

INSTITUTO NACIONAL PARA LA EVALUACION DE LA EDUCACION

CRITERIOS técnicos y de procedimiento para el análisis de los instrumentos de evaluación, el proceso de calificación y la emisión de resultados de la evaluación con fines de diagnóstico del personal docente y técnico docente que ingresó en el ciclo escolar 2015-2016 al término de su primer año escolar en Educación Básica (EB) y Educación Media Superior (EMS).

Al margen un logotipo, que dice: Instituto Nacional para la Evaluación de la Educación.- México.

CRITERIOS TÉCNICOS Y DE PROCEDIMIENTO PARA EL ANÁLISIS DE LOS INSTRUMENTOS DE EVALUACIÓN, EL PROCESO DE CALIFICACIÓN Y LA EMISIÓN DE RESULTADOS DE LA EVALUACIÓN CON FINES DE DIAGNÓSTICO DEL PERSONAL DOCENTE Y TÉCNICO DOCENTE QUE INGRESÓ EN EL CICLO ESCOLAR 2015-2016 AL TÉRMINO DE SU PRIMER AÑO ESCOLAR EN EDUCACIÓN BÁSICA (EB) Y EDUCACIÓN MEDIA SUPERIOR (EMS).

Con fundamento en lo dispuesto en los artículos 3o., fracción IX de la Constitución Política de los Estados Unidos Mexicanos; 22, 26, 28, fracción X; 38, fracciones VI, IX y XXII de la Ley del Instituto Nacional para la Evaluación de la Educación; y en los Lineamientos para llevar a cabo la evaluación con fines de diagnóstico del personal docente y técnico docente que ingresó en el ciclo escolar 2015-2016 al término de su primer año escolar en Educación Básica y Media Superior. LINEE-05-2016, la Junta de Gobierno aprueba emitir los siguientes criterios técnicos y de procedimiento para el análisis de los instrumentos de evaluación, el proceso de calificación y la emisión de resultados de la evaluación con fines de diagnóstico del personal docente y técnico docente que ingresó en el ciclo escolar 2015-2016 al término de su primer año escolar en educación básica (EB) y educación media superior (EMS).

Los presentes criterios técnicos y de procedimiento tienen como finalidad establecer los referentes necesarios para garantizar la validez, confiabilidad y equidad de los resultados de los procesos de evaluación implicados. Su contenido se organiza en cuatro apartados: 1) Características generales de los instrumentos; 2) Criterios técnicos para el análisis e integración de los instrumentos de evaluación; 3) Procedimiento para el establecimiento del punto de corte y estándar de desempeño en las áreas de los exámenes de opción múltiple; y 4) Resultado de la evaluación diagnóstica. En la parte final, se integra un Anexo técnico con información detallada de los aspectos técnicos que se consideran en los distintos apartados del documento.

Definición de términos

Para los efectos del presente documento se emplean las siguientes definiciones:

- I. **Alto impacto:** Se indica cuando los resultados del instrumento tienen consecuencias importantes para las personas o las instituciones; por ejemplo, en los procesos de admisión o certificación.
- II. **Calificación:** Proceso de asignación de una puntuación o nivel de desempeño logrado a partir de los resultados de una medición.
- III. **Confiabilidad:** Cualidad de las mediciones obtenidas con un instrumento, que se caracterizan por ser consistentes y estables cuando éste se aplica en distintas ocasiones.
- IV. **Constructo:** Elaboración teórica formulada para explicar un proceso social, psicológico o educativo y cuya adecuada descripción permite que sea susceptible de ser observable o medible.
- V. **Correlación punto biserial:** Medida de consistencia que se utiliza en el análisis de reactivos, indica si hay una correlación entre el resultado de un reactivo con el resultado global del examen.
- VI. **Criterio de evaluación:** Indicador de un valor aceptable sobre el cual se puede establecer o fundamentar un juicio de valor sobre el desempeño de una persona.
- VII. **Desempeño:** Resultado obtenido por el sustentante en un proceso de evaluación o en un instrumento de evaluación educativa.
- VIII. **Dificultad de un reactivo:** Indica la proporción de personas que responden correctamente el reactivo de un examen.
- IX. **Distraectores:** Opciones de respuesta incorrectas del reactivo de opción múltiple, que probablemente serán elegidas por los sujetos con menor dominio en lo que se evalúa.
- X. **Dominio:** Conjunto de conocimientos, habilidades, destrezas, actitudes u otros atributos que tienen las siguientes propiedades: límites, extensión y definición. También se puede aplicar a contenidos, procedimientos u objetos.

- XI. Educación básica:** Tipo de educación que comprende los niveles de preescolar, primaria y secundaria en todas sus modalidades, incluyendo la educación indígena, la especial y la que se imparte en los centros de educación básica para adultos.
- XII. Educación media superior:** Tipo de educación que comprende el nivel de bachillerato, los demás niveles equivalentes a éste, así como la educación profesional que no requiere bachillerato o sus equivalentes.
- XIII. Error estándar de medida:** Desviación estándar de una distribución hipotética de errores de medida de una población.
- XIV. Escala:** Procedimiento para asignar números, puntuaciones o medidas a objetos o sucesos con propiedades específicas a partir de reglas definidas.
- XV. Escalamiento:** Proceso a través del cual se construye una escala que facilita la interpretación de la calificación que obtienen los sustentantes en uno o varios instrumentos de evaluación, colocando las puntuaciones de los distintos instrumentos o formas a una escala común.
- XVI. Especificaciones de tareas evaluativas o reactivos:** Descripción detallada de las características relevantes que se espera hagan los sujetos al sustentar el instrumento de evaluación y que es posible observar a través de las tareas evaluativas o los reactivos. Tienen el papel de guiar a los comités académicos en la elaboración y validación de las tareas evaluativas o de los reactivos y que éstos cuenten con los elementos necesarios para construirlos alineados al objeto de medida o constructo que se desea evaluar a través del instrumento.
- XVII. Estándar:** Principio de valor o calidad en la conducción y uso de los procedimientos de evaluación y que son acordados por expertos en evaluación.
- XVIII. Evaluación:** Acción de emitir juicios de valor sobre un objeto, sujeto o evento que resultan de comparar los resultados de una medición u observación con un referente previamente establecido.
- XIX. Examen:** Instrumento de evaluación que se emplea para identificar el nivel de dominio de los sustentantes sobre un constructo específico.
- XX. Instrumento de evaluación:** Procedimiento de recolección de datos que suelen tener distintos formatos, atendiendo a la naturaleza de la evaluación, por ejemplo, instrumentos de selección de respuesta, instrumentos de respuesta construida, cuestionarios, observaciones, portafolios, entre otros.
- XXI. Instrumento de evaluación referido a un criterio:** Instrumento que permite comparar el desempeño de las personas evaluadas, con un estándar preestablecido.
- XXII. Medición:** Proceso de asignación de valores numéricos a atributos de las personas, características de objetos o eventos de acuerdo con reglas específicas que permitan que sus propiedades puedan ser representadas cuantitativamente.
- XXIII. Muestra:** Subconjunto de la población de interés que refleja las variables medidas en una distribución semejante a las de la población.
- XXIV. Multirreactivo:** Conjunto de reactivos de opción múltiple que están vinculados a un planteamiento general, por lo que este último es indispensable para poder resolverlos.
- XXV. Objeto de medida:** Conjunto de características o atributos que se miden en el instrumento de evaluación.
- XXVI. Parámetro estadístico:** Número que resume un conjunto de datos que se derivan del análisis de una cualidad o característica del objeto de estudio.
- XXVII. Perfil:** Conjunto de características, requisitos, cualidades o aptitudes que deberá tener el sustentante a desempeñar un puesto o una función descritos específicamente.
- XXVIII. Puntuación:** Número de aciertos obtenidos en un instrumento de evaluación.
- XXIX. Reactivo:** Unidad básica de medida de un instrumento de evaluación que consiste en una pregunta o instrucción que requiere una respuesta del sujeto.
- XXX. Sesgo:** Error en la medición de un atributo (por ejemplo, conocimiento o habilidad), debido a una variable no controlada, como las diferencias culturales o lingüísticas de las personas evaluadas.
- XXXI. Validez:** Juicio valorativo integrador sobre el grado en que los fundamentos teóricos y las evidencias empíricas apoyan la interpretación de las puntuaciones de los instrumentos de evaluación.

1. Características generales de los instrumentos

La evaluación con fines de diagnóstico del personal docente y técnico docente que ingresó en el ciclo escolar 2015-2016 al término de su primer año escolar en EB y EMS, es un proceso integrado que incluye varios instrumentos que dan cuenta de los diferentes aspectos que se describen en los Perfiles, parámetros e indicadores establecidos por la autoridad educativa. A continuación, se describen sucintamente cada uno de ellos.

Educación básica

- Reporte de cumplimiento de responsabilidades profesionales.
- Examen de conocimientos y habilidades docentes que favorecen el aprendizaje de los alumnos.
- Cuestionario de características y necesidades formativas de los docentes y técnicos docentes de nuevo ingreso.

Reporte de cumplimiento de responsabilidades profesionales.

Este instrumento tiene como propósito obtener información acerca del cumplimiento de las responsabilidades profesionales de los docentes y técnicos docentes de nuevo ingreso y de su participación en la organización y el funcionamiento de la escuela. Es respondido por el director escolar o el supervisor de zona, según sea el caso.

Examen de conocimientos y habilidades docentes que favorecen el aprendizaje de los alumnos

Este instrumento tiene como propósito aportar información acerca de los conocimientos y las habilidades que los docentes y técnicos docentes de nuevo ingreso tienen para favorecer el aprendizaje de los alumnos. Se caracteriza por presentar un conjunto de situaciones cuya solución demanda la puesta en juego de conocimientos y habilidades relacionadas con la práctica docente.

Cuestionario de características y necesidades formativas de los docentes y técnicos docentes de nuevo ingreso

Este instrumento tiene como propósito recabar información relacionada con las características profesionales y necesidades de formación de los docentes. Con este instrumento se recupera la experiencia adquirida por el docente o técnico docente en el trabajo directo con los alumnos y en la relación con sus colegas, directivos, padres de familia y comunidad; se trata de conocer el contexto en que desarrolla su práctica; la forma como se ha desempeñado para lograr que sus alumnos aprendan y sus necesidades de formación profesional. Se integra en tres apartados: características profesionales de los docentes y técnicos docentes de nuevo ingreso, trabajo en el aula y en la escuela, y necesidades de formación y desarrollo profesional.

Educación media superior

- Reporte de cumplimiento de responsabilidades profesionales.
- Examen de casos de competencias didácticas.
- Cuestionario de características y necesidades formativas de los docentes y técnicos docentes de nuevo ingreso.

Reporte de cumplimiento de responsabilidades profesionales

Este instrumento tiene como propósito recopilar información para identificar el grado de cumplimiento de las responsabilidades profesionales del personal docente y técnico docente, particularmente en lo que se refiere a su participación en el funcionamiento de la escuela y en órganos colegiados, intercambio de experiencias con colegas y su vinculación con la comunidad. El informe es respondido por el director del docente o técnico docente.

Examen de casos de competencias didácticas

Este instrumento tiene como propósito evaluar las competencias docentes que favorecen el aprendizaje de los estudiantes, la colaboración en la escuela y el vínculo con la comunidad escolar, a partir del análisis de situaciones vinculadas con la práctica docente, contextualizadas en distintos escenarios. Está diseñado a partir de reactivos relativos a situaciones cotidianas contextualizadas en distintos escenarios de la práctica docente en la educación media superior; son reactivos asociados que permiten la valoración de la capacidad del docente para plantear alternativas para su resolución a situaciones o problemas de la práctica cotidiana.

Cuestionario de características y necesidades formativas de los docentes y técnicos docentes de nuevo ingreso

Este instrumento tiene como propósito identificar las necesidades de formación de los docentes y técnicos docentes a fin de diseñar y ofertar programas y apoyos para su desarrollo profesional. Consta de tres secciones: a) características profesionales del docente y técnico docente de nuevo ingreso, b) trabajo en el aula y en la escuela, y c) necesidades de formación y desarrollo profesional.

2. Criterios técnicos para el análisis e integración de los instrumentos de evaluación

Un aspecto fundamental que debe llevarse a cabo antes de emitir cualquier resultado de un proceso de evaluación es el análisis psicométrico de los instrumentos que integran la evaluación, con el objetivo de verificar que cuentan con la calidad técnica necesaria para proporcionar resultados confiables, acordes con el objetivo que se persigue.

Las técnicas empleadas para el análisis de un instrumento dependen de su naturaleza, de los objetivos específicos para el cual fue diseñado, así como del tamaño de la población evaluada. Sin embargo, en todos los casos, debe aportarse información sobre la dificultad y discriminación de sus reactivos o tareas evaluativas, así como la precisión de las áreas que conforman el instrumento, los indicadores de consistencia interna o estabilidad del instrumento, los cuales, además de los elementos asociados a la conceptualización del objeto de medida, forman parte de las evidencias que servirán para valorar la validez de la interpretación de sus resultados. Estos elementos, deberán reportarse en el informe o manual técnico del instrumento.

Con base en los resultados de estos procesos de análisis deben identificarse las tareas evaluativas o los reactivos que cumplen con los criterios psicométricos especificados en este documento para integrar el instrumento, para calificar el desempeño de las personas evaluadas, con la mayor precisión posible.

Para llevar a cabo el análisis de los instrumentos de medición utilizados en el proceso de evaluación, es necesario que los distintos grupos de sustentantes de las entidades federativas queden equitativamente representados, dado que la cantidad de sustentantes por tipo de evaluación en cada entidad federativa es notoriamente diferente. Para ello, se definirá una muestra de sustentantes por cada instrumento de evaluación que servirá para analizar el comportamiento estadístico de los instrumentos y orientar los procedimientos descritos más adelante que son previos para la calificación. Para conformar dicha muestra, cada entidad federativa contribuirá con 500 sustentantes como máximo, elegidos aleatoriamente. Si hay menos de 500 sustentantes, todos se incluirán en la muestra (OECD; 2002, 2005, 2009, 2014). Si no se realizara este procedimiento, las decisiones sobre los instrumentos de evaluación se verían fuertemente influenciadas, indebidamente, por el desempeño mostrado por aquellas entidades que se caracterizan por tener un mayor número de sustentantes.

Sobre la conformación de los instrumentos de evaluación

Con la finalidad de obtener puntuaciones de los aspirantes con el nivel de precisión requerido para los propósitos de los concursos, los instrumentos de evaluación deberán tener las siguientes características:

Exámenes de opción múltiple:

- Los instrumentos de evaluación deberán tener, al menos, 60 reactivos efectivos para calificación y contar con al menos 20 reactivos al nivel de desagregación de las áreas de contenidos.
- Las especificaciones de reactivos deberán integrarse por una definición operacional del contenido específico a evaluar, un reactivo ejemplo y la bibliografía en la que se sustenta el reactivo.
- En todos los casos, deberá documentarse el procedimiento que se siguió para determinar la estructura del instrumento y la cantidad de reactivos que lo conforman, a fin de justificar la relevancia (ponderación) de los contenidos específicos evaluados en el mismo.
- Para el diseño de los casos asociados a un sólo reactivo debe cuidarse la extensión, a fin de que se incorpore únicamente la información indispensable para resolverlo.
- Para el diseño de los casos con formato de multirreactivo, deberá verificarse que: a) todos los reactivos necesiten del planteamiento general para ser contestados; b) los reactivos evalúen conocimientos o habilidades complejas, no de reconocimiento; c) los reactivos sean independientes entre sí, esto es, que para poder responderse no requieran de la información incorporada en alguno de ellos, o bien, de la respuesta dada a algún otro.

Criterios y parámetros estadísticos

Los instrumentos empleados para estos procesos de evaluación deberán atender los siguientes criterios (Cook y Beckman 2006; Downing, 2004; Stemler y Tsai, 2008) con, al menos, los valores de los parámetros estadísticos indicados a continuación:

En el caso de los instrumentos de evaluación basados en reactivos de opción múltiple:

- La respuesta correcta deberá tener una dificultad clásica de 10% a 90% y una correlación punto biserial corregida igual o mayor que 0.15.
- Los distractores deberán tener correlaciones punto biserial negativas.
- La confiabilidad del instrumento deberá ser igual o mayor que 0.80; la confiabilidad para cada una de las áreas que constituyen los instrumentos y a partir de los cuales se establecerán los puntos de corte deberá ser igual o mayor a 0.75.

En caso de que algún instrumento no satisfaga los criterios y parámetros estadísticos antes indicados, la Junta de Gobierno del Instituto determinará lo conducente, buscando salvaguardar la estructura del instrumento que fue aprobada por el Consejo Técnico y atendiendo al marco jurídico aplicable.

En el caso del Reporte de cumplimiento de responsabilidades profesionales y del Cuestionario de características y necesidades formativas de los docentes y técnicos docentes de nuevo ingreso, para cada una de las escalas que lo constituyen:

- La correlación entre cada reactivo con la puntuación global de la escala deberá ser igual o mayor que 0.30.
- La confiabilidad del constructo medido a través de la escala deberá ser igual o mayor que 0.80.

Adicionalmente, para estos instrumentos, conformados fundamentalmente por escalas tipo Likert, se debe generar evidencia de que los constructos se integran conforme a lo esperado, esto es: a) los reactivos se integran a las dimensiones previstas en el diseño del instrumento; b) hay una correlación positiva y significativa entre las distintas escalas que integran el instrumento; c) existe la posibilidad de implementar un modelo de medición a los datos; d) es posible valorar la dimensión del constructo latente y, si es factible, e) se verifique que no hay un comportamiento diferencial de los reactivos o las escalas entre subpoblaciones o grupos (Muraki, 1999; Wu y Adams, 2007; Bentler, 2006; Masters, 1982).

Si se diera el caso de que en algún instrumento no se cumpliera con los criterios y parámetros estadísticos antes indicados, la Junta de Gobierno del Instituto determinará lo que procede, buscando salvaguardar el constructo del instrumento que fue aprobado por el Consejo Técnico y atendiendo al marco jurídico aplicable.

3. Procedimiento para el establecimiento del punto de corte y estándar de desempeño en las áreas de los exámenes de opción múltiple.

Un paso crucial en el desarrollo y uso de los instrumentos de evaluación de naturaleza criterial, como es el caso de los que se utilizarán para este proceso de evaluación, es el establecimiento del punto de corte que divide el rango de calificaciones para diferenciar entre niveles de desempeño. Para estos procesos de evaluación diagnóstica, se establecerán los puntos de corte a partir de las áreas de conocimiento que constituyen los instrumentos de evaluación de opción múltiple; lo anterior con la finalidad de proporcionar información más detallada sobre los resultados de la evaluación, toda vez que el objetivo central de la misma está en proporcionar retroalimentación al docente para que establezca una ruta de fortalecimiento profesional.

En los instrumentos de evaluación de tipo criterial, la calificación obtenida por cada sustentante se contrasta con un estándar de desempeño establecido por un grupo de expertos que describe el nivel de competencia requerido para algún propósito determinado, es decir, los conocimientos y las habilidades que, para cada instrumento de evaluación, se consideran indispensables para un desempeño adecuado en la función profesional docente. En este sentido el estándar de desempeño delimita el marco interpretativo de las puntuaciones obtenidas en un instrumento por los sustentantes. El procedimiento para el establecimiento del punto de corte y el estándar de desempeño incluye tres fases:

Primera fase

Para contar con un marco de referencia común a los distintos instrumentos de evaluación, se establecen los descriptores genéricos de los niveles de desempeño, cuya función es orientar a los comités académicos en el trabajo del desarrollo de los descriptores específicos para cada área de conocimientos y del instrumento en su conjunto. Para todos los exámenes de opción múltiple se utilizarán únicamente dos niveles de desempeño para las áreas: Nivel I y Nivel II. Los descriptores genéricos para las diferentes áreas de los distintos exámenes y cada nivel se indican en las Tablas 1a y 1b.

Tabla 1a. Descriptores genéricos de los niveles de desempeño para las áreas de los exámenes de conocimientos y habilidades docentes que favorecen el aprendizaje de los alumnos, en EB

| Nivel de desempeño | Descriptor por área de conocimiento |
|--------------------|--|
| Nivel I | El docente muestra conocimientos básicos sobre el desarrollo y aprendizaje de sus alumnos, el currículo, los principios filosóficos, fundamentos legales y finalidades de la educación pública mexicana; asimismo cuenta con conocimientos elementales para organizar su intervención docente con base en los propósitos, contenidos y enfoques didácticos de la educación básica. Por otra parte, reconoce algunos conceptos clave del currículo, así como las características generales de la organización de actividades didácticas y el funcionamiento general de una escuela; muestra habilidades elementales para resolver situaciones de su práctica profesional y para establecer vínculos con la comunidad en la que se encuentra la escuela. |
| Nivel II | El docente muestra conocimientos consistentes sobre el desarrollo y aprendizaje de sus alumnos, el currículo, los principios filosóficos, fundamentos legales y finalidades de la educación pública mexicana; organiza y sistematiza su intervención docente con base en los propósitos, contenidos y enfoques didácticos de la educación básica y el reconocimiento de la diversidad cultural de sus alumnos. Asimismo, resuelve situaciones de su práctica profesional, construye ambientes favorables para la sana convivencia y el aprendizaje de sus alumnos, participa en el funcionamiento eficaz del centro educativo y establece vínculos con la comunidad en la que se encuentra la escuela. |

Tabla 1b. Descriptores genéricos de los niveles de desempeño para las áreas de los exámenes de casos de competencias didácticas, en EMS.

| Nivel de desempeño | Descriptor por área de conocimiento |
|--------------------|---|
| Nivel I | El docente muestra conocimientos elementales acerca de las características de aprendizaje de sus estudiantes, del modelo basado en competencias y de los propósitos de la asignatura correspondiente; también cuenta con conocimientos básicos para diseñar planes de trabajo que incorporan estrategias y técnicas orientadas al desarrollo de competencias vinculadas con el contexto social de los estudiantes, así como estrategias de evaluación pertinentes para la obtención de información sobre el desempeño de los estudiantes, y estrategias de inclusión, de autoconocimiento, de corresponsabilidad y de resolución de conflictos. |
| Nivel II | El docente muestra conocimientos adecuados acerca de las características de aprendizaje de sus estudiantes, del modelo basado en competencias y de los propósitos de la asignatura correspondiente; diseña planes de trabajo que incorporan estrategias y técnicas orientadas al desarrollo de competencias, vinculadas con el contexto social de los estudiantes, define estrategias de evaluación pertinentes para la obtención de información sobre el desempeño de los estudiantes, e implementa estrategias de inclusión, de autoconocimiento, de corresponsabilidad y de resolución de conflictos. |

Segunda fase

En esta fase se establece el punto de corte a partir del cual se distinguen los niveles de desempeño, Nivel I y Nivel II. En este proceso participan los comités académicos correspondientes a cada instrumento de evaluación. Dichos comités se deberán conformar con especialistas que han participado en el diseño de los instrumentos, garantizando que esté representada la diversidad cultural en que se desenvuelve la acción educativa del país. En todos los casos, sus miembros deberán ser capacitados específicamente para ejercer su mejor juicio profesional a fin de identificar cuál es la puntuación requerida para que el sustentante alcance un determinado nivel o estándar de desempeño.

Los insumos que tendrán como referentes para el desarrollo de esta actividad serán: la documentación que describe la estructura de los instrumentos, las especificaciones y los ejemplos de tareas evaluativas o reactivos incluidos en las mismas. En todos los casos, el punto de corte se referirá a la ejecución típica o esperable de un sustentante hipotético, con un desempeño mínimamente aceptable para el nivel de desempeño II. Para ello, se deberá determinar, para cada tarea evaluativa o reactivo considerado en cada una de las áreas del instrumento, cuál es la probabilidad de que dichos sustentantes hipotéticos lo respondan correctamente y, con base en la suma de estas probabilidades, establecer la calificación mínima requerida o punto de corte (Angoff, 1971).

Una vez establecido el punto de corte, que divide el rango de calificaciones para diferenciar el Nivel I y Nivel II, considerando el conjunto de reactivos que, en cada caso, el sustentante hipotético es capaz de responder, se deberán describir los conocimientos y las habilidades específicos que están implicados en cada nivel de desempeño, es decir, lo que dicho sustentante conoce y es capaz de hacer.

Tercera fase

En la tercera fase se llevará a cabo un ejercicio de retroalimentación a los miembros de los comités académicos con el propósito de contrastar sus expectativas sobre el desempeño de la población evaluada, considerando la distribución de sustentantes que se obtiene en cada nivel de desempeño al utilizar el punto de corte definido en la segunda fase, a fin de determinar si es necesario realizar algún ajuste en la decisión tomada con anterioridad (Beuk, 1984).

Esta tercera fase se efectuará solamente para aquellos instrumentos de evaluación en los que el tamaño de la población evaluada sea igual o mayor a 100 sustentantes. Si la población es menor a 100 sustentantes, el punto de corte será establecido de acuerdo con lo descrito en la segunda fase.

4. Resultado de la evaluación diagnóstica

A continuación, se presentan dos subapartados: en el primero se describen los procedimientos para calificar los resultados de los sustentantes en cada instrumento; en el segundo se detallan los procedimientos para la obtención del resultado global.

Calificación de los resultados obtenidos por los sustentantes en los distintos instrumentos que constituyen el proceso de evaluación

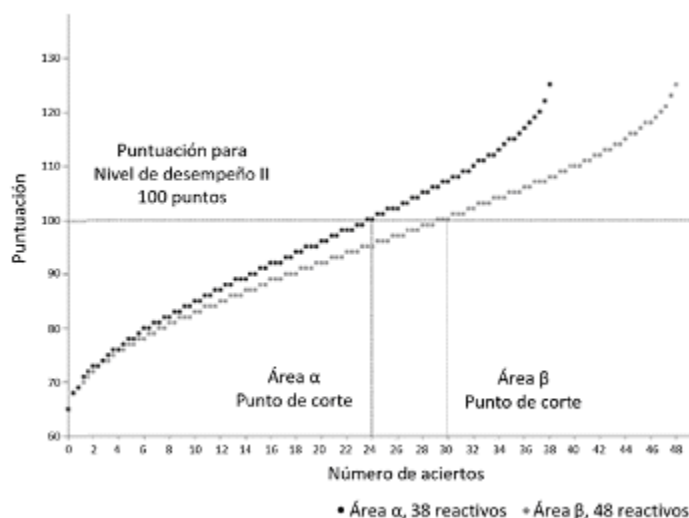
En cada plan de evaluación es indispensable definir la escala en la que se reportarán los resultados de los sustentantes. Existen varios tipos de escalas de calificación; en las escalas referidas a norma, las calificaciones indican la posición relativa del sustentante en una determinada población. En las escalas referidas a criterio, cada calificación en la escala representa un nivel particular de desempeño referido a un estándar previamente definido en un campo de conocimiento o habilidad específicos.

El escalamiento que se llevará a cabo en este proceso de evaluación permitirá construir una métrica común para las áreas de todos los instrumentos que se administrarán. Consta de dos transformaciones: la primera, denominada doble arcoseno, permite estabilizar la magnitud de la precisión de las puntuaciones a lo largo de la escala; la segunda transformación es lineal y ubica el punto de corte del nivel de desempeño II en un mismo valor para las áreas de todos los exámenes: puntuación de 100 en esta escala (cuyo rango va de 60 a 170 puntos¹).

Al utilizar esta escala, diferente a las escalas que se utilizan para reportar resultados de aprendizaje en el aula (de 5 a 10 o de 0% a 100%, donde el 6 o 60% de aciertos es aprobatorio), se evita que se realicen interpretaciones equivocadas de los resultados obtenidos, en virtud de que cada calificación representa un nivel particular de desempeño respecto a un estándar previamente definido, el cual puede implicar un número de aciertos diferente en cada caso.

En la siguiente gráfica se puede observar el número de aciertos obtenidos en dos áreas de longitudes diferentes y con puntos de corte distintos que, a partir del escalamiento, es posible graficar en una misma escala, trasladando el punto de corte a 100 puntos, aun cuando en cada área de conocimientos el punto de corte refiera a un número de aciertos diferente. En este ejemplo la distribución de las puntuaciones va de 65 a 125 puntos.

¹ Pueden encontrarse ligeras variaciones en este rango debido a que la escala es aplicable a múltiples instrumentos con características muy diversas, tales como la longitud, el tipo de instrumento y su nivel de precisión, diferencias entre los puntos de corte que atienden a las particularidades de los contenidos que se evalúan, entre otras. Por otra parte, para realizar el escalamiento, el sustentante debe, al menos, haber alcanzado un acierto en el examen; en caso contrario, se reportará como cero y obtendrá NI. Para mayores detalles sobre los procesos que se llevan a cabo para el escalamiento de las puntuaciones, consultar el anexo técnico.



Calificación del resultado global y escala en que se reportará

El resultado global de la evaluación se definirá considerando únicamente los instrumentos de evaluación constituidos por reactivos de opción múltiple, a saber:

- Educación básica: Examen de conocimientos y habilidades docentes que favorecen el aprendizaje de los alumnos.
- Educación media superior: Examen de casos de competencias didácticas.

Como ya se mencionó, el **Reporte de cumplimiento de responsabilidades profesionales** tiene por función aportar información sobre el cumplimiento del docente en sus funciones y es proporcionado por el directivo escolar o la autoridad inmediata que corresponda.

Asimismo, el **Cuestionario de características y necesidades formativas de los docentes y técnicos docentes de nuevo ingreso** aportará información relevante sobre la opinión de los docentes respecto a lo que, desde su perspectiva, consideran que es fundamental para trazar una ruta de formación profesional.

Estos dos instrumentos no serán considerados para la calificación global, pero se dará retroalimentación al docente sin afectar su calificación. De no contar con esta información, quedará asentado en el informe individual de resultados.

El resultado global del instrumento de evaluación estará dado por la combinación posible de los resultados de las áreas que lo componen:

Resultado global en el instrumento de evaluación

Esperado

En esta categoría se encuentran los docentes que alcanzaron el nivel de desempeño II en las dos áreas que constituyen el instrumento de evaluación. Los docentes que obtuvieron este resultado global muestran un nivel de dominio sólido y organizado de los conocimientos y las habilidades contemplados en el instrumento que se consideran indispensables para un adecuado desempeño docente, mismos que pueden consolidar a través de procesos de formación profesional.

Básico

En esta categoría se encuentran los docentes que alcanzaron el nivel de desempeño II en una de las áreas que constituyen el instrumento de evaluación y en la otra el nivel de desempeño I. Los docentes que obtuvieron este resultado global muestran un nivel de dominio adecuado en alguno de los aspectos que se evalúan; sin embargo, presenta deficiencias que resultan conveniente subsanar a través de procesos de formación profesional.

A desarrollar

En esta categoría se encuentran los docentes que alcanzaron el nivel de desempeño I en las dos áreas que constituyen el instrumento de evaluación. Los docentes que obtuvieron este resultado global muestran un nivel de dominio elemental en los distintos aspectos que se evalúan a través del instrumento, por lo que es indispensable fortalecer sus conocimientos y habilidades a través de procesos de formación profesional.

Los docentes recibirán retroalimentación sobre los instrumentos de evaluación, para que conozcan sus fortalezas y áreas de oportunidad que contribuyan a definir su trayectoria de actualización profesional.

El resultado “No se presentó a la evaluación”

Para el caso en que el docente no sustente el examen de conocimientos o el de casos de competencias didácticas, su resultado global será “No se presentó a la evaluación” y sólo se le asignará “NP: no presentó”.

Anexo técnico

El propósito de este anexo es detallar los aspectos técnicos específicos de los distintos procedimientos que se han enunciado en el cuerpo del documento, así como brindar mayores elementos para su entendimiento y fundamento metodológico. Es importante destacar que el establecimiento de puntos de corte se considera a partir de las áreas de conocimiento que constituyen los instrumentos de opción múltiple, razón por la cual los procesos metodológicos aquí descritos deben considerar esta acotación.

Métodos para establecer puntos de corte y niveles de desempeño

Método de Angoff

El método de Angoff está basado en los juicios de los expertos sobre los reactivos y contenidos que se evalúan a través de las áreas de los exámenes. De manera general, el método considera que el punto de corte se define a partir de la ejecución promedio de un sustentante hipotético que cuenta con los conocimientos, las habilidades o destrezas que se consideran indispensables para la realización de una tarea en particular; los jueces estiman, para cada pregunta, cuál es la probabilidad de que dicho sustentante acierte o responda correctamente.

Procedimiento

Primero se juzgan algunas preguntas, con tiempo suficiente para explicar las razones de las respuestas al grupo de expertos y que les permite homologar criterios y familiarizarse con la metodología. Después, se le solicita a cada juez que estime la probabilidad mínima de que un sustentante conteste correctamente un reactivo, el siguiente y así hasta concluir la totalidad de los reactivos. Posteriormente se calcula el puntaje esperado (*raw score*: la suma de estas probabilidades multiplicadas por uno -toda vez que cada reactivo vale un punto). Las decisiones de los jueces se promedian obteniendo el punto de corte. La decisión del conjunto de jueces pasa por una primera ronda para valorar sus puntos de vista en plenaria y puede modificarse la decisión hasta llegar a un acuerdo en común.

Método de Beuk

En 1981, Cess H. Beuk propuso un método para establecer estándares de desempeño, el cual busca equilibrar los juicios de expertos basados solamente en las características de los instrumentos de evaluación, lo que mide y su nivel de complejidad, con los juicios que surgen del análisis de resultados de los sustentantes una vez que un instrumento de evaluación es administrado.

Procedimiento

En el cuerpo del documento se señalaron tres fases para el establecimiento del punto de corte. Para completar la tercera fase, es necesario recolectar con antelación las respuestas a dos preguntas dirigidas a los integrantes de los distintos comités académicos especializados involucrados en el diseño de las evaluaciones y en otras fases del desarrollo del instrumento. Estas preguntas son:

a) ¿Cuál es el mínimo nivel de conocimientos o habilidades que un sustentante debe tener para aprobar el área del instrumento de evaluación? (Expresado como porcentaje de aciertos del área, k).

b) ¿Cuál es la tasa de aprobación de sustentantes que los jueces estiman que aprueben el área del instrumento? (Expresado como porcentaje, v).

Para que los resultados de la metodología a implementar sean estables e integren diferentes enfoques que contribuyan a la diversidad cultural, se deberán recolectar las respuestas de, al menos, 30 especialistas integrantes de los diferentes comités académicos que hayan participado en el diseño de los instrumentos.

Adicionalmente, se debe contar con la distribución de los sustentantes para cada posible punto de corte, con la finalidad de hacer converger el juicio de los expertos con la evidencia empírica.

Los pasos a seguir son los siguientes:

1. Se calcula el promedio de k (\bar{k}), y de v (\bar{v}). Ambos valores generan el punto A con coordenadas (\bar{k}, \bar{v}) , (ver siguiente figura).

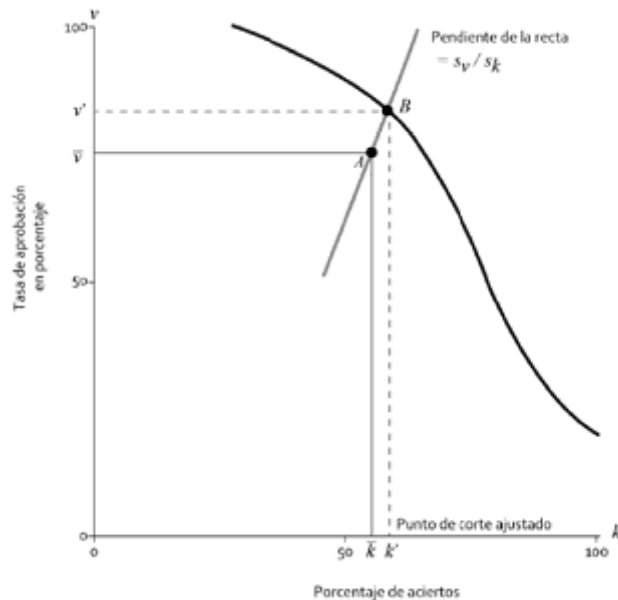
2. Para cada posible punto de corte se grafica la distribución de los resultados obtenidos por los sustentantes en cada área del instrumento de evaluación.

3. Se calcula la desviación estándar de k y v (s_k y s_v).

4. A partir del punto A se proyecta una recta con pendiente s_v/s_k hasta la curva de distribución empírica (del paso 2). El punto de intersección entre la recta y la curva de distribución es el punto B. La recta se define como: $v = (s_v/s_k)(k - \bar{k}) + \bar{v}$.

El punto B, de coordenadas (k', v') , representa los valores ya ajustados, por lo que k' corresponderá al punto de corte del estándar de desempeño.

El método asume que el grado en que los expertos están de acuerdo es proporcional a la importancia relativa que los expertos dan a las dos preguntas, de ahí que se utilice una línea recta con pendiente s_v/s_k .



Escalamiento de las puntuaciones

El escalamiento (Wilson, 2005) se llevará a cabo a partir de las puntuaciones crudas (cantidad de aciertos) de los sustentantes, y se obtendrá una métrica común para todas las áreas de los instrumentos de evaluación, que va de 60 a 170 puntos aproximadamente, ubicando el primer punto de corte (nivel de desempeño II) para todas las áreas de los instrumentos en los **100 puntos**. El escalamiento consta de dos transformaciones:

- Transformación doble arcoseno que estabiliza la magnitud de la precisión que se tiene para cada punto de la escala. **De no implementarla, para cada área del instrumento se tendría que estimar el error estándar de medida para todas y cada una de las puntuaciones de la escala.**
- Transformación lineal que ubica el primer punto de corte en 100 unidades y define el número de distintos puntos en la escala (el rango de las puntuaciones) con base en la confiabilidad del área de cada instrumento, por lo que, a mayor confiabilidad, habrá más puntos en la escala (Shun-Wen Chang, 2006).

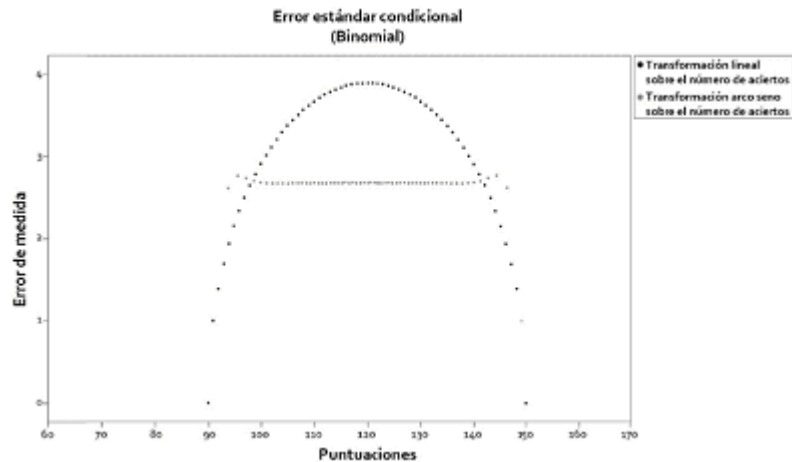
Para cuantificar el nivel de precisión de las puntuaciones de cada área del instrumento, se utilizará el método delta (Kendall y Stuart, 1977), que calcula los errores estándar de medición condicionales, que se describe ulteriormente en este anexo.

Finalmente, es importante destacar que, para realizar el escalamiento, el sustentante debió alcanzar, al menos, un acierto en el área del instrumento de evaluación en cuestión. De no ser así, se reportará como cero y el resultado será N I.

Procedimiento para la transformación doble arcoseno

En los casos de los exámenes de opción múltiple, deberá calcularse el número de respuestas correctas que haya obtenido cada sustentante en cada área del instrumento de evaluación. Los reactivos se calificarán como correctos o incorrectos de acuerdo con la clave de respuesta correspondiente. Si un sustentante no contesta un reactivo o si selecciona más de una alternativa de respuesta para un mismo reactivo, se calificará como incorrecto.

Como se observa en la gráfica (Won-Chan, Brennan y Kolen, 2000), con excepción de los valores extremos, el error estándar de medición se estabiliza a lo largo de la distribución de las puntuaciones observadas, a diferencia de la transformación lineal de las puntuaciones crudas.



Para estabilizar la varianza de los errores estándar de medición a lo largo de la escala, se utilizará la función c :

$$c(k_i) = \frac{1}{2} \left\{ \arcsen \sqrt{\frac{k_i}{K+1}} + \arcsen \sqrt{\frac{k_i+1}{K+1}} \right\} \quad (1)$$

Donde:

i se refiere a un sustentante

k_i es el número de respuestas correctas que el sustentante i obtuvo en el área

K es el número de reactivos del área

Procedimiento para la transformación lineal

La puntuación mínima aceptable que los sustentantes deben tener para ubicarse en el nivel de desempeño II (N II) en las áreas de los instrumentos de evaluación, se ubicará en el valor 100. Para determinarla se empleará la siguiente ecuación:

$$P_i = A * c(k_i) + B \quad (2)$$

Donde $A = \frac{Q}{[c(K) - c(0)]}$, $B = 100 - A * c(PC)$, Q es la longitud de la escala, $c(K)$ es la función c evaluada en K , $c(0)$ es la misma función c evaluada en cero y PC es el punto de corte (en número de aciertos) que se definió para establecer los niveles de desempeño y que corresponde al mínimo número de aciertos que debe tener un sustentante para ubicarlo en el nivel de desempeño II.

Debido a la longitud de las áreas y las confiabilidades esperadas, Q tomará el valor de 60 (Kolen y Brennan, 2014). Las puntuaciones P_i deben redondearse al entero más próximo, utilizando el criterio de que puntuaciones con cinco décimas o más, suben al siguiente entero.

Cálculo de las puntuaciones de los contenidos específicos de primer nivel en las áreas de los instrumentos de evaluación

Para calcular las puntuaciones del sustentante (i) en los contenidos específicos del primer nivel, se utilizará la puntuación ya calculada para el área (P_i), el número de aciertos del área (k_i), y el número de aciertos de cada uno de los contenidos específicos que conforman el área (k_{Aji}). Las puntuaciones de los contenidos específicos (P_{Aji}) estarán expresadas en números enteros y su suma deberá ser igual a la puntuación total del área (P_i).

Si el área está conformada por dos contenidos específicos, primero se calculará la puntuación del contenido específico 1 (P_{A1i}), mediante la ecuación:

$$P_{A1i} = P_i * \frac{k_{A1i}}{k_i} \quad (3)$$

El resultado se redondeará al entero inmediato anterior con el criterio de que puntuaciones con cinco décimas suben al siguiente entero. La otra puntuación del contenido específico del primer nivel (P_{A2i}) se calculará como:

$$P_{A2i} = P_i - P_{A1i} \quad (4)$$

Para las áreas de los instrumentos de evaluación con más de dos contenidos específicos, se calculará la puntuación de cada uno siguiendo el mismo procedimiento, empleando la ecuación (3) para los primeros. La puntuación del último contenido específico, se calculará por sustracción como complemento de la puntuación del área, el resultado se redondeará al entero positivo más próximo. De esta manera, si el área consta de j contenidos específicos, la puntuación del j -ésimo contenido específico será:

$$P_{Aji} = P_i - \sum_{k=1}^{j-1} P_{Aki} \quad (5)$$

En los casos donde el número de aciertos de un conjunto de contenidos específicos del área del instrumento sea cero, no se utilizará la fórmula (3) debido a que no está definido el valor de un cociente en donde el denominador tome el valor de cero. En este caso, el puntaje deberá registrarse como cero.

Procedimiento para el error estándar condicional. Método delta

Dado que el error estándar de medición se calcula a partir de la desviación estándar de las puntuaciones y su correspondiente confiabilidad, dicho error es un "error promedio" de toda el área del instrumento. Por lo anterior, se debe implementar el cálculo del error estándar condicional de medición (CSEM), que permite evaluar el error estándar de medición (SEM) para puntuaciones específicas, por ejemplo, el punto de corte.

Para cuantificar el nivel de precisión de las puntuaciones de las áreas del instrumento, se utilizará el método delta, (Muñiz, 2003), que calcula los errores estándar de medición condicionales. Para incluir la confiabilidad de las áreas del instrumento de medición se usa un modelo de error binomial, para el cálculo del error estándar condicional de medición será:

$$\sigma(X) = \sqrt{\frac{1 - \alpha}{1 - KR21} \left[\frac{X(n - X)}{n - 1} \right]}$$

Donde:

X es una variable aleatoria asociada a los puntajes

n es el número de reactivos del área del instrumento

KR21 es el coeficiente de Kuder-Richardson.

α es el coeficiente de confiabilidad de Cronbach, KR-20 (Thompson, 2003):

$$\alpha = \frac{n}{n - 1} \left(1 - \frac{\sum_{j=1}^n \sigma_j^2}{\sigma_X^2} \right)$$

$\sum_{j=1}^n \sigma_j^2$ = suma de las varianzas de los n reactivos

σ_X^2 = varianza de las puntuaciones en el área del instrumento

Para calcular el error estándar condicional de medición de la transformación P_i , se emplea el Método delta, el cual establece que si $P_i = g(X)$, entonces un valor aproximado de la varianza de $g(X)$ está dado por:

$$\sigma^2(P_i) \doteq \left(\frac{dg(X)}{dX} \right)^2 \sigma^2(X)$$

De ahí que:

$$\sigma(P_i) \doteq \frac{dg(x)}{dx} \sigma(x)$$

Aplicando lo anterior al doble arcoseno tenemos lo siguiente:

$$\sigma(P_i) \doteq \frac{A}{2} \left[\frac{1}{2(k+1) \left(\sqrt{\frac{x}{k+1}} \right) \left(\sqrt{1 - \frac{x}{k+1}} \right)} + \frac{1}{2(k+1) \left(\sqrt{\frac{x+1}{k+1}} \right) \left(\sqrt{1 - \frac{x+1}{k+1}} \right)} \right] \sigma(x)$$

Donde $\sigma(x)$ es el error estándar de medida de las puntuaciones crudas y $\sigma(P_i)$ el error estándar condicional de medición, de la transformación P_i , que ya incorpora la confiabilidad.

Para los puntajes que se les aplique la equiparación $x_e = b_1x + b_0$, con b_1 como pendiente y b_0 como ordenada al origen, el procedimiento es análogo, y el error estándar condicional de medición para la transformación $P_{ie} = A * c(x_e) + B$, que ya incorpora la confiabilidad, está dado por:

$$\sigma(P_{ie}) = \frac{A}{2} \left[\frac{b_1}{2(k+1) \left(\sqrt{\frac{x_e}{k+1}} \right) \left(\sqrt{1 - \frac{x_e}{k+1}} \right)} + \frac{b_1}{2(k+1) \left(\sqrt{\frac{x_e+1}{k+1}} \right) \left(\sqrt{1 - \frac{x_e+1}{k+1}} \right)} \right] \sigma(x_e)$$

Donde x_e son las puntuaciones equiparadas, las cuales son una transformación de las puntuaciones crudas, por lo que el error estándar de medida de dicha transformación se define como:

$$\sigma(x_e) = b_1 * \sigma(x)$$

La ventaja de llevar a cabo la transformación doble arco seno es que se estabiliza la magnitud de la precisión que se tiene para cada punto de la escala (Brennan, 2012; American College Testing, 2013; 2014a; 2014b). Esto permite atender al estándar 2.14 de los Estándares para las Pruebas Educativas y Psicológicas de la *American Educational Research Association et. al.*, 2014, que establece que los errores estándar de medida condicionales deben reportarse en varios niveles de puntuación, *a menos que haya evidencia de que el error estándar es constante a lo largo de la escala*, lo cual ocurre en este caso, al implementar la transformación doble arco seno.

El dato obtenido del error estándar condicional deberá reportarse en la misma escala en que se comunican las calificaciones de los sustentantes e incorporarse en el informe o manual técnico del instrumento (estándar 2.13 de los Estándares para las Pruebas Educativas y Psicológicas de la *American Educational Research Association et. al.*, 2014).

Referencias

- American College Testing, (2013). *ACT Plan Assessments Technical Manual*, Iowa City, IA: Author.
- American College Testing, (2014a). *ACT Assessments Technical Manual*, Iowa City, IA: Author.
- American College Testing, (2014b). *ACT QualityCore Assessments Technical Manual*, Iowa City, IA: Author.
- American Educational Research Association (AERA), American Psychological Association (APA) y National Council on Measurement in Education (NCM). (2014). *Standards for educational and psychological testing*. Washington, D.C.: Author.
- Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), *Educational Measurement* (2nd ed.). Washington, DC: American Council on Education.
- Beuk C. H. (1984). A Method for Reaching a Compromise between Absolute and Relative Standards in Examinations. *Journal of Educational Measurement*, 21 (2) p. 147-152.
- Brennan, R. L. (2012). Scaling PARCC Assessments: Some considerations and a synthetic data example en: <http://parconline.org/about/leadership/12-technical-advisory-committee>.
- Cook D. A. y Beckman T. J. (2006). Current Concepts in Validity and Reliability for Psychometric Instruments: Theory and Application. *The American Journal of Medicine* 119, 166.e7-166.e16
- Downing, S.M. (2004). Reliability: On the reproducibility of assessment data. *Med Educ*; 38(9):1006-1012.
- Holland, P. W., & Strawderman, W. E. (2011). How to average equating functions, if you must. In A. A. von Davier (Ed.), *Statistical models for test equating, scaling, and linking* (pp. 89–107). New York, NY: Springer
- Jonsson, A. & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review* 2: 130–44.

Kendall, M. & Stuart, A. (1977). *The advanced theory of statistics, Vol. 1: Distribution theory*. 4ª Ed. New York, NY: MacMillan.

Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices* (3rd ed.). New York, NY: Springer-Verlag.

Masters, G. (1982). A Rasch model for Partial Credit Scoring. *Psychometrika*-vol. 47, No. 2.

Muñiz, J. (2003). *Teoría clásica de los test*. Madrid: Ediciones pirámide.

Muraki, E. (1999). Stepwise Analysis of Differential Item Functioning Based on Multiple-Group Partial Credit Model. *Journal of Educational Measurement*.

OECD (2002). PISA 2000 *Technical Report*, PISA, OECD Publishing.

OECD (2005). PISA 2003 *Technical Report*, PISA, OECD Publishing.

OECD (2009). PISA 2006 *Technical Report*, PISA, OECD Publishing.

OECD (2014). PISA 2012 *Technical Report*, PISA, OECD Publishing.

Rezaei, A. R. & Lovorn, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing Writing* 15 (1.) 18–39.

Shun-Wen, Ch. (2006). Methods in Scaling the Basic Competence Test, *Educational and Psychological Measurement*, 66 (6) 907-927

Stellmack, M. A., Konheim-Kalkstein, Y. L., Manor, J. E., Massey, A. R., & Schmitz, J. A. P. (2009). An assessment of reliability and validity of a rubric for APA-style introductions, *Teaching of Psychology*, 36, 102-107.

Stemler, E. & Tsai, J. (2008). Best Practices in Interrater Reliability Three Common Approaches in *Best practices in quantitative methods* (pp. 89–107). SAGE Publications, Inc.

Thompson, B. (ed.) (2003). *Score reliability. Contemporary thinking on reliability issues*. SAGE Publications, Inc.

Wilson, M. (2005). *Constructing measures. An item response modeling approach*. Lawrence Erlbaum Associates, Publishers.

Won-Chan, L., Brennan, R. L., & Kolen, M. J. (2000). Estimators of Conditional Scale-Score Standard Errors of Measurement: A Simulation Study. *Journal of Educational Measurement*, 37(1), 1-20.

Wu, M. & Adams, R. (2007). *Applying the Rasch Model to Psycho-social measurement. A practical approach*. Melbourne: Educational measurement solutions.

TRANSITORIOS

Primero. Los presentes Criterios entrarán en vigor al día siguiente de su publicación en el Diario Oficial de la Federación.

Segundo. Los presentes Criterios, de conformidad con los artículos 40 y 48 de la Ley del Instituto Nacional para la Evaluación de la Educación, deberán hacerse del conocimiento público a través de la página de Internet del Instituto www.inee.edu.mx

Ciudad de México, a diecinueve de agosto de dos mil dieciséis.- Así lo aprobó la Junta de Gobierno del Instituto Nacional para la Evaluación de la Educación en la Vigésima Sesión Extraordinaria de dos mil dieciséis, celebrada el diecinueve de agosto de dos mil dieciséis. Acuerdo número SEJG/20-16/04, R. La Consejera Presidenta, **Sylvia Irene Schmelkes del Valle**.- Rúbrica.- Los Consejeros: **Eduardo Backhoff Escudero, Teresa Bracho González, Margarita María Zorrilla Fierro**.- Rúbricas.

El Director General de Asuntos Jurídicos, **Agustín E. Carrillo Suárez**.- Rúbrica.

(R.- 436757)