

Revisión de experiencias de evaluación docente



Comisión Directiva del INEE: Javier Lasida (presidente), Guillermo Fossati y Pablo Caggiani

Directora del Área Técnica: Carmen Haretche

Directora de la Unidad de Evaluación Curricular: Andrea Rajchman

La autora de este documento es Cecilia Emery.

Montevideo, 2023

ISBN: 978-9915-9543-1-8

© Instituto Nacional de Evaluación Educativa (INEEd)

Edificio Los Naranjos, planta alta, Parque Tecnológico del LATU

Av. Italia 6201, Montevideo, Uruguay

(+598) 2604 4649 – 2604 8590

ineed@ineed.edu.uy

www.ineed.edu.uy

Cómo citar: INEE (2023). *Revisión de experiencias de evaluación docente*, Recuperado de <https://www.ineed.edu.uy/images/publicaciones/Revision-de-experiencias-de-evaluacion-docente.pdf>

Este informe trata de niños, adolescentes y adultos mujeres y varones. El uso del masculino genérico obedece a un criterio de economía de lenguaje y procura una lectura más fluida, sin ninguna connotación discriminatoria.

Consideraciones generales

Una educación de calidad resulta indispensable tanto para promover el desarrollo y bienestar de las personas, como para el crecimiento económico y la reducción de las desigualdades sociales. En virtud de ello, los Gobiernos han realizado esfuerzos importantes en las últimas décadas orientados al aumento de la cobertura educativa. La mayoría de los países de América Latina ha alcanzado una cobertura universal en la educación primaria, mientras que el 77% de los niños de 4 a 5 años y el 80% de los adolescentes de 15 a 17 años asisten a un centro educativo¹.

Entre los diversos esfuerzos para lograrla, uno de los principales factores es la calidad de los docentes. Estudios antecedentes han mostrado que los maestros y profesores son el recurso escolar más importante para la promoción del desarrollo de los aprendizajes. Algunos análisis han concluido que las intervenciones que involucran a los docentes son particularmente importantes para la mejora educativa (Ganimian y Murnane, 2016; Kremer, Brannen, y Glennerster, 2013; McEwan, 2015; Popova y Evans, 2016). Adicionalmente, existe evidencia de que el impacto de las buenas prácticas de los docentes trasciende los aprendizajes en la escuela, llegando incluso a influir en otros ámbitos, como son las tasas de egreso, el aumento de la probabilidad de acceder y culminar estudios superiores y los ingresos laborales, entre otros (Cruz-Aguayo, Hincapié, y Rodríguez, 2020).

Es por esto que uno de los factores que subyacen a los logros de algunos sistemas educativos — como los de Finlandia, Japón, Dinamarca, Suiza y República de Corea— sea contar con una política articulada de selección, capacitación y desarrollo profesional de los docentes, lo que hace que la profesión de educador sea socialmente respetada y valorada.

La literatura especializada señala que los sistemas de evaluación deben tener como principal objetivo el desarrollo y potenciación tanto de los conocimientos disciplinares como de las habilidades pedagógicas de los docentes. Sin embargo, en términos de implementación la tarea no parece fácil. Tanto en países desarrollados como en vías de desarrollo, los estudios muestran cómo, en un mismo país (con igual marco regulatorio y administrativo), existen variaciones significativas en la calidad de los maestros. Estas diferencias pueden estar presentes incluso dentro de una misma escuela, en donde las prácticas pedagógicas son heterogéneas incluso entre docentes (Araujo, Carneiro, Cruz-Aguayo, y Schady, 2016; Bruns y Luque, 2015; Chetty, Friedman, y Rockoff, 2014; Hanushek y Rivkin, 2010).

Así, las prácticas de evaluación docente pueden contribuir a identificar variaciones de desempeño entre los maestros y profesores. Adicionalmente, el uso adecuado de sus resultados proporciona información valiosa para el aprovechamiento de fortalezas y la búsqueda de superación de las dificultades en el ejercicio de la profesión. Sin embargo, para el cumplimiento de estos propósitos y lograr que las evaluaciones a los docentes sean realmente aceptadas y utilizadas, se requiere que sean válidas y confiables. Los autores advierten que **la información que surja de ellas debe ser objetiva y estar efectivamente asociada a su quehacer cotidiano, de manera que permita que tanto los docentes como las autoridades educativas comprendan cuáles son las fortalezas y debilidades en la labor.**

Durante décadas se observó la realidad educativa desde el supuesto de que el impacto sobre las posibilidades de éxito académico de las condiciones socioeconómicas y culturales del entorno o los

¹ Datos tomados de la base de datos CIMA del Banco Interamericano de Desarrollo (<https://cima.iadb.org/es>).

estudiantes era tan fuerte que poco se podía hacer desde los centros educativos para contrarrestarlas o revertirlas. El resultado de este enfoque —principalmente desde los cincuenta a los ochenta— señalaba la capacidad explicativa del contexto socioeconómico y cultural sobre los logros de la gestión escolar. Sin embargo, en los últimos treinta años los sistemas educativos han redoblado los esfuerzos hacia el mejoramiento de la calidad educativa y se identificó al “desempeño profesional del maestro” como uno de los factores más influyentes en el salto cualitativo de la gestión escolar (Schulmeyer, 2002).

Este relevamiento constituye un acercamiento general a la temática y se divide en tres secciones. La primera indaga en la conceptualización de la práctica de evaluación a los docentes y su relevancia. La segunda presenta algunos instrumentos de evaluación docente más frecuentemente empleados. Finalmente, la tercera sección reseña algunos aspectos de los procesos de evaluación docente a nivel global: los implementados por 11 sistemas que, de acuerdo con sus logros en pruebas estandarizadas internacionales, ofrecen a sus estudiantes una educación de calidad y cómo son las evaluaciones a los docentes en 13 países de América Latina y el Caribe.

La evaluación docente: concepto y pertinencia

Evaluar es una de las prácticas más extendidas en el ámbito educativo. Las evaluaciones están presentes de manera constante, tanto a nivel de sistema como de centros, docentes y estudiantes. En términos conceptuales amplios, y dado que fenómenos tales como los procesos de aprendizaje o los logros educativos no resultan directamente observacionales, evaluar implica construir oportunidades —bajo la forma de instrumentos— para el relevamiento de evidencias que den cuenta del estado de desarrollo o avance de los procesos objeto de análisis.

Los propios profesores y maestros son visitados y evaluados por sus directores, supervisores o inspectores. Algunas de estas evaluaciones tienen como finalidad orientar al docente para realizar mejor su trabajo; otras sí asignan al docente una calificación, que normalmente tiene importancia para su avance en la carrera. A veces una misma evaluación busca cumplir con ambos propósitos. Estas evaluaciones también se realizan a través de diversos procedimientos e instrumentos: observación del docente trabajando en clase con sus alumnos, una entrevista al docente, revisión de la planificación del curso, etc. [...] Los docentes también son evaluados en muchos países para acceder a su plaza. En algunos países esto se hace a través de un conjunto de pruebas teóricas y prácticas que se denominan “concurso”. Como resultado de estas pruebas suelen ocurrir dos cosas distintas: en primer lugar, se distingue entre los docentes que han logrado los requisitos mínimos para acceder a la plaza y aquellos que no lo han logrado. En segundo lugar, se ordena a los candidatos en función del resultado obtenido, de modo que los mejor evaluados tendrán prioridad a la hora de elegir plaza. En otros países, esta labor de selección es realizada directamente por los directores de los centros educativos y, a veces, por agencias especializadas en evaluación y certificación de docentes. Los instrumentos de evaluación pueden ser muy diferentes: una prueba de conocimientos sobre la disciplina que el docente va a enseñar, la observación del docente dictando una clase o una entrevista (Ravela, 2006, p. 18).

Si bien todos los tipos de evaluación contribuyen a la mejora del sistema educativo, para el mejoramiento de su calidad es necesario no solo realizar buenas pruebas estandarizadas de logros del sistema, sino también mejorar las evaluaciones docentes; el modo en que estos evalúan a los alumnos en las aulas; las evaluaciones de selección de directores e inspectores, y las de planes, programas e innovaciones, entre otros aspectos. Con independencia del objeto a ser analizado, cabe señalar que las evaluaciones educativas se distinguen, principalmente, por cuáles sean los propósitos o fines que persiguen (Ravela, 2006). Según esto, se identifican seis tipos de evaluación que se exponen en la tabla 1.

Tabla 1

CARACTERÍSTICAS DE LAS EVALUACIONES SEGÚN SUS PROPÓSITOS Y FINES

Propósitos y fines de la evaluación	Principales características
Acreditación o certificación	Se proponen dejar constancia o reconocer que un individuo o institución posee ciertas características o cualidades.
Ordenamiento para la selección	Se orientan a ordenar candidatos a ocupar cargos vacantes (docentes, de dirección o de supervisión) en función de su capacidad para la tarea. Suelen tener poca visibilidad, pero son de enorme importancia dado que determinan la calidad de los profesionales que tendrán responsabilidades en el sistema educativo.
Tomas de decisiones “blandas” o decisiones de mejora	Se las conoce como “evaluaciones formativas” y tienen por principal propósito servir de base para emprender acciones de mejora de aquello que ha sido evaluado. En la evaluación docente su finalidad es brindar orientación profesional para el mejor desarrollo de la enseñanza y del relacionamiento con los alumnos. En todos estos casos estas evaluaciones no tienen una consecuencia formal específica para el evaluado.
Tomas de decisiones “duras”	Estas evaluaciones están dirigidas a tomar decisiones que tienen consecuencias importantes y directas para un programa o institución.
Establecimiento de incentivos	Son evaluaciones que proponen incentivos para la mejora del desempeño de individuos e instituciones En la mayoría de los sistemas educativos los incrementos salariales dependen de la antigüedad o el acceso a cargos superiores, pero no hay premios económicos para los buenos docentes, por lo que faltan estímulos para un mejor desempeño y permanecer en el aula. Por ello, se establecen incentivos económicos vinculados a aspectos como los esfuerzos de capacitación académica, el cumplimiento de obligaciones en materia de puntualidad y asiduidad en la asistencia, etc.
Rendición de cuentas y responsabilización	Estas evaluaciones —también denominada “responsabilidad por los resultados” — se proponen “rendir cuentas” de lo que se ha logrado. En este contexto surgen las evaluaciones estandarizadas, tanto a nivel país como a nivel internacional, para hacer “socialmente visibles” los resultados de la labor educativa.

Fuente: elaboración propia a partir de Ravela (2006).

Nota: se marcan con fondo verde aquellas que aplican a las evaluaciones de docentes.

Instrumentos de evaluación docente

Actualmente las discusiones sobre políticas efectivas de formación docente, así como sobre medidas de promoción para la retención de los docentes en las aulas, se ubican en el centro del debate educativo. Una de las principales es la evaluación docente. En tanto fenómeno relativamente reciente, su investigación resulta escasa, incipiente y, principalmente, concentrada en casos particulares (Cruz-Aguayo et al., 2020). Como se desarrolla más adelante, son varios los países que han llevado a cabo esfuerzos significativos en materia de tiempo y recursos a fin de mejorar las condiciones de la carrera docente.

Una primera inquietud surge al analizar **cómo podría la evaluación contribuir a los propósitos de atraer, formar y retener en las aulas a los mejores docentes para la mejora de los aprendizajes y logros de los estudiantes**. Los autores coinciden en la importancia del diseño e implementación de evaluaciones de desempeño docente a partir de instrumentos que otorguen información válida y confiable.

En primer término, se estima que un instrumento de evaluación docente proporciona información válida cuando sus resultados permiten identificar con claridad a los educadores que tienen un impacto positivo en el aprendizaje de los estudiantes. Adicionalmente, un instrumento o un conjunto de instrumentos otorga información confiable cuando los resultados obtenidos de cada docente “reflejan su desempeño típico en clase y no dependen del día en particular en que la información fue recolectada o de la persona que estuvo a cargo de esa recolección” (Cruz-Aguayo et al., 2020, p. 11).

Las evaluaciones válidas y confiables resultan de particular importancia para que las autoridades educativas puedan conocer cuáles son las fortalezas y debilidades del plantel docente y así impulsar acciones que las potencien o las superen. Además, de la validez y confiabilidad de las evaluaciones depende, en parte, que el cuerpo docente acepte la evaluación, confíe en sus resultados y actúe en consecuencia. Por eso, los estándares internacionales sobre evaluaciones educativas (AERA, APA, y NCME, 2014) señalan que es clave producir información y análisis de resultados basados en los principios de confiabilidad y validez.

Estas evaluaciones requieren, dado la complejidad del quehacer de los docentes, de una **combinación de distintos instrumentos articulados en sistemas integrales de evaluación docente**. A continuación, se presentan algunos de los instrumentos más comunes junto a algunas consideraciones sobre su efectividad según las investigaciones disponibles.

Pruebas estandarizadas de logro a estudiantes

Si se parte del supuesto de que el objetivo final de una adecuada implementación de la evaluación docente es la mejora de los aprendizajes de los estudiantes, resulta esperable que diversos sistemas educativos utilicen instrumentos basados en pruebas estandarizadas a los alumnos como elemento constitutivo de la evaluación docente:

de hecho, el 97% de los docentes incluidos en la muestra del Estudio Internacional sobre Enseñanza y Aprendizaje (TALIS, por sus siglas en inglés) de 2013 reporta que alguna medida de desempeño de

sus estudiantes es utilizada como parte de las evaluaciones docentes que se llevan a cabo en su escuela” (Cruz-Aguayo et al., 2020, p. 13).

Entre los instrumentos más comunes que utilizan este tipo de información se observan las **medidas de eficiencia docente de valor agregado** y los **percentiles de progreso académico del estudiante**. En el primer caso, se utilizan los resultados previos de pruebas cognitivas aplicadas a los alumnos para estimar, controlando por un conjunto de variables contextuales, el resultado que se espera de cada uno de ellos. Así, la diferencia entre el valor esperado en la prueba y el que efectivamente obtuvo cada estudiante se agrega para todos los estudiantes a cargo de cada docente, de lo que resulta el valor agregado del docente. En el segundo caso, las medidas de crecimiento en percentil comparan el aumento del logro de un estudiante con el de otros que en años anteriores tuvieron medidas de desempeño similares. A partir de ello, se estima una medida de efectividad del docente basada en el crecimiento en el percentil promedio de sus alumnos (Cruz-Aguayo et al., 2020).

La inclusión de pruebas estandarizadas a los estudiantes como instrumento de evaluación docente se funda en la consideración de, principalmente, tres aspectos que se estiman positivos:

1. estas pruebas son efectivas en la medición del aprendizaje de los estudiantes —uno de los objetivos más importantes de la tarea docente—, aunque no capturen todos los objetivos que se propone la institución educativa;
2. la literatura especializada observa que las medidas de valor agregado de los docentes son buenos predictores de medidas de valor agregado en el futuro (Araujo et al., 2016; Kane, McCaffrey, Miller, y Staiger, 2013), y
3. se ha observado que las medidas de valor agregado docente guardan correlación con el desarrollo individual de sus estudiantes a largo plazo, como el acceso a la educación superior, el nivel salarial alcanzado en la adultez o el embarazo adolescente (Chetty et al., 2014).

A pesar de las fortalezas que esta estrategia parece tener, se han identificado importantes limitaciones en sistemas educativos que la utilizan con base de la evaluación docente (Haretche, 2011):

1. algunas de ellas se asocian a aspectos de la población medida (la falta de aleatoriedad en la asignación de estudiantes entre grupos y escuelas, el bajo número de estudiantes que podría tener un grupo, la volatilidad por el cambio de composición grupal de un año a otro, entre otros);
2. otras limitaciones se asocian a dificultades de discriminación (cómo distinguir el efecto del docente actual de la influencia o efecto de los anteriores, o cómo identificar la contribución de un docente en particular), y
3. considerar los resultados de los estudiantes como parte de las evaluaciones docentes plantea dilemas en torno al concepto de buena enseñanza: “es posible que un docente con un buen valor agregado haya sido bueno no en enseñar el currículum sino en preparar a los alumnos para responder pruebas estandarizadas. También es posible que quien haya enseñado un currículum alineado con la prueba obtenga mejor valor agregado que otro que enseñó un currículum más rico e incluye materiales que no son evaluados por la prueba” (Haretche, 2011, p. 14).

Medidas de objetivos de aprendizaje

Este instrumento de evaluación se compone de los objetivos de aprendizaje que los docentes establecen al inicio de un año lectivo respecto de los logros académicos que esperan que sus alumnos obtengan. El superior jerárquico del docente supervisa y aprueba estos objetivos y, una vez culminado el año escolar, evalúa junto a él cuál es el progreso efectivamente alcanzado por los alumnos durante el año de trabajo. En función del progreso establecido, el docente recibe una calificación de su desempeño (Cruz-Aguayo et al., 2020).

Se considera una ventaja de esta estrategia, en comparación con las pruebas estandarizadas a los estudiantes, que puede aplicarse a todos los docentes del sistema, con independencia de la asignatura o el grado en que cumplen funciones. Sin embargo, su nivel de estandarización y objetividad es bajo, dado que, en definitiva, la medición queda sujeta a la autonomía de cada docente en su definición y a la rigurosidad de los superiores al evaluar su cumplimiento. Aunque la literatura especializada indica que la implementación de este instrumento es subjetiva, insume mucho tiempo de docentes y superiores y es débil la correlación que guarda con los resultados de los estudiantes en pruebas estandarizadas, igualmente es uno de los dispositivos más utilizados en gran parte de los sistemas educativos y es significativo el peso que se le asigna dentro de la calificación final que se les da a los docentes (Lachlan-Haché, Matlach, Guiden, y Castro, 2015).

Observación estandarizada de aula

La observación del trabajo docente en el aula es uno de los instrumentos más antiguos de todos: “incluso antes de la implementación de sistemas de evaluación con múltiples instrumentos y con distintas consecuencias en todo el mundo, los docentes siempre han sido observados de una forma u otra por superiores jerárquicos” (Cruz-Aguayo et al., 2020, p. 16). Sin embargo, en los últimos años ha cambiado la estructura con la que se llevan a cabo estas observaciones. Un ejemplo de esto son las rúbricas específicamente diseñadas para la observación. Estas pautas requieren definiciones previas sobre lo que se espera que un buen docente deba saber y hacer, tomando en cuenta su contexto, el tiempo y los recursos disponibles para realizarla.

Los autores señalan la importancia de ciertas definiciones en la construcción del instrumento, como el número de observaciones que se hará a cada docente, su tiempo de duración, si serán coordinadas o sin aviso, quién será el encargado de efectuarlas y el tipo de entrenamiento que se le dará a cada observador o codificador, entre otros aspectos (Donaldson y Papay, 2015; Steinberg y Donaldson, 2016). Ello impacta tanto en la validez y la confiabilidad de la información recolectada por medio de este instrumento (Cohen y Goldhaber, 2016) como en el tiempo y los costos de su implementación.

A diferencia de las medidas a partir de los objetivos de aprendizaje, la observación en aula es un instrumento más aceptado por los docentes. Aunque estudios antecedentes muestran que estas medidas están correlacionadas con el desempeño estudiantil en pruebas estandarizadas (Rockoff y Speroni., 2010), aún no existe suficiente evidencia rigurosa sobre sus resultados (Cohen y Goldhaber, 2016).

Portafolio docente

El portafolio docente es un instrumento posterior al de observación de aula e implica la compilación de materiales por parte del docente que registra y da cuenta del proceso de enseñanza que lleva adelante durante el año lectivo. Este instrumento posibilita una aproximación global al desempeño mediante la incorporación de las planificaciones, evaluaciones, reflexiones sobre la propia práctica y cómo esta se adapta a las necesidades y procesos de los alumnos.

En algunos países, como en Chile: “el portafolio está compuesto de tres módulos: i) descripción de la práctica docente y reflexión acerca de su quehacer docente, ii) una clase grabada y iii) descripción del trabajo colaborativo entre docentes y reflexión al respecto. Se trata de un instrumento que, al parecer, resulta particularmente útil para la autorreflexión” (Cruz-Aguayo et al., 2020, p. 18). Sin embargo, son pocas las investigaciones que han evaluado su efectividad.

Encuesta a estudiantes y otros actores

Recientemente se han considerado las encuestas a estudiantes como un instrumento complementario en las evaluaciones docentes. Bien diseñadas, las encuestas pueden tener potencial en términos de confiabilidad, dado que acopian información proveniente de varios actores y son mucho menos costosas que otros instrumentos de evaluación.

Si bien no están exentas de algunos problemas de sesgo por deseabilidad social, estudios señalan que son un instrumento válido (Benton y Cashin, 2011), al igual que otros tales como las entrevistas a directores o inspectores y entrevistas o encuestas a las familias.

Pruebas estandarizadas a docentes

Las pruebas estandarizadas aplicadas a docentes son una herramienta que, en función de su diseño, permiten identificar su conocimiento pedagógico o disciplinar:

el uso de este tipo de instrumento es común dentro de los sistemas educativos de América Latina y el Caribe, particularmente en las evaluaciones asociadas al ingreso a la carrera y en aquellas requeridas para ascender. Es un instrumento común también en los sistemas educativos dentro de Estados Unidos donde es un requisito para la certificación docente (Cruz-Aguayo et al., 2020, p. 19).

Una ventaja de estas pruebas es su implementación, más sencilla y menos costosa que la de los otros instrumentos. Asimismo, es el único tipo de evaluación para el cual hay evidencia causal positiva que muestra cómo su implementación puede efectivamente mejorar el desempeño de los estudiantes (Brutti y Sánchez, 2017; Estrada, 2019; Ome, 2012).

Cómo se evalúa a los docentes en los sistemas educativos: relevamiento de experiencias

La evaluación docente en 11 sistemas que obtienen de los mejores logros en las pruebas internacionales (OCDE)

La multiplicidad de diseños de evaluación docente observados en los distintos países evidencia que no existe una única respuesta de los sistemas a este desafío de la calidad educativa. Adicionalmente, cada caso cuenta con un contexto propio y sus formas de evaluación docente expresan las posibilidades, objetivos y prioridades particulares que cada nación aspira a lograr mediante su implementación.

Cabe señalar que las iniciativas de evaluación docente impulsadas por países que obtienen buenos resultados en las pruebas estandarizadas internacionales no necesariamente son trasladables a sistemas educativos que, como los de América Latina y el Caribe, presentan características diversas y propias, como restricciones presupuestales, distinta experiencia en materia de implementación o particularidades culturales y de su cuerpo docente, entre otras. Sin embargo, el análisis de sus particularidades puede ilustrar y brindar aportes relativos al diseño, implementación y consecuencias que surgen de estos sistemas (Cruz-Aguayo et al., 2020).

El presente relevamiento reseña algunas características generales de los sistemas de evaluación docente de Australia, Ontario (Canadá), Reino Unido, Nueva Zelanda, Singapur, República de Corea y Estonia. Estos países y provincia fueron seleccionados en virtud del desempeño obtenido en pruebas estandarizadas internacionales o porque la literatura especializada señala que cuentan con aspectos novedosos en sus diseños². Adicionalmente se consideran los sistemas escolares públicos de Denver y Columbia, junto a los estados de Nuevo México y Tennessee (Estados Unidos). Estos se incorporaron a efectos de ilustrar sistemas de evaluación docente que combinan una mayor cantidad de instrumentos en simultáneo, incluso medidas de valor agregado.

La información disponible sobre los sistemas de evaluación de cada país y provincia es heterogénea y fragmentada. En términos muy generales, Ontario se propone como uno de los objetivos centrales de la evaluación el desarrollo profesional a través de una prueba diagnóstica que, posteriormente, otorga al docente una retroalimentación de sus prácticas en el aula. En el Reino Unido, donde la responsabilidad de la evaluación docente recae sobre cuerpos colegiados a nivel de cada centro educativo, el énfasis está puesto, principalmente, en la obtención de niveles de excelencia y la satisfacción laboral del cuerpo docente.

Australia distingue cuatro estadios en la carrera docente (graduado, competente, de alto nivel y líder), cuyo ascenso depende del éxito en las evaluaciones, que son de carácter voluntario y pueden realizarse de manera interna en los centros educativos o a través de agencias externas en el caso de querer certificar los dos niveles superiores (de alto nivel y líder). El sistema de evaluación australiano

² Aunque Finlandia se posiciona entre los países con mejores desempeños en materia de aprendizajes, no cuenta con un sistema de evaluación docente ni tampoco con un sistema de inspección escolar (que se suprimió en la década de 1990). Por esto, los factores que explican el buen desempeño profesional docente y la confianza que la sociedad deposita en los educadores se deben, entre otros aspectos culturales, a la formación profesional de base y no a la posterior evaluación o supervisión de su desempeño.

define para cada uno de ellos estándares de desempeño con relación a siete aspectos: conocer a los estudiantes y cómo estos aprenden; conocer los contenidos y cómo enseñarlos; planificar e implementar una enseñanza y un aprendizaje eficaces; crear y mantener entornos de aprendizaje seguros y de apoyo; evaluar, retroalimentar e informar sobre el aprendizaje de los estudiantes; desarrollo profesional, y compromiso con colegas, padres/cuidadores y la comunidad (INEE, 2017).

Nueva Zelanda es el más reciente de los casos seleccionados: la implementación de sus evaluaciones docentes data de 2018. Si bien el diseño e instrumentación corresponde a cada centro, la administración de la educación a nivel central debe proporcionar orientaciones generales. Una situación similar se observa en Estonia, donde la legislación obliga a aplicar evaluaciones docentes en cada institución educativa, pero estas no están centralizadas a nivel de sistema.

En Singapur, el *Plan de Desarrollo Profesional del Servicio Educativo* está conformado por tres componentes: el Sistema Mejorado de Manejo del Desempeño, el Sistema de Reconocimiento a través de bonificaciones monetarias y un sistema de carrera con tres opciones distintas (carrera de docente de aula, carrera de líder educativo y carrera de especialista experimentado). Este sistema se basa en competencias que definen aquellos conocimientos, aptitudes y habilidades que cada docente debe demostrar en ciertas etapas de su carrera. Entre sus principales propósitos la evaluación docente de Singapur se orienta a: fomentar el mejoramiento profesional del docente a través de la reflexión sobre sus logros; mejorar la calidad educativa que se ofrece a los alumnos, y vincular el sistema de evaluación docente con las decisiones sobre el salario, las bonificaciones y la progresión en la carrera (Cruz-Aguayo et al., 2020).

En el caso de República de Corea, la evaluación es obligatoria para todos los docentes que cumplan más de dos meses de trabajo y se basa en tres componentes: la evaluación de desempeño (que valora cualidades del trabajo por medio de encuestas a estudiantes y pares y mediante la observación de aula), el sistema de incentivos (que determina los salarios de los docentes de acuerdo con su desempeño) y la evaluación para el desarrollo profesional (que brinda retroalimentación a los docentes) (Cruz-Aguayo et al., 2020).

Estados Unidos, por otra parte, implementa múltiples instrumentos complementarios a los efectos de caracterizar el desempeño de maestros y profesores, lo que lo hace un sistema de evaluación docente costoso (Stecher et al., 2018). En muchos casos los resultados tienen consecuencias en términos laborales o salariales.

Una primera aproximación a las características de estos sistemas de evaluación permite señalar que, aunque con diversas estructuras, **todos los países y la provincia considerados en el análisis presentan formulaciones de estándares de desempeño docente propios. Esto es, cuentan con descripciones explícitas de lo socialmente deseable con relación al rol de maestros y profesores: qué debe saber y saber hacer un buen docente.** Estas orientaciones constituyen un marco fundamental, tanto para el desarrollo de la carrera, como de la mejora continua del quehacer cotidiano. Asimismo, también resultan un referente indispensable de los sistemas de evaluación docente de cada país.

De los relevados, el sistema de estándares más detallado es el de Singapur, donde cada etapa de la carrera docente se corresponde con un conjunto propio de competencias debidamente especificadas y evaluadas. En Ontario, se establecen 16 competencias, que incluyen conceptos tales como el compromiso del docente con el aprendizaje, el bienestar de los estudiantes y el conocimiento

profesional. **Al igual que ocurre en Singapur, en Ontario el nivel de exigencia varía según el momento de la carrera profesional en la que se encuentre cada docente.** Por ejemplo, a los principiantes se les evalúa únicamente en ocho de las 16 competencias y de su adecuado cumplimiento depende su avance al siguiente nivel (Cruz-Aguayo et al., 2020).

En el caso de Nueva Zelanda, si bien presenta marcos de desempeño a nivel central, estos varían de acuerdo con el contexto para ciertas escuelas o poblaciones específicas (por ejemplo, centros educativos donde los cursos son dictados en maorí). También Estonia cuenta con estándares de desempeño docente, aunque su aplicación no es obligatoria en la evaluación:

La definición de estos estándares y su inclusión dentro de las evaluaciones son importantes, ya que además de guiar a los propios docentes con respecto a lo que se espera de ellos, contribuyen a que los instrumentos, los actores responsables de la evaluación y los resultados que emanan sean claros, imparciales y coincidan con los objetivos de excelencia educativa únicos de cada país (Cruz-Aguayo et al., 2020, p. 35).

Un segundo aspecto a observar en los distintos países es la periodicidad con la que realizan evaluaciones a los docentes, cuál es su obligatoriedad y quién está a cargo de su aplicación. En los sistemas que pertenecen a Estados Unidos estas instancias son obligatorias, anuales e instrumentadas por las autoridades de cada institución educativa. Sin embargo, para la aplicación de algunos instrumentos (como sucede en Nuevo México) participan también evaluadores externos (Cruz-Aguayo et al., 2020). Esta es una cualidad de la aplicación que permite mejorar la imparcialidad y consistencia a los resultados, al tiempo que exige de actores que estén adecuadamente entrenados para la tarea.

Las variaciones en torno a estos aspectos son mayores entre los demás sistemas. En Ontario es obligatoria la evaluación de los docentes al menos una vez cada cinco años —aunque estos deban realizar un plan de aprendizaje anual—, de acuerdo con lo que disponga el director del centro donde trabajan. Tanto en Reino Unido como en Nueva Zelanda las evaluaciones docentes son anuales y es potestad de los directores la definición de la periodicidad, según lo consideren más adecuado para sus docentes y centros.

Singapur y República de Corea tienen evaluaciones docentes anuales y su estructura está mucho más centralizada en una única administración:

En Singapur, aunque los docentes son evaluados por su supervisor directo de manera regular, la evaluación formal la lleva a cabo un panel de supervisores compuesto por el director, el subdirector y los jefes de departamento, quienes los califican de manera colectiva. En República de Corea, los directores son los encargados de asegurar que las directrices impartidas por el Ministerio de Educación en torno a las evaluaciones se cumplan. Finalmente, en Estonia la evaluación interna de la institución debe llevarse a cabo cada tres años, aunque cada año debe elaborarse un reporte de progreso interno (Cruz-Aguayo et al., 2020, p. 36).

Adicionalmente, los docentes de Estonia pueden aplicar de manera voluntaria a evaluaciones que les permiten ascender en la carrera (definidas por los centros educativos), así como también a procesos de certificación (definidos por la asociación de docentes del país).

Sumado a la formulación de estándares de desempeño y a la periodicidad y obligatoriedad de las evaluaciones de los docentes, cabe observar qué instrumentos son los que emplean algunos de los

países que obtienen los mejores logros en las pruebas estandarizadas a nivel internacional. La tabla 2 sintetiza los instrumentos empleados en las evaluaciones docentes de los países, ciudades y regiones administrativas mencionados y sus diversas combinaciones.

Tabla 2
INSTRUMENTOS APLICADOS EN LAS EVALUACIONES DOCENTES DE SISTEMAS EDUCATIVOS SELECCIONADOS POR SUS BUENOS LOGROS EN LAS EVALUACIONES INTERNACIONALES

PAÍS	Observación de clase	Portafolio	Resultados de exámenes a estudiantes	Objetivos de aprendizaje estudiantil	Encuestas a estudiantes o padres	Autoevaluación	Otros
Australia							
Ontario (Canadá)							
Reino Unido							
Nueva Zelanda							
Singapur							
República de Corea							
Estonia							
Denver (EE. UU.)							
Columbia (EE. UU.)							
Nuevo México (EE. UU.)							
Tennessee (EE. UU.)							

Fuente: reelaborado y ampliado a partir de Cruz-Aguayo et al. (2020).

Se observa que en los sistemas seleccionados de Estados Unidos las evaluaciones docentes suelen incluir una multiplicidad de instrumentos, lo que no ocurre en la mitad de los demás sistemas de la Organización para la Cooperación y el Desarrollo Económicos (OCDE) de este relevamiento. Los docentes en Nueva Zelanda, además de ser observados en el aula, deben completar una autoevaluación que dé cuenta de lo que efectivamente están alcanzando con relación a los estándares exigidos en el país. Por otra parte, Singapur y República de Corea articulan en sus evaluaciones docentes el uso de al menos cuatro instrumentos. Para estos dos países, el sistema de evaluación vigente incluye, además de la observación en clase, el análisis de portafolios docentes, la

autoevaluación y los resultados de exámenes estudiantiles (Singapur) y encuestas al alumnado (República de Corea) (Cruz-Aguayo et al., 2020).

Aunque en todos los procesos de evaluación docente se incluyen también reuniones e instancias de intercambio con directores o supervisores, más que instrumentos específicamente considerados, se estiman como oportunidades para la obtención de información la coordinación de actividades de observación o la entrega de resultados finales de la propia evaluación; en virtud de ello no han sido mencionadas como instrumentos en sí mismos.

Más allá de la heterogeneidad de modalidades y herramientas seleccionadas por los distintos países y sistemas, se constata que la observación de aula es el instrumento más comúnmente empleado al momento de evaluar a los docentes. Esta técnica, que permite una aproximación a las interacciones, relaciones y métodos de enseñanza a medida que ocurre en las aulas, sirve como evidencia directa para identificar fortalezas y áreas específicas del proceso de enseñanza que deberían mejorarse.

Aunque estas observaciones pueden realizarse por medio de grabaciones o en vivo, los videos no han sido muy utilizados, aunque se pueden encontrar ejemplos en los sistemas de Estados Unidos y Chile.

Una de las principales limitaciones de esta técnica es que su validez requiere de un alto entrenamiento de los observadores, sumado a los costos que implica el procesamiento de la información a partir de los diversos aspectos a observar.

En Ontario (Canadá) el director de la institución educativa debe realizar una reunión previa a la observación para explicarle al docente los detalles de la instancia y luego una reunión posterior, donde se analizan los resultados, calificaciones generales y se plantean las estrategias a seguir. En el caso del Reino Unido, la reglamentación solo exige la observación de una clase y una entrevista con el docente en donde se definen los objetivos laborales. En Nueva Zelanda, el sistema estipula que se lleven a cabo al menos dos observaciones anuales y se cuenta con sugerencias a nivel central para orientar la evaluación.

En Singapur y República de Corea el instrumento de observación en clase está más detallado y estandarizado. En Singapur las evaluaciones se llevan a cabo al final del año escolar mediante el supervisor y un panel de superiores que hayan trabajado con el docente (como el director, el subdirector o los jefes de departamento), quienes a través de diferentes protocolos evalúan el potencial del docente con los instrumentos diseñados para tal fin. También es interesante el caso de Columbia (EE. UU.), donde al comienzo se exigía que todos los docentes fuesen calificados por observadores externos dos veces al año, pero esta modalidad debió ser abolida en 2016 por los altos costos presupuestales que implicaba (Cruz-Aguayo et al., 2020).

Como se mencionó al inicio del informe (ver tabla 1), tanto la definición de los propósitos de las evaluaciones realizadas como el uso de los resultados obtenidos son fundamentales para la mejora de la calidad educativa en general y del cuerpo docente en particular.

En la tabla 3 se resumen las dimensiones en las que tiene impacto la calificación final obtenida por los docentes en los sistemas de evaluación observados. Las consecuencias más observadas se vinculan al ingreso percibido por la función mediante bonificaciones o incrementos salariales, conjuntamente a posibilidades de ascenso.

Tabla 3

PRINCIPALES CONSECUENCIAS A PARTIR DE LOS RESULTADOS DE LA EVALUACIÓN DOCENTE EN LOS SISTEMAS EDUCATIVOS SELECCIONADOS POR SUS BUENOS LOGROS EN LAS EVALUACIONES INTERNACIONALES

PAÍS	Diagnóstica	Bonificaciones /aumento en salarios	Ascenso	Formación permanente	Permanencia en la función docente
Australia					
Ontario (Canadá)					
Reino Unido					
Nueva Zelanda					
Singapur					
República de Corea					
Estonia					
Denver (EE. UU.)					
Columbia (EE. UU.)					
Nuevo México (EE. UU.)					
Tennessee (EE. UU.)					

Fuente: reelaborado a partir de Cruz-Aguayo et al. (2020).

En Singapur y República de Corea se observa que los resultados se vinculan al desarrollo profesional mediante propuestas de formación permanente o a la estabilidad laboral, lo que indica que los docentes que obtienen bajas calificaciones quedan sujetos a programas de formación continua o incluso a la remoción de su función cuando estos resultados se obtienen de manera reiterada. Solo en Ontario la evaluación cumple una función exclusivamente diagnóstica, lo que permite identificar el alcance de los objetivos de desarrollo profesional de los docentes (Cruz-Aguayo et al., 2020).

En Nueva Zelanda la clasificación final está vinculada a la progresión salarial y a la revisión anual de los objetivos profesionales. En República de Corea, por otra parte, los resultados buscan identificar docentes de alto y bajo rendimiento para brindarles acceso a oportunidades de investigación o a programas de desarrollo profesional. Finalmente, los resultados de los tres sistemas de evaluación de Singapur definen aspectos relacionados con la promoción dentro de la carrera docente,

bonificaciones anuales por desempeño e incluso con la posibilidad de despido para aquellos que no cumplan con los estándares requeridos en la profesión (Cruz-Aguayo et al., 2020).

Más allá de las consecuencias que las evaluaciones tienen para con los docentes, algunos sistemas también las toman en cuenta para la definición de otros aspectos de la política educativa. En el Reino Unido y Estonia, por ejemplo, los resultados de las evaluaciones docentes son uno de los principales insumos a partir de los que se definen las necesidades de desarrollo profesional, la progresión de la carrera docente y los niveles de retribuciones correspondientes a cada uno de los niveles de logro profesional³.

La evaluación docente en 13 países de América Latina y el Caribe

En la actualidad varios países de América Latina y el Caribe realizan esfuerzos para desarrollar sus sistemas de evaluación docente. En esta sección se presentan los casos de Argentina, Bolivia, Chile, Colombia, Costa Rica, El Salvador, Guatemala, México, Paraguay, Perú, República Dominicana y Uruguay, con el afán de observar el rumbo que toma la evaluación de los docentes en la región. Como se ha mencionado anteriormente, la información disponible sobre este aspecto de los sistemas educativos es parcial, fragmentada y no siempre cuenta con el grado de actualización necesario, por lo que se plantean en perspectiva comparada aquellos aspectos para los que hay datos disponibles.

A partir de las posibilidades que ofrece la información recabada, esta sección se articula partiendo de cómo son las evaluaciones docentes en los distintos países en tres momentos de la carrera: la evaluación de aspirantes o de ingreso a la carrera docente (concurso docente); la evaluación de la formación inicial o durante el período de prueba, como se lo llama según el caso, y la evaluación del ejercicio o desempeño profesional. Dado que estas evaluaciones se aplican en distintas etapas de la trayectoria profesional, es esperable que tanto sus objetivos como diseños e implementaciones varíen, no solamente entre países, sino también entre sistemas educativos dentro de un mismo país (como sucede en Argentina, por ejemplo, donde el gobierno de la educación es provincial).

La evaluación de ingreso a la carrera, proceso generalmente conocido como “concurso docente”, se realiza en casi todos los sistemas educativos seleccionados. Esta instancia define si un individuo puede acceder de manera efectiva a una plaza dentro del sistema público. La segunda evaluación establece si los docentes noveles, recientemente iniciados en sus funciones, cuentan efectivamente con las habilidades pedagógicas y disciplinares requeridas para su permanencia en el rol. Esta es la evaluación menos común de todas entre los países seleccionados. En tercer lugar, se consideran las evaluaciones realizadas a educadores que, en el ejercicio de sus funciones, son regidos por los estatutos docentes de cada sistema o país.

³ En Uruguay estos aspectos están contemplados en el [Estatuto del funcionario docente](#) de la Administración Nacional de Educación Pública.

Tabla 4

TIPOS DE EVALUACIONES CONTEMPLADAS EN LOS SISTEMAS DE EVALUACIÓN DOCENTE DE AMÉRICA LATINA Y EL CARIBE Y SU MARCO LEGAL

País	Marco legal	Evaluación		
		Aspirantes/ de ingreso	Formación inicial	Ejercicio profesional/ desempeño
Argentina	Ley n.º 14473, Estatuto del Docente (1958) y otras normas jurisdiccionales derivadas de ellos Evaluación Educativa: la construcción de una política federal 2016-2019	Sí	No	Sí
Bolivia	Ley n.º 1565, Reforma educativa de 1994 Normas generales para la Gestión Educativa. Resolución Ministerial n.º 001/2022	Sí	Sí	Sí
Chile	Reglamento del artículo 70 del decreto con fuerza de ley n.º 1, 1996, Ministerio de Educación			Sí
Colombia	Ley General de Educación n.º 115 (1994) y Ley n.º 715 de 2001 Manual de evaluación de desempeño, Ministerio de Educación Nacional, 2003	Sí	No	Sí
Costa Rica	Estatuto de Servicio Civil	Sí	No	Sí
El Salvador	Acuerdo ejecutivo n.º 15, 1913 Decreto n.º 917 Ley General de Educación de 1996	Sí	Sí	No
Guatemala	Decreto n.º 1485 del Congreso de la República Ley de Dignificación y Catalogación del Magisterio Nacional de fecha 1961	No	No	Sí
México	Ley General de Educación, 1992 Ley General del Servicio Profesional Docente, 2013	Sí	Sí	Sí
Paraguay	Evaluación del desempeño docente, Ministerio de Educación y Cultura, 2011 Resolución Ministerial n.º 83/2022, Ministerio de Educación y Ciencias	No	No	Sí
Perú	Ley de profesorado n.º 24029 y su modificación Ley n.º 28044 Evaluación ordinaria del desempeño docente, Ministerio de Educación, 2020	Sí	Sí	Sí
República Dominicana	Ordenanza I de 1996	Sí	Sí	Sí
Uruguay	Estatuto del funcionario docente de 1988	Sí	No	Sí

Fuente: reelaboración actualizada y ampliada a partir de Schulmeyer (2002).

El relevamiento permite observar que en la mayoría de los países seleccionados de la región se llevan a cabo evaluaciones de ingreso a la carrera docente (concursos), siendo en algunos casos obligatorios para el ingreso a la docencia. En la mayoría de los casos las evaluaciones son de oposición (prueba de conocimiento disciplinar o pedagógico) o presentación de méritos, como principales instrumentos. Sin embargo, no se encontró información detallada sobre la estructura de los concursos para cada caso.

Tabla 5
INSTRUMENTOS APLICADOS EN PAÍSES SELECCIONADOS EN AMÉRICA LATINA Y EL CARIBE CON EVALUACIONES PERIÓDICAS Y OBLIGATORIAS QUE UTILIZAN MÚLTIPLES INSTRUMENTOS

PAÍS	Reporte del director o superior jerárquico	Evaluación de pares	Autoevaluación	Portafolio	Observación en clase	Examen	Encuestas a estudiantes o padres de familia	Otros
Chile								
El Salvador								
México								
Paraguay (anual)								
Perú								
República Dominicana								
Uruguay (anual)								

Fuente: reelaborado a partir de Cruz-Aguayo et al. (2020).

Nota: solamente se listan los países que aplican de manera periódica y obligatoria una multiplicidad de instrumentos, según la información encontrada.

En algunos sistemas la evaluación docente cumple, complementariamente, una función de fiscalización ante dificultades que enfrentan los países en el ámbito de la formación inicial de los docentes o, incluso, del perfil poblacional de quienes aspiran a la profesión en la región (Elacqua et al., 2018). En ese sentido, continuar con los procesos de mejora y de implementación adecuada de esta evaluación podría traer beneficios en la calidad de la fuerza docente en la región. Aunque la evidencia aún es escasa, estudios para Colombia (Brutti y Sánchez, 2017; Ome, 2012) y México (Estrada, 2019) sugieren que estos sistemas de selección que establecen evaluaciones al ingreso de la carrera están teniendo algunos impactos positivos en los países que los implementan. Algunos autores indican que el nivel de motivación también es central para un buen desempeño en el sector público (Banuri, Keefer, y de Walque, 2018).

Adicionalmente, los primeros años de la carrera cumplen un papel central en el desarrollo de las capacidades de los docentes (Ávalos, 2016; Jensen, Sandoval-Hernández, Knoll, y Gonzalez, 2012). Actualmente, la mayoría de los sistemas relevados, incluso Uruguay, contempla en su legislación o

implementa algún tipo de iniciativa de inducción docente⁴. Aunque el grado de instrumentación o continuidad de estos programas no resulta claro en los datos parciales encontrados, se estima que su existencia constituye de por sí un paso importante en atención a la evidencia de que la correcta inserción de los docentes noveles y el apoyo a los recién ingresados en la carrera resulta fundamental para la calidad del ejercicio profesional posterior. La revisión permite observar que, de los sistemas seleccionados, la evaluación de los docentes al finalizar el período de prueba, en los casos que la hubiere, es la menos común de todas.

En la mayoría de los países la normativa establece que los docentes deben ser periódicamente evaluados y de manera obligatoria. La periodicidad más frecuente es la anual, mientras que en pocos casos la asiduidad varía entre tres y cuatro años, dependiendo de los resultados que los docentes obtengan. La gran mayoría de los sistemas en los que la evaluación a los docentes es anual se basan en la aplicación de un único instrumento compuesto por formularios estandarizados o preestablecidos que el director, inspector o superior jerárquico deben completar (Cruz-Aguayo et al., 2020).

El tercer tipo de evaluación mencionada es la que habilita los ascensos en la carrera docente. Estos ascensos horizontales se distinguen de aquellos en los que los docentes ascienden a cargos administrativos que revisten una mayor jerarquía, como son los de dirección, coordinación o inspección (ascensos verticales).

La descripción de los procesos de diseño e implementación de cada instrumento en cada sistema trasciende el objetivo exploratorio del presente relevamiento. Sin embargo, la revisión de distintos autores y fuentes permite identificar algunas cuestiones a tener en cuenta. No necesariamente todo lo que se establece en el ámbito normativo sobre las evaluaciones docentes se cumple o se puede cumplir. En la tabla 5 se observa que son menos los países que han podido implementar efectivamente evaluaciones docentes, aunque sus legislaciones las contemplen. Este fenómeno también se observa en Uruguay: aunque los estatutos de funcionarios docentes prevén la visita de aula y fijan observaciones, los hechos distan de lo que la legislación reglamenta. La situación presenta diferencias por subsistema: en la Encuesta Nacional Docente realizada en 2015 por el Instituto Nacional de Evaluación Educativa (INEEd), el 65% de los maestros de escuelas públicas manifiesta haber recibido visitas de un inspector en 2015, mientras que esta proporción desciende al 14% en secundaria (INEEd, 2017). En secundaria muchos de los docentes manifiestan que nunca fueron evaluados por un inspector de asignatura. Una situación similar se observa en Ecuador: aunque el reglamento menciona que la evaluación docente debe incluir siete instrumentos, en 2017 únicamente se había logrado la implementación de manera censal de uno de ellos, el examen de saberes disciplinares.

⁴ En el caso de Uruguay existe el Proyecto de acompañamiento a docentes y educadores noveles del Uruguay que es llevado adelante por el Consejo de Formación en Educación y el Programa para la Mejora de la Calidad en la Educación Media y en la Formación en Educación (<http://novelesuruguay.cfe.edu.uy/index.php/proyecto-noveles-del-uruguay>).

Tabla 6

CARACTERÍSTICAS DE LAS EVALUACIONES DOCENTES PARA LOS SISTEMAS EDUCATIVOS DE AMÉRICA LATINA Y EL CARIBE SELECCIONADOS

País	Evaluación		
	1. De aspirantes/de ingreso		
	¿Para qué?	¿Qué?	¿Cómo?
Argentina	No se encontró dato	No se encontró dato	No se encontró dato
Bolivia	Selección y diagnóstico para seguimiento de cohorte	Razonamiento verbal y matemático, aptitud para la docencia, lengua originaria: quechua, aimara y guaraní para los Institutos Normales bilingües	Pruebas escritas de selección múltiple y de respuesta única Entrevista que examina el manejo oral de lengua originaria
Colombia	No se encontró dato	No se encontró dato	No se encontró dato
Costa Rica	No se encontró dato	No se encontró dato	No se encontró dato
El Salvador	No se encontró dato	No se encontró dato	No se encontró dato
México	Regulación del ingreso y diagnóstico Para la asignación de plazas docentes	Desarrollo de habilidades básicas: habilidad verbal, habilidad matemática y habilidad para el razonamiento formal	Pruebas escritas de selección múltiple
Paraguay	En el caso de Paraguay se aplica una prueba para diagnosticar la calidad de la formación recibida por los futuros docentes en los Institutos de Formación Docente. Los resultados aportan información a proyectos de autoevaluación institucional de los Institutos de Formación Docente	Matemática, lengua guaraní, lengua castellana, Ciencias Sociales, Ciencias Naturales y Ciencias y Tecnología. Además, se aplican instrumentos exploratorios de desarrollo personal del aspirante	Censo del último curso de formación (tercer año) en todos los Institutos de Formación Docente, públicos y privados. Los examinadores toman una prueba por día, luego las hojas se leen por lector óptico y las de redacción se corrigen manualmente con profesores especializados
Perú	En el caso de Perú esta prueba determina el ingreso a la institución de formación. También determina la necesidad de pasar o no por un ciclo de nivelación	La mayoría evalúa conocimientos generales, razonamiento verbal y matemático	Pruebas escritas. En algunos casos se complementa con una entrevista personal
República Dominicana	Determinar si los aspirantes poseen los requisitos mínimos en términos cognoscitivos de aptitudes y actitudes	Nivel académico Aptitud Actitud Convivencia	Análisis de expedientes. Aplicación de pruebas de rendimiento académico. Aplicación de pruebas de cociente intelectual, intereses y personalidad

País	Evaluación		
	2. De formación inicial		
	¿Para qué?	¿Qué?	¿Cómo?
Bolivia	Seguimiento a los institutos y universidades de la calidad de la formación docente y hacer ajustes en procura de la mejora Promoción y trámite de título en provisión nacional de egresados de Unidades móviles del programa de profesionalización docente	Ejes de formación del Programa de formación inicial: gestión institucional, pedagogía, técnicas de enseñanza y realidad nacional	Pruebas escritas de selección múltiple y de respuesta única Entrevista que examina el manejo oral de lengua originaria Se aplica durante el tercer, cuarto y sexto semestre de ejercicio
El Salvador	Para certificar	Contenidos programáticos	Prueba escrita
México	Para dar seguimiento a los futuros docentes, como retroalimentación a las instituciones de formación docentes	Desarrollo de habilidades básicas: verbal, matemática y para el razonamiento formal Dominio de contenidos específicos de la especialidad	Pruebas escritas de selección múltiple
Perú	Seguimiento y monitoreo de la formación docente Toma de decisiones sobre situaciones que no estén funcionando	Gestión del director y personal de la institución Impacto de la capacitación de los docentes en su práctica Desarrollo de la práctica de los docentes Articulación investigación-práctica	Fichas de registro etnográfico y encuestas
República Dominicana	Determinar si los aspirantes poseen los requisitos mínimos en términos cognoscitivos de aptitudes y actitudes	Nivel académico Aptitud Actitud Convivencia	Análisis de expedientes Aplicación de pruebas de rendimiento académico Aplicación de pruebas de cociente intelectual, intereses y personalidad

País	Evaluación		
	3. Ejercicio profesional/desempeño		
	¿Para qué?	¿Qué?	¿Cómo?
Argentina	Ascensos de jerarquía Aumento en la asignación de horas cátedra	Títulos Antecedentes (capacitación, publicaciones, postgrados) Desempeño	Análisis de documentación Observación de clases.
Bolivia	No se encontró dato	No se encontró dato	No se encontró dato

Chile	Para que los docentes de aula que se desempeñan en establecimientos que dependen de los municipios o de los Servicios Locales de Educación a lo largo del país se fortalezcan en la profesión docente y contribuyan a mejorar la calidad de la educación	Todos los instrumentos se construyen a partir de los dominios, criterios y descriptores señalados en el Marco para la Buena Enseñanza (MBE)	Portafolio: planificación, evaluación y reflexión; clase grabada; y descripción de experiencia de trabajo colaborativo Autoevaluación Entrevista de un evaluador par Informe de referencia de terceros (superiores)
Colombia	Orientar acciones de mejoramiento profesional Identificar qué debe mejorar el educador como persona y profesional Contribuir a la construcción de identidad y definición del concepto de educador Generar una cultura de evaluación basada en la práctica reflexiva Contribuir al mejoramiento de la calidad y la autonomía institucional Facilitar información sistemática y fiable para la toma de decisiones	Cómo se percibe el educador a sí mismo Cómo asume el educador su papel pedagógico Cómo se vincula el educador con los estudiantes Cómo asume el educador el contexto y cultura institucional	Desarrollada en fases Fase preparatoria: la comunidad educativa del plantel hace ajustes a la pauta analítica, sobre el mapa del saber, del saber ser y del saber hacer Fase de obtención de información: aplicación de instrumentos (fuentes personales de integrantes de la comunidad educativa o cuerpos colegiados de la institución que puedan dar testimonio sobre el desempeño del evaluado; documentos, registros y testimonios escritos personales e institucionales, como innovaciones educativas, resultados significativos en los aprendizajes de los estudiantes, desarrollo de proyectos pedagógicos, informes; observaciones de aula, y entrevistas). Fase de evaluación previa, análisis de información cualitativa Fase de evaluación final, calificación cuantitativa y juicio evaluativo
Costa Rica	Para la mejora y contar con retroalimentación	Procedimientos de enseñanza, técnicas, métodos, presentación y objetivos de los programas	Cuestionario y observación de aula (se usa una escala de 1 a 5)
El Salvador	No se encontró dato	No se encontró dato	No se encontró dato
Guatemala	No se encontró dato	No se encontró dato	No se encontró dato
México	Para recompensar a los mejores docentes a través de estímulos	La formación profesional, que comprende dominio del marco	Pruebas escritas de selección múltiple

	económicos Para generar diagnósticos de retroalimentación docente que permitan el diseño de cursos de actualización.	legal genérico, de la reglamentación específica del área o nivel educativo, de los materiales curriculares y de la especialidad	
Paraguay	No se encontró dato	No se encontró dato	No se encontró dato
Perú	No se encontró dato	No se encontró dato	No se encontró dato
República Dominicana	Determinar competencias en el desempeño y para la aplicación de incentivos salariales	Desempeño en el aula, cumplimiento de normas, relaciones con directivos, colegas y estudiantes	Opinión de estudiantes, directores de centros y orientadores
Uruguay	Para calificar el desempeño del docente La calificación incide en el escalafón	Capacidad técnico-pedagógica Conducción del proceso de enseñanza y aprendizaje Orientación dada al curso, su planificación y desarrollo Clima de trabajo, cooperación e iniciativa Respeto al alumno y promoción de su capacidad de autodeterminación Posibilidades de desarrollo del trabajo creativo	Observación de aula

Fuente: reelaboración a partir de Schulmeyer (2002).

Nota: para cada tipo de evaluación docente se presentan las características únicamente de los países que las prevén en sus sistemas educativos.

Se observa que en las evaluaciones de ingreso a la carrera docente y de formación inicial la mayoría de los países aplican pruebas escritas como principal o único instrumento, mientras que para evaluar el desempeño o ejercicio profesional de los docentes no solamente se combina una mayor diversidad de instrumentos, sino que el más frecuentemente implementado es la observación de aula.

Tabla 7
SISTEMAS DE EVALUACIÓN DOCENTE DE AMÉRICA LATINA Y EL CARIBE SELECCIONADOS QUE CUENTAN CON ESTÁNDARES DE DESEMPEÑO

PAÍS	ESTÁNDARES DE DESEMPEÑO DOCENTE
Argentina	No tiene
Bolivia	No tiene
Chile	Tiene
Colombia	Ciertos estándares definidos para implementación de evaluaciones de desempeño
Costa Rica	No tiene

El Salvador	No tiene
Guatemala	Tiene
México	Tiene
Paraguay	Ciertos estándares definidos para implementación de evaluaciones de desempeño
Perú	Tiene
República Dominicana	Tiene
Uruguay	No tiene

Fuente: reelaboración a partir de Schulmeyer (2002).

La formulación de estándares de desempeño docente en cada país constituye una definición fundamental y necesaria para la efectividad de los esfuerzos de evaluación docente. Estos son una herramienta común a los tres tipos de evaluaciones planteadas, aunque lo que se espera de los docentes puede variar según el contexto; es por esto que igualmente se deben definir indicadores específicos por nivel, modalidad, idioma de enseñanza, entre otros.

Adicionalmente, en la mayoría de los países relevados, tanto los que obtienen buenos logros en las pruebas internacionales como los pertenecientes a América Latina y el Caribe, uno de los principales instrumentos de los sistemas de evaluación docente es la observación de aula. Para cada caso varía el grado de información disponible sobre las especificaciones y las características de la pauta empleada.

En el entendido de que los sistemas de evaluación docente deben ser válidos, confiables y transparentes, lo que implica que todos los actores tengan claro qué es lo que un docente debe saber y debe saber hacer, la formación inicial y la formación permanente y en servicio deben ser diseñadas para asegurar que los maestros y profesores en ejercicio logren efectivamente el cumplimiento de los estándares socialmente definidos en su desempeño diario. Se debe perseguir, de manera articulada, un diseño e implementación de las evaluaciones que permita identificar si un docente en ejercicio cumple los estándares de desempeño de manera adecuada. Finalmente, cabe señalar la importancia de que los docentes conozcan lo que el sistema espera de ellos, las diferencias entre estas expectativas de logro y sus actuaciones profesionales, y los recursos a disposición para la mejora continua.

A partir del presente relevamiento exploratorio se evidencia la relevancia para un sistema de evaluación docente de avanzar en la formulación tanto de estándares o niveles de desempeño docente como de instrumentos que, alineados a estos, proporcionen evidencia confiable sobre las fortalezas y necesidades del cuerpo docente.

En el caso de Uruguay, donde la evaluación de los docentes contempla únicamente la observación de aula y reportes de parte de los superiores, avanzar en la definición de un marco de referencia que defina qué se espera de los educadores en cada nivel o etapa del desarrollo profesional podría incluso contribuir a definir mejor aquellos aspectos a ser ponderados en las visitas del cuerpo directivo o inspectivo, las evaluaciones de ingreso (concursos), etc.

A su vez, en el contexto actual de la transformación educativa que se está llevando a cabo en este momento en Uruguay, que incluye modificaciones en los planes y programas de formación docente, resulta pertinente la revisión de los procesos de evaluación de docente en lo que refiere a sus finalidades, momentos e instrumentos, en consonancia con las definiciones de estándares de desempeño.

Síntesis de aspectos principales

1. La gran mayoría de los sistemas analizados cuenta con evaluaciones docentes, muchas de las cuales combinan más de un instrumento. Los países (o provincias y estados según el caso) se distinguen entre los que las orientan teniendo como referencia la definición de estándares de desempeño y los que no.
2. Todos los casos observados de sistemas educativos con buenos logros integran en la evaluación docente la observación de aula como una de las principales herramientas. Los autores señalan que de este instrumento debe cuidarse especialmente su validez y confiabilidad. En virtud de ello existen variantes entre observadores, número de observaciones y frecuencias, entre otros.
3. El uso de instrumentos basados en resultados obtenidos por los propios estudiantes en exámenes estandarizados es una herramienta mucho más común en los sistemas educativos de Estados Unidos que en el resto de los casos observados, lo que se explica por la existencia de financiamientos específicos para tales efectos en ese país. Los autores señalan que este tipo de herramientas pueden ser adecuadas para identificar la efectividad de los docentes en lo que hace a los logros obtenidos por los estudiantes en materia disciplinar. Sin embargo, también se han identificado limitaciones importantes en su uso. Es un instrumento que exige inversiones de gran porte; actualmente no se observan en los países de América Latina y el Caribe.
4. En los sistemas seleccionados de América Latina y el Caribe las evaluaciones de ingreso a la carrera docente (concurso) y de desempeño profesional son las más frecuentes. Una proporción menor de países prevé evaluaciones de formación inicial o durante el período de prueba.
5. Se ha observado que los sistemas de evaluación docente cumplen, complementariamente a los objetivos establecidos para con el docente evaluado, funciones de monitoreo, formulación de política y retroalimentación de las iniciativas de formación docente de varios sistemas.
6. Las pruebas estandarizadas de evaluación docente para identificar su conocimiento pedagógico o disciplinar son el único tipo de evaluación para el cual hay evidencia causal positiva que muestra cómo su implementación puede efectivamente mejorar el desempeño de los estudiantes.
7. Para las evaluaciones de ingreso a la carrera docente y de formación inicial en América Latina y el Caribe el instrumento más empleado son las pruebas escritas. Sin embargo, para evaluar el desempeño o ejercicio profesional de los docentes los países combinan una mayor diversidad de instrumentos y el más frecuentemente implementado es la observación de aula.
8. En algunos sistemas de América Latina y el Caribe se observa una distancia entre las disposiciones establecidas en los marcos regulatorios y las evaluaciones docentes que efectivamente se llevan a cabo.

9. Tan importante como la evaluación docente es el uso que se haga de sus resultados. En la mayoría de los sistemas observados derivan de una misma evaluación tanto las consecuencias formativas como aquellas relacionadas con incentivos y ascensos.
10. Todo sistema de evaluación docente requiere de la existencia de estándares sobre lo que se espera de un docente tanto en términos de conocimientos disciplinares como pedagógicos. Asimismo, debe estar clara su vinculación con el desarrollo de la carrera docente.

Bibliografía

- AERA, APA, y NCME. (2014). *Standards for educational and psychological testing*. Washington D.C.: American Educational Research Association.
- Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., y Schady, N. (2016). Teacher Quality and Learning Outcomes in Kindergarten. *The Quarterly Journal of Economics*, 131(3), 1415–1453.
- Ávalos, B. (2016). Learning from Research on Beginning Teachers. En J. Loughran & M. L. Hamilton (Eds.), *International Handbook of Teacher Education* (pp. 487-522). Nueva York: Springer.
- Banuri, S., Keefer, P., y de Walque, D. (2018). *Love the Job... or the Patient? Task vs. Mission-Based Motivations in Health Care* (N.º 964). <https://doi.org/10.18235/0001496>
- Benton, S. L., y Cashin, W. E. (2011). *Student Ratings of Teaching: A Summary of Research and Literature* (N.º 50). Recuperado de https://www.ideaedu.org/idea_papers/student-ratings-of-teaching-a-summary-of-research-and-literature/
- Bruns, B., y Luque, J. (2015). *Great Teachers: How to Raise Student Learning in Latin America and the Caribbean*. Recuperado de <https://openknowledge.worldbank.org/handle/10986/20488>
- Brutti, Z., y Sánchez, F. (2017). *Does Better Teacher Selection Lead to Better Students? Evidence from a Large-Scale Reform in Colombia* (N.º 11). Recuperado de <https://repositorio.uniandes.edu.co/handle/1992/8694>
- Chetty, R., Friedman, J. N., y Rockoff, J. E. (2014). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review*, 104(9), 2593–2632. <https://doi.org/10.1257/aer.104.9.2593>
- Cohen, J., y Goldhaber, D. (2016). Building a More Complete Understanding of Teacher Evaluation Using Classroom Observations. *Educational Researcher*, 45(6), 378–387. <https://doi.org/10.3102/0013189X166594>
- Cruz-Aguayo, Y., Hincapié, D., y Rodríguez, C. (2020). *Profesores a prueba: Claves para una evaluación docente exitosa*. <https://doi.org/10.18235/0002149>
- Donaldson, M. L., y Papay, J. P. (2015). Teacher evaluation for accountability and development. En H. F. Ladd & M. E. Goertz (Eds.), *Handbook of Research in Education Finance and Policy* (pp. 174–193). Nueva York: Routledge.
- Elacqua, G., Hincapie, D., Vegas, E., Alfonso, M., Montalva, V., y Paredes, D. (2018). *Profesión profesor en América Latina: ¿ Por qué se perdió el prestigio docente y cómo recuperarlo?* Inter-American Development Bank.
- Estrada, R. (2019). Rules versus Discretion in Public Service: Teacher Hiring in Mexico. *Journal of Labor Economics*, 37(2), 545-579.
- Ganimian, A. J., y Murnane, R. J. (2016). Improving education in developing countries: Lessons from rigorous impact evaluations. *Review of Educational Research*, 86(3), 719–755. Recuperado de <https://journals.sagepub.com/doi/10.3102/0034654315627499>
- Hanushek, E. A., y Rivkin, S. G. (2010). Generalizations about Using Value-Added Measures of Teacher Quality. *American Economic Review*, 100(2), 267–271.
- Haretche, C. (2011). *Valor agregado en educación. Qué es y cómo se usa* (pp. 1-15). pp. 1-15. Recuperado de <https://www.thedialogue.org/analysis/boletin-25-valor-agregado-en-educacion-que-es-y-como-se-usa/>
- INEE. (2017). *Evaluación del desempeño de docentes, directivos y supervisores en educación básica y media superior de México. Análisis y evaluación de su implementación 2015-2016. Informe final*. Recuperado de <https://www.inee.edu.mx/wp-content/uploads/2019/01/P1F209.pdf>
- Jensen, B., Sandoval-Hernández, A., Knoll, S., y Gonzalez, E. J. (2012). *The Experience of New Teachers: Results from TALIS 2008*. Recuperado de <http://dx.doi.org/10.1787/9789264120952-en>
- Kane, T. J., McCaffrey, D. F., Miller, T., y Staiger, D. O. (2013). *Have We Identified Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment* (pp. 1-49). pp. 1-49. Recuperado de <https://eric.ed.gov/?id=ED540959>
- Kremer, M., Brannen, C., y Glennerster, R. (2013). The challenge of education and learning in the developing world. *Science*, 340(6130), 297–300.
- Lachlan-Haché, L., Matlach, L., Guiden, A., y Castro, M. (2015). *What We Know About SLOs: An*

- Annotated Bibliography of Research on and Evaluations of Student Learning Objectives* (pp. 1-43). pp. 1-43. Recuperado de <https://www.air.org/resource/what-we-know-about-slos-annotated-bibliography-research-and-evaluations-student-learning>
- McEwan, P. J. (2015). Improving Learning in Primary Schools of Developing Countries: A Meta-Analysis of Randomized Experiments. *Review of Educational Research*, 85(3), 353–394. <https://doi.org/10.3102/0034654314553127>
- Ome, A. (2012). *The Effects of Meritocracy for Teachers in Colombia* (pp. 1-51). pp. 1-51. Bogotá: Fedesarrollo.
- Popova, D., y Evans, A. (2016). What Really Works to Improve Learning in Developing Countries? An Analysis of Divergent Findings in Systematic Reviews. *World Bank Research Observer*, 31(2), 242-270. Recuperado de <https://openknowledge.worldbank.org/handle/10986/29308>
- Ravela, P. (2006). *Para comprender las evaluaciones educativas* (1.ª ed.). Santiago de Chile: PREAL.
- Rockoff, J. E., y Speroni., C. (2010). Subjective and Objective Evaluations of Teacher Effectiveness. *The American Economic Review*, 100(2).
- Schulmeyer, A. (2002). *Estado actual de la evaluación docente en trece países de América Latina. Trabajo presentado a la Conferencia Regional UNESCO: “El Desempeño de los Maestros en América Latina y el Caribe: Nuevas Prioridades”* (pp. 1-34). pp. 1-34. Recuperado de <https://unesdoc.unesco.org/ark:/48223/pf0000135506?posInSet=2&queryId=N-EXPLORE-9642004c-736b-4a26-a35c-2f3384923f13>
- Stecher, B. M., Holtzman, D. J., Garet, M. S., Hamilton, L. S., Engberg, J., Steiner, E. D., ... Chambers, J. (2018). *Improving Teaching Effectiveness: Final Report: The Intensive Partnerships for Effective Teaching Through 2015–2016*. Recuperado de https://www.rand.org/pubs/research_reports/RR2242.html
- Steinberg, M. P., y Donaldson, M. L. (2016). The New Educational Accountability: Understanding the Landscape of Teacher Evaluation in the Post-NCLB Era. *Education Finance and Policy*, 11(3), 340–359.