



Orientaciones Técnicas para el desarrollo de la Evaluación Diagnóstica inicial en pedagogías

Departamento de Evaluación,
Medición y Registro Educacional
Vicerrectoría de Asuntos Académicos
Universidad de Chile

Enero 2019

CONTENIDO

I.	Introducción	3
II.	Aspectos fundamentales	4
III.	Aspectos técnicos	6
IV.	Experiencias analizadas	8
V.	Referencias	10

I. Introducción

En este documento se presentan las Orientaciones Técnicas para el desarrollo de las evaluaciones diagnósticas para estudiantes de pedagogía en Chile. La Ley 20.903 sobre el desarrollo profesional docente incluyó la obligatoriedad, a partir del año 2017, de evaluaciones diagnósticas al ingreso de todas las carreras de pedagogía. La ley entrega autonomía en la implementación de estas evaluaciones, las cuales guardan un enorme potencial para mejorar la formación inicial docente. Es por esto que este documento espera ser una guía que apoye el diseño, implementación y uso de estas evaluaciones. Esta dirigido tanto para formadores de futuros profesores como también para los y las tomadores de decisiones de las universidades formadoras de docentes.

Estas orientaciones forman parte del proyecto FONIDE N°1700009 “Estándares técnicos para el diagnóstico inicial de los futuros docentes. Una propuesta basada en evidencia” desarrollado por el Departamento de Evaluación, Medición y Registro Educativo de la Universidad de Chile, en alianza con la Universidad de O’Higgins y la Pontificia Universidad Católica de Valparaíso. Este proyecto estudió las experiencias de implementación y usos de las evaluaciones diagnósticas aplicadas el año 2017 en 5 universidades. La metodología del estudio fue mixta y centró su análisis en tres focos principales, a saber, propósitos y usos de la evaluación, administración en la aplicación, y contenidos y comportamiento psicométrico de las pruebas (Giaconi, et al. 2019).

La estructura y orientaciones principales planteadas se definieron en base a los resultados del proyecto FONIDE mencionado anteriormente y a los Estándares para Pruebas Educativas y Psicológicas (en adelante, Los Estándares) (AERA, APA y NCME, 2018) que son un referente a nivel mundial para lograr la calidad de las evaluaciones.

Este documento se divide en tres grandes secciones. En primer lugar, se encuentran los Aspectos Fundamentales, guiados por las preguntas “¿qué medir?” y “¿qué hacer con las mediciones?”, cuyo principal interés es orientar los procesos de elaboración y definición de los objetivos de la Evaluación Diagnóstica. En segundo lugar, se encuentran los Aspectos Técnicos, guiados por las preguntas sobre “¿cómo medir?” y “¿qué aspectos técnicos y procedimentales tener en cuenta?”, las cuales buscan orientar al lector interesado en lo que respecta a la implementación de las pruebas, desde la fase de construcción hasta los reportes de resultados. En tercer lugar, están las recomendaciones específicas relacionadas con los dominios evaluados en las universidades incluidas en el estudio FONIDE. Estos dominios corresponden a matemáticas, lenguaje (lectura y escritura) y aspectos no disciplinares.

II. Aspectos fundamentales

Definir el propósito de la evaluación

Las expectativas que recaen en las evaluaciones diagnósticas son muy altas y, por lo mismo, los instrumentos de evaluación que se diseñen para tal propósito tienen un rol gravitante, tanto para la recolección de evidencia sobre habilidades y conocimientos claves para el éxito de los estudiantes que ingresan a pedagogía, como también para el logro del perfil de egreso establecido por cada universidad formadora. Bajo este contexto, es importante que estos instrumentos cumplan con ciertas características fundamentales que permitan entregar evidencias de validez. Ésta se define en los estándares como “grado en que la evidencia y la teoría respaldan las interpretaciones de los puntajes de una prueba para usos propuestos de las pruebas. La validez es, por tanto, la consideración más fundamental al desarrollar y evaluar pruebas” (AERA, APA y NCME, 2018, p. 11).

En ese sentido, es fundamental establecer claramente los propósitos y usos específicos que se darán a los resultados que informen estas evaluaciones, los cuales deben ser conocidos y compartidos por todos quienes participen de la evaluación. El propósito que se defina para estas evaluaciones es también una guía para el diseño y confección de los instrumentos y, por tanto, es el primer paso en el proceso de construcción. Además, tener claridad respecto a los propósitos es la única manera de poder evaluar en qué grado las interpretaciones de una prueba son adecuadas para los usos definidos.

Para definir el propósito de las evaluaciones resulta relevante que cada universidad tenga en cuenta las características de los estudiantes que ingresan a las carreras de pedagogía, así como también el perfil de egreso establecido por cada carrera y las competencias docentes que se espera que tengan los futuros profesores. La evidencia de la investigación realizada dio cuenta que la definición de los propósitos de la evaluación era uno de los puntos más débiles en los casos estudiados (Giacconi et al., 2019). Se deben generar las instancias de diálogo y reflexión al interior de las universidades formadoras de docentes para definir cuáles serían los objetivos a evaluar en los estudiantes a su ingreso.

Establecer los usos posibles (conjunto de acciones o medidas) a partir de los resultados de la evaluación

Los usos previstos guardan relación con la obtención de información que permita a las instituciones realizar diversas acciones con la cohorte de estudiantes de pedagogía de primer año. Tal como indica la Ley 20.903, se espera que los resultados de estas evaluaciones tengan un carácter referencial y formativo, a partir de los cuales se puedan diseñar y establecer acciones de nivelación y acompañamiento. La evidencia del estudio FONIDE muestra que tales acciones existen en las universidades, pero no se encuentran aún vinculadas a los resultados de la evaluación, sino que responden a otros criterios (Giacconi et al., 2019, p. 32). Y al mismo tiempo, los usos de la información que proporcionan los instrumentos, según las propias instituciones, no están del todo definidos y establecidos.

Por eso, una de los principales objetivos que deben perseguir estos instrumentos es ser una referencia efectiva para la toma de decisiones. En el mediano plazo, este compromiso puede favorecer a que los formadores de profesores realicen ajustes pertinentes en sus estrategias de enseñanza, sobre todo en los cursos introductorios. Desde luego, los programas de acompañamiento y nivelación también podrían fortalecerse con información pertinente de este tipo.

Por ello, es recomendable precisar para qué se necesita la información que puedan entregar estas evaluaciones, y en función de eso velar porque la evaluación diagnóstica se encuentre integrada dentro de un marco de acciones y medidas de fortalecimiento de la formación inicial docente.

Generar información pertinente y que pueda ser considerada y abordada en el transcurso de la formación inicial

Es relevante que la información que recaben estos instrumentos sea pertinente, novedosa y enfocada en aspectos del aprendizaje o de la formación que puedan ser intervenidos en el transcurso de la formación inicial.

Sobre lo primero, se debería velar porque lo que se mida entregue información nueva, que no sea posible de obtener a partir de la información que proporcionen otros instrumentos y/o evaluaciones. Un buen ejemplo es evitar que la evaluación se base en modelos como la PSU, dado que ese instrumento ya está informando sobre esas competencias y conocimientos del estudiante. Por su parte, lo que se necesita saber al ingreso de la carrera dice más relación con aptitudes específicas del estudiante orientadas al buen desempeño en sus estudios de pedagogía. Algunas preguntas que pueden guiar esta búsqueda son:

- ❖ ¿qué es lo que tienen que conocer las instituciones del alumnado para lograr el perfil de egreso?;
- ❖ ¿qué es lo que tienen que conocer las instituciones del alumnado para lograr la retención y graduación?

Sobre lo segundo, es fundamental que lo que se evalúe sea factible de trabajar y monitorear en el transcurso de la formación inicial. Por ejemplo, pesquisar ciertas variables asociadas a los procesos de aprendizaje que guardan alta relación con el éxito académico, o aquellas que se encuentren directamente vinculadas al perfil de egreso de la institución (ej., competencias docentes, sello formativo o valórico, creencias sobre la enseñanza y el aprendizaje).

III. Aspectos técnicos

Diseño y desarrollo de las pruebas

Los procesos de diseño y construcción de los instrumentos deben ser anteceditos por una discusión que esclarezca los propósitos y usos de la evaluación, tal y como se dijo anteriormente. En ese sentido, estos procesos deben estar guiados por estas definiciones iniciales, y en ningún caso pueden descuidar esa fase. Esta tarea no necesariamente debe recaer en el personal técnico, sino que pueden ser definidas con antelación por un equipo académico, y que luego sean implementadas por el equipo responsable. Es importante registrar la definición del constructo evaluado, así como también los objetivos de la evaluación en un Marco de Evaluación (AERA, APA y NCME, 2018, p. 11). El Marco de Evaluación es un documento central y articulador de una evaluación.

Estos instrumentos deben además contar con evidencias de validez, confiabilidad e imparcialidad que permitan respaldar las interpretaciones de los puntajes de las evaluaciones para los propósitos y usos. Este tipo de evidencia se describe en la próxima orientación.

Validación

Las evidencias de validez son un aspecto central para una evaluación, ya que permiten justificar su uso. Para eso es importante tener en cuenta las cinco fuentes de evidencia de validez establecidas por los Estándares (AERA, APA y NCME, 2018, pp. 28-33): 1) evidencia basada en el contenido, 2) evidencia basada en los procesos de respuesta, 3) evidencia desde la estructura interna, 4) evidencia en relación con otras variables, y 5) evidencia basada en las consecuencias de la evaluación. La evidencia de validez que es necesario recolectar depende de la evaluación y sus propósitos y usos. De todas maneras, cabe considerar la evidencia de validez basada en procesos de respuesta como una alternativa sugerente para los investigadores que quieran conocer en detalle las razones del comportamiento de determinados ítems. En algunos casos del proyecto FONIDE, los tomadores de decisiones y equipos constructores consideraban que los ítems eran muy fáciles, pero luego los resultados daban cuenta que el ítem era en verdad bastante difícil para la población evaluada (Giaconi, et al. 2019). Esa situación puede esclarecerse con investigación orientada específicamente a conocer cómo los estudiantes se enfrentan a ciertos ítems.

Finalmente, es conveniente obtener evidencia de validez de los instrumentos de manera continua, esto implica que los instrumentos deben ser constantemente piloteados y evaluados. Esto es necesario por dos razones. La primera es que las poblaciones de estudiantes de pedagogía pueden cambiar y los instrumentos pueden necesitar ajuste. La segunda es que los criterios técnicos que pueden orientar la construcción y validación de instrumentos no son estáticos, sino que se encuentran sujetos a los propósitos y usos definidos por cada institución. Un buen ejemplo es que los ítems no necesariamente deben discriminar y ubicar a los estudiantes en una curva de desempeño, ya que esto es así cuando el instrumento es de selección y tiene consecuencias. En el caso de las evaluaciones diagnósticas, los ítems pueden trabajarse de otro modo, considerando, por ejemplo, que un ítem con baja tasa de respuestas correctas es en sí mismo un hallazgo, al igual que un ítem que tenga muchas respuestas correctas.

Puntuación, administración y presentación de reportes e interpretación

El resguardo de la validez de un instrumento de evaluación se juega en las distintas etapas de su diseño, desarrollo, aplicación y análisis (AERA, APA y NCME, 2018). Respecto a la puntuación y administración, se deben elegir formatos de prueba de acuerdo al constructo evaluado y sus propósitos y usos (ej., respuesta abierta o selección múltiple, papel o digital, etc.), aunque es recomendable que las pruebas sean aplicadas en computador, ya que eso puede ayudar a automatizar los procesos de corrección y favorecer una pronta publicación de los resultados que sirva de insumo tanto a los equipos directivos, profesores y a los mismos estudiantes. La idea es que una evaluación de estas características de cuenta del panorama de la población de estudiantes en el primer semestre del año aplicado. A su vez, hay que tener en consideración adecuaciones especiales para estudiantes en situación de discapacidad.

Por otro lado, es importante que la administración sea estandarizada y para lograr esto hay que generar protocolos y capacitar a los equipos que desarrollen esta tarea. Se aconseja elaborar protocolos que detallen los procesos, duraciones, materiales necesarios, así como también que sirvan de base para la capacitación de los responsables de la aplicación de las pruebas.

Conformación del equipo técnico responsable de la evaluación

Realizar evaluaciones de esta envergadura supone también inversión en recursos humanos capacitados técnicamente en materias de medición en evaluación, administración y aplicación de las evaluaciones. Es deseable que se constituya un equipo responsable de la Evaluación Diagnóstica con competencias que le permitan diseñar o adaptar instrumentos y generar evidencias de validez de manera regular, en esto se incluye la administración de los instrumentos. Por otro lado se necesita que esté equipo procure que los usos previstos para el instrumento se vinculen con los resultados efectivos de la evaluación, lo que incluye reportar los resultados de manera que resulten funcionales a las acciones de acompañamiento y nivelación.

IV. Experiencias analizadas

La evidencia del proyecto FONIDE permite caracterizar los instrumentos por el dominio evaluado. Los dominios evaluados por todas o la mayoría de las universidades¹ participantes del estudio fueron los siguientes.

- 1) Lenguaje
- 2) Matemática
- 3) Aspectos no disciplinares, orientados a caracterizar al estudiante en lo relativo al aprendizaje y creencias docentes.

Un aspecto que destaca en los dominios evaluados, es que la mayoría no son específicos para estudiantes de pedagogía. Se sugiere considerar las competencias específicas que tienen que desarrollar los docentes que están relacionadas con el quehacer en el aula. En los siguientes apartados se describen orientaciones basadas en las experiencias de medición de cada dominio analizadas en el estudio FONIDE.

Evaluar competencias de Lenguaje (Lectura y Escritura)

Tanto la comprensión lectora como la producción escrita son aspectos transversales y muy relevantes de evaluar en los futuros docentes. En el caso de Lectura, hay que considerar no repetir lo ya evaluado en la PSU o SIMCE, y más bien enfocarse en la comprensión de textos que sean relevantes en la formación inicial docente.

En Escritura, es relevante que las tareas sean motivantes y situadas, de modo que el estudiante pueda realizar el trabajo en un contexto ecológico y atingente. Un buen referente a tener en consideración son las pruebas que miden la escritura a través de preguntas sobre sus creencias docentes u opiniones sobre alguna situación específica que pueda generarse en contextos de enseñanza. Además, hay que poner atención también a los procesos de construcción de rúbricas de evaluación, así como también los tiempos de corrección, a menudo más lentos que los de pruebas estandarizadas con ítems de selección múltiple.

Evaluar competencia matemática

La evidencia del proyecto FONIDE muestra que las instituciones usaron instrumentos que evaluaban habilidades básicas en matemáticas, competencias matemáticas y conocimientos matemáticos (Giacconi et al., 2019).

Asimismo, es relevante que se diseñen evaluaciones ajustadas a las realidades de cada programa formativo. Por ejemplo, pueden ser constructos distintos los que se pretendan evaluar en pedagogía básica versus aquellos de pedagogía en educación media con mención en matemáticas, dadas las necesidades específicas de cada programa.

1 En el estudio se encontraron instrumentos que evaluaban otras disciplinas, como por ejemplo inglés o música, estos instrumentos no fueron analizados por no ser temas comunes para las universidades de la muestra.

Por último, en este campo es menester tener presente un enfoque de género. La evidencia del proyecto FONIDE muestra que en ciertos casos hubo diferencias estadísticamente significativas en el desempeño de matemáticas entre hombres y mujeres. Además, existe evidencia que muestra que los profesores son modelos que pueden reproducir o mitigar estereotipos. Por esto, es necesario que las universidades consideren estos aspectos a lo largo de la formación docente.

Evaluar aspectos no disciplinares

La evidencia muestra que se midieron dos conjuntos de variables no disciplinares: 1) aquellas relacionadas con el aprendizaje (ej., metacognición y estrategias de aprendizaje), pertinentes para un conjunto amplio de estudiantes de educación superior, y 2) aquellas relacionadas con la enseñanza y el rol del profesor (ej., disposiciones a la docencia, y creencias sobre la enseñanza y el aprendizaje) que son adecuadas de medir en el caso de los estudiantes de carreras de pedagogía (Giaconi et al., 2019). Estos dominios muestran ser relevantes de medir en el diagnóstico, pero de forma complementaria a los instrumentos mencionados anteriormente. Para la mayoría de estas medidas, es importante tener presente que la deseabilidad social es un fenómeno muy recurrente y problemático, por lo tanto, los equipos constructores deben tener especial cuidado tanto en la construcción de ítems o tareas, como también en la interpretación de los puntajes o niveles de logro asociados.

Asimismo, y considerando que los docentes en ejercicio representan un modelo normativo y social para los estudiantes en el aula, estos instrumentos pueden ser una potencial oportunidad para conocer las visiones o sesgos implícitos en los estudiantes que ingresan a las pedagogías. De esta manera se avanza también en comprender que los profesionales de la educación no sólo deben acreditar un buen desempeño, sino también demostrar competencias para transmitir visiones favorables a la equidad a niños y jóvenes.

V. Referencias

- ❶ American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2018). Estándares para Pruebas Educativas y Psicológicas (M. Lieve, Trans.). Washington, DC: American Educational Research Association. (Original work published 2014)
- ❷ Giaconi, V., Varas, L., Ravest, J., Martin, A., Gómez, G., Queupil, J.P., Díaz, K. (2019) Fortaleciendo la Formación Inicial Docente: experiencia universitaria en la implementación de la Evaluación Diagnóstica Inicial para Pedagogías. Proyecto FONIDE 170009. Santiago: MINEDUC



Av. José Pedro Alessandri 685, Ñuñoa - Santiago, Chile

 /demre.uchile  /demre_uchile  /DEMREuchile  /demre.uchile  www.demre.cl