para cada criterio. Este señalamiento se realizará marcando los cuadros para cada caso.
  3. En caso de encontrar algún error o comentario que implique un cambio parcial o total de algún reactivo de opción múltiple, se le pide por favor deje sus comentarios claramente plasmados para proceder a la revisión de la pregunta y tomar en cuenta su sugerencia de mejora.
- ¡Muchas gracias por su apoyo!

Información general sobre el experto		Respuestas
1.	Instrumento que está evaluando (disciplina y grado escolar al que corresponde):	
2.	Nombre del profesor / experto:	
3.	Fecha de la evaluación del instrumento (dd/mm/aaaa):	
4.	Nombre del entrevistador:	
5.	Nombre de la escuela en la que labora:	
6.	Clave de la escuela:	
7.	Grupo a los que da clases:	
8.	Materia(s) que imparte en el caso de ser profesor de secundaria:	
9.	Fecha de nacimiento (dd/mm/aaaa):	
0.	Género	
1.	Formación profesional de licenciatura:	
2.	Institución en la que estudió la licenciatura:	
3.	Formación profesional de maestría (si aplica):	
4.	Institución en la que estudió la maestría:	

° del reactivo

**Criterios a través de los cuales evaluar cada pregunta (parte "B").**

<p>1</p>	<p>Claridad en el lenguaje a lo largo de la pregunta, opción correcta y distractores:</p> <p><input type="checkbox"/> Deficiente</p> <p><input type="checkbox"/> Mala</p> <p><input type="checkbox"/> Regular</p> <p><input type="checkbox"/> Buena</p> <p><input type="checkbox"/> Muy buena</p>	<p>Existe una mejor respuesta que la respuesta correcta:</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó sí favor de escribir la opción correcta:</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>El reactivo es apropiado para el plan de estudio al que corresponde:</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó <b>No</b> favor de escribir el nivel correcto:</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>El reactivo cumple con todos los criterios de forma (redacción, ortografía, respuesta correcta y distractores vinculados al tema, etc.)</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó <b>No</b> favor de escribir la sugerencia de cambio o error encontrado:</p> <p>_____</p> <p>_____</p> <p>_____</p>
<p>2</p>	<p>Claridad en el lenguaje a lo largo de la pregunta, opción correcta y distractores:</p> <p><input type="checkbox"/> Deficiente</p> <p><input type="checkbox"/> Mala</p> <p><input type="checkbox"/> Regular</p> <p><input type="checkbox"/> Buena</p> <p><input type="checkbox"/> Muy buena</p>	<p>Existe una mejor respuesta que la respuesta correcta:</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó sí favor de escribir la opción correcta:</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>El reactivo es apropiado para el plan de estudio al que corresponde:</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó <b>No</b> favor de escribir el nivel correcto:</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>El reactivo cumple con todos los criterios de forma (redacción, ortografía, respuesta correcta y distractores vinculados al tema, etc.)</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó <b>No</b> favor de escribir la sugerencia de cambio o error encontrado:</p> <p>_____</p> <p>_____</p> <p>_____</p>

° del reactivo

**Criterios a través de los cuales evaluar cada pregunta (parte "B").**

<p>3</p>	<p>Claridad en el lenguaje a lo largo de la pregunta, opción correcta y distractores:</p> <p><input type="checkbox"/> Deficiente</p> <p><input type="checkbox"/> Mala</p> <p><input type="checkbox"/> Regular</p> <p><input type="checkbox"/> Buena</p> <p><input type="checkbox"/> Muy buena</p>	<p>Existe una mejor respuesta que la respuesta correcta:</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó si favor de escribir la opción correcta:</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>El reactivo es apropiado para el plan de estudio al que corresponde:</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó No favor de escribir el nivel correcto:</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>El reactivo cumple con todos los criterios de forma (redacción, ortografía, respuesta correcta y distractores vinculados al tema, etc.)</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó No favor de escribir la sugerencia de cambio o error encontrado:</p> <p>_____</p> <p>_____</p> <p>_____</p>
<p>4</p>	<p>Claridad en el lenguaje a lo largo de la pregunta, opción correcta y distractores:</p> <p><input type="checkbox"/> Deficiente</p> <p><input type="checkbox"/> Mala</p> <p><input type="checkbox"/> Regular</p> <p><input type="checkbox"/> Buena</p> <p><input type="checkbox"/> Muy buena</p>	<p>Existe una mejor respuesta que la respuesta correcta:</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó si favor de escribir la opción correcta:</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>El reactivo es apropiado para el plan de estudio al que corresponde:</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó No favor de escribir el nivel correcto:</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>El reactivo cumple con todos los criterios de forma (redacción, ortografía, respuesta correcta y distractores vinculados al tema, etc.)</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó No favor de escribir la sugerencia de cambio o error encontrado:</p> <p>_____</p> <p>_____</p> <p>_____</p>

° del reactivo

**Criterios a través de los cuales evaluar cada pregunta (parte "B").**

<p>5</p>	<p>Claridad en el lenguaje a lo largo de la pregunta, opción correcta y distractores:</p> <p><input type="checkbox"/> Deficiente</p> <p><input type="checkbox"/> Mala</p> <p><input type="checkbox"/> Regular</p> <p><input type="checkbox"/> Buena</p> <p><input type="checkbox"/> Muy buena</p>	<p>Existe una mejor respuesta que la respuesta correcta:</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó si favor de escribir la opción correcta:</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>El reactivo es apropiado para el plan de estudio al que corresponde:</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó No favor de escribir el nivel correcto:</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>El reactivo cumple con todos los criterios de forma (redacción, ortografía, respuesta correcta y distractores vinculados al tema, etc.)</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó No favor de escribir la sugerencia de cambio o error encontrado:</p> <p>_____</p> <p>_____</p> <p>_____</p>
<p>6</p>	<p>Claridad en el lenguaje a lo largo de la pregunta, opción correcta y distractores:</p> <p><input type="checkbox"/> Deficiente</p> <p><input type="checkbox"/> Mala</p> <p><input type="checkbox"/> Regular</p> <p><input type="checkbox"/> Buena</p> <p><input type="checkbox"/> Muy buena</p>	<p>Existe una mejor respuesta que la respuesta correcta:</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó si favor de escribir la opción correcta:</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>El reactivo es apropiado para el plan de estudio al que corresponde:</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó No favor de escribir el nivel correcto:</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>El reactivo cumple con todos los criterios de forma (redacción, ortografía, respuesta correcta y distractores vinculados al tema, etc.)</p> <p><input type="checkbox"/> Si</p> <p><input type="checkbox"/> No</p> <p>Si marcó No favor de escribir la sugerencia de cambio o error encontrado:</p> <p>_____</p> <p>_____</p> <p>_____</p>



° del reactivo

**Criterios a través de los cuales evaluar cada pregunta (parte "B").**

<p>7</p>	<p>Claridad en el lenguaje a lo largo de la pregunta, opción correcta y distractores:</p> <p><input type="checkbox"/> Deficiente</p> <p><input type="checkbox"/> Mala</p> <p><input type="checkbox"/> Regular</p> <p><input type="checkbox"/> Buena</p> <p><input type="checkbox"/> Muy buena</p>	<p>Existe una mejor respuesta que la respuesta correcta:</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> <p>Si marcó sí favor de escribir la opción correcta:</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>El reactivo es apropiado para el plan de estudio al que corresponde:</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> <p>Si marcó No favor de escribir el nivel correcto:</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>El reactivo cumple con todos los criterios de forma (redacción, ortografía, respuesta correcta y distractores vinculados al tema, etc.)</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> <p>Si marcó No favor de escribir la sugerencia de cambio o error encontrado:</p> <p>_____</p> <p>_____</p> <p>_____</p>
<p>8</p>	<p>Claridad en el lenguaje a lo largo de la pregunta, opción correcta y distractores:</p> <p><input type="checkbox"/> Deficiente</p> <p><input type="checkbox"/> Mala</p> <p><input type="checkbox"/> Regular</p> <p><input type="checkbox"/> Buena</p> <p><input type="checkbox"/> Muy buena</p>	<p>Existe una mejor respuesta que la respuesta correcta:</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> <p>Si marcó sí favor de escribir la opción correcta:</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>El reactivo es apropiado para el plan de estudio al que corresponde:</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> <p>Si marcó No favor de escribir el nivel correcto:</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>El reactivo cumple con todos los criterios de forma (redacción, ortografía, respuesta correcta y distractores vinculados al tema, etc.)</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> <p>Si marcó No favor de escribir la sugerencia de cambio o error encontrado:</p> <p>_____</p> <p>_____</p> <p>_____</p>

° del reactivo

**Criterios a través de los cuales evaluar cada pregunta (parte "B").**

9	<p>Claridad en el lenguaje a lo largo de la pregunta, opción correcta y distractores:</p> <p><input type="checkbox"/> Deficiente</p> <p><input type="checkbox"/> Mala</p> <p><input type="checkbox"/> Regular</p> <p><input type="checkbox"/> Buena</p> <p><input type="checkbox"/> Muy buena</p>	<p>Existe una mejor respuesta que la respuesta correcta:</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> <p>Si marcó sí favor de escribir la opción correcta:</p> <p>_____</p> <p>_____</p>	<p>El reactivo es apropiado para el plan de estudio al que corresponde:</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> <p>Si marcó No favor de escribir el nivel correcto:</p> <p>_____</p> <p>_____</p>	<p>El reactivo cumple con todos los criterios de forma (redacción, ortografía, respuesta correcta y distractores vinculados al tema, etc.)</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> <p>Si marcó No favor de escribir la sugerencia de cambio o error encontrado:</p> <p>_____</p> <p>_____</p>
10	<p>Claridad en el lenguaje a lo largo de la pregunta, opción correcta y distractores:</p> <p><input type="checkbox"/> Deficiente</p> <p><input type="checkbox"/> Mala</p> <p><input type="checkbox"/> Regular</p> <p><input type="checkbox"/> Buena</p> <p><input type="checkbox"/> Muy buena</p>	<p>Existe una mejor respuesta que la respuesta correcta:</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> <p>Si marcó sí favor de escribir la opción correcta:</p> <p>_____</p> <p>_____</p>	<p>El reactivo es apropiado para el plan de estudio al que corresponde:</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> <p>Si marcó No favor de escribir el nivel correcto:</p> <p>_____</p> <p>_____</p>	<p>El reactivo cumple con todos los criterios de forma (redacción, ortografía, respuesta correcta y distractores vinculados al tema, etc.)</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> <p>Si marcó No favor de escribir la sugerencia de cambio o error encontrado:</p> <p>_____</p> <p>_____</p>

## Apéndice C

### Prueba Preliminar de 3° de primaria en la asignatura de Historia.

#### EXAMEN QUE SE APLICA A 4TO grado de primaria.

- 1 ccss\_3prim\_r\_001  
País que se compone de 31 estados y un distrito federal.
  - a. México.
  - b. Belice.
  - c. Guatemala
  - d. El Salvador.
  
- 2 ccss\_3prim\_r\_002  
En México, tanto los estados como la federación sostiene tres poderes:
  - a. Legislativo, ejecutivo y judicial.
  - b. País estado y del municipio.
  - c. Ciudad, colonia y barrio.
  - d. Dirección, subdirección y supervisor.
  
- 3 ccss\_3prim\_r\_003  
Son comunidades que se caracterizan porque cuentan con medios de comunicación y transporte más rápidos y avanzados, además de servicios públicos como agua potable, luz eléctrica y pavimentación.
  - a. Urbanas.
  - b. Internacionales.
  - c. Rurales.
  - d. Estatales.
  
- 4 ccss\_3prim\_r\_004  
Son comunidades que se caracterizan por tener actividades basadas en la agricultura, la ganadería, la pesca y en la elaboración de artesanías:
  - a. Rurales.
  - b. Urbanas.
  - c. Internacionales.
  - d. Estatales.
  
- 5 ccss\_3prim\_r\_005  
Son lugares donde se guardan y exhiben objetos apreciados por su belleza o antigüedad para que sean conocidos. En algún momento dichos objetos fueron de gran utilidad:
  - a. Museos.
  - b. Galerías.
  - c. Escuelas.
  - d. Archivos.
  
- 6 ccss\_3prim\_r\_006  
La historia, para su estudio, se divide en:
  - a. Etapas o épocas.
  - b. Meses y años.
  - c. Años y siglos.
  - d. Ayer y hoy.

- 7 ccss\_3prim\_r\_007  
Son dos de las fuentes de información que necesita un historiador para elaborar la historia de un pueblo:
- Objetos y documentos.
  - Comidas y cultura.
  - Materiales y documentos.
  - Objetos y comercio.
- 8 ccss\_3prim\_r\_008  
Fue la primera cultura que se estableció en Tabasco:
- Olmeca.
  - Maya.
  - Tolteca.
  - Azteca.
- 9 ccss\_3prim\_r\_009  
Los olmecas se establecieron en La Venta lo que actualmente es el municipio de:
- Huimanguillo.
  - Nacajuca
  - Centla
  - Paraíso
- 10 ccss\_3prim\_r\_0010  
Recibían este nombre las personas que obligaban a los olmecas a trabajar la tierra para entregarles parte de su cosecha como tributo:
- Sacerdotes – magos.
  - Artesanos - artistas
  - Guerreros - asesinos
  - Agricultores - esclavos
- 11 ccss\_3prim\_r\_011  
Fue una de las dos clases sociales de la cultura olmeca, estaba constituida por los sacerdotes-magos, comerciantes y guerreros.
- Primera.
  - Segunda.
  - Alta.
  - Baja.
- 12 ccss\_3prim\_r\_012  
En el arte olmeca sobresalen las esculturas talladas en piedra llamadas:
- Cabezas colosales.
  - Pirámides del sol y la luna.
  - Pirámides mayas.
  - Atlantes.

- 13 ccss\_3prim\_r\_013  
La cultura \_\_\_\_\_ fue otra de las grandes culturas prehispánicas que se asentó en Tabasco.
- Maya.
  - Azteca.
  - Tolteca.
  - Zapoteca.
- 14 ccss\_3prim\_r\_014  
Era la clase social considerada la menos favorecida entre los mayas:
- Esclavos.
  - Agricultores.
  - Gobernantes.
  - Comerciantes.
- 15 ccss\_3prim\_r\_015  
En sus inicios la religión maya celebraba ceremonias sencillas de adoración a:
- La naturaleza.
  - La luna.
  - La muerte.
  - El rayo.
- 16 ccss\_3prim\_r\_016  
Los mercaderes de estas culturas tenían como punto de reunión el territoriotabasqueño
- Azteca y maya.
  - Maya y teotihuacano.
  - Azteca y Teotihuacano.
  - Maya y zapoteca.
- 17 ccss\_3prim\_r\_017  
Fueron un grupo descendientes de los mayas y que habitaban en Tabasco:
- Chontales.
  - Zapotecas.
  - Mixtecas.
  - Tlaxcaltecas.
- 18 ccss\_3prim\_r\_018  
Municipio donde el Capitán Hernán Cortés llevó a cabo su primer gran batalla en tierras Tabasqueñas:
- Centla.
  - Cunduacan
  - Tenochtitlan.
  - Teotihuacan.
- 19 ccss\_3prim\_r\_019  
Se le llamó así a la unión de los nativos de la región, de los españoles y los negros africanos traídos como esclavos.
- Mestizaje.
  - Colonización.
  - Encomienda.
  - Conquista.

- 20 ccss\_3prim\_r\_020  
Consistía en el reparto de tierras entre los conquistadores, quienes tenían como obligación proteger y educar en la religión cristiana a los nativos.
- a. Encomienda.
  - b. Mestizaje.
  - c. Colonización.
  - d. Conquista.
- 21 ccss\_3prim\_r\_021  
En la sociedad colonial era el grupo que gobernaba y tenía muchos privilegios y obligaban a los indígenas a pagar tributo.
- a. Españoles.
  - b. Mestizos.
  - c. Indígenas.
  - d. Guerreros.
- 22 ccss\_3prim\_r\_022  
Fue el grupo social que surgió de la unión los españoles e indígenas, al pasar el tiempo fueron considerados como un grupo aparte e inferior.
- a. Mestizo.
  - b. Indígena.
  - c. Españoles.
  - d. Guerreros.
- 23 ccss\_3prim\_r\_023  
Fueron el grupo que siguieron viviendo en sus poblados, en la misma forma en que lo hacían antes de la llegada de los españoles.
- a. Indígena.
  - b. Pardos.
  - c. Guerreros.
  - d. Mestizos.
- 24 ccss\_3prim\_r\_024  
Fueron el resultado de la unión de dos personas de diferentes grupos étnicos, generalmente una era negra.
- a. Pardos.
  - b. Guerreros.
  - c. Mestizos.
  - d. Españoles.

**Apéndice D**  
**Sabanas de datos de la prueba de Historia, que se aplico a 4° de**  
**Primaria**

Alumnos	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19	R20	R21	R22	R23	R24	Suma de correctas	Calificación	
	A1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	9	37.5
A2	1	1	1	1	1	1	0	0	0	0	0	1	1	0	1	0	0	0	1	1	1	0	0	0	12	50	
A3	1	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	1	0	1	0	8	33.33
A4	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	1	1	1	1	1	0	14	58.33	
A5	1	1	1	0	1	0	1	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	9	37.5	
A6	1	1	1	1	1	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	1	0	0	10	41.67	
A7	1	1	1	1	1	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	1	0	0	10	41.67	
A8	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	0	0	1	1	1	1	1	1	20	83.33	
A9	1	1	0	0	1	1	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	7	29.17	
A10	1	1	0	0	1	0	0	0	0	0	1	1	0	1	1	0	0	0	1	1	1	0	1	0	11	45.83	
A11	1	0	0	1	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	1	1	0	1	0	11	45.83	
A12	1	1	0	0	1	1	0	0	0	1	1	0	0	0	0	1	0	0	1	1	1	0	0	0	10	41.67	
A13	1	1	0	0	1	0	1	0	0	1	1	0	0	0	0	1	0	0	0	1	1	0	0	0	9	37.5	
A14	1	1	0	0	1	1	0	0	1	0	1	1	0	0	1	0	0	0	1	1	0	1	0	0	11	45.83	
A15	1	1	1	1	1	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	1	0	1	1	12	50	
A16	1	1	1	1	1	0	1	0	0	1	0	1	1	1	1	0	0	0	0	0	1	1	1	1	15	62.5	
A17	1	0	0	0	1	1	0	1	0	1	0	1	0	0	1	0	0	0	1	1	1	0	0	0	10	41.67	
A18	1	1	0	0	1	0	1	0	0	1	1	1	0	0	0	1	0	0	0	0	0	1	0	0	9	37.5	
A19	1	1	1	0	1	0	0	0	0	0	1	0	1	1	1	0	0	0	0	0	0	1	0	0	9	37.5	
A20	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	4	16.67	
A21	1	0	0	1	1	0	0	1	0	0	0	0	1	1	1	0	0	1	1	0	1	1	0	1	12	50	
A22	1	0	1	0	1	0	1	1	0	0	1	0	0	0	1	0	0	1	0	1	0	1	1	0	10	41.67	
A23	1	1	0	0	1	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0	1	1	0	0	11	45.83	
A24	1	1	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	6	25	
A25	1	1	0	0	1	1	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	7	29.17	
A26	1	1	0	0	0	0	1	0	0	0	1	0	0	1	0	1	0	0	0	1	1	0	0	0	8	33.33	
A27	1	1	1	1	1	0	1	0	0	0	0	1	1	0	0	1	0	0	0	1	0	0	1	1	13	54.17	
A28	1	1	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0	16	66.67	
A29	1	1	0	0	1	1	0	0	1	0	1	1	1	0	0	1	0	1	0	1	0	0	1	1	13	54.17	
A30	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	10	41.67	
A31	1	1	0	0	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	0	1	1	1	1	13	54.17	
A32	1	1	1	0	1	1	1	1	0	1	0	0	1	1	1	1	0	1	1	0	1	0	0	1	16	66.67	
A33	1	1	0	0	1	1		0	0	1	0	1	1	0	0	1	0	0	1	0	1	0	0	0	10	41.67	
A34	1	1	0	0	1	1	1	0	0	0	1	1	0	0	1	0	0	0	0	0	0	1	0	0	9	37.5	
A35	1	0	0	0	1	0	1	0	0	0	1	1	0	0	0	0	0	1	0	0	1	1	0	0	8	33.33	
A36	1	0	0	1	1	0	0	0	0	0	1	1	0	1	1	0	0	1	0	1	1	1	1	1	13	54.17	
A37	0	0	1	1	1	0	0	1	0	1	0	0	0	1	0	1	1	1	1	1	0	1	1	1	14	58.33	
A38	0	1	1	1	0	0	0	1	1	1	0	0	1	1	0	1	0	0	1	1	0	1	1	1	14	58.33	
A39	1	0	1	0	1	0	0	1	1	0	0	0	1	1	1	0	0	0	0	0	0	1	0	1	0	10	41.67
A40	1	1	0	0	1	1	1	0	0	0	1	1	0	0	1	0	0	1	0	0	1	0	1	0	11	45.83	
	38	30	18	16	37	17	14	13	7	12	17	22	17	22	22	16	2	8	14	20	31	13	17	11			
Indice de dificultad	95	75	45	40	92	43	35	33	18	30	42	55	43	55	55	40	5	20	35	50	78	33	43	28			

Estadísticos que se obtuvieron con esta información:

Medidas de tendencia central

Media Aritmética: 45.08

Mediana: 41.6667

Moda: 41.6667

Medidas de Dispersión

Valor Mínimo obtenido: 16.67

Valor Máximo obtenido: 83.33

Desviación estándar: 12.6082

Varianza: 154.99

Índices de discriminación

Grupo de menor rendimiento

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19	R20	R21	R22	R23	R24	Suma de correctas	Calificación de suma	
A1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	9	37.50	
A3	1	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	1	0	1	0	8	33.33
A5	1	1	1	0	1	0	1	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	9	37.50
A9	1	1	0	0	1	1	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	7	29.17
A13	1	1	0	0	1	0	1	0	0	1	1	0	0	0	0	1	0	0	0	1	1	0	0	0	0	9	37.50
A18	1	1	0	0	1	0	1	0	0	1	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0	9	37.50
A19	1	1	1	0	1	0	0	0	0	0	1	0	1	1	1	0	0	0	0	0	1	0	0	0	0	9	37.50
A20	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	4	16.67
A24	1	1	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	6	25.00
A25	1	1	0	0	1	1	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	7	29.17
A26	1	1	0	0	0	0	1	0	0	0	1	0	0	1	0	1	0	0	0	1	1	0	0	0	0	8	33.33
A35	1	0	0	0	1	0	1	0	0	0	1	1	0	0	0	0	1	1	0	0	1	1	0	0	0	8	33.33
Suma	12	9	3	1	10	3	5	4	1	3	6	3	3	5	4	5	1	1	0	3	7	2	2	0	0		

Grupo de mayor rendimiento

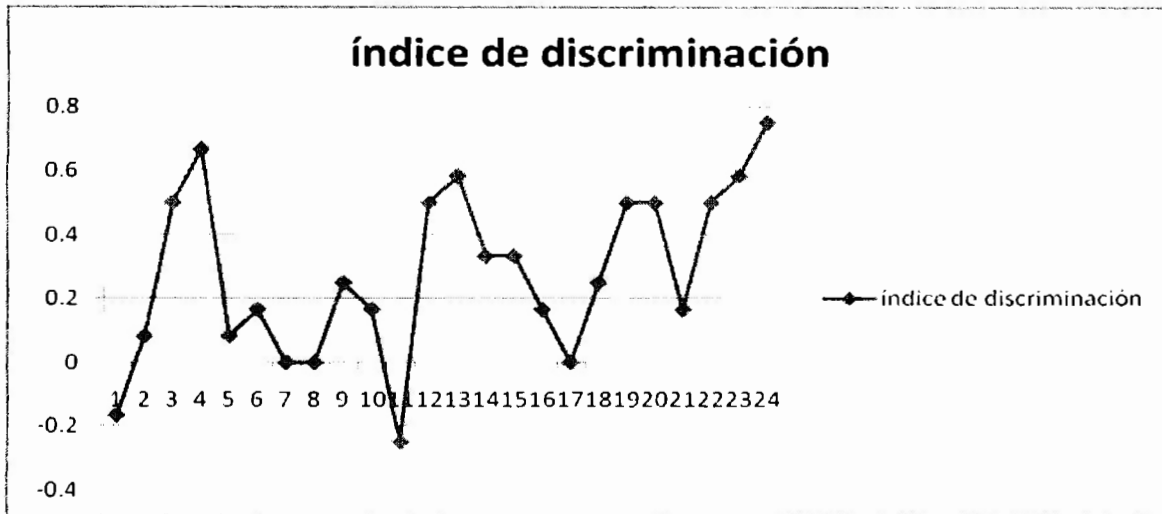
	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19	R20	R21	R22	R23	R24	Suma de correctas	Calificación	
A2	1	1	1	1	1	1	0	0	0	0	0	1	1	0	1	0	0	0	1	1	1	0	0	0	12	50	
A4	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	1	1	1	1	1	0	14	58.33	
A8	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	0	0	1	1	1	1	1	1	20	83.33	
A16	1	1	1	1	1	0	1	0	0	1	0	1	1	1	1	0	0	0	0	0	1	1	1	1	1	15	62.50
A27	1	1	1	1	1	0	1	0	0	0	0	1	1	0	0	1	0	0	0	1	1	0	1	1	1	13	54.17
A28	1	1	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0	16	66.67	
A29	1	1	0	0	1	1	0	0	1	0	1	1	1	0	0	1	0	1	0	1	0	0	1	1	13	54.17	
A31	1	1	0	0	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	0	1	1	1	1	13	54.17	
A32	1	1	1	0	1	1	1	0	1	0	0	0	1	1	1	1	0	1	1	0	1	1	0	0	1	16	66.67
A36	1	0	0	1	1	0	0	0	0	0	1	1	0	1	1	0	0	1	0	1	1	1	1	1	13	54.17	
A37	0	0	1	1	1	0	0	1	0	1	0	0	0	1	0	1	1	1	1	1	0	1	1	1	14	58.33	
A38	0	1	1	1	0	0	0	1	1	1	0	0	1	1	0	1	0	0	1	1	0	1	1	1	14	58.33	
Suma	10	10	9	9	11	5	5	4	4	5	3	9	10	9	8	7	1	4	6	9	9	8	9	9	0		



## Resultados de Índice Discriminación y Gráfico



-0.2 0.1 0.5 0.7 0.08 0.2 0 0 0.3 0.2 -0.3 0.5 0.6 0.3 0.3 0.2 0 0.3 0.5 0.5 0.2 0.5 0.6 0.8



Resultados de aplicación del método de alfa Cronbach en 4°.

Alumnos	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19	R20	R21	R22	R23	R24
A1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0
A2	1	1	1	1	1	1	0	0	0	0	0	1	1	0	1	0	0	0	1	1	1	0	0	0
A3	1	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	1	0	1	0
A4	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	1	0	0	1	1	1	1	1	0
A5	1	1	1	0	1	0	1	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
A6	1	1	1	1	1	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	1	0	0
A7	1	1	1	1	1	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	1	0	0
A8	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	0	0	1	1	1	1	1	1
A9	1	1	0	0	1	1	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
A10	1	1	0	0	1	0	0	0	0	0	1	1	0	0	1	1	0	0	0	1	1	1	0	1
A11	1	0	0	1	1	0	0	0	0	0	1	1	0	1	1	0	0	0	1	1	1	0	1	0
A12	1	1	0	0	1	1	0	0	0	1	1	0	0	0	0	1	0	0	1	1	1	0	0	0
A13	1	1	0	0	1	0	1	0	0	1	1	0	0	0	0	1	0	0	0	1	1	0	0	0
A14	1	1	0	0	1	1	0	0	1	0	1	1	0	0	0	1	0	0	1	1	0	1	0	0
A15	1	1	1	1	1	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	1	0	1	1
A16	1	1	1	1	1	0	1	0	0	1	0	1	1	1	1	0	0	0	0	0	1	1	1	1
A17	1	0	0	0	1	1	0	1	0	1	0	1	0	0	1	0	0	0	1	1	1	0	0	0
A18	1	1	0	0	1	0	1	0	0	1	1	1	0	0	0	1	0	0	0	0	1	0	0	0
A19	1	1	1	0	1	0	0	0	0	0	1	0	1	1	1	0	0	0	0	0	1	0	0	0
A20	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
A21	1	0	0	1	1	0	0	1	0	0	0	0	1	1	1	0	0	1	1	0	1	1	0	1
A22	1	0	1	0	1	0	1	1	0	0	1	0	0	0	1	0	0	1	0	0	1	1	0	0
A23	1	1	0	0	1	1	1	0	0	1	1	1	0	1	0	0	0	0	0	1	1	0	0	0
A24	1	1	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0
A25	1	1	0	0	1	1	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0
A26	1	1	0	0	0	0	1	0	0	0	1	0	0	1	0	1	0	0	0	1	1	0	0	0
A27	1	1	1	1	1	0	1	0	0	0	0	1	1	0	0	1	0	0	0	1	1	0	1	1
A28	1	1	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0
A29	1	1	0	0	1	1	0	0	1	0	1	1	1	1	0	0	1	0	1	0	1	0	0	1
A30	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0
A31	1	1	0	0	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	0	1	1	1	1
A32	1	1	1	0	1	1	1	1	0	1	0	0	1	1	1	1	0	1	1	0	1	0	0	1
A33	1	1	0	0	1	1	0	0	0	1	0	1	1	0	0	1	0	0	1	0	1	0	0	0
A34	1	1	0	0	1	1	1	0	0	0	1	1	0	0	1	0	0	0	0	0	1	0	0	0
A35	1	0	0	0	1	0	1	0	0	0	1	1	0	0	0	0	1	0	0	1	1	0	0	0
A36	1	0	0	1	1	0	0	0	0	0	1	1	0	1	1	0	0	1	0	1	1	1	1	1
A37	0	0	1	1	1	0	0	1	0	1	0	0	0	1	0	1	1	1	1	1	0	1	1	1
A38	0	1	1	1	0	0	0	1	1	1	0	0	1	1	0	1	0	0	1	1	0	1	1	1
A39	1	0	1	0	1	0	0	1	1	0	0	0	1	1	1	0	0	0	0	0	1	0	1	0
A40	1	1	0	0	1	1	1	0	0	0	1	1	0	0	1	0	0	1	0	0	1	0	1	0
Suma	38	30	18	16	37	17	14	13	7	12	17	22	17	22	22	16	2	8	14	20	31	13	17	11
Suma total	434																							
Var. K	0.000	0.0	0.5	0.5	0.0	0.5	0.5	0.5	0.0	0.0	0.5	0.5	0.0	0.0	0.5	0.0	0.0	0.5	0.0	0.0	0.0	0.5	0.0	0.0
Suma var.	24																							
Alfa de Cron	364.500																							
	0.16710																							
	078																							