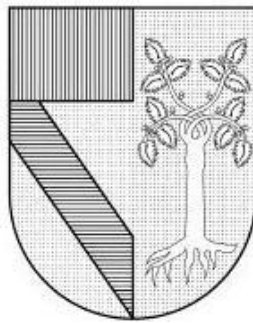


UNIVERSIDAD PANAMERICANA

ESCUELA DE PEDAGOGÍA



“Diseño y elaboración de la prueba de matemáticas
para bachilleratos tecnológicos”

INFORME DE ACTIVIDAD PROFESIONAL

Q U E P R E S E N T A

ERNESTO EDUARDO MORALES MALDONADO

P A R A O B T E N E R E L G R A D O D E :

MAESTRO EN EDUCACIÓN

**DIRECTOR DEL PROGRAMA ACADÉMICO: Dra. María del Carmen
García Higuera**

**DIRECTOR DEL INFORME DE ACTIVIDAD PROFESIONAL: Dra. Mónica
del Carmen Meza Mejía**

ÍNDICE

Contenido	Pág.
Introducción	3
1. Conceptos fundamentales de la evaluación	5
1.1. Tipos y modalidades de las pruebas	5
1.2. ¿Qué es una prueba?	6
1.3. Taxonomía de Benjamín Bloom	7
1.4. Definición de medición	8
1.5. Definición de evaluación	10
1.5.1. Momentos de la evaluación	13
1.5.2. Modalidades de la evaluación	15
1.5.3. Evaluación educativa	16
1.5.4. Evaluación del aprendizaje	17
1.5.5. Evaluación institucional	18
1.6. Objetividad	18
1.7. Validez	18
1.8. Confiabilidad	21
1.9. El Maestro en Educación: sus implicaciones y funciones en el diseño y elaboración de la prueba	22
2. Antecedentes del Centro Nacional de Evaluación para la Educación Superior, A.C. (CENEVAL)	24
2.1. Orígenes	24
2.2. Misión	25
2.3. Principios del Centro	25
2.4. Perspectivas	25
2.5. Actividades principales	26
2.6. Servicios	26
2.7. Estructura	29
2.8. Conformación de la Asamblea General de Asociados y de su Consejo Directivo	30
2.9. Exámenes que opera el CENEVAL	31

Contenido	Pág.
3. Descripción de actividades profesionales en el diseño de la prueba de Matemáticas para Bachilleratos Tecnológicos Federales	34
3.1. Características de la Prueba de Matemáticas para Bachilleratos Tecnológicos Federales	35
3.2. Objetivo de la prueba	35
3.3. Población objetivo	35
3.4. Características	36
3.5. Alcances	36
3.6. Administración	37
3.7. Resultados	37
4. Etapas para la construcción de una prueba	38
4.1. Antecedentes de la prueba	38
4.2. Diseño	43
4.2.1. Perfil referencial	43
4.2.2. Especificaciones	48
4.2.3. Estructura de la prueba	50
4.2.4. Validación del perfil referencial y estructura de la prueba	51
4.3. Construcción	51
4.3.1. Verificación	53
4.3.2. Revisión cualitativa. Revisión técnica	53
4.3.3. Revisión cualitativa. Validación externa	54
4.3.4. Revisión cualitativa. Revisión de estilo	57
4.3.5. Revisión cuantitativa. Piloteo de reactivos	57
4.3.6. Revisión cuantitativa. Calibración de reactivos	58
5. Ensamble	60
5.1. Integración de la versión operativa	60
5.2. Revisión editorial	61
6. Aplicación	61
7. Calificación	61
7.1. Establecimiento de puntos de corte y definición de niveles de dominio	62
8. Reportes	64
8.1. Guía para el sustentante	64
Conclusiones	65
Bibliografía	66

Introducción

El propósito de este informe de actividad profesional es realizar una descripción de las funciones e implicaciones del Maestro en Educación en el diseño y elaboración de un examen de opción múltiple para el nivel medio superior.

Dicho trabajo está compuesto por cuatro capítulos. En el primero, “Conceptos fundamentales de la evaluación”, se abordan: tipos y modalidades de las pruebas, ¿qué es una prueba, taxonomía de Benjamín Bloom?, definiciones de medición y evaluación, momentos de la evaluación, modalidades de la evaluación, validez, confiabilidad y evaluación del aprendizaje.

En el segundo capítulo, “Antecedentes del Centro Nacional de Evaluación para la Evaluación Superior (CENEVAL)”, se explica el origen, la misión, los principios, las perspectivas, la estructura, los servicios y exámenes del Centro.

El tercer capítulo describe las actividades del Maestro en Educación en el diseño del examen de opción múltiple para el nivel medio superior.

En el cuarto y último capítulo se abordan las etapas de la construcción de la Prueba de Matemáticas para Bachilleratos Tecnológicos Federales, sus antecedentes, diseño, ensamble, alcances y a quién va dirigido.

La inquietud de realizar este trabajo sobre el diseño y elaboración de un examen de opción múltiple de nivel medio superior radicó en que hay actividades profesionales que corresponden al Maestro en Educación que es relevante darles a conocer y documentar. Algunas implicaciones y funciones están asociadas con algunas asignaturas abordadas en el plan de estudios de la Maestría en Educación de la Universidad Panamericana.

Por otro lado, aunque el CENEVAL cuenta con 22 años de experiencia y tiene definidos tanto los mecanismos de operación como los procedimientos y la metodología validada, hasta la fecha no hay antecedente de un trabajo que se centre en las implicaciones y funciones en el diseño y elaboración de un examen de opción múltiple de nivel medio superior.

Sobra decir que este trabajo resulta de la autoevaluación de mi desempeño y actividades como Maestro en Educación y que es una oportunidad para promover el interés de otros profesionales por el diseño y la elaboración de un examen de opción múltiple.

1. Conceptos fundamentales de la evaluación

1.1. Tipos y modalidades de las pruebas

Los instrumentos como herramientas utilizadas para recolectar información nos ayudan a medir, lo cual constituye una actividad presente en la práctica docente, la cual a su vez nos aproxima al monitoreo y evaluación del proceso educativo. ¿Qué instrumento y cuáles son sus características para usarse en la recolección de datos, con fines más allá de la medición, dando cuenta del monitoreo y la evaluación?¹

Los instrumentos de evaluación más utilizados para recolectar datos son:

1. Encuesta
2. Preguntas de interpretación abierta y preguntas cerradas
3. Encuestas efectivas
4. Debates de grupos focales
5. Reactivos de opción múltiple

Los reactivos de opción múltiple, especialmente por su versatilidad y características psicométricas, son los reactivos más comunes en los exámenes modernos. No se sabe quién elaboró el primer reactivo de opción múltiple, pero desde el punto de vista de la evaluación académica fue algo fortuito. Los reactivos de opción múltiple son los más versátiles de todos los reactivos objetivos ya que pueden utilizarse para medir logros de aprendizajes simples y complejos en todos los niveles y en todas las áreas temáticas.

¹ García-Garro, Alberto, "Instrumentos de evaluación", *Revista Mexicana de Anestesiología*, vol. 30, núm. 3, julio-septiembre de 2007, pp. 158-164

Responder de manera correcta un reactivo de opción múltiple requiere habilidad para discriminar y no solo capacidad para reconocer o recordar la respuesta correcta. Los diseñadores de exámenes tienden a elaborar reactivos que miden la comprensión y el pensamiento más que medir el conocimiento directo de la materia.

Los componentes indispensables de un reactivo de opción múltiple son:

- Base del reactivo. Estímulo en forma de pregunta, enunciado incompleto, imagen, una combinación de todas, etc., al cual debe responder el examinado.
- Distractores. Opciones de las cuales se debe seleccionar la respuesta correcta. Los reactivos pueden constar de cuatro o cinco alternativas.
- Respuesta correcta. Opción que responde correctamente a la base de la pregunta.

La totalidad de los exámenes elaborados por el CENEVAL son de opción múltiple, estos pueden evaluar objetivos de todos los niveles de complejidad y permiten un diagnóstico más aproximado.²

1.2. ¿Qué es una prueba?

Una prueba es un instrumento de medición compuesto de reactivos, diseñado para medir el nivel de ejecución de un individuo o de un grupo en un dominio o constructo dado.

² Arredondo Galván, Pico Contreras y Palencia Gómez. *Nuevo Manual de Didáctica de las Ciencias Histórico-Sociales*, Limusa, 1997, p. 140

Se distinguen las pruebas normativas o construidas con referencia a normas, y las criterioles o construidas con referencia a criterios,³ cada una de ellas se utiliza de acuerdo con los distintos efectos y decisiones que requiera la institución, estas persiguen propósitos diferentes, se diseñan y elaboran bajo principios y metodologías distintas; por lo tanto, el significado y alcance de sus resultados tampoco son los mismos. A continuación presento brevemente una breve descripción de cada una.

- a) Normativa: Los resultados de aprendizaje se interpretan comparando el desempeño de cada educando con el logrado por los demás miembros del grupo.
- b) Criterial: Se realiza la comparación del aprendizaje del alumno con un parámetro establecido de antemano, bajo el supuesto de que en cualquier grupo todos y cada uno de sus miembros deben alcanzar el aprendizaje pretendido.

1.3. Taxonomía de Benjamín Bloom

El empleo de una taxonomía es parte esencial en el diseño de una prueba. Una taxonomía es un modelo clasificado que, dentro de la gran complejidad que involucra, simplifica la categorización de los elementos o atributos que distinguen a cada uno de ellos. Dos de sus características más significativas para hacer un uso correcto de ella son las siguientes: Para poder alcanzar una categoría es necesario pasar por las de los niveles inferiores y que todos los enunciados o componentes se redacten forzosamente, de manera específica, unívoca para cada contenido y nivel de tal forma que permita organizar los contenidos.

³ Backhoff E. Peón Z. y Sánchez M. (2005). *Manual técnico de construcción de reactivos*. México. INNE CENEVAL, *Estándares de calidad para instrumentos de evaluación educativa*, México D.F, 2000

En 1956, Benjamin Bloom elaboró una taxonomía que parte de la idea de describir los procesos de pensamiento desde niveles inferiores (asociados con la memoria y evocación) hasta de pensamiento de alto orden (análisis, síntesis y evaluación). Cuenta con seis categorías o niveles, divididas a su vez en subcategorías definidas indistintamente en términos de productos y procesos. Esta taxonomía se le reconoce cuatro criterios de clasificación involucrados, los cuales son:

1. Psicológico. Explica que el primer nivel es de tipo memorístico se avanza en la comprensión y formación de nuevos conceptos, su movilización ante situaciones concretas para generalizar el conocimiento, el paso por el proceso de pensamiento convergente y divergente hasta llegar a la toma de decisiones.
2. Pedagógico. El aprendizaje ocurre siguiendo una exposición tradicional por parte del docente (primer nivel de conocimiento) hasta el análisis de casos y estimación de errores (en los niveles más altos del análisis, síntesis y evaluación).
3. Lógico inclusivo. Cualquier nivel involucra a los inferiores.
4. Un criterio objetivo. Facilita la identificación de un nivel de manera unívoca.

1.4. Definición de medición

Muchas veces se suele confundir el concepto de medición con el de evaluación, de ninguna manera son lo mismo, de hecho la medición es tan solo una de las actividades de la evaluación.

A continuación presento algunas definiciones de medición asociadas a las ciencias exactas, las ciencias sociales y la educación y, finalmente, asociadas a la medición.

Asociadas a las ciencias exactas:

- Asignación de numerales o números a objetos y sucesos de acuerdo con reglas lógicamente aceptables⁴.
- Es la asignación de numerales a objetos o acontecimientos según reglas⁵

Asociada a las Ciencias Sociales y la Educación:

- La medida es un modo de expresión que implica la idea de comparar. Si se trata de cantidades, estas se miden comparándolas con la unidad de medida que se adopte. Medir los rendimientos de una prueba significa simplemente cuantificarlos en términos numéricos, o sea, ponerlos a nivel comparativo con la unidad de medida [...]. Según Gronlund, la medición implica: **a)** señalar y definir el medio para medir; **b)** definir la cualidad o el atributo que se habrá de medir; **c)** determinar las operaciones para que el atributo se manifieste, y **d)** aplicar los procedimientos necesarios para traducir las observaciones en enunciados cuantitativos⁶.

Hay dos puntos en esta definición que merecen atención especial. Antes que nada debe quedar claro que el objeto de la medición siempre es una propiedad y nunca una cosa o persona que la posee: No se mide al hombre, más bien se mide su estatura, su peso o su tensión arterial; no se define una habitación, lo correcto es medir la superficie o el perímetro; se mide la altura de un edificio, pero no el edificio mismo.

⁴ Guilford, J.P.(1984): Psychometric Methods. Nueva York: McGraw Hill Book Company.

⁵ Kerlinger, Fred N. (1975), **Investigación del comportamiento. Técnicas y metodología**. México, Interamericana.

⁶ Carreño Huerta, F (1979), **Enfoques y principios de la evaluación**. México, Trillas pág. 67

En el contexto de las pruebas se puede considerar la definición que sigue: Medir es asignar números a la conducta de acuerdo con reglas. La medición responde a la pregunta: ¿cuánto?, es decir, proporciona una descripción de la ejecución de una persona; no nos dice nada sobre el valor de dicha ejecución. Sin embargo, cuando se interpreta esta, se le atribuye cierto valor. En otras palabras, si la medición dice: “Se obtienen 5 aciertos”, la evaluación dice su esos aciertos son algo suficiente, bueno satisfactorio, etcétera⁷.

Aparentemente medir es una cosa sencilla, sin embargo, esta se debe realizar con unidades estandarizadas y aceptadas internacionalmente. Una forma sencilla de explicar este concepto es usar el ejemplo de un metro, como unidad de medida. No es lo mismo medir la estatura de una persona con un metro y a otras con cuartas.

Respecto al diseño de un instrumento de evaluación, en este caso de un examen de opción múltiple, es básico entender que la medición es solo un paso seguido de una valoración, una emisión de juicio de valor que servirá para tomar decisiones. Lo más peculiar es que tanto la medición como la evaluación, aunque sean confundidas frecuentemente, siempre estarán juntas.

1.5. Definición de evaluación

Una vez definido el concepto de medición y aclarada la diferencia entre la medición y la evaluación, presentaré algunas definiciones.

⁷ CENEVAL. Material de apoyo para el Taller de Elaboración de Reactivos de Opción Múltiple, México D.F, 1999.

Según la Dra. Sara Rosa Medina y la Dra. Marveya Villalobos⁸ “El término evaluación incluye varias acepciones: valorar, enjuiciar, comparar, controlar. Cada una de estas se suele identificar con fines muy diversos. Pero la evaluación es el proceso continuo, integral y participativo que permite identificar una problemática, analizarla y explicarla mediante información relevante. Como resultado, proporciona juicios de valor que sustentan la consecuente toma de decisiones”.

Básicamente esta definición supone dos operaciones fundamentales: la de obtener datos objetivos –medición- y la de interpretarlos. La extensión y la complejidad de esta segunda operación dependen del grado de significado que posea la unidad con que se hizo la medición.

A continuación presento algunas definiciones de medición asociadas al aprendizaje y al proceso educativo:

Definiciones asociadas al aprendizaje:

- Es el proceso de obtención de evidencias que permiten juzgar el logro de objetivos previamente determinados⁹.
- Es un modo de expresar el juicio subjetivo que merece una actitud compleja usando los diferentes contenidos o manifestaciones de tal actividad y los diferentes grados de apreciación, para poder determinar con cierta precisión el valor que atribuimos a la realidad de que se trata. La evaluación es una forma de expresión más precisa que la simple descripción, justamente porque sitúa la realidad dentro de una escala de valores objetivamente
- En una reunión sistemática de evidencias a fin de determinar si en realidad se producen ciertos cambios en los alumnos y establecer también el grado de cambio en cada alumno. Ello implica que la evaluación debe ser entendida como: 1. un método para adquirir y procesar la evidencia para mejorar el

⁸ Medina Martínez, Sara, Villalobos Pérez-Cortés, Elvia, Evaluación Institucional. Universidad Panamericana, México, 2006.

⁹ Ahumada A. Pedro. (1983). Principios y procedimientos de evaluación educacional. Ediciones Universitarias de Valparaíso. p.101;

aprendizaje del estudiante y la enseñanza; 2. algo que abarca una gran cantidad de evidencias más allá del habitual examen final; 3. una ayuda para aclarar las metas y objetivos más importantes de la educación y como un proceso para determinar el grado en que los estudiantes evolucionan en las formas deseadas; 4. un sistema de control de calidad en el que puede determinarse en cada etapa del proceso de enseñanza y aprendizaje si ese proceso es eficaz o no, y si no lo es, qué cambios deben efectuarse para asegurar su eficacia antes de que sea demasiado tarde; 5. un instrumento de práctica educativa que permita establecer si ciertos procedimientos alternativos son igualmente eficaces o no para alcanzar un conjunto de metas educacionales¹⁰.

Por ejemplo, para evaluar la estatura de un alumno en la clase de educación física se considera lo que se mide de altura en relación con el tipo de deporte que se desea practicar (basquetbol o equitación).

Lo fundamental al respecto es evaluar bien, con base en evidencias confiables y válidas, en forma sistematizada, con la participación de los implicados y con la asunción de las responsabilidades respectivas.¹¹ Más adelante me detendré en otros conceptos que son eje en la construcción de un instrumento de evaluación: confiabilidad y validez.

Por otra parte, para que la evaluación se lleve a cabo, hay actores que ejecutan las acciones, con base en todo un sustento teórico-metodológico. También cabe destacar que todas las pruebas tienen un componente de varianza o error adjudicable al evaluador y no al estudiante; por eso se debe verificar si la prueba es pertinente al grupo de estudiantes al que está siendo aplicada: índice de validez e índice de confiabilidad.

¹⁰ Bloom, B., et al. *Taxonomía de los objetivos de la educación: la clasificación de las metas educacionales : manuales I y II.*, 1975, pp 22-23.

¹¹ Gago, Antonio, *Apuntes acerca de la evaluación educativa*, SEP, México, 2002.

Hemos visto que la evaluación excede a la medición; aun cuando se basa en ella, el elemento distintivo entre una y otra es precisamente la interpretación de los resultados que varían en función del elemento respecto a la medición, según la naturaleza del fenómeno evaluado.

1.5.1. Momentos de la evaluación

Los momentos de la evaluación son: diagnóstica o inicial, sumativa o formativa y final, y, según Juan Domingo Santibáñez Riquelme, sirven para lo siguiente:

Diagnóstica o inicial: Este tipo de evaluación permite satisfacer varios objetivos y se usa en distintos momentos.

La evaluación como diagnóstico, utilizada antes de la enseñanza, satisface dos finalidades:

1. Identificar a los alumnos que no poseen las habilidades necesarias de nivelación o recuperación que permita dejarlos en condiciones adecuadas para poder trabajar y lograr eficientemente el objetivo que se pretende.
2. Identificar a los alumnos que ya han alcanzado los objetivos de enseñanza propuestos y que, al quedar libres de esas tareas, pueden dedicarse a profundizar en el tema objeto de estudio, a avanzar hacia el logro de otros objetivos.

Formativa: Cuando durante el proceso educativo se realizan evaluaciones tendientes a remediar o a reorientar diferencias y dificultades detectadas, tanto en la enseñanza como en los aprendizajes que se están desarrollando, entonces se está en presencia de la función formativa de la evaluación.

Algunas de las funciones de este tipo de evaluación son:

1. Proporciona al profesor una retroalimentación constante acerca de la calidad de sus enseñanzas.
2. Cuando se utiliza frecuentemente, ayuda al alumno a establecer su propio ritmo de estudio y le informa sobre la eficiencia de su trabajo.
3. Ayuda al alumno a identificar sus deficiencias y errores, dándole la oportunidad de esforzarse con el fin de alcanzar el dominio de los aprendizajes que se esperan de él.

Cabe destacar que para que las informaciones recogidas por cualquier medio evaluativo posean un carácter formativo deben entregarse a los estudiantes con el máximo detalle.

Se espera que una consecuencia inmediata de la evaluación formativa sea la pronta superación de los obstáculos detectados durante los procesos de enseñanza-aprendizaje.

Sumativa o final: Su principal característica es que se efectúa con el fin de proporcionar una calificación que testimonie el juicio valorativo del profesor sobre el desempeño final observado en el alumno al conducir un periodo escolar de duración determinada.

Este tipo de evaluación es una de las más comunes en el CENEVAL, ya que permite a las instituciones comprobar el producto del aprendizaje es decir, lo que resultó del trabajo, tanto del estudiante como del maestro.

1.5.2. Modalidades de la evaluación

Las modalidades de participación en la evaluación responden a la consideración de que esta no debe ser responsabilidad de un solo actor sino consensuada entre los sujetos involucrados. Esto implica una planificación y ejecución en conjunto por parte de los actores.

Algunas de estas modalidades se presentan en la siguiente tabla:

Modalidad	Descripción
Autoevaluación	<p>Se caracteriza porque cada estudiante evalúa sus propios procesos de aprendizaje, para ello previamente se explicitan de manera consensuada los criterios de logro que permitan un proceso de reflexión tendiente a la autocrítica con miras a la propia superación.</p> <p>La autoevaluación surge como un proyecto propio del participante donde cada uno decide las acciones que le permiten tomar decisiones sobre su práctica para trabajar sobre ella y objetivarla con responsabilidad y compromiso</p>
Coevaluación	<p>Es un proceso de interacción social que se espera se genere en un clima de respeto y de valorización de la opinión de los pares, en donde prevalezcan criterios y principios de honestidad, comprensión, aceptación, flexibilidad y apertura a la crítica, está fundamentada en el logro de la perfectibilidad del proceso, esta generalmente funciona como complemento de la autoevaluación.</p>

Modalidad	Descripción
Evaluación externa o evaluación unidireccional	Es la que hace el docente o un jurado invitado (criterio de experto) sobre los logros del estudiante. Esta es la estrategia tradicionalmente utilizada para evaluar el rendimiento del estudiante en el aula.
Evaluación multidireccional	Consiste en la evaluación del estudiante por parte de los diferentes miembros del grupo incluido el docente, invitados externos y el propio estudiante. Este proceso permite triangular contrastar la información para asegurar su validez y confiabilidad por juicio de observadores.

1.5.3. Evaluación educativa

La evaluación educativa se concibe como una estrategia útil para el mejoramiento de la calidad educativa, es una actividad continua e integrada que tiene por objetivo proporcionar la mayor cantidad de información para valorarla y tomar decisiones pertinentes respecto al objeto o sujeto de evaluación. Por ello, la evaluación puede realizarse al sistema en su conjunto a cualquiera de sus componentes: las instituciones, los programas, los docentes, los estudiantes, entre otros esta evaluación cumple con la función de ofrecer información significativa para:

- Conocer la calidad del proceso educativo
- Fundamentar la toma de decisiones educativas
- Mejorar la planeación del proceso educativo
- Decidir sobre la certificación de sujetos
- Determinar la acreditación de programas

1.5.4. Evaluación del aprendizaje

La evaluación del aprendizaje en los Estados Unidos de Norteamérica antes de 1850 había descansado principalmente en el examen oral, en donde el maestro hacía una pregunta y el alumno señalaba. No había uniformidad ni en las preguntas formuladas a los diferentes alumnos, ni en las respuestas de estos. No había posibilidad de comparaciones, respecto a las tareas, o a las evaluaciones.

Durante la segunda mitad del siglo XIX, los exámenes orales realizados por jurados examinadores visitantes fueron reemplazados por “exámenes escritos formalmente establecidos” como base para la promoción o admisión a una escuela, academia o universidad. Esto es muy importante ya que es en donde el examen venido de fuera, cedió el puesto a la evaluación hecha por el maestro en clase. La técnica usada era el ensayo que el alumno respondía con sus propias palabras.

El examen escrito tenía ventajas sobre el oral, por ejemplo: a) presentar las mismas tareas a cada miembro del grupo, b) dejar que cada alumno trabajara durante todo el periodo del examen. Aunque ya era un avance significativo, la apreciación de las respuestas de cada individuo continuaba siendo altamente subjetiva y “dependía de las normas y prejuicios de cada calificador”. Ya en 1900 comenzaron a desarrollarse y generalizarse los exámenes calificados objetivamente, en los cuales una clave preestablecida podía aplicarse rutinaria y uniformemente a las respuestas dadas por cada alumno, también en ese año surgió la idea de una norma de aprovechamiento para una edad o grado, con la cual pudiese compararse el rendimiento de cualquier individuo o grupo.

1.5.5. Evaluación institucional

La evaluación institucional es un proceso continuo, integral y participativo que permite identificar una problemática, analizarla y explicarla mediante información relevante. Como resultado proporciona juicios de valor que sustenta la consecuente toma de decisiones.

1.6. Objetividad

Se refiere a que un instrumento tiene independencia respecto a la persona que lo aplica o califica y que cumple con una serie de requerimientos de diseño para asegurar que la aplicación, calificación e interpretación de los resultados obtenidos no generen discrepancias de apreciación atribuibles a los evaluadores.

1.7. Validez

La validez es medir lo que se pretende medir. Es curioso que para algo que suena tan sencillo se requiera todo un procedimiento, pero efectivamente así es; por ejemplo, en el diseño de un instrumento de evaluación, al momento de definir el objeto de medida, lo que se busca es validez. Un instrumento de evaluación debe tener idealmente más de un tipo de validez.

A continuación presento algunas clasificaciones:

- a) **Validez racional o de contenido:** Se refiere a que el examen al menos en su aspecto deber contener una muestra adecuada y representativa de ítems, tanto en términos de los conocimientos específicos que se propone medir como de las clases de capacidades o comprensiones que debe reflejar tal conocimiento.

- b) **Validez concurrente:** Cuando las puntuaciones del examen se correlacionan razonablemente bien con algunos criterios contemporáneos de conducta, es necesario demostrar que la población de uno es comparable en todos los aspectos pertinentes a la muestra en la que se validó el instrumento. La dificultad de la validez concurrente es que el criterio seleccionado sea pertinente, apropiado, confiable y válido.
- c) **Validez predictiva:** Cuando la conducta de criterio con la cual se relacionan las puntuaciones obtenidas en el examen son una medida futura de ejecución, estamos hablando de validez predictiva, es decir, esta se presenta cuando las puntuaciones obtenidas se correlacionan significativamente con las calificaciones posteriores.
- d) **Validez de construcción:** Este tipo de validez tiene que ver con las cualidades o atributos psicológicos que un examen mide, este tipo de validez se basa en inferencias lógicas justificables que proceden de testimonios experimentales o de otra clase.

Todas las pruebas son indicadores imperfectos de cualidades o habilidades que tratan de medir, no son la excepción, así que se pueden mejorar.

La validez frecuentemente recibe críticas porque las medidas psicológicas y educativas tienden a ser indirectas y de carácter inferencial, en lugar de basarse en muestras conductuales directas del rasgo o capacidad de que se trate, porque un examen para medir el aprendizaje supone simplemente la capacidad para resolver de manera acertada a un conjunto específico. La razón es porque entre más indirecta o inferencialmente esté relacionada una puntuación del examen con el rasgo o capacidad que mida, tanto más importante será el problema de la validez.

Un instrumento de evaluación del aprendizaje, mide si el dominio de cierto cuerpo de conocimientos es lo suficientemente estable, claro y bien organizado, como para reflejar la estructura de ideas de cierta disciplina o subdisciplina.

Para ser objetivo en mis planteamientos, me gustaría citar uno de los muchos factores que hay que cuidar frecuentemente y que desafortunadamente es un punto débil hablando de un instrumento de evaluación del aprendizaje porque estas tienden a perder validez cuando tienen ítems que presuponen el conocimiento de materiales que no se incluyen ordinariamente dentro del ámbito de la disciplina o subdisciplina que están destinados a medir.

La validez de una prueba de ejecución depende en parte de lo bien que compruebe las capacidades verdaderas que le son exigidas a un individuo en aquellas ejecuciones de la vida real.

La validación es una de las preocupaciones primordiales de un evaluador puesto que si una prueba no es válida para el objetivo que fue preparado, los resultados no significan lo que se cree que significan.

No se puede subestimar la esencia del propósito para el que se elabora o se utiliza una prueba. Uno de sus problemas más comunes es su mala utilización: utilizar una prueba con un propósito para el que no se diseñó y para lo cual se desconoce su validez. Lo que no quiere decir que la prueba no sirva para más de un propósito, pero el objetivo es solo uno.¹²

Finalmente me gustaría comentar que el problema de la validez concurrente o predictiva se ilustra perfectamente en el siguiente ejemplo: Un examen de opción múltiple sobre educación vial puede mostrar buena validez de contenido, pero obviamente tendrá menos validez concurrente y predictiva en relación con las maneras de manejar que una prueba en carretera, lo que quiero resaltar es que los exámenes de opción múltiple, con todo y los valiosos que son, no sustituyen a los exámenes directos y prácticos, por eso es que su uso debe ser analizado y reflexionado previamente.

¹² J. Charles Alderson., Caroline Claphan y Dianne Wall. *Exámenes de idiomas: elaboración y evaluadores*, 1998.

1.8. Confiabilidad

Para emplear con confianza cualquier instrumento de evaluación, este debe mostrar un grado satisfactorio de precisión o confiabilidad, es decir, debe producir puntuaciones consistentes consigo mismas en diferentes momentos. Un ejemplo muy claro y que me ayuda mucho cuando quiero explicar más claramente la confiabilidad es el de un termómetro clínico que produce tres lecturas distintas, de 37, 39.0 y 34°, en un mismo paciente, no se puede considerar confiable.

La confiabilidad consiste en asegurar la pertinencia y permanencia tanto del procedimiento como de las estrategias y métodos utilizados para evaluar el aprendizaje.

Por definición, un instrumento de evaluación no confiable no puede ser válido, ya que el grado de confiabilidad necesario depende del uso que se le vaya a dar a las puntuaciones del examen; por ejemplo, si las puntuaciones se van a utilizar para la evaluación individual, será necesario un grado mucho mayor de confiabilidad que si fuera el caso de estimación gruesa o con fines de investigación.

La extensión de un examen es el factor aislado más importante que influye en su confiabilidad. Cuanto más corta sea una prueba tanto más probable será que las situaciones de ella resulten influidas por el muestreo al azar o por los factores de la situación. Por otro lado, no concederles tiempo suficiente a los estudiantes para que completen el examen tiene el mismo efecto en la confiabilidad que reducir el número de ítems, ya que la confiabilidad de una prueba disminuye también por la imprecisión y subjetividad de la manera de calificar y por la presencia de ítems que tengan baja fuerza discriminatoria

Existen tres tipos de coeficientes para presentar la confiabilidad:

- a) La repetición de la prueba con tiempo suficiente para que el evaluado olvide los ítems, calculando el coeficiente de correlación que permite conocer el grado de confiabilidad de dicha prueba.
- b) Las formas paralelas, técnica similar a la anterior, pero con pruebas no idénticas sino con ítems equivalentes.
- c) División en mitades, con el cual se calcula un coeficiente de consistencia interna, mediante la correlación entre las puntuaciones de ambas mitades.

Específicamente en la evaluación del aprendizaje, la calidad depende, fundamentalmente, de cómo se aborde la validez y la confiabilidad, cualidades esenciales que deben estar presentes en el desarrollo del proceso de recoger y analizar la información conducente a garantizar una mayor confianza sobre las conclusiones emitidas.

1.9. El Maestro en Educación: sus implicaciones y funciones en el diseño y elaboración de la prueba

El maestro en educación tiene como objetivo de trabajo la reflexión e intervención sobre el comportamiento humano en situaciones educativas mediante el desarrollo de las capacidades de las personas, grupos e instituciones. Se entiende el término educativo en el sentido más amplio de formación y desarrollo personal y colectivo.

Desarrolla su actividad profesional principalmente en el marco de los sistemas sociales dedicados a la educación en todos sus diversos niveles y modalidades, tanto en los sistemas reglados, no reglados, formales e informales, y durante todo el ciclo vital de la persona. Asimismo interviene en todos los procesos pedagógicos que afectan al aprendizaje o evaluación del aprendizaje, o que de este se derivan, independientemente de su origen personal, grupal, social, de salud, etcétera, responsabilizándose de las implicaciones educativas de su intervención profesional y coordinándose, si procede, con otros profesionales.

Definir conductualmente las actividades del maestro implica especificar en forma objetiva todas las conductas que debe desarrollar como profesional aplicado y como investigador.

Esto abarca desde el observar, definir variables cuantitativamente, escribir reportes, programar variables, analizar literatura previa, saber dar instrucciones, saber interactuar con otros en situaciones naturales, analizar datos, representar en formas diversas un mismo fenómeno o conjunto de observaciones, evaluar la adecuación de un concepto a un grupo de datos, saber discriminar la aplicabilidad de una técnica a una situación determinada, etcétera.

2. Antecedentes del Centro Nacional de Evaluación para la Evaluación Superior, A.C. (CENEVAL)

2.1. Orígenes

El Centro tiene su origen en los resolutivos de la Asamblea General de la Asociación Nacional de Universidades e Instituciones Superiores (ANUIES), en su XXV sesión Ordinaria en abril de 1993. La Asamblea General, a fin de atender a su vez una petición del entonces secretario de Educación Pública (Dr. Ernesto Zedillo Ponce de León) acordó proponer a la Coordinación Nacional para la Planeación de la Educación Superior (Conpes) la creación de un organismo no gubernamental que tuviera a su cargo, con base en una serie de especificaciones, el diseño y operación de los exámenes nacionales.

El 2 de febrero de 1994 encomendó al Mtro. Antonio Gago Huguet la dirección de los trabajos necesarios a fin de proceder de inmediato a constituir formalmente la asociación civil que se abocaría. El 28 de abril de 1994 se protocolizó ante el notario público el acta constitutiva del CENEVAL, en la cual se formalizó su personalidad jurídica como una asociación civil.

A partir de 1994, se construyen y validan los primeros instrumentos para evaluar no solo a quienes ingresan y egresan de licenciatura. Son además a quienes ingresan a la educación media superior, con lo cual también se atendió una propuesta de los rectores y directores agrupados en las ANUIES.

2.2. Misión¹³

Contribuir a mejorar la calidad de la educación media superior y superior y programas especiales mediante el fomento de una cultura de la evaluación y la realización de evaluaciones externas de los aprendizajes logrados en cualquier etapa de los procesos educativos, de manera independiente, con los más altos estándares de calidad y en forma adicional a las que llevan a cabo las instituciones a las que prestan servicios

2.3. Principios del Centro

- **Imparcialidad.** Los exámenes no favorecen a ningún organismo o institución en particular.
- **Calidad técnica.** Los instrumentos de medición se alinean a *rigurosos estándares* de calidad para garantizar su validez y confiabilidad.
- **Dinamismo.** Los exámenes se revisan y actualizan frecuentemente a fin de *ajustar los contenidos* a los cambios de las áreas disciplinares o profesionales.
- **Confidencialidad.** Las instituciones usuarias tienen el derecho de que sus *resultados no sean revelados* sin su consentimiento.
- **Transparencia.** Todos los procesos están *abiertos a escrutinio*, siempre y cuando no se contrapongan al principio de confidencialidad.

2.4. Perspectivas

Las razones y perspectivas son evaluar, certificar y acreditar, no se trata de alguna moda o una manía de tecnócratas, de neoliberales o de globalización. Las funciones mencionadas anteriormente tienen su razón de ser en una aspiración social incuestionable y sus raíces no están en alguna coyuntura económica. Lo anterior sería, desde mi punto de vista ofender a los educadores y desvirtuar a las IES.

¹³ Manual de organización, CENEVAL, 2015

Como institución eminentemente técnica y con casi quince años de experiencia, el CENEVAL ha sistematizado y perfeccionado los procedimientos y procesos de diseño, construcción, aplicación, calificación y reporte de los resultados de los exámenes, siempre apegado a estrictos estándares internacionales de calidad que guían el quehacer de agencias de evaluación, como los establecidos por la *International Association for the Evaluation of Educational Achievement*, *American Psychological Association*, *Educational Testing Service* y *The Joint Committee on Standards for Educational Evaluation*.

La metodología que utiliza el CENEVAL se basa en tres etapas para la elaboración de la prueba. En el siguiente capítulo abordo cada una de ellas.

2.5. Actividades principales

- Elaboración y aplicación de instrumentos de medición de conocimientos y habilidades.
- Difusión de los resultados que arrojan las pruebas.
- Promoción de la cultura de la evaluación.
- Diseño de estrategias de evaluación.

2.6. Servicios

Para el logro de sus objetivos, el CENEVAL desarrolla las siguientes actividades genéricas:

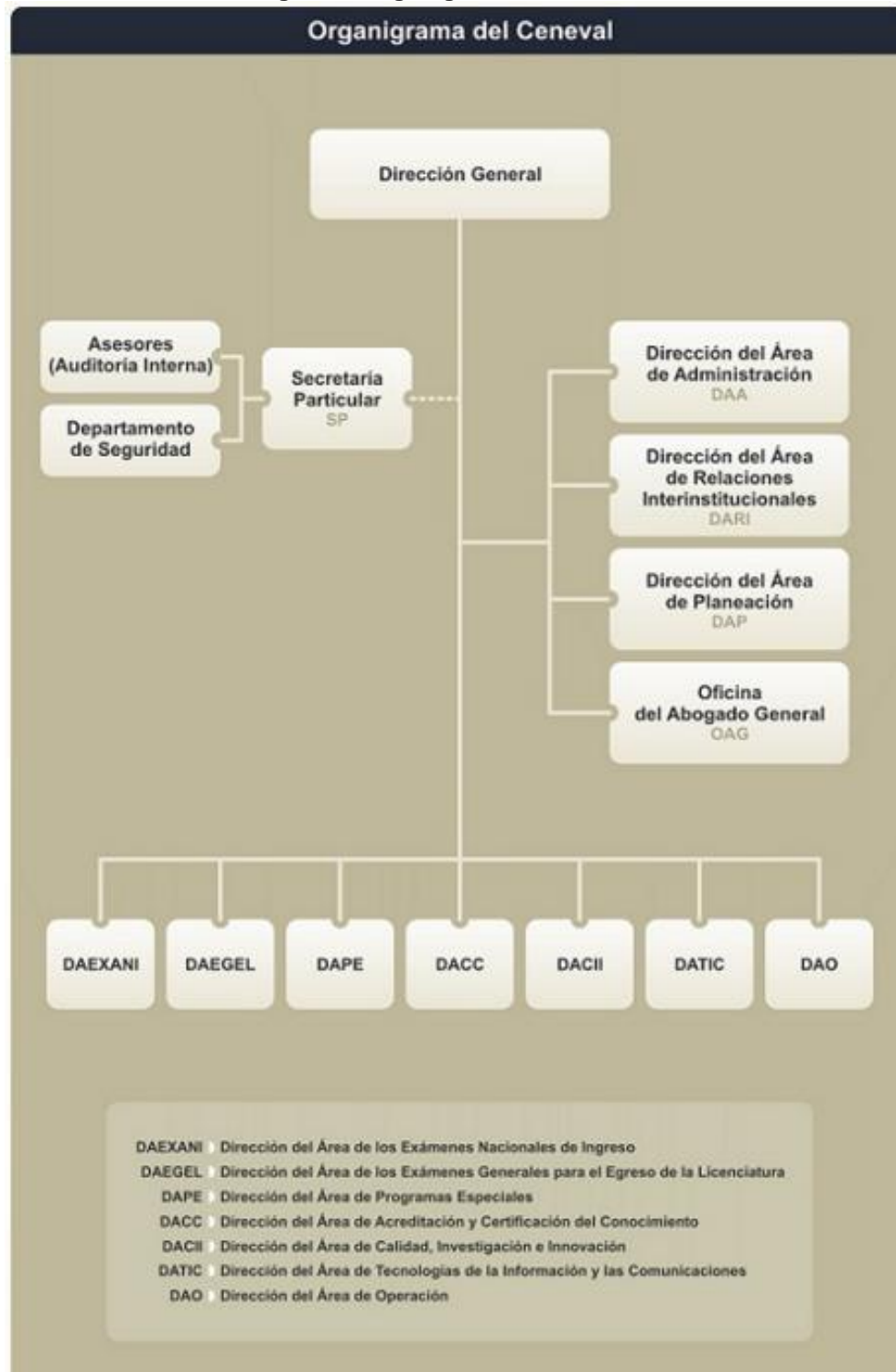
- Diseñar, elaborar, aplicar, calificar, validar y establecer la confiabilidad de exámenes para evaluar los resultados de aprendizaje logrados por estudiantes y egresados del sistema de educación media superior y superior en México.
- Aportar elementos externos que complementen los procesos de evaluación interna.

- Asesoría en materia de evaluación educativa a las instituciones para la preparación de los estudiantes en sus exámenes, el seguimiento de sus egresados, la interpretación de los resultados y la determinación de las medidas que conviene adoptar para el mejoramiento de la calidad educativa.
- Capacitación de evaluadores, tanto los requeridos por los exámenes del CENEVAL como otros con el fin de consolidar grupos institucionales que ayuden a desconcentrar posteriormente las tareas de evaluación educativa.
- Generar información útil que permita establecer parámetros de comparación entre instituciones y subsistemas educativos en el ámbito educativo regional y nacional. Tal información puede servir de base para realizar estudios e investigaciones que permitan la toma de decisiones fundamentadas.
- Difusión de sus funciones, programas y actividades.
- Realizar publicaciones, incluyendo la difusión de información nacional e internacional sobre evaluación.
- Identificación de personal especializado tanto en evaluación educativa como en las diversas áreas académicas que serán objeto de las evaluaciones.
- Identificación de instituciones internacionales prestigiadas en este campo e intercambio de experiencias, así como participación en proyectos internacionales.
- Organización de seminarios especializados.
- Experimentación en medios telemáticos para probar otras modalidades de la evaluación, de la capacitación, de la prestación de servicios, etcétera.
- Vinculación con las otras tareas evaluativas que realiza la Conaeva (autoevaluación institucional, interinstitucional y del conjunto del sistema de educación superior).
- Asesoría a colegios de profesionales y otros organismos interesados en exámenes de refrendo en algunas áreas especializadas.
- Fortalecimiento de la base documental y bibliográfica de la institución y comercialización de algunos de estos servicios.

- Búsqueda de fondos nacionales e internacionales para apoyar sus actividades.
- Promover programas académicos formales en el campo de la evaluación.
- Favorecer el desarrollo de la cultura de la evaluación y el establecimiento de estándares de calidad.

2.7. Estructura¹⁴

Figura 1. Organigrama del Ceneval



Fuente: <http://www.ceneval.edu.mx/ceneval-web/content.do?page=1704>

¹⁴ Manual organizacional, CENEVAL, 2016

2.8. Conformación de la Asamblea General de Asociados y del Consejo Directivo

El Centro se rige a través de la Asamblea General de Asociados y de sus Consejo Directivo, constituido de la siguiente manera:

Instituciones miembros de la Asamblea General de Asociados y del Consejo Directivo del CENEVAL A.C.

Asociaciones e instituciones educativas

Asociación Nacional de Universidades e Instituciones de Educación Superior, A.C. (ANUIES)1

Federación de Instituciones Mexicanas Particulares de Educación Superior, A.C. (FIMPES)1

Instituto Politécnico Nacional (IPN)1

Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM)

Universidad Autónoma del Estado de México (UAEM)1

Universidad Autónoma de San Luis Potosí (UASLP)

Universidad Autónoma de Yucatán (UADY)1

Universidad Nacional Autónoma de México (UNAM)2

Universidad Popular Autónoma del Estado de Puebla (UPAEP)1

Universidad Tecnológica de México (UNITEC)

Asociaciones y colegios de profesionales

Barra Mexicana Colegio de Abogados, A.C.1

Barra Colegio Nacional de Actuarios, A.C. (CNA)

Colegio Nacional de Psicólogos, A.C. (CNP)

Federación de Colegios y Asociaciones de Médicos Veterinarios y Zootecnistas de México, A.C.

Instituto Mexicano de Contadores Públicos, A.C. (IMCP)1

Organizaciones sociales y productivas

Academia de Ingeniería, A.C. 1

Academia Mexicana de Ciencias, A.C.

Academia Nacional de Medicina, A.C.

Fundación ICA, A.C.1

Autoridades educativas gubernamentales

Secretaría de Educación Pública1

2.9. Exámenes que opera el CENEVAL¹⁵

Exámenes Nacionales de Ingreso, EXANI

- EXANI-I
- EXANI-II
- EXANI-III
- PRE EXANI-I

Exámenes Generales para el Egreso de la Licenciatura, EGEL

- Arquitectura
- Biología
- Ciencia Política y Administración Pública
- Ciencias Agrícolas
- Ciencias Computacionales
- Ciencias de la Comunicación
- Comercio/Negocios Internacionales
- Contaduría
- Diseño Gráfico
- Enfermería
- Informática
- Ingeniería Civil
- Ingeniería Computacional
- Ingeniería de Software

¹⁵ Manual organizacional, CENEVAL, 2016

- Ingeniería Eléctrica
- Ingeniería Electrónica
- Ingeniería en Alimentos
- Ingeniería Industrial
- Ingeniería Mecánica
- Ingeniería Mecánica Eléctrica
- Ingeniería Mecatrónica
- Ingeniería Química
- Medicina General
- Medicina Veterinaria y Zootecnia
- Nutrición
- Odontología
- Psicología
- Química
- Química Clínica
- Químico Farmacéutico Biólogo

Exámenes Diagnósticos de Licenciatura

Examen Intermedio de Licenciatura en:

- Ciencias Básicas de Ingenierías
- Negocios (EXIL-Negocios)

Acuerdos 286 y 357

- ACREDITA-SEC (Acuerdo 286 Secundaria)
- ACREDITA-BACH (Acuerdo 286 Bachillerato)
- EGAL Exámenes Generales para la Acreditación de la Licenciatura (Acuerdo 286 Licenciatura)
- EGAL-EPRIM (Acuerdo 286 Licenciatura en Educación Primaria)
- EGAL-EPRE (Acuerdo 357 Licenciatura en Educación Preescolar)
- Examen General para la Acreditación como Técnico Superior Universitario en Comercialización Inmobiliaria (EGATSU-CI)

Otros exámenes

Exámenes UPN

- Licenciatura en Educación
- Examen General de Conocimientos de la Licenciatura en Educación Preescolar y Licenciatura en Educación Primaria para el Medio Indígena
- Examen para Profesionales Técnicos en Enfermería.
- Examen General para el Egreso del Profesional Técnico en Enfermería, EGEPT-ENFER

Exámenes de Certificación

- Uniforme de Certificación de la Contaduría Pública, EUC
- Único de Certificación de la Calidad del Actuario, EUCCA

Exámenes de Evaluación de Lengua Inglesa con Fines Profesionales

- EGAL-EIN Examen General de Conocimientos y Habilidades para la Acreditación de la Licenciatura en Enseñanza del Inglés
- Proceso de Acreditación de los Docentes de Inglés
- APTIS
- TKT (Teaching Knowledge Test) Cambridge ESOL

Examen de Certificación de Español

- Certificación del Español como Lengua Extranjera, ECELE
- Expresión Escrita en Español (Expresé)

Evaluación de Competencias Docentes

- Proceso de Evaluación de Competencias Docentes para la Educación Media Superior (ECODEMS)

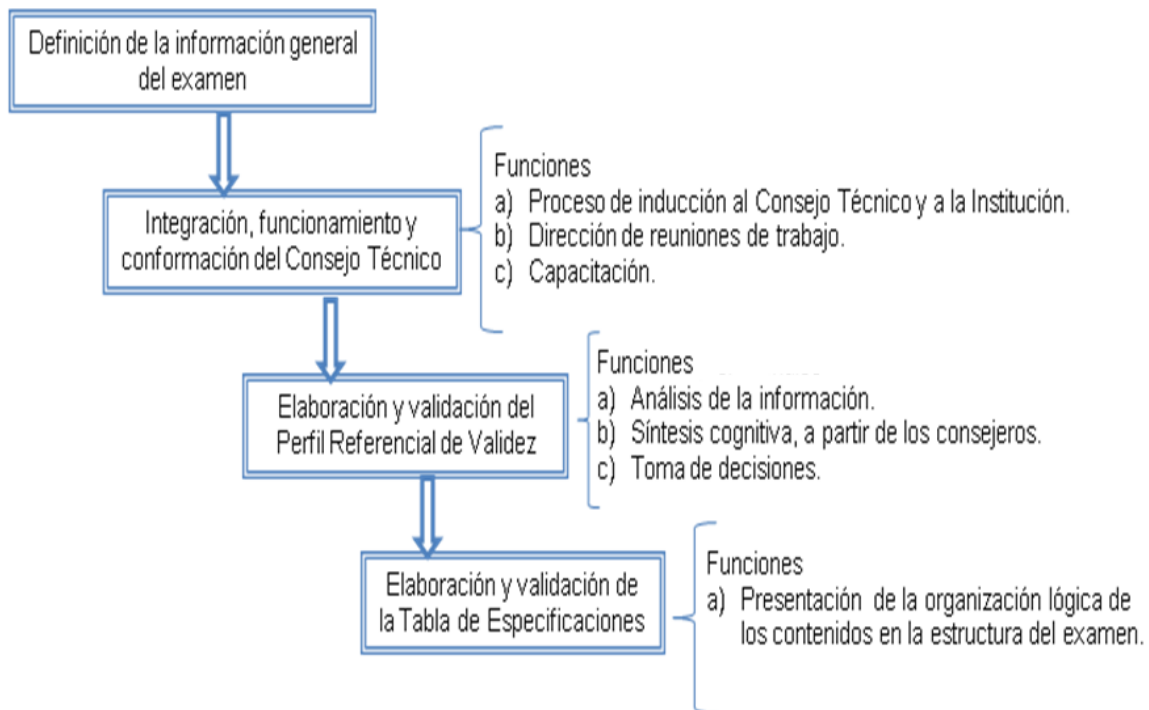
Sistema Procesal Penal Acusatorio

- Examen Diagnóstico de Conocimientos y Habilidades en el Sistema Procesal Penal Acusatorio (EXSIPA)

3. Descripción de actividades del Maestro en Educación en el diseño de la Prueba de Matemáticas para Bachilleratos Tecnológicos Federales

El proceso de construcción del examen está dividido por etapas. A continuación lo presento, ya que creo que es útil visualizarlo en su totalidad. La siguiente descripción de las actividades del Maestro en Educación en las etapas de diseño y elaboración.

Figura 2. Proceso de construcción de un examen



Fuente: Prontuario de normas técnicas y estándares del Ceneval, 2011

3.1. Características generales de la Prueba de Matemáticas para Bachilleratos Tecnológicos Federales

Aquellas actividades relacionadas directamente con las implicaciones y funciones profesionales que realicé durante el diseño del examen y que me involucran como Maestro en Educación se presentan a continuación y están ordenadas según las etapas del diseño de la prueba.

La primera actividad es la definición de la información general del examen, cuyo propósito es definir cuál sería el objetivo del examen, sus características principales y a quién se dirigiría el examen o población objetivo. Esta información fue definida con las autoridades del CENEVAL y la institución que solicita el servicio de manera conjunta.

La Prueba de Matemáticas para Bachilleratos Tecnológicos Federales es una evaluación diagnóstica que mide los conocimientos y habilidades de Álgebra, Geometría y Trigonometría, Cálculo, Geometría Analítica, Probabilidad y Estadística y Matemáticas Aplicadas establecidos en los programas de estudio del bachillerato, acorde con la Reforma Integral de la Educación Media Superior (RIEMS).

3.2. Objetivo de la prueba

El objetivo de la prueba es evaluar los conocimientos y habilidades matemáticas adquiridas en la instrucción escolar por los jóvenes que concluyen el bachillerato tecnológico en las instituciones públicas de educación media superior.

3.3. Población objetivo

Esta prueba está dirigida a los alumnos que estudian en planteles de la Dirección General de Educación Tecnológica Industrial (DGETI) y de la Dirección General de Educación Tecnológica Agropecuaria (DGETA).

3.4. Características

La Prueba de Matemáticas para Bachilleratos Tecnológicos Federales es una prueba objetiva, ya que está construida bajo una metodología que garantiza validez y confiabilidad de los resultados. Asimismo, es una prueba criterial. Este tipo de evaluaciones se caracterizan porque:

- a)** Parten de la determinación de los conocimientos y habilidades que debe poseer un sustentante. Estos conocimientos y habilidades se delimitan en un perfil referencial (PR), con el cual es posible establecer de manera clara y precisa aquello que se pretende medir.
- b)** Son de carácter absoluto, es decir, el desempeño de un individuo es independiente de lo que responden otros individuos, a diferencia de la *evaluación referente a la norma*, la cual compara el resultado del individuo con los resultados de una población o grupo a los que pertenece. La evaluación criterial permite identificar con precisión el rendimiento de un sustentante en función de su desempeño al resolver la prueba y con relación a un criterio previamente fijado.
- c)** El establecimiento de categorías y de puntos de corte. De acuerdo con el rendimiento de un sustentante en la prueba, este podrá ser ubicado en una categoría establecida: insuficiente, elemental, bueno o excelente.

Asimismo, la prueba es sensible a la instrucción, lo cual implica que se apega al currículo establecido para el año de estudio correspondiente y puede detectar, a lo largo del tiempo, variaciones en los conocimientos y habilidades adquiridos por los estudiantes.

3.5. Alcances

La Prueba de Matemáticas para Bachilleratos Tecnológicos Federales es uno de los instrumentos del Proyecto de Alineación de Incentivos (*ALI*), por lo cual sus resultados:

- a) Solamente son válidos para los grupos de población incorporados al proyecto.

- b) No pueden generalizarse al subsistema de educación tecnológica, ni a otros subsistemas de educación media superior.

Finalmente, cabe recalcar que el diseño y propósitos de la prueba se limitan a la emisión de un diagnóstico general sobre el nivel de dominio en el que se ubica el sustentante, respecto a los conocimientos y habilidades evaluados. Las decisiones que se toman a partir de los resultados deben considerar cuidadosamente los objetivos, estrategias y procedimientos del proyecto (*ALI*), así como el contexto de la evaluación y las condiciones de cada escuela.

3.6. Administración

La Prueba de Matemáticas para Bachilleratos Tecnológicos Federales se aplica al término del ciclo escolar 2009-2010, a una muestra de alumnos seleccionados por la SEMS y su equipo de investigadores de la Universidad de Pennsylvania, acorde con los fines del proyecto (*ALI*). La responsabilidad en la organización logística y aplicación recae en la SEMS. La prueba se aplica tanto a alumnos de los grupos control como a alumnos de grupos experimentales.

3.7. Resultados

Se emiten distintos tipos de reportes, según los propósitos del proyecto de al (*ALI*):

1. Alumnos con tratamiento experimental e incentivos individuales.
2. Alumnos con tratamiento experimental e incentivos individuales y por grupo.

Los reportes indican el nivel de dominio obtenido por el sustentante al aplicar la Prueba de Matemáticas para Bachilleratos Tecnológicos Federales: insuficiente, elemental, bueno y excelente. También informan al sustentante el monto del incentivo a que se ha hecho acreedor.

4. Etapas para la construcción de la prueba

4.1. Antecedentes de la prueba

En abril de 2008, la (SEMS) de la SEP solicitó el apoyo del CENEVAL para la elaboración de un instrumento que evaluara las habilidades y conocimientos en matemáticas de los estudiantes de bachillerato.

El proyecto fue llevado a cabo por la SEMS, en coordinación con un equipo de investigadores de la Universidad de Pennsylvania, e intento probar la eficacia de los estímulos para mejorar el aprendizaje. Para ello prevé la dotación de incentivos a alumnos, docentes y personal administrativo, en función del rendimiento alcanzado por los estudiantes en la prueba.

Para ese efecto, el proyecto prevé la organización de grupos control y experimentales. Los primeros no recibirán información sobre el proyecto, en tanto que los segundos se dividirán de la siguiente manera: Tratamiento 1 (T1): entrega de incentivos para alumnos; Tratamiento 2 (T2): entrega de incentivos a alumnos y docentes, y Tratamiento 3 (T3): entrega de incentivos a alumnos, docentes y personal administrativo.

De acuerdo con el proyecto, el instrumento debía apegarse al currículo del bachillerato, informar sobre los conocimientos y habilidades adquiridos por los alumnos en cada grado escolar y, en el caso de la población de tercer grado, informar sobre el dominio de contenidos de los tres grados escolares. Asimismo, el instrumento debía permitir la comparación de resultados respecto al nivel de dominio obtenido por los estudiantes en la prueba *Enlace Básica, Tercero de Secundaria* y la realización de estudios de seguimiento sobre sus avances.

Debido a la multiplicidad de programas de estudio para el bachillerato (más de 200), así como las diferencias entre las modalidades que comprende la educación media superior (bachillerato general, bachillerato tecnológico y bachillerato con profesional técnico), el CENEVAL presentó una propuesta a la SEMS y a su equipo de investigadores de la Universidad de Pennsylvania consistente en la construcción de tres instrumentos, uno para cada grado escolar, enfocados al área de Matemáticas y para el subsistema de bachillerato tecnológico, a fin de acotar la evaluación en cuanto a contenidos y población de estudio.

La prueba de primer grado comprendería las asignaturas de primero y segundo semestres, Álgebra, Geometría y Trigonometría. La de segundo grado comprendería las asignaturas de tercero y cuarto semestres, Geometría Analítica y Cálculo Diferencial. La prueba de tercer grado comprendería las asignaturas de Probabilidad y Estadística, y Matemáticas Aplicadas.

Las pruebas se aplicarían tanto a los alumnos de los grupos control y experimentales al término de los ciclos escolares 2008-2009, 2009-2010 y 2010-2011, y se construirían considerando la coexistencia de los programas de estudio de la Reforma Curricular 2004-2007 y la Reforma Integral de la Educación Media Superior (RIEMS), vigente a partir del ciclo escolar 2008-2009.

Así, para los dos primeros años de aplicación, la prueba de primer grado se diseñaría tomando como marco de referencia los contenidos de la Reforma Integral de la Educación Media Superior (RIEMS), en tanto que las de segundo y tercer grados considerarían los contenidos comunes a ambas reformas.

Para el tercer año de aplicación, ciclo escolar 2010-2011, las pruebas de primero y segundo grados se construirían de la misma forma, en tanto que la prueba de tercer grado se modificaría: comprendería los contenidos comunes a ambas reformas en Probabilidad y Estadística, y se substituiría Matemáticas Aplicadas por Cálculo Integral, nueva asignatura de la RIEMS a impartirse en ese ciclo escolar.

Las pruebas de Matemáticas para Bachilleratos Tecnológicos Federales se construyen como una prueba objetiva, con referencia a criterio y sensible a la instrucción.

Los resultados de la prueba son válidos solamente para la población interviniente y no permiten realizar generalizaciones sobre el desempeño del subsistema tecnológico, de la educación media superior o de las entidades federativas que participan.

La construcción de la prueba de tercer grado del ciclo escolar 2008-2009 fue reportada en el manual técnico con fecha diciembre de 2009; la prueba de tercer grado del ciclo escolar 2009-2010 es motivo de este manual técnico.

Su construcción implicó la realización de diversos procesos que se integran en las siguientes etapas: 1) diseño, 2) construcción, 3) verificación, 4) ensamble, 5) aplicación, 6) calificación y 7) reportes, las cuales se alinean con la metodología del CENEVAL y cumplen con estándares de calidad internacionales. Las tres últimas etapas estuvieron a cargo de la SEMS y su equipo de investigadores de la Universidad de Pennsylvania.

En cada una de estas etapas intervinieron cuerpos colegiados, como el Consejo Consultivo y los Comités Académicos, conformados por especialistas en el área de estudio de la prueba, que avalaron y legitimizaron cada proceso.

A continuación, se informa sobre el personal que integra cada cuerpo colegiado que intervino en la elaboración de la prueba, así como sus funciones.

- **Consejo Consultivo.** Integrado por las autoridades de la Subsecretaría de Educación Media Superior y el equipo de investigadores de la Universidad de Pennsylvania. Es el encargado de vigilar el diseño, instrumentación y análisis de la prueba, impulsando acciones que permitan su pertinencia y congruencia con los objetivos propuestos. Valora la estructura de la prueba y sus contenidos; revisa los resultados del estudio piloto y propone modificaciones a partir de ellos; evalúa técnica y administrativamente la forma de aplicación de la prueba. Valida y propone los puntos de corte, define los reportes que se deriven y sus usos. Propone la estructura y contenido de los cuestionarios de contexto y los mecanismos para su aplicación.
- **Comité Académico de Diseño del Perfil Referencial.** Integrado por docentes especialistas en los contenidos y habilidades de las asignaturas de Matemáticas del nivel medio superior con conocimiento y experiencia en el subsistema de bachillerato tecnológico. Es el responsable de la elaboración del perfil referencial y la estructura de la prueba. Define sus contenidos y elabora las especificaciones.
- **Comité Académico de Validación del Perfil Referencial.** Integrado por docentes especialistas en los contenidos y habilidades de las asignaturas de Matemáticas del nivel medio superior, con conocimientos y experiencia en el subsistema de bachillerato tecnológico. Verifica el diseño del perfil referencial y estructura de la prueba. Confirma la pertinencia, claridad, suficiencia y consistencia de los contenidos y las especificaciones.

- **Comité Académico de Elaboración de Reactivos.** Integrado por docentes y especialistas en los contenidos y habilidades de las asignaturas de Matemáticas del nivel medio superior, con conocimientos y/o experiencia en el subsistema de bachillerato tecnológico. Elabora los reactivos a partir de las especificaciones definidas por el Comité de Diseño. Se apega a las delimitaciones y a los lineamientos establecidos para la elaboración de reactivos.
- **Comité Académico de Validación de Reactivos.** Integrado por docentes y especialistas en los contenidos y habilidades de las asignaturas de Matemáticas del nivel medio superior, con conocimiento y/o experiencia en el subsistema de bachillerato tecnológico. Garantiza que los reactivos elaborados correspondan a las especificaciones y delimitaciones establecidas, vigila la claridad, precisión consistencia y suficiencia del contenido.
- **Comité Académico de Puntos de Corte.** Integrado por docentes y especialistas en los contenidos y habilidades de las asignaturas de Matemáticas del nivel medio superior, con conocimiento y/o experiencia en el subsistema de bachillerato tecnológico. Determina las características de los sustentantes en cada nivel de dominio, analiza las características de las pruebas y sus resultados, y aplica los procedimientos para determinar puntos de corte que permitan diferenciar a los que saben de los que no.

4.2. Diseño

Durante el diseño de un instrumento de evaluación se definen los elementos que constituyen el marco de referencia en que se fundamentará todo el proceso, como son el perfil referencial (PR) y la estructura de la prueba. Un PR es el referente básico para la elaboración de la prueba y constituye el material fundamental en que se sustenta la validez del instrumento. Por su parte, la estructura de la prueba es una representación esquemática que indica las partes constitutivas del instrumento y su ponderación.

El CENEVAL reconoce la valiosa contribución que brindan los miembros de los consejos técnicos y comités académicos con la finalidad de lograr la consolidación de estos. Se emitió un reglamento en el que se especifican sus funciones, integración y formas de organización. Esta normativa ha dado lugar a la renovación de estos cuerpos colegiados.

4.2.1. Perfil Referencial

El (PR) es un conjunto estructurado de contenidos y especificaciones que señala de manera clara y precisa aquello que se pretende medir.

Para la Prueba de Matemáticas para Bachilleratos Tecnológicos Federales se consideró el perfil referencial elaborado para la prueba correspondiente a 2009. Como parte del Proyecto (ALI), el Consejo Consultivo solicitó que la prueba 2011 evaluara los mismos conocimientos y habilidades valorados con la prueba del año anterior y así posibilitar su equiparación.

El PR de la Prueba de Matemáticas para Bachilleratos Tecnológicos Federales, parte de la siguiente conceptualización de la Matemática desarrollada por los integrantes del Comité de Diseño, así como de una concepción del aprendizaje planteada en el programa de estudios.

La matemática es uno de los lenguajes universales que permite desarrollar nuevos conocimientos y utilizar el pensamiento y razonamiento lógico en situaciones de la vida, en un contexto histórico.

En el Bachillerato Tecnológico, la matemática es una herramienta que promueve y genera el uso de conceptos, estrategias, procedimientos y recursos para la resolución de problemas numéricos, geométricos, estadísticos y probabilísticos en el ámbito científico y tecnológico.

Desde la perspectiva de la Reforma Integral de la Educación Media Superior (RIEMS), el aprendizaje es primordialmente un proceso durante el cual el educando recorre un camino y en ese recorrido va dando cuenta de sus aprendizajes a partir de diversos productos que puede elaborar en distintos momentos del proceso.

Además de que el aprendizaje es un producto observable y medible, también es un producto subjetivo, ya que el educando es un sujeto que durante la construcción del conocimiento desarrolla no solo su dimensión intelectual, sino sus dimensiones afectiva y física (SEP: 2008-5).

El campo de la aplicación de la matemática es muy amplio y brinda herramientas fundamentales para emplearse en las áreas de conocimiento y sobre todo en el contexto cotidiano. Esto da la pauta para abordar problematizaciones como problemas, hechos, fenómenos naturales y sociales, entre otros, a partir de los contenidos de las asignaturas de la Matemática (SEP: 2008-26).

Además de lo anterior, para la elaboración del PR se consideraron los siguientes propósitos de Matemáticas establecidos en la RIEMS, así como los de las asignaturas correspondientes al grado escolar en cuestión.

Los propósitos educativos de la Matemática se explicitan a través de competencias específicas, las cuales se describen a continuación:

1. Argumentar razonamientos, a través del lenguaje verbal y matemático, con métodos numéricos, gráficos, analíticos y variacionales, de tal manera que explicita y justifique la solución de distintos problemas.
2. Utilizar, representar e interpretar modelos matemáticos mediante la aplicación de procedimientos aritméticos, algebraicos, geométricos y variacionales, para la comprensión de situaciones reales o formales.
3. Comparar dos o más variables o números de tal manera que se establezca o analice su relación y permita determinar, estimar o aproximar el comportamiento de fenómenos del entorno social o natural.
4. Utilizar las mediciones con instrumentos para representar y contrastar matemáticamente las dimensiones espaciales, estructurales y funcionales del ambiente que nos rodea.
5. Operar y relacionar de diferentes maneras las formas de eslabonamiento entre los símbolos o procesos matemáticos y aplicarlos al estudio de fenómenos específicos.
6. Valorar las Matemáticas como una ciencia que está en evolución, y como obra del ser humano, que ha permitido el estudio de los fenómenos de su entorno físico y abstracto.
7. Reconocer la pertinencia y aplicar el enfoque determinista o aleatorio al estudio de procesos o fenómenos.

El perfil referencial cubre las siguientes áreas en la prueba y sus propósitos son:

- **Álgebra**

Desarrollar la capacidad del razonamiento matemático haciendo uso del lenguaje algebraico, a partir de la resolución de problemas de la vida cotidiana, dentro y fuera del contexto matemático, representados en modelos donde se aplican conocimientos y conceptos algebraicos en un clima de colaboración y respeto.

- **Geometría y Trigonometría**

Desarrollar la capacidad de la orientación espacial mediante el análisis y representación de problemas que implican figuras geométricas en un clima de participación y responsabilidad.

- **Geometría Analítica**

Desarrollar las capacidades del razonamiento matemático y la orientación espacial, mediante la resolución de problemas que implican modelos matemáticos representados en el plano cartesiano, en un ambiente propicio para el aprendizaje colaborativo.

- **Cálculo**

Desarrollar la capacidad del razonamiento matemático, mediante el análisis e interpretación de las relaciones entre dos variables que provienen de problematizaciones surgidas de la actividad humana y de los fenómenos naturales, en un ambiente propicio para el aprendizaje colaborativo.

- **Probabilidad y Estadística**

Desarrollar la capacidad del razonamiento matemático utilizando las herramientas básicas de la estadística descriptiva y de la teoría de la probabilidad para muestrear, procesar y comunicar información social y científica para la toma de decisiones.

- **Matemáticas Aplicadas**

Los estudiantes aplicarán conceptos básicos de Matemáticas en el tratamiento específico de problemáticas sociales, económicas y de la ciencia y la tecnología, permitiendo así fortalecer sus conocimientos.

En cuanto a las subáreas, el PR comprende los conceptos fundamentales de los programas de estudio asociados a cada una de las áreas:

- **Álgebra**
 - Lenguaje algebraico
 - Ecuaciones
- **Geometría y Trigonometría**
 - Geometría
 - Trigonometría
- **Geometría Analítica**
 - Sistemas coordenados
 - Lugares Geométricos
- **Cálculo**
 - Funciones
 - Límites
 - Derivadas
- **Probabilidad y Estadística**
 - Probabilidad
 - Estadística descriptiva
- **Matemáticas Aplicadas**
 - Teoría elemental del muestreo
 - Teoría de estimación estadística

Como parte de la metodología de elaboración del PR, estos conceptos fundamentales se desagregan en un conjunto de conocimientos y habilidades específicos, asegurando un muestreo de los contenidos relevantes de cada área y subárea de la prueba.

El PR de la Prueba de Matemáticas para Bachilleratos Tecnológicos Federales incluye el nivel cognitivo con que se evalúa cada uno de los conocimientos y habilidades específicos establecidos. Para ello, comprende una taxonomía con los siguientes niveles: reconocimiento, desarrollo y transferencia.

1. Reconocimiento

Las habilidades de este grupo se refieren a tareas que permiten determinar si el sustentante conoce los conceptos y notaciones matemáticas, así como las formas de representación numérica, simbólica y gráfica.

En este grupo de procesos, el sustentante realiza principalmente tareas de recuperación inmediata de información.

2. Desarrollo

Las habilidades de este grupo incluyen tareas que permiten determinar si el sustentante conoce y aplica las técnicas para la resolución de operaciones o funciones dadas.

En este grupo de procesos el sustentante también debe obtener equivalencias y traducir de un lenguaje a otro; asimismo, debe identificar los procedimientos, las reglas o las operaciones necesarias para resolver una situación determinada e interpretar los resultados presentados en lenguaje numérico, simbólico o gráfico.

3. Transferencia

Las habilidades de este grupo incluyen la resolución de problemas o situaciones similares a la vida real.

Los problemas de este tipo plantean exigencias en su comprensión y requieren que el sustentante reconozca y aplique la técnica matemática correspondiente con el fin de solucionar problemas numéricos, simbólicos y gráficos.

4.2.2 Especificaciones

Las especificaciones forman parte del PR y tienen como propósito proporcionar un marco normativo, claro y significativo, que aporte a los elaboradores de reactivos los elementos necesarios para interpretar, definir, delimitar y determinar la profundidad y alcance de los contenidos y procesos cognitivos por ser evaluados, y así generar reactivos efectivos e interpretaciones válidas.

Adicionalmente, las especificaciones sirven para comunicar a los profesores, investigadores, autoridades educativas y sociedad en general, qué es lo que mide cada reactivo, así como la cobertura de conocimientos y habilidades evaluadas.

Para la Prueba de Matemáticas para Bachilleratos Tecnológicos Federales, Primer Grado, 2011, el Comité Académico de Diseño elaboró, desde 2008, las especificaciones respectivas. Para este efecto, se analizaron los conocimientos y habilidades registrados en el PR, el nivel cognitivo asignado y la posible estructura del reactivo por presentar a los sustentantes. Las especificaciones así diseñadas deben servir de base para la elaboración de reactivos.

La capacitación respecto al proceso de elaboración de especificaciones fue proporcionada en las sesiones del Comité de Diseño del PR. La información básica es la siguiente:

Cada especificación consta de dos partes fundamentales: acción y condición. La acción se delimita con un verbo en infinitivo que refleja el manejo de conocimientos o habilidades específicos que se espera por parte del sustentante.

La condición se refiere a las situaciones, los elementos o las informaciones que deben proporcionarse al sustentante para que realice la acción solicitada.

Cada especificación se acompaña de una delimitación en la cual se señala el tipo de situaciones, datos o informaciones que se deben proporcionar en el reactivo.

4.2.3. Estructura de la prueba

Las especificaciones de la prueba también permiten definir el número de reactivos que deben asignarse para la evaluación de cada conocimiento o habilidad establecidos en el PR y, por lo tanto, la estructura final de la prueba.

Cuadro 1. Estructura de la Prueba					
Área	Subárea	Reactivos	Procesos cognitivos		
			Reconocimiento	Desarrollo	Transferencia
Álgebra	Lenguaje Algebraico	20	10	9	1
	Ecuaciones	18	6	9	3
Subtotal		38	16	18	4
Geometría y Trigonometría	Geometría	23	12	5	6
	Trigonometría	19	7	9	3
Subtotal		42	19	14	9
Geometría Analítica	Sistemas Coordenados	18	4	8	6
	Lugares Geométricos	27	2	16	9
Subtotal		45	6	24	15
Cálculo Diferencial	Funciones	7	2	5	0
	Límites	7	2	5	0
	Derivada	16	3	10	3
Subtotal		30	7	20	3
Probabilidad y Estadística	Estadística Descriptiva	44	5	20	7
	Probabilidad	16	1	6	6
Subtotal		45	6	26	13
Matemáticas Aplicadas	Teoría Elemental de Muestreo	22	3	17	2
	Teoría de Estimación Estadística	3	1	0	2
Subtotal		25	4	17	4
Total		225	58	119	48

Fuente: Manual Técnico del examen de Bachilleratos Tecnológicos

Como podrá observarse, la estructura contempla como áreas y subáreas a las asignaturas y conceptos fundamentales del grado escolar en cuestión. Para cada una de esas se anota el total de reactivos. De acuerdo con los procesos cognitivos establecidos, se anota la cantidad de reactivos asignada.

4.2.4. Validación del perfil referencial y estructura de la prueba

La estructura de la prueba y el PR fueron revisados y validados por el Comité de Validación y consideró los siguientes criterios:

-
- Apego de la estructura de la prueba con las áreas y subáreas establecidas en el programa de estudio
- Pertinencia de los conocimientos y habilidades por evaluar respecto a los propósitos del programa de estudios
- Suficiencia de los conocimientos y habilidades específicos por evaluar
- Pertinencia de los procesos cognitivos para la evaluación de conocimientos y habilidades específicos
- Claridad de las especificaciones
- Correspondencia de las especificaciones con los conocimientos y habilidades específicos por evaluar
- Dificultad de los reactivos en las pruebas piloto
- Correlación punto biserial de los reactivos en las pruebas piloto
- Dominio de los conocimientos y habilidades por los estudiantes
- Abordaje de los conocimientos y habilidades en el salón de clases
- Claridad de los reactivos

4.3. Construcción

La actividad principal en la etapa de construcción de una prueba es la elaboración de reactivos. En el caso de esta Prueba de Matemáticas para Bachilleratos Tecnológicos Federales, el CENEVAL convocó a especialistas en matemáticas para integrar el Comité de Elaboración respectivo.

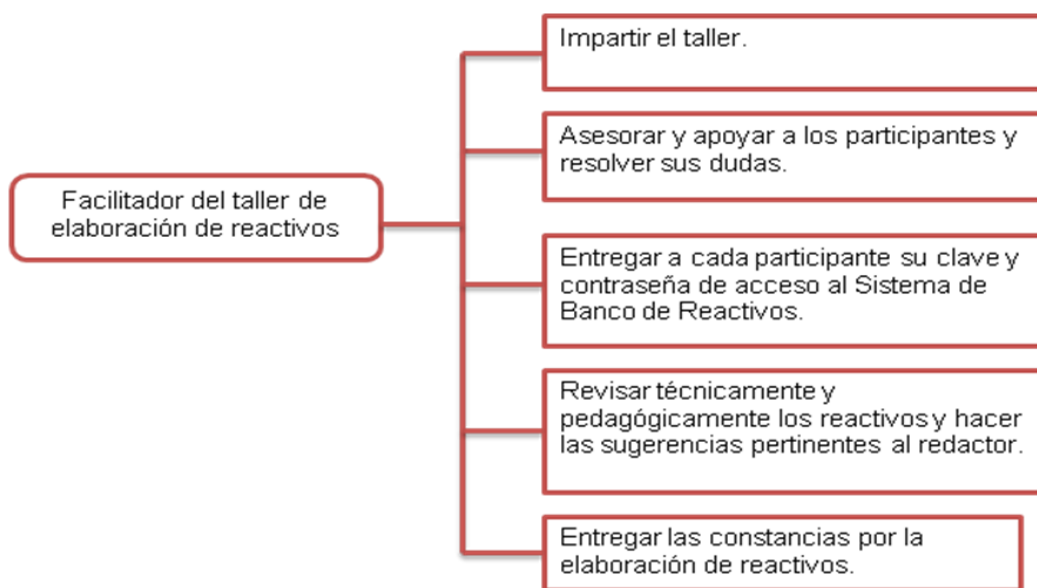
Un requisito indispensable para participar en el comité fue que sus integrantes contaran con experiencia en el nivel medio superior y tuvieran los conocimientos sobre las áreas que evalúa la prueba.

La elaboración de los reactivos de la Prueba de Matemáticas para Bachilleratos Tecnológicos Federales partió de las necesidades que presentan el PR que son tres veces el número de reactivos que pide.

Durante este proceso apoyé en la impartición de cinco talleres de capacitación. En total se contó con 26 elaboradores que construyeron más de 350 reactivos.

La capacitación se basó en la exposición del marco institucional del CENEVAL, algunos aspectos teóricos de la evaluación educativa, las etapas para la construcción de una prueba objetiva y la información general sobre las pruebas de Matemáticas para la prueba. De manera comentada y con base en ejemplos, se expusieron los lineamientos institucionales para la elaboración de reactivos.

Figura 3. Diagrama de taller de elaboración de reactivos



Fuente: Prontuario de normas técnicas y estándares del Ceneval, 2011.

Dada la capacitación, los especialistas elaboraron los reactivos, según las especificaciones asignadas. Los especialistas del CENEVAL registraron los reactivos en el Banco de Reactivos para Asesores Externos (BRAE).

4.3.1. Verificación

El objetivo de esta fase fue verificar la calidad de los reactivos que se elaboraron. Todos los reactivos pasaron por una revisión cualitativa y una cuantitativa.

Revisión cualitativa:

- Revisión técnica de reactivos
- Validación de reactivos
- Revisión de estilo

Revisión cuantitativa:

- Piloteo de reactivos
- Calibración

4.3.2. Revisión cualitativa. Revisión técnica

Consiste en verificar que los reactivos correspondan a la especificación del PR, al proceso cognitivo por evaluar y a los lineamientos institucionales para la elaboración de reactivos. Asimismo, comprende la revisión del contenido.

Esta actividad la realiza personal especializado del CENEVAL considerando lo siguiente:

1. **Base del reactivo.** Se revisó que incluyera la información necesaria para poder contestarlo adecuadamente, que cumpliera con alguno de los formatos establecidos y que fuera independiente, es decir, que la información contenida en un reactivo no sugiriera la solución de otro.

Opciones de respuesta. Se verificó que los reactivos tuvieran solo una respuesta correcta y tres distractores. En lo referente a la respuesta correcta, esta tuvo que ser incuestionable y resolver el planteamiento de manera satisfactoria, mientras que los distractores tuvieron que ser plausibles y totalmente incorrectos. Esto se corroboró mediante la revisión de la “argumentación” a cada opción de respuesta.

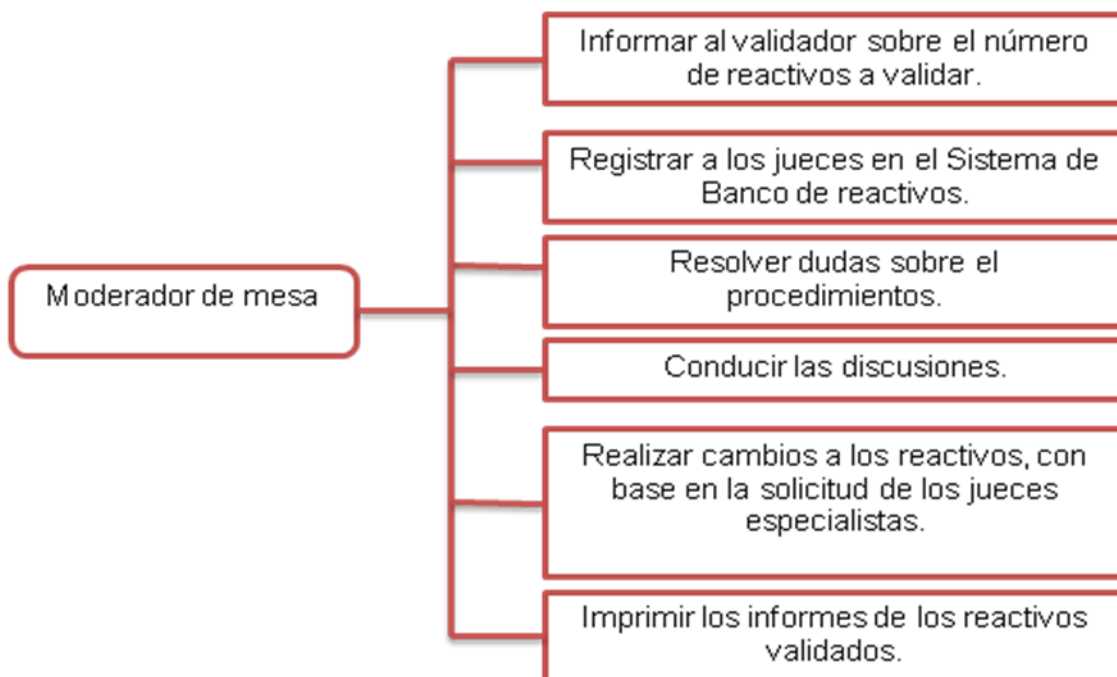
En caso de que algún reactivo no cumpliera con los criterios señalados, se proporcionó retroalimentación al elaborador para su modificación.

4.3.3. Revisión cualitativa. Validación externa

Una vez concluida la elaboración de reactivos, aquellos que cumplieron con los lineamientos técnicos del CENEVAL fueron sometidos a un proceso de validación con el fin de garantizar que correspondieran al contenido definido en el PR y representaran, de manera clara y precisa, los elementos esenciales de los conocimientos y habilidades evaluados en la prueba. Cada reactivo fue validado por lo menos por dos y máximo por especialistas.

Para llevar a cabo este proceso, los especialistas recibieron una capacitación basada en los aspectos institucionales del CENEVAL, la elaboración de pruebas, las características de las pruebas de Matemáticas para el Proyecto de **ALI** y los criterios para la validación de reactivos.

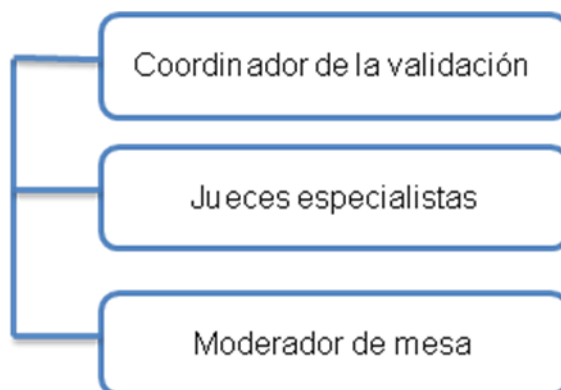
Figura 4. Diagrama de validación de reactivos



Fuente: Prontuario de normas técnicas y estándares del Ceneval, 2011.

Se organizaron mesas de trabajo coordinadas por un representante del CENEVAL, cuya tarea consistió en informar a los especialistas sobre el número de reactivos por validar, resolver preguntas sobre el procedimiento, realizar los cambios necesarios a los reactivos a solicitud de los especialistas e imprimir los informes de los reactivos validados.

Figura 5. Diagrama de Integración de mesas de trabajo



Fuente: Prontuario de normas técnicas y estándares del Ceneval, 2011.

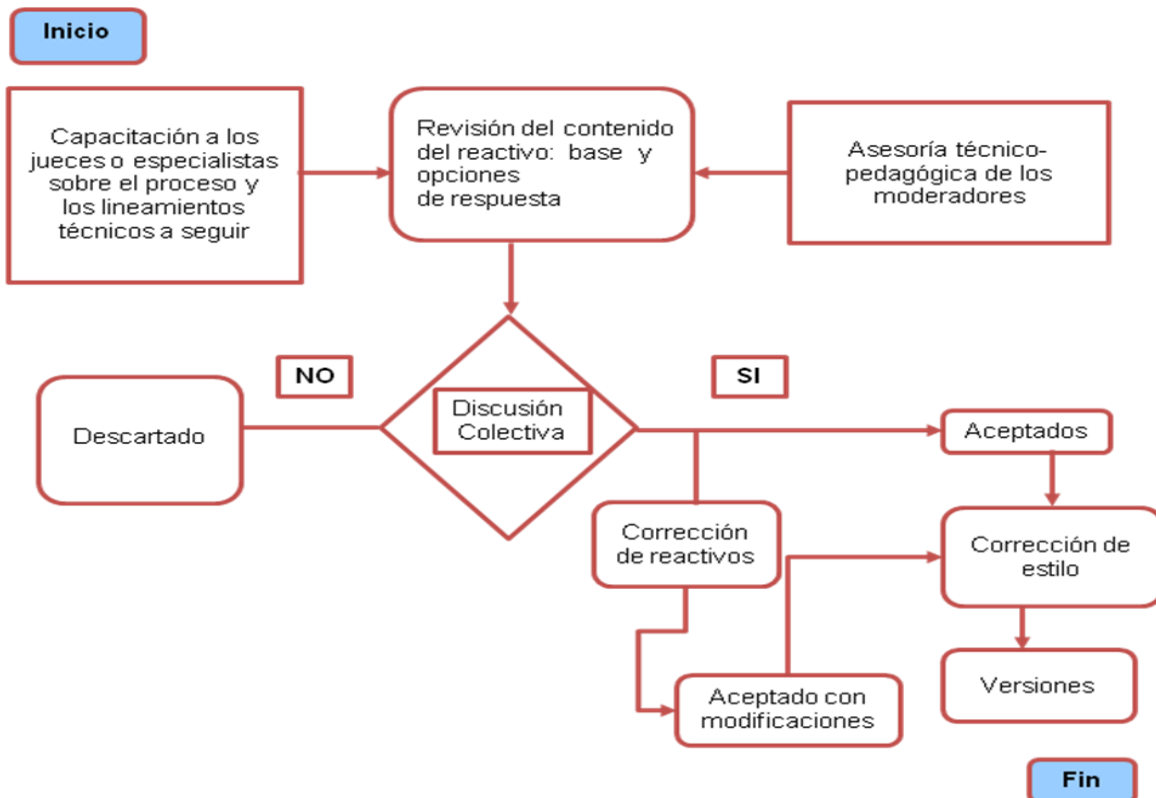
Por su parte, los especialistas se encargaron de leer y analizar cada reactivo de manera individual, discutir colectivamente su contenido, emitir un dictamen y, en su caso, hacer propuestas de mejora.

Figura 6. Diagrama de especialistas



Fuente: Prontuario de normas técnicas y estándares del Ceneval, 2011.

Figura 7. Diagrama de taller de validación de reactivos



Fuente: Prontuario de normas técnicas y estándares del Ceneval, 2011.

4.3.4. Revisión cualitativa. Revisión de estilo

La revisión de estilo de los reactivos se llevó a cabo por los expertos del CENEVAL y consistió en asegurar una presentación clara, sistemática y homogénea de cada reactivo. Se revisó el empleo correcto y uniforme de la puntuación, la ortografía, las abreviaturas y la presentación de tablas y otros apoyos gráficos, así como las fórmulas y notaciones. Las correcciones derivadas de la revisión se incorporaron a los reactivos capturados en el sistema BRAE.

4.3.5. Revisión cuantitativa. Piloteo de reactivos

Posterior al proceso de validación, se llevó a cabo el ensamble de las versiones para el estudio piloto de los reactivos. Para esto, se consideraron únicamente los reactivos que en la validación resultaron aceptados. Después de hacer las correcciones se ensamblaron los instrumentos.

La propuesta de piloteo fue avalada por el Consejo Consultivo de la prueba. En esta se determinó la población sustentante, el número de versiones y el total de reactivos que las integran.

La población sustentante se constituyó con los alumnos de bachilleratos tecnológicos. Para asegurar que la población había cursado las asignaturas respectivas, tanto la SEMS como sus investigadores propusieron aplicar los reactivos de primer grado a los estudiantes de segundo grado.

En cuanto al número de reactivos por pilotear, se propuso probar tres veces la cantidad de reactivos de la prueba y ensamblar cuadernillos cortos que permitieran agilizar la resolución del examen a los sustentantes y evitar el cansancio. En total se pilotearon 240 reactivos seleccionados.

Los reactivos por pilotear se incluyeron en seis versiones de 40 reactivos cada una. Todas las versiones contenían tanto reactivos de todas las áreas antes mencionadas.

4.3.6. Revisión cuantitativa. Calibración de reactivos

La calibración de los reactivos permitió valorar el comportamiento estadístico de los reactivos y obtener sus características psicométricas para seleccionar aquellos que conformarían la prueba operativa de 2011.

El análisis de los reactivos se elaboró con base en la teoría clásica (TC) y la teoría de la respuesta al ítem (TRI), de la cual se usó el modelo logístico de dos parámetros.

En el análisis de los reactivos, con base en la Teoría Clásica, se distinguen dos indicadores:

1. *Índice de dificultad.* Se entiende como la proporción de personas que responden correctamente un reactivo de una prueba. Entre mayor sea esta proporción, menor será su dificultad. Lo que quiere decir que se trata de una relación inversa: a mayor dificultad del reactivo, menor será la proporción de personas que respondan correctamente. Los valores que se consideraron aceptables en este parámetro se encuentran en una dificultad entre 10 y 90.
2. *Correlación punto biserial o coeficiente de discriminación.* Se calcula para determinar el grado en que las habilidades que mide la prueba, también las mide el *reactivo*. Su importancia radica en que permite conocer la tendencia de las respuestas a los reactivos con respecto a la escala de la cual forman parte. Se puede esperar que quien tuvo una puntuación alta en toda la prueba deberá tener altas probabilidades de contestar correctamente el reactivo.
3. También se debe esperar lo contrario, es decir, que quien tuvo bajas puntuaciones en la prueba deberá tener pocas probabilidades de contestar correctamente el reactivo. Así, un buen reactivo debe discriminar entre aquellos que obtuvieron buenas calificaciones en la prueba y aquellos que obtuvieron bajas calificaciones. La correlación punto biserial puede presentar valores de -1 a 1. El valor que se consideró aceptable para los reactivos de la prueba fue ≥ 0.001 .

En la teoría de respuesta al ítem los reactivos se analizan con base en modelos logísticos de uno, dos y tres parámetros. Para el análisis de los reactivos de la prueba se utilizó el modelo logístico de dos parámetros. Dichos parámetros hacen alusión a los índices de discriminación y dificultad.

4. *Índice de dificultad (b)*. Indica la posición del ítem en la escala de aptitud. Cuando más grande es el valor de dificultad, mayor es la aptitud requerida para que el *examinado* tenga una probabilidad de resolver correctamente el ítem. Un reactivo es aceptable si sus valores son ≥ -1.65 y ≤ 2.0 .

5. *Índice de discriminación (a)*. Indica la cualidad que tiene el reactivo de diferenciar a los sustentantes que dominan el conocimiento de aquellos que no. Para considerar aceptables los reactivos deben presentar una discriminación ≥ 0.45 .

La diferencia entre la teoría clásica y la teoría de respuesta al ítem reside en que esta última se centra en las propiedades de los reactivos individuales y no en las propiedades globales de la prueba.

5. Ensamble

5.1. Integración de la versión operativa

Dada la petición del Consejo Consultivo, la prueba debería evaluar los mismos conocimientos y habilidades que pide el PR.

Debido a lo anterior, la integración de la Prueba de Matemáticas para Bachilleratos Tecnológicos Federales se llevó a cabo bajo los siguientes criterios:

- Apegarse al perfil referencial establecido para la prueba, sus contenidos, especificaciones y estructura. Se efectuó una selección de reactivos, representando cada una de las especificaciones establecidas en el PR, así como la estructura de la prueba por área, subárea y global.
- Contar con calidad técnica y, en ese sentido con los mejores resultados estadísticos
- Representar al conjunto de reactivos de toda la prueba y, por tanto al constructo evaluado
- Incluir 70% de reactivos ya piloteados. Este tipo de reactivos debía cumplir con las siguientes características:
 - Haber sido aplicados previamente para analizar su calidad
 - Contar con una calidad técnica acorde con los parámetros establecidos
 - Corresponder al conjunto de reactivos no cubiertos por los reactivos ancla

Dada la aplicación de los criterios anteriores se llevó a cabo una revisión cualitativa de todos los reactivos seleccionados por sus estadísticos considerando:

- Su apego a la especificación establecida en el perfil referencial de la prueba
- Su apego a los atributos establecidos en cada delimitación
- La claridad en la redacción

5.2. Revisión editorial

Los cuadernillos de la prueba se sometieron a una revisión por parte de los expertos del CENEVAL con la finalidad de asegurar su calidad editorial. Las correcciones se efectuaron tanto en la versión impresa como en el sistema BRAE.

6. Aplicación

La aplicación de la Prueba de Matemáticas para Bachilleratos Tecnológicos Federales se realiza en el mes de mayo de cada año, en más de 80 planteles de la (SEMS) de 24 estados de la República. La impresión de los cuadernillos, la organización de la aplicación y su ejecución estuvo a cargo de la SEMS-SEP. La asignación de alumnos a los grupos control y experimentales fue responsabilidad de la SEMS y sus investigadores de la Universidad de Pensilvania.

Para responder, los sustentantes utilizaron una hoja de respuestas diseñada de manera coordinada por la SEMS y el CENEVAL. Previa impresión de las hojas de respuesta, el CENEVAL realizó pruebas que garantizaran la lectura precisa de la información.

7. Calificación

En esta fase se procesó la información recolectada durante la fase de aplicación, a través de las siguientes etapas:

- Lectura de las hojas de respuesta
- Calibración de reactivos
- Definición de puntos de corte y de niveles de dominio
- Equiparación
- Calificación de sustentantes

7.1. Establecimiento de puntos de corte y definición de niveles de dominio

Para clasificar a los sustentantes y emitir los reportes de resultados es indispensable establecer estándares que favorezcan la toma de decisiones. El establecimiento de estándares comprende dos procesos principales: determinación de niveles de dominio (ND) y establecimiento de puntos de corte.

En el caso de la prueba de Matemáticas para Bachilleratos Tecnológicos Federales Primer Grado, desde 2009 se conformó un Comité de Puntos de Corte integrado con especialistas en la materia, que habían participado previamente en otros grupos colegiados como en los comités de diseño y de validación del perfil referencial, y los comités de diseño y validación de reactivos.

El comité de puntos de corte sesionó el 13 y 14 de julio de 2009, considerando los siguientes niveles de dominio. La descripción de estos niveles partió de una propuesta general del CENEVAL y se redactó finalmente por el Comité de Puntos de Corte.

Nivel de dominio 1, Insuficiente

El contenido es muy sencillo. Los contenidos y habilidades de este nivel pueden ser demostrados por los alumnos, con fuertes carencias en las asignaturas de Matemáticas al concluir el grado escolar.

Nivel de dominio 2, Regular

El contenido es sencillo. Los contenidos y habilidades de este nivel pueden ser demostrados por los alumnos que presentan un repertorio básico de la asignatura al concluir el grado escolar.

Nivel de dominio 3, Bueno

El contenido es complejo. Los contenidos y habilidades de este nivel pueden ser demostrados por los alumnos que presentan un buen repertorio de la asignatura al concluir el grado escolar.

Nivel de dominio 4, Excelente

El contenido es muy complejo. Los contenidos y habilidades de este nivel pueden ser demostrados por los alumnos que presentan un repertorio completo de la asignatura al concluir el grado escolar.

Respecto a los métodos de puntos de corte, en 2009 el comité elaboró propuestas con los métodos Angoff, Hofstee, Bookmark y Norma, mismas que se entregaron a los investigadores de la Universidad de Pennsylvania para fines de calificación y diagnóstico de los sustentantes.

Es preciso mencionar que posterior a estas propuestas los investigadores de la Universidad de Pennsylvania, determinaron las puntuaciones a considerar y los diagnósticos finales. Para ello, consideraron las calificaciones obtenidas y la distribución de sustentantes por niveles de dominio en la prueba *Enlace Básica, Tercero de Secundaria*, así como los recursos y objetivos del Proyecto Alineación de Incentivos (ALI).

Asimismo, los investigadores de la Universidad de Pennsylvania establecieron las siguientes descripciones para los niveles de dominio:

1. Insuficiente. Necesitas adquirir los conocimientos y desarrollar las habilidades de las asignaturas evaluadas.
2. Elemental. Requieres fortalecer la mayoría de los conocimientos y desarrollar las habilidades de las asignaturas evaluadas.
3. Bueno. Posees un nivel de dominio adecuado de los conocimientos y las habilidades de las asignaturas evaluadas.
4. Excelente. Posees un alto nivel de dominio de los conocimientos y las habilidades de las asignaturas evaluadas

8. Reportes

Los reportes de resultados se diseñaron por el Consejo Consultivo, en función de los usos de información de la Prueba.

En el Anexo II se muestra un ejemplo de los reportes de resultados que se entregan a los sustentantes, según los propósitos del Proyecto.

- Alumnos con tratamiento experimental e incentivos individuales
- Alumnos con tratamiento experimental e incentivos individuales y por grupo.

Los reportes indican los datos de identificación del sustentante y de la escuela a la que pertenece. Asimismo, indican el nivel de dominio obtenido por el sustentante al aplicar la prueba: insuficiente, elemental, bueno o excelente y el nivel de dominio que obtuvo en la prueba *Enlace Básica, Tercero de Secundaria*. También informan al sustentante el monto del incentivo a que se ha hecho acreedor.

8.1 Guía para el sustentante

En esta prueba no procede la elaboración de una guía para el sustentante. La Prueba de Matemáticas para Bachilleratos Tecnológicos Federales, Primer Grado es un instrumento solicitado por la SEMS para fines de investigación.

Conclusiones

Es importante destacar que la participación en el diseño del examen me ha permitido desplegar y articular los conocimientos y práctica adquiridos durante mi formación académica como maestro en educación.

Finalmente, es importante destacar que la realización de este proyecto me ha permitido desplegar y articular los conocimientos y práctica adquiridos durante mi formación académica en la maestría a la vez que he profundizado en un área emergente de la pedagogía como es la evaluación de la intervención educativa, ya que considero que mi participación como maestro en educación dentro del equipo de trabajo para el diseño y la construcción de este examen, me permitió aplicar tanto el saber, como el saber hacer de la pedagogía para el mejor desarrollo del examen, así como aportar elementos para el análisis en torno al deber ser del fenómeno educativo, enfrentándome a nuevos retos y posibilidades de aprendizaje que amplían la visión de mi ejercicio pedagógico en el ámbito de la educación media y superior.

CENEVAL:

- a) Para el CENEVAL, el hecho de que el yo como maestro en educación participe en el diseño y elaboración de este examen de opción múltiple, beneficia a la institución porque en nuestro perfil profesional tenemos presente permanentemente lo humano durante los procesos, tanto los exámenes van dirigidos a las personas como los reactivos en sí son como un ser vivo, además de que contamos con la preparación y el contexto educativo para desarrollar dicha actividad.

- b) La colaboración entre el CENEVAL con diferentes instituciones, ha permitido el desarrollo de instrumentos que proporcionan elementos para la valoración del ejercicio profesional y que constituyen el primer ejercicio en México de profesionalización y estandarización de instrumentos para la certificación nacional del ejercicio profesional con la participación de especialistas en evaluación educativa y del ejercicio de la profesión.
- c) Quiero resaltar la importancia de las acciones de evaluación a cargo del CENEVAL.
- d) Con respecto a la responsabilidad política y social del centro, de la confianza que otras instituciones le han otorgado por realizar su trabajo.
- e) Creo que el CENEVAL tiene el enorme compromiso de redoblar esfuerzos para mejorar la calidad de los productos y servicios que brinda a sus usuarios y detectar fortalezas y áreas de oportunidad que permitan una mejora constante.
- f) El CENEVAL reconoce la valiosa contribución que brindan los miembros de los consejos técnicos y comités académicos, con la finalidad de lograr la consolidación de los primeros, se emitió un reglamento en el que se especifican sus funciones, integración y formas de organización. Esta normativa ha dado lugar a la renovación de estos cuerpos colegiados.

Bibliografía

- ARREDONDO GALVÁN, Pico Contreras, Palencia Gómez. (1997). Nuevo Manual de Didáctica de las Ciencias Histórico-Sociales. Editorial Limusa, 1997, pp. 140.
- BLOOM, B (1990). Taxonomía de los objetivos de la educación: clasificación de las metas educacionales. Buenos Aires. El Ateneo
- CENEVAL (1994-2001). La primera Etapa: México.
- CENEVAL (1999). Estatuto del CENEVAL. México.
- CENEVAL (2000). Manual de estándares para la construcción de instrumentos. México.
- CENEVAL (2011). Prontuario de normas técnicas y estándares del Ceneval. México
- GAGO, A. (2000). Apuntes acerca de la evaluación educativa. México. SEP pág. 41
- GARCÍA GARRO, Alberto (2007). Instrumentos de evaluación, Revista Mexicana de Anestesiología, Vol. 30 No. 3 Julio-Septiembre, pp. 158-164.
- MEDINA, MARTÍNEZ, Sara Rosa (1996), Educación y Modernidad, El bachillerato en México ante los Desafíos del Tercer Milenio Universidad Autónoma de San Luis Potosí
- MEDINA, MARTÍNEZ, Sara Rosa y VILLALOBOS, PÉREZ-CORTÉS, Marveya (2006). Evaluación Institucional, pp. 29, 30. Publicaciones Cruz
- VIDAL, R. y J. Tristán (2006). Estándares de calidad para pruebas objetivas. Colombia. Cooperativa Editorial Magisterio.
- VIDAL, R. (2000). Manual Técnico del Centro Nacional de Evaluación para la Educación Superior. México. CENEVAL.