

# Consideraciones sobre la validez y la justicia en las evaluaciones del desempeño docente



# Índice

Cuestionamientos e inquietudes sobre la evaluación docente y su relación con la validez y la justicia.....	5
Requisitos de una evaluación válida.....	8
Claridad sobre qué y a quién se quiere evaluar.....	8
Contenidos congruentes con lo que se quiere evaluar.....	11
Implementación de las evaluaciones con accesibilidad, transparencia y trato digno a los sustentantes.....	15
Resultados de la aplicación congruentes con los supuestos con los que se diseñó la evaluación.....	16
Estandarización.....	18
Reflexiones finales.....	19





## Cuestionamientos e inquietudes sobre la evaluación docente y su relación con la validez y la justicia

Las preocupaciones y demandas de los maestros del país sobre la Evaluación del desempeño docente, compartidas por otros sectores como académicos, líderes de opinión y organizaciones sociales, se han concentrado principalmente alrededor de cuatro asuntos: la estandarización de la evaluación, la pertinencia del tipo de instrumentos y de su contenido, la contextualización de éstos y la implementación del modelo de evaluación.

1. Se cuestiona que una evaluación estandarizada pueda medir y valorar la práctica docente —compleja por definición—, así como hacerlo con justicia mediante el reconocimiento de la diversidad de contextos socioeconómicos y culturales y de las diferentes condiciones escolares.
2. Se duda de que el tipo de instrumentos empleado y sus contenidos cubran dicha complejidad y respondan a la variedad de perfiles docentes; se juzga que no incluyen la sustancia del trabajo docente y que dejan fuera lo verdaderamente importante.
3. Se afirma que, como lo manda la ley, no están considerados los contextos que condicionan el quehacer y los resultados de los docentes.
4. Finalmente, se alzan quejas sobre los mecanismos de implementación de las acciones evaluativas y sobre el trato que reciben los docentes en las diferentes fases del proceso.

Estas críticas guardan relación con la incertidumbre respecto del impacto que puede tener la evaluación sobre las trayectorias profesionales de los docentes.

Desde la perspectiva técnica, los cuestionamientos se pueden resumir de la siguiente manera: ¿el modelo de evaluación (los perfiles, parámetros e indicadores; los instrumentos y su calificación; la interpretación de resultados y la retroalimentación que se ofrece) es adecuado para tomar decisiones sobre la permanencia, la promoción y el reconocimiento de los docentes en su profesión?

En el lenguaje de la evaluación, esta pregunta alude a la validez y a la justicia del modelo. Validez y justicia son dos caras de una misma preocupación: la primera es la cara técnica mientras que la segunda, la social. La validez de una evaluación guarda relación directa con la justicia en la medida en que conlleva consecuencias para las personas.<sup>1</sup>

<sup>1</sup> American Educational Research Association/American Psychological Association/National Council on Measurement. (2014). *Standards for Educational and Psychological Testing*. Washington: Author; Newton, Paul E. & Shaw, Stuart D. (2014). *Validity in Educational & Psychological Assessment*. Los Angeles: Cambridge Assessment/SAGE.



La validez se puede definir de varias formas; una con amplia aceptación entre los especialistas es la que la describe como “...el grado en el cual la evidencia y la teoría sustentan las interpretaciones de las puntuaciones para los usos propuestos de las pruebas”.<sup>2</sup> Esta definición introduce una idea central: la validez no es un atributo constante de un instrumento, sino uno propio de las interpretaciones y usos, que debe someterse a revisión continua, cada vez que se aplican los instrumentos, se interpretan sus resultados y se toman decisiones. Quienes diseñan, aplican y utilizan evaluaciones tienen la responsabilidad no sólo de estar atentos a posibles amenazas a la validez reportadas por los evaluados u otros actores interesados, sino de vigilar y analizar continuamente el proceso de evaluación para identificar riesgos que pudieran ir apareciendo.

Considerando esta responsabilidad, en ejercicio de su autonomía y en cumplimiento de su mandato constitucional y de sus atribuciones legales, el Instituto Nacional para la Evaluación de la Educación (INEE) ha realizado un proceso de validación de cada etapa del diseño y desarrollo de la Evaluación del desempeño docente; ha sido receptivo a los planteamientos provenientes de diferentes sectores y, además, ha implementado diversos estudios para detectar elementos que pudieran afectar su validez. Estas acciones, que constituyen una evaluación de la evaluación, han quedado sistematizadas en los siguientes documentos:

- *Informe técnico del Modelo de Evaluación del Desempeño del Servicio Profesional Docente.* Dirección General para la Evaluación de Docentes y Directivos, Unidad de Evaluación del Sistema Educativo Nacional. México, agosto de 2016.
- *Evaluación del desempeño desde la experiencia de los docentes. Consulta con docentes que participaron en la primera Evaluación del desempeño en 2015.* Dirección General de Directrices para la Mejora de la Educación, Unidad de Normatividad y Política Educativa del Instituto Nacional para la Evaluación de la Educación. México, junio 2016.
- *Informe de supervisión de la Evaluación del desempeño docente y del personal con funciones de docencia, dirección y supervisión en educación básica y media superior. Ciclo escolar 2015-2016.* Dirección General de Lineamientos para las Evaluaciones, Unidad de Normatividad y Política Educativa del Instituto Nacional para la Evaluación de la Educación. México, junio 2016.
- *Evaluación de desempeño para la permanencia de docentes, directivos y supervisores en educación básica y media superior de México. Análisis y evaluación de su implementación 2015-2016.* Oficina Regional de Educación de la UNESCO para América Latina y el Caribe (OREALC), a solicitud del Instituto Nacional para la Evaluación de la Educación. Santiago de Chile, enero 2016.
- *Evaluación internacional de los procesos de evaluación de ingreso y promoción al Servicio Profesional Docente en educación básica y educación media superior en México, 2014-2015. Informe evaluativo.* Acuerdo de colaboración entre el Instituto Nacional para la Evaluación de la Educación de México y el IPEE-UNESCO. Buenos Aires, abril 2015.

<sup>2</sup> American Educational Research Association/American Psychological Association/National Council on Measurement. (2014). *Standards for Educational and Psychological Testing*. Washington: Author.



En el presente documento se da respuesta a las inquietudes de los docentes y de la sociedad acerca de la validez de la evaluación en el marco del SPD, específicamente la que refiere al desempeño. Para ello, se hace una breve revisión de los principales requisitos de una evaluación válida y justa y se detalla la manera en que la Evaluación del desempeño docente ha buscado cumplir esos requisitos. Con base en los estudios mencionados, se recuperan los aspectos del modelo que requieren ser modificados y se señalan algunas líneas de trabajo y acciones que se tienen previstas para continuar mejorando su validez. En esa revisión aparecen las nociones de contextualización, estandarización y pertinencia que han estado presentes en la opinión pública.



# Requisitos de una evaluación válida

## Claridad sobre qué y a quién se quiere evaluar

Un primer requisito de la validez es explicitar:

- Lo que se quiere evaluar (el constructo) y sobre qué población.
- La interpretación que se hará de los resultados de un instrumento o conjunto de instrumentos.
- Los fundamentos para hacer esa interpretación y para el uso que se hará de los resultados.

En el marco de la Ley General del Servicio Profesional Docente (LGSPD), el uso de los resultados de las evaluaciones comprende la toma de decisiones sobre:

- a) El ingreso de aspirantes al Servicio Profesional Docente (Artículo 21).
- b) La permanencia en el Servicio Profesional Docente (Título Segundo, Capítulo VIII).
- c) La promoción y el reconocimiento de quienes ya forman parte de dicho sistema (Título Segundo, Capítulos IV, V y VI).
- d) Los apoyos conducentes a la mejora del desempeño de los docentes (Título Cuarto, Capítulo I).

El objeto o constructo a evaluar ha sido expresado en los Perfiles, parámetros e indicadores (PPI) dados a conocer por la Secretaría de Educación Pública (SEP) en documentos específicos para cada nivel educativo y función.<sup>3</sup> Los PPI en su conjunto constituyen el Marco General de una Educación de Calidad (LGSPD Artículo 4, fracción XVII) y son el referente para la Evaluación del desempeño, ya que definen las características de un buen docente en términos de conocimientos, habilidades y responsabilidades profesionales orientadas a una enseñanza de calidad. Los perfiles publicados representan el primer paso en la construcción de prácticas de evaluación y de formación de docentes orientadas a fortalecer la eficacia escolar y una enseñanza de calidad, en la medida en que permiten traducir los grandes ideales educativos en acciones viables y concretas.<sup>4</sup>

<sup>3</sup> Secretaría de Educación Pública (SEP) (2015). Perfil, parámetros e indicadores para Docentes y Técnicos Docentes. Evaluación del desempeño. Ciclo escolar 2015-2016. México: Autor.

SEP. (2016). Perfil, parámetros e indicadores para Docentes y Técnicos Docentes en Educación Básica. Ciclo escolar 2016-2017. México: Autor.

SEP. (2015). Perfil, parámetros e indicadores para personal con funciones de Dirección y Supervisión. Ciclo escolar 2015-2016. México: Autor.

SEP. (2016). Perfil, parámetros e indicadores para personal con funciones de Dirección y Supervisión en Educación Básica. Ciclo escolar 2016-2017. México: Autor.

SEP. (2015). Perfil, parámetros e indicadores para Docentes y Técnicos Docentes en Educación Media Superior. Ciclo escolar 2015-2016. México: Autor.

SEP. (2016). Perfil, parámetros e indicadores para Docentes y Técnicos Docentes en Educación Media Superior. Ciclo escolar 2016-2017. México: Autor.

<sup>4</sup> Leyva, Y., Conzuelo, S., y Navarro, M. (2015). La buena enseñanza. Hacia la consolidación de un marco de la buena enseñanza en México. Perfil, parámetros e indicadores. *RED. Revista de evaluación para docentes y directivos*, 1(1).





Los PPI fueron formulados por la SEP mediante un proceso en el que participaron docentes frente a grupo, directores de escuela, supervisores, jefes de sector, asesores técnico pedagógicos, responsables de los niveles educativos y autoridades educativas locales en cada entidad federativa, quienes generaron propuestas. A partir de éstas se definieron perfiles compuestos de dimensiones, parámetros e indicadores. Se identificaron cinco grandes dimensiones para los docentes de educación básica (EB) y cinco para los de educación media superior (EMS), que responden a las particularidades de cada nivel (ver figura 1).

**Figura 1** Dimensiones de los perfiles docentes de educación básica y educación media superior

Educación básica		Educación media superior	
Dimensión 1	Un docente que conoce a sus alumnos, sabe cómo aprenden y lo que deben aprender.	Dimensión 1	Un docente que adapta los conocimientos sobre la disciplina que imparte y los procesos de enseñanza-aprendizaje de acuerdo con las características de los estudiantes.
Dimensión 2	Un docente que organiza y evalúa el trabajo educativo y realiza una intervención didáctica pertinente.	Dimensión 2	Un docente que planifica los proceso de formación, enseñanza aprendizaje y evaluación atendiendo al enfoque basado en competencias, y los ubica en contextos disciplinares, curriculares y sociales.
Dimensión 3	Un docente que se reconoce como profesional que mejora continuamente para apoyar a los alumnos en su aprendizaje.	Dimensión 3	Un docente que organiza y desarrolla su formación continua a lo largo de su trayectoria profesional.
Dimensión 4	Un docente que asume las responsabilidades legales y éticas inherentes a su profesión para el bienestar de los alumnos.	Dimensión 4	Un docente que vincula el contexto socio-cultural y escolar con el proceso de enseñanza-aprendizaje.
Dimensión 5	Un docente que participa en el funcionamiento eficaz de la escuela y formenta su vínculo con la comunidad para asegurar que todos los alumnos concluyan con éxito su escolaridad.	Dimensión 5	Un docente que construye ambientes de aprendizaje autónomo y colaborativo atendiendo el marco normativo y ético.

Estos dos perfiles generales de la buena docencia son sólo un paso en el proceso de definir el objeto de evaluación. Cuando se especifica lo que se quiere evaluar sobre una población y los usos que se darán a los resultados de esa evaluación, se debe plantear también la siguiente pregunta: ¿se pueden identificar características de subpoblaciones que, de no tomarse en cuenta, podrían dar lugar a errores en la medición? En una población se puede encontrar una gran variedad de subpoblaciones a partir de diferentes criterios, no todos relevantes para lo que se busca evaluar. En el caso de la evaluación docente, las características de algunas subpoblaciones —como aquellas definidas por su afición a los deportes o su estado civil— no son relevantes para el objeto a evaluar; sin embargo, hay otras que exigen ser consideradas como, por ejemplo, el nivel educativo en que trabajan, las modalidades, los tipos de servicio, las asignaturas y las especialidades docentes. Es así que estos diferentes ámbitos de la actuación docente dieron lugar a 33 perfiles en EB y a 6 en EMS (ver tablas 1 y 2).



**Tabla 1** Perfiles docentes de educación básica

Figura	Perfiles, parámetros e indicadores	Total
Docentes (Nacionales)	<b>Por nivel educativo:</b> Educación Preescolar Educación Primaria Educación Secundaria	15
	<b>Por modalidad, especialidad o servicio educativo:</b> Educación Indígena: Preescolar indígena Educación Indígena: Primaria indígena Educación Especial Educación Física Educación Telesecundaria Segunda Lengua: Inglés Asignatura Francés Tecnología. Educación Secundaria Artes: Artes Visuales. Educación Secundaria Artes: Danza. Educación Secundaria Artes: Música. Educación Secundaria Artes: Teatro. Educación Secundaria	
Docentes (Complementarios)	Lengua indígena Segunda Lengua: Inglés Asignatura Francés Asignatura Estatal Artes Tecnología	6
Técnicos Docentes	Maestro de taller de Lectura y Escritura. Educación Preescolar, Primaria y Secundaria Acompañante de Música. Educación Preescolar Maestro de enseñanza Artística. Educación Primaria Maestro de taller. Educación Primaria Promotor de TIC. Educación Primaria Maestro de Música. Educación Indígena Maestro de taller. Educación Indígena Tutor en Albergue Rural. Educación Primaria Acompañante de Música. Educación Especial Maestro de taller. Educación Especial Maestro de taller. Misiones Culturales Maestro de aula de medios. Educación Secundaria	12
<b>TOTAL</b>		<b>33</b>

**Tabla 2** Perfiles docentes de educación media superior

Figura	Perfiles, parámetros e indicadores	Total
Docentes	Campo Disciplinar Matemáticas Campo Disciplinar Ciencias experimentales Campo Disciplinar Humanidades Campo Disciplinar Ciencias sociales Campo Disciplinar Comunicación	5
Técnico Docente	Desempeño de Técnico Docente	1
<b>TOTAL</b>		<b>6</b>

Esta diversidad de perfiles constituye un importante avance para el Sistema Educativo Nacional en la medida en que orienta la realización de evaluaciones justas y transparentes, favorece la legitimación social de la profesión docente y sirve de guía puntual para procesos de formación inicial y continua.



En el marco de sus atribuciones, el INEE realizó el proceso de validación de todos los PPI mediante comités especializados en los que participaron docentes y personal con funciones de dirección y supervisión de los diferentes niveles y modalidades de EB y EMS. Los criterios de validación utilizados fueron los de congruencia, pertinencia, generalización, suficiencia y claridad.<sup>5</sup> Como resultado de la revisión, el Instituto emitió recomendaciones a la Coordinación Nacional del Servicio Profesional Docente (CNSPD) sobre aspectos técnicos de la construcción de los PPI para que cumplieran de manera suficiente con los criterios de validación.

Para mantener la vigencia y pertinencia de los PPI se requiere no sólo su revisión a la luz de los resultados de las evaluaciones sino también escuchar la voz de los actores involucrados. Al realizar esas acciones, el INEE ha reconocido tres principales líneas de trabajo para una mejor definición del objeto a evaluar:

1. Clarificar la redacción de algunos parámetros e indicadores.
2. Fortalecer la congruencia entre ambos.
3. Ajustar algunos indicadores para que reflejen mejor el quehacer docente en sus respectivos ámbitos, y que representen mejor la diversidad de las prácticas de enseñanza.

Para apoyar a la SEP en la mejora de estos aspectos, en marzo de 2015 el INEE elaboró el documento “Referentes técnicos para la elaboración de parámetros e indicadores derivados de los distintos perfiles asociados al SPD” y brindó asesoría a la CNSPD para mejorar los procesos de participación social y aumentar la representatividad en los comités para el desarrollo de los parámetros e indicadores.

## Contenidos congruentes con lo que se quiere evaluar

Además de especificar qué y a quiénes se quiere evaluar, una evaluación válida debe contar con instrumentos de medición acordes a estos propósitos. En su desarrollo es importante tener en cuenta dos consideraciones:

- a) ¿Lo que se quiere evaluar (las cinco dimensiones de los PPI) está suficientemente representado en los instrumentos? Ninguna evaluación puede medir todos los aspectos de un constructo ni hacerlo de manera completa, no importa si es estandarizada o no. Por eso, el objetivo es que los instrumentos recojan una muestra representativa del quehacer docente.
- b) ¿Los instrumentos están midiendo aspectos ajenos a lo que se quiere medir? No debe exigirse a los evaluados una habilidad o conocimiento que no es parte del objeto a medir. Por ejemplo, un examen de conocimientos didácticos de primaria no debe incluir reactivos o tareas evaluativas que demanden nociones de ingeniería, pues esos conocimientos no son parte de las cinco dimensiones de la docencia de calidad.

¿De qué manera los instrumentos de la evaluación docente cumplen con estas consideraciones? Como se mencionó, los perfiles tienen elementos de práctica, de conocimientos y habilidades y de contribución a resultados. Para la Evaluación del desempeño se ha decidido considerar, “además del dominio de los conocimientos disciplinares, curriculares y didácticos, diversos elementos de la práctica profesional, tomando en cuenta la diversidad de contextos sociales y culturales en los que

<sup>5</sup> Instituto Nacional para la Evaluación de la Educación. (2015). *Lineamientos para llevar a cabo la evaluación del desempeño de quienes realizan funciones de docencia, dirección y supervisión en educación básica y media superior*. LINEE-05-2015, Artículo 8.



se lleva a cabo”.<sup>6</sup> En los documentos Etapas, aspectos, métodos e instrumentos (EAMI),<sup>7</sup> se establece la importancia de que la evaluación recupere las condiciones laborales de los maestros, a través de su contexto sociocultural y los rasgos particulares de la escuela y el aula en la que desarrollan su práctica profesional. Asimismo, se describen las características del proceso en su conjunto, las etapas e instrumentos que se emplearán, y se especifican los aspectos a evaluar en correspondencia con los PPI.

En los EAMI se establecen las etapas para evaluar el desempeño del docente, como se enuncian en la tabla 3.

**Tabla 3** Etapas de la Evaluación del desempeño docente en educación básica y media superior

Educación básica	Educación media superior
<ul style="list-style-type: none"> <li>• <b>Etapa 1.</b> Informe de cumplimiento de responsabilidades profesionales</li> <li>• <b>Etapa 2.</b> Expediente de evidencias de enseñanza</li> <li>• <b>Etapa 3.</b> Examen de conocimientos y competencias didácticas que favorecen el aprendizaje de los alumnos</li> <li>• <b>Etapa 4.</b> Planeación didáctica argumentada</li> <li>• <b>Etapa 5.</b> Evaluación complementaria. Segunda Lengua: Inglés (Para los docentes de Secundaria que imparten la Asignatura Segunda Lengua: Inglés, se realizará una quinta etapa.)</li> </ul>	<ul style="list-style-type: none"> <li>• <b>Etapa 1.</b> Informe de cumplimiento de responsabilidades</li> <li>• <b>Etapa 2.</b> Expediente de evidencias de enseñanza</li> <li>• <b>Etapa 3.</b> Evaluación de conocimientos actualizados y de las competencias didácticas que favorecen el aprendizaje y el logro de las competencias de los estudiantes</li> <li>• <b>Etapa 4.</b> Planeación didáctica argumentada o presentación de constancia de CERTIDEMS o ECODEMS</li> </ul>

A continuación, se describe el contenido de cada etapa.

**Etapa 1:** Informe de cumplimiento de responsabilidades profesionales. El director de la escuela o, en su caso, el supervisor, emite un informe sobre aspectos relacionados con el trabajo del docente en la escuela, la atención a alumnos, la promoción de ambientes favorables para la sana convivencia, la colaboración en la escuela e integración con sus pares, la vinculación con padres de familia, todo ello con la intención de identificar el grado de cumplimiento de las responsabilidades profesionales del docente o técnico docente que son inherentes a su profesión. El instrumento tuvo carácter diagnóstico; careció de efecto para la calificación global en 2015-2016, pero su contenido fue parte de la retroalimentación.

**Etapa 2:** Expediente de Evidencias de Enseñanza. El docente recopila una muestra de los trabajos realizados por sus alumnos, como evidencia de su práctica de enseñanza; la analiza y explica como parte del ejercicio cotidiano de reflexión que hace sobre su práctica y sobre los efectos que ésta tiene en el aprendizaje de sus alumnos. Los aspectos que se evalúan están relacionados con la consideración del contexto de la escuela y de los estudiantes; con los contenidos del currículo y las actividades de aprendizaje y con los resultados de los alumnos y de su propia práctica.

<sup>6</sup> Instituto Nacional para la Evaluación de la Educación. (2015). *Lineamientos para llevar a cabo la evaluación del desempeño de quienes realizan funciones de docencia, dirección y supervisión en educación básica y media superior. LINEE-05-2015*, Artículos 9 y 34.

<sup>7</sup> Secretaría de Educación Pública. (2015). *Etapas, aspectos, métodos e instrumentos. Proceso de evaluación del desempeño docente. Educación básica*. México. Autor.

SEP (2015). *Etapas, aspectos, métodos e instrumentos. Proceso de evaluación del desempeño docente. Educación media superior*. México. Autor.



**Etapa 3:** Examen de conocimientos y competencias didácticas (EB) y Evaluación de conocimientos actualizados y de las competencias didácticas que favorecen el aprendizaje y el logro de las competencias de los estudiantes (EMS). A partir de la resolución de situaciones hipotéticas de la práctica educativa, se evalúan los conocimientos y competencias didácticas que el docente pone en juego para propiciar el aprendizaje de los alumnos, la colaboración en la escuela y el vínculo con los padres de familia y la comunidad. Estos exámenes son los mismos para los docentes de un nivel, modalidad, tipo de servicio, disciplina o especialidad porque se asume que todos deben tener una base mínima común de conocimientos y habilidades.

**Etapa 4:** Planeación Didáctica Argumentada. El docente elabora una planeación didáctica como muestra del ejercicio cotidiano de su práctica y un escrito en el que argumenta sobre el propósito, la estructura, el contenido y los resultados esperados de dicha planeación, y reflexiona acerca de lo que espera que aprendan sus alumnos y la forma en que evaluará lo aprendido.

Este esquema con etapas, aspectos, métodos e instrumentos para diferentes ámbitos del quehacer docente hace posible que cada persona sea evaluada con, al menos, cuatro instrumentos que constituyen lo que se ha llamado su trayecto evaluativo.

El conjunto de los instrumentos ha posibilitado evaluar aspectos de todas las dimensiones contenidas en los PPI: desde aquellos que rescatan la responsabilidad profesional, el dominio de los contenidos y las prácticas didácticas, hasta los que se ocupan de la planeación de la enseñanza, la generación de ambientes propicios para el aprendizaje en el aula y el cumplimiento de la normalidad mínima de operación escolar. Particularmente, el Expediente de evidencias de enseñanza y la Planeación didáctica argumentada son fortalezas de la evaluación por su congruencia con lo que se quiere evaluar. Estas etapas permiten que los docentes contextualicen su quehacer y fundamenten las decisiones de su intervención a partir de: a) caracterizar a sus estudiantes y sus condiciones sociales, culturales, económicas; b) describir los planteles en los que laboran, y c) mostrar la forma en que su práctica docente es pertinente para ese contexto. Las evidencias y argumentaciones presentadas por los sustentantes son revisadas y calificadas por evaluadores capacitados y certificados por el INEE, que fueron o son docentes del mismo nivel, tipo, modalidad educativa, asignatura o función que el docente a quien evalúan. Utilizan rúbricas que los instruyen a considerar no sólo la corrección técnica de los procesos de enseñanza documentados por las evidencias de enseñanza y de la planeación didáctica argumentada, sino su adecuación al contexto en el que labora el docente y que fue caracterizado por él mismo. Dos evaluadores revisan las evidencias de enseñanza y la planeación didáctica argumentada de cada docente de manera ciega. En caso de no haber coincidencia entre los dos, interviene un tercero que resuelve la discrepancia evitando perjudicar al docente evaluado.<sup>8</sup>

<sup>8</sup> Las normas para el proceso de selección, capacitación y certificación de evaluadores, así como los criterios para la calificación de instrumentos se pueden consultar en:

Instituto Nacional para la Evaluación de la Educación. (2015). *Lineamientos para la certificación de evaluadores que participarán en la evaluación del desempeño en educación básica y media superior para el ciclo escolar 2015-2016*. LINEE-02-2015.

INEE. (2015). *Lineamientos para la selección y capacitación de evaluadores que participarán en la evaluación del desempeño en educación básica y media superior para el ciclo escolar 2015-2016*. LINEE-03-2015.

INEE. (2015). *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 del desempeño de quienes realizan funciones de docencia en Educación Básica, 2015-2016*.

INEE. (2015). *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 del desempeño de quienes realizan funciones de docencia en Educación Media Superior, 2015-2016*.



Es importante resaltar que las evidencias de enseñanza y la planeación argumentada no solicitan de los docentes evidencias de un determinado nivel de logro de resultados en el aprendizaje de conocimientos, habilidades y actitudes por parte de los estudiantes, ni de otros efectos de su acción en el aula o en la comunidad escolar, como podría ser el clima del salón de clase. Esto es así debido a que los resultados del trabajo docente se ven influidos por el contexto de una manera muy distinta que la práctica. Mientras que ésta debe adecuarse al contexto, los resultados dependen en buena medida del contexto social, económico y cultural de los estudiantes.

En la evaluación 2015-2016 se utilizaron 77 instrumentos. De ellos, 74 fueron diseñados ad hoc para el SPD: 58 de opción múltiple y 16 rúbricas para calificar los instrumentos de respuesta construida. Además, se utilizaron los exámenes Certificación de Competencias Docentes para la Educación Media Superior (CERTIDEMS) y Evaluación de Competencias Docentes para la Educación Media Superior (ECODEMS) —cuyos resultados fueron considerados para cubrir lo medido por la planeación didáctica argumentada—, y el de Aptitud en la lengua inglesa para los docentes de esa asignatura.

El Instituto realizó el proceso de validación de los EAMI verificando el cumplimiento de los criterios de:<sup>9</sup> a) congruencia, que se refiere a la correspondencia de los aspectos, etapas y métodos con los indicadores seleccionados; b) pertinencia, que permite reconocer si el método propuesto para la evaluación es útil, adecuado y procedente. En la validación de los reactivos y tareas evaluativas de los instrumentos se incluyó, además, el criterio de ausencia de posibles sesgos por cuestiones lingüísticas, de género, cultura, religión, nivel socioeconómico y acceso a la tecnología. Cuando se detectaron componentes que no cumplían con estos criterios, se solicitó su corrección antes de validarlos.

El cumplimiento del conjunto de los criterios utilizados por el INEE en la revisión técnica de los instrumentos contribuye a la validez de éstos. Dichos criterios se basan en estándares de calidad técnica propuestos por organismos de evaluación educativa nacionales e internacionales.

La vigilancia sobre el proceso de evaluación ha mostrado que, aunque se han tenido logros importantes en la definición de los EAMI, todavía hay un campo de mejora en aspectos como:

1. La articulación de las visiones de las autoridades con las de los docentes en el informe de cumplimiento de responsabilidades profesionales.
2. El equilibrio en la distribución de los indicadores entre los diferentes instrumentos de cada trayecto evaluativo, pues ahora están concentrados principalmente en los exámenes objetivos.
3. La representatividad de los exámenes objetivos respecto de lo que se requiere en cada ámbito del quehacer docente tanto en EB como en EMS.
4. La precisión de las rúbricas para que permitan diferenciar mejor los niveles de desempeño.
5. La vinculación de las etapas 2 y 4 en el proceso de evaluación, tal como ocurre en las prácticas docentes en la escuela.
6. El grado de especificidad de los criterios para la detección de posibles sesgos por género, lengua, cultura y condiciones socioeconómicas.

El INEE ha formulado recomendaciones pertinentes a la SEP para mejorar estos y otros aspectos y continúa revisando y perfeccionando sus procedimientos de validación.

<sup>9</sup> Instituto Nacional para la Evaluación de la Educación. (2015). Criterios técnicos para el desarrollo y uso de instrumentos de evaluación educativa, 2014-2015. México: Autor.



## Implementación de las evaluaciones con accesibilidad, transparencia y trato respetuoso a los sustentantes

La forma en que se implementa una evaluación puede contribuir a asegurar su validez o puede dar lugar a barreras para que los evaluados demuestren su competencia en el constructo a medir. Las amenazas a la validez pueden originarse en diferentes aspectos y fases de la Evaluación del desempeño docente.

Por ejemplo, si por diferentes circunstancias los sustentantes no disponen del tiempo que se ha establecido para contestar un instrumento, podrían obtenerse mediciones insuficientes sobre su dominio del objeto a evaluar. De igual manera, si los docentes evaluados en una sede no tienen acceso a agua y alimentos a lo largo de la jornada de aplicación, se les estaría demandando algo más que sus habilidades y conocimientos. La falta de información sobre cómo deben prepararse los docentes para la evaluación también puede ser un factor que les impida demostrar sus competencias cabalmente. Algunas de las acciones que deben realizar los evaluados conllevan un nivel de conocimiento y familiaridad con el uso de las tecnologías de información y comunicación que, si bien está contenido en los indicadores de los perfiles docentes, no ha estado incluido en la formación de algunos maestros. Los docentes con discapacidad visual no podrían leer y dar respuesta a los instrumentos. Aquellos con discapacidad motora corren el riesgo de no poder acceder con oportunidad a las aulas de evaluación. Asimismo, un trato irrespetuoso por parte de los conductores de la evaluación puede afectar el estado emocional de los maestros y, por tanto, su desempeño en la misma.<sup>10</sup>

Se ha documentado ampliamente que ha habido diferencias notables en la implementación de la evaluación docente y, por tanto, en la experiencia de los sustentantes en las distintas entidades y sedes: la notificación de que alguien será evaluado no siempre ha llegado con oportunidad; el ambiente físico y social de algunas sedes ha sido incómodo y, en ocasiones, hostil; los procedimientos de registro y de acceso han dificultado el ingreso a las aulas; las instrucciones de los aplicadores han sido, a veces, poco claras; y el equipo de cómputo no siempre ha funcionado de manera regular. Estas situaciones han provocado molestias considerables en algunos docentes. Sin embargo, los análisis realizados por el INEE no arrojan evidencia de que las diferencias de funcionamiento entre las sedes hayan puesto en desventaja a quienes fueron evaluados en aquellas que tuvieron problemas.

Pero también se han identificado medidas tomadas por las autoridades educativas locales para apoyar a quienes tenían menos habilidades digitales: asesoría para la entrega de evidencias, centros de ayuda digital, grupos de apoyo y otras. Por su parte, la autoridad federal, en la *Guía para el coordinador de sede de aplicación SEP*, previó las siguientes acomodaciones:<sup>11</sup>

- Evaluados con discapacidad visual: apoyo de un aplicador que leería las preguntas en voz alta y registraría las respuestas, así como tiempo adicional.
- Evaluados con discapacidad motora: asignación a las aulas en planta baja.

Actualmente se está estudiando la posibilidad de hacer otras acomodaciones para éstas y otras subpoblaciones.

<sup>10</sup>Debe decirse que, incluso si el mal trato no tuviera efecto alguno en el desempeño, seguiría siendo inadmisibles en tanto representa una falta de respeto a su dignidad.

<sup>11</sup>Herramientas y procedimientos para dar acceso a la evaluación a las personas con alguna discapacidad sin alterar la medición del constructo de interés.



Los estudios que el INEE ha realizado le han permitido hacer recomendaciones específicas a las autoridades educativas para corregir las problemáticas surgidas en la implementación. De hecho, los mismos estudios muestran que ya en la Evaluación del desempeño realizada en julio de 2016 a quienes ingresaron al Servicio dos años antes, se introdujeron correcciones a los procedimientos y fue posible apreciar una mejoría moderada en la percepción de los evaluados respecto de la expresada por los participantes en la evaluación de 2015. Este incremento de la satisfacción ocurre en casi todos los elementos del proceso de evaluación, tanto los previos a la aplicación (notificación, utilidad de las guías, entre otros), como los que se llevan durante ésta (las instrucciones, el trato del aplicador, el funcionamiento del equipo de cómputo y otros).

A pesar de que los docentes reportan mayor satisfacción con el proceso, su opinión sigue siendo poco favorable en algunos rubros, en especial, la duración de las aplicaciones y la claridad con la que se plantean las preguntas, así como los casos de los exámenes de conocimientos y competencias didácticas y la suficiencia y funcionamiento de los equipos de cómputo). Por ello, el INEE continúa recogiendo y analizando información para aportar a la mejora de la logística y a la prevención de barreras a una evaluación válida y justa.

### Resultados de la aplicación congruentes con los supuestos con los que se diseñó la evaluación

Al diseñar los instrumentos de una evaluación se está planteando una hipótesis acerca de que la selección de contenidos a medir, los instrumentos a aplicar, así como sus reactivos y tareas evaluativas, son una forma adecuada de medir lo que se desea evaluar y tomar las decisiones correspondientes. Una evaluación válida debe aportar elementos para juzgar el grado en que esta hipótesis se verifica, lo que se logra habitualmente mediante análisis estadísticos de los resultados. En una evaluación como la del desempeño docente, las evidencias mínimas requeridas se refieren a los siguientes aspectos:

- El conjunto de reactivos tiene un grado de dificultad conveniente para el uso que se dará a los resultados, así como una probabilidad adecuada de discriminación o capacidad para distinguir entre sustentantes con alto y bajo desempeño.
- La confiabilidad de las puntuaciones cumple con los criterios técnicos convencionales.
- Las puntuaciones obtenidas no muestran tener sesgo derivado de factores externos al desempeño.<sup>12</sup>
- El procedimiento para determinar puntos de corte está documentado y es consistente. Es decir, las acciones realizadas para definir los puntajes que representan los diferentes niveles de desempeño, en especial el insuficiente, están claramente expresadas y siguen buenas prácticas en la materia.

<sup>12</sup>Esto complementa las consideraciones conceptuales y lógicas utilizadas para el diseño y desarrollo de los instrumentos, y contribuye a la revisión de posibles sesgos en los que se incurrió durante dicho desarrollo.





Con base en los análisis de dificultad, discriminación, confiabilidad y otros que se han realizado sobre los instrumentos aplicados, el Instituto ha determinado la cancelación de algunos de ellos. Las puntuaciones obtenidas en dichos instrumentos no tuvieron efecto en la calificación de los docentes evaluados. Por lo demás, los indicadores psicométricos han sido adecuados.

Como se ha dicho en este documento, una instancia evaluadora debe examinar continuamente los resultados de sus instrumentos para identificar aspectos que atenten contra la validez. Por lo anterior, el INEE tiene prevista una agenda de estudios de validez para continuar el fortalecimiento del desarrollo de evaluaciones justas y contextualizadas. Actualmente, se desarrolla uno sobre los instrumentos de respuesta construida.



## Estandarización

En los párrafos previos se ha aludido de diferentes maneras a las nociones de pertinencia y contextualización de las evaluaciones, así como a las de trato digno y accesibilidad a la evaluación. Enseguida se discute la noción de estandarización, considerando que podrá comprenderse mejor a partir del análisis previo.

Contrariamente a lo que se piensa, la estandarización de instrumentos es un medio para la justicia en las evaluaciones. La estandarización busca ofrecer a todos los evaluados las mismas oportunidades para plasmar su capacidad, de manera que sus resultados sean comparables y se eviten injusticias. Para que ello ocurra, se deben usar los mismos reactivos o tareas evaluativas, dar las mismas instrucciones, proporcionar el mismo tiempo para responder, establecer condiciones ambientales similares, entre otras medidas. Sin embargo, poner en las mismas condiciones a personas diferentes puede requerir de adaptaciones, por la forma en que algunas de ellas responden al instrumento, sin alterar el concepto de estandarización. Por ejemplo, un docente con debilidad visual puede necesitar una prueba distinta de la de sus colegas; esa prueba debe ser igual en los contenidos, pero estar impresa en macrotipo y contar con diversas adecuaciones de diseño. Esto se llama técnicamente *acomodación*, como se vio antes en este documento. Se puede afirmar, sin caer en contradicción, que la correcta estandarización de una prueba puede necesitar de este tipo de adaptaciones para poner en verdadera igualdad de condiciones a todos los sustentantes.

Por otra parte, hay situaciones en las que las características de un grupo demandan adaptaciones tales que el objeto a medir se modificaría tanto que ya no sería el mismo. Por ejemplo, el quehacer de las educadoras de preescolar indígena difiere cualitativamente del de las de preescolares generales. Lo mismo se puede decir de los docentes de Educación Física y de otras especialidades. En esos casos, el uso de un solo instrumento estandarizado para estos y otros grupos de maestros resultaría en injusticias y en falta de comparabilidad con el resto de sus colegas, por lo que se optó por tener perfiles e instrumentos de evaluación diferenciados.

Por lo examinado en este documento, se puede afirmar que la evaluación del desempeño docente es una evaluación estandarizada y también una diversificada, en la que hay varias adaptaciones para poner en igualdad de condiciones a diferentes participantes: la diversificación de perfiles, la posibilidad de que cada evaluado caracterice el contexto para el cual su práctica es pertinente, las acomodaciones para docentes con debilidad visual y con discapacidad motora, la eliminación de reactivos que pueden introducir sesgo hacia algunos grupos, y otras medidas. Los 74 instrumentos de la evaluación docente desarrollados para el SPD están diversificados para atender a la gran variedad de ámbitos educativos e incluyen tareas evaluativas que permiten introducir el contexto, pero, al interior de cada ámbito, están estandarizados para asegurar la comparabilidad y evitar injusticias.

El grado en que deben diversificarse los instrumentos y las acomodaciones que deben introducirse para asegurar la justicia de una evaluación no pueden definirse de una vez y para siempre. Las sucesivas aplicaciones de una evaluación pueden ir mostrando la conveniencia de nuevas acomodaciones e incluso, de rediseñar sustancialmente instrumentos para hacerlos más pertinentes a las prácticas docentes.



## Reflexiones finales

En este documento se revisaron algunos de los requisitos para que una evaluación sea válida y la manera en que se aplicaron a la Evaluación del desempeño. Se expuso información sobre la forma en que el INEE ha buscado cubrirlos y sobre las tareas pendientes para continuar revisando la validez de la evaluación. El propósito del documento ha sido responder a las inquietudes sobre la pertinencia y contextualización de las evaluaciones, así como a las quejas sobre el trato que han recibido algunos docentes en las sedes.

Una responsabilidad central del Instituto es asegurar la validez de la evaluación, es decir, el grado en que la teoría y la evidencia empírica justifican las interpretaciones de sus resultados y las decisiones que se toman a partir de ellas. Para atenderla adecuadamente, el INEE ha asumido la tarea de recoger, sistematizar y analizar continuamente información que le permita sustentarla. Ha introducido mejoras y correcciones puntuales en los dos años de operación de las evaluaciones docentes y, como se ha mencionado, ha hecho recomendaciones a la SEP. Pero, además, está replanteando el modelo de evaluación docente y programando una serie de acciones para la mejora continua de su validez.

Todos los esfuerzos que el Instituto realice para asegurar este rasgo fundamental de las evaluaciones redundarán en decisiones justas para los docentes y benéficas para el aprendizaje de los estudiantes de México. •

