

# LA EVALUACIÓN NACIONAL DIAGNÓSTICA (END): PERCEPCIONES Y CONSIDERACIONES PARA LA FORMACIÓN INICIAL DOCENTE DESDE ESTUDIANTES Y DIRECTORES DE CARRERAS<sup>1</sup>

Andrea Ruffinelli<sup>2</sup>, Tatiana Cisternas<sup>3</sup>, Carla Förster<sup>4</sup>,  
Francisca Donoso<sup>5</sup>

## RESUMEN

El artículo explora conocimientos, valoraciones, expectativas y recomendaciones de estudiantes de pedagogía y sus direcciones de carrera, acerca de la evaluación centralizada y estandarizada que se les aplica en Chile (Evaluación Nacional Diagnóstica, END) para medir la calidad de la formación inicial, considerando que los evaluados han sido un actor invisibilizado en esta política. Mediante un diseño mixto consecutivo se aplicaron entrevistas individuales, grupos focales y una encuesta. Los hallazgos sugieren elementos novedosos para revisar la política. Los actores muestran adhesión a una evaluación nacional, pero desde una visión de ésta como herramienta para la mejora. Cuestionan el instrumento vigente por los límites de la estandarización y la ausencia de evaluación de desempeños prácticos. Critican la falta de información sobre el instrumento y advierten la necesidad de mayor descripción de los resultados para que puedan ser interpretados y utilizados para la mejora. Se concluye sobre la necesidad de revisar los múltiples objetivos que persigue esta evaluación y se propone una reflexión acerca del sentido y alcance de los mecanismos evaluativos para fortalecer la formación inicial del profesorado.

Conceptos clave: Formación Inicial Docente, políticas educativas, evaluación, rendición de cuentas, educación superior, Evaluación Nacional Diagnóstica (END).

## **THE NATIONAL DIAGNOSTIC EVALUATION (END): PERCEPTIONS AND CONSIDERATIONS FOR INITIAL TEACHER TRAINING FROM STUDENTS AND CAREER DIRECTORS**

### ABSTRACT

*The article explores the knowledge, assessments, expectations and recommendations of student teachers and their career managers regarding the centralized and standardized evaluation applied to them in Chile (National Diagnostic Evaluation, NDT) to assess the quality of initial*

---

1 Investigación financiada por el Grupo de Estudios sobre Carrera Docente del Centro de Medición MIDE UC de la Pontificia Universidad Católica de Chile.

2 Universidad Alberto Hurtado, Santiago, Chile. Contacto: aruffine@uahurtado.cl

3 Universidad Alberto Hurtado, Santiago, Chile. Contacto: tcistern@uahurtado.cl

4 Universidad de Talca, Talca, Chile. Contacto: carla.forster@utalca.cl

5 Universidad Metropolitana de Ciencias de la Educación, Santiago, Chile. Contacto: francisca.donos@umce.cl

*training, considering that those evaluated have been largely overlooked in this policy. Utilizing a consecutive mixed design, we conducted individual interviews, focus groups and a survey. The findings offer novel insights for revising the policy. While the participants express support for a national evaluation, they understand it primarily as a tool for improvement. Critiques center around the current instrument's limitations in terms of standardization and the absence of practical performance evaluation. They also highlight concerns about insufficient information on the instrument and emphasize the need for a detailed description of results to facilitate interpretation and use for improvement. In conclusion, the study advocates for a comprehensive review of the multiple objectives of this evaluation, prompting reflection on the meaning and scope of evaluative mechanisms to strengthen initial teacher training.*

*Key concepts: Initial Teacher Training, Educational Policies, Accountability, Higher Education, National Diagnostic Evaluation (END).*

## 1. El contexto chileno de formación inicial docente y su evaluación

En los años 90 las naciones debieron elegir entre dos tipos de reformas educativas: aquellas enfocadas en la equidad y el desarrollo de habilidades, y las basadas en la rendición de cuentas y los incentivos (Darling-Hammond, 2012). Las primeras promueven un sistema educativo desafiante, con un currículo reflexivo basado en habilidades complejas y colaborativas, respaldado por docentes altamente calificados y sistemas de evaluación continua. El segundo tipo de reforma promueve la autonomía y autorregulación profesional, pero condiciona a los docentes a través de estrategias de mercado, como exámenes nacionales, pago por mérito, reducción de la inversión en formación docente y desvinculación de la academia, atracción de personas sin preparación docente a la enseñanza y estimulación de la competencia entre pares, que premia o castiga docentes y escuelas según sus resultados en los exámenes estandarizados, estrechando el currículum y desvalorizando al docente (Feldfeber, 2007). Chile es uno de los países que asumió este último enfoque (Darling-Hammond, 2012) y que prontamente inició un camino de matices en que las políticas se orientaron hacia un tensionado doble objetivo: control mediante rendición de cuentas y apoyo para la mejora (Ávalos, 2014).

En el concierto internacional, diversos sistemas se fueron plegando hacia regulaciones basadas en estándares y rendición de cuentas de distinto tipo. Así, algunos países evalúan la FID, pero, mientras en Canadá y Singapur controlan distintos momentos de esta formación, en Chile se evalúa durante el año previo al egreso, como requisito de titulación. Algunos países certifican el ingreso a la profesión y otros certifican el desarrollo profesional (Darling-Hammond et al., 2017; Ingvarson y Rowell, 2017; Zeichner y Holland, 2016). Entre los países que evalúan la FID también se encuentran Reino Unido (Qualified Teacher Status, QTS), Estados Unidos (edTPA) y Australia (LANTITE), pero, a diferencia de Chile, lo hacen al egreso, con sistemas de altas consecuencias individuales (Ingvarson y Rowell, 2017; Vaillant, 2018), por ser habilitantes para el desempeño (Cohen et al., 2020; Wilson y Goff, 2019); mientras los

resultados de la END en Chile no tienen consecuencias individuales para los estudiantes.

En Chile, la evaluación a la FID responde al escenario que se fragua durante la segunda mitad de los años 90, momento en que la privatización de la educación y la falta de responsabilidad del Estado provocaron un aumento del 200% en la matrícula de carreras de pedagogía en 10 años (Cox, Meckes y Bascopé, 2010), a pesar de la baja valoración social y remuneración de la profesión (Cabezas y Claro, 2011). Esto se debió, en parte, a los casi inexistentes requisitos de admisión para estas carreras que, durante la primera década de los años 2000, continuaron enmarcadas en una educación superior que operaba como un mercado que se desarrollaba en muy heterogéneas condiciones de exigencias de ingreso, egreso y oportunidades formativas, sobre un escenario básico de precariedad (Ruffinelli, 2017).

En este contexto, en 2008 se lanzó el Programa INICIA, para medir, voluntaria, centralizada y estandarizadamente, la calidad de la formación y ofrecer apoyo para la mejora, aunque finalmente solo se mantuvo la prueba, el énfasis en la medición. Se criticó su aplicación al final de la formación, su enfoque en la rendición de cuentas en lugar del rol formativo, su carácter no representativo y su incapacidad para dar cuenta del desempeño profesional, al ser un instrumento de lápiz y papel con formato de selección múltiple (Montecinos, 2014; Ávalos, 2014). A pesar de las críticas, los resultados de la prueba INICIA fueron utilizados para clasificar programas formativos, con el objetivo de persuadir a las instituciones formadoras de obtener buenos resultados para atraer más y mejores estudiantes, asumiendo el Estado solo un rol de informador.

Entre 2010 y 2020 se produjo un giro hacia la revalorización social de la profesión docente, con la promulgación de la Ley 20.903 en 2016, que creó el Sistema de Desarrollo Profesional Docente (SDPD). Este sistema reconoce que el desarrollo profesional comienza en la Formación Inicial Docente (FID), establece derechos como la inducción profesional y la formación continua, y crea un sistema de reconocimiento y evaluación del progreso docente.

El SDPD fortalece la regulación de la profesión docente en Chile, pero también enfatiza la rendición de cuentas. Aborda la FID profundizando en las estrategias de estandarización, evaluación y control, definiendo requisitos de ingreso, acreditación obligatoria y estándares formativos. La ley indica que para acreditar los programas es necesario aplicar dos evaluaciones diagnósticas: una al inicio de la carrera y otra al final, realizada por el Ministerio de Educación de Chile (MINEDUC), la Evaluación Nacional Diagnóstica (END) que reemplaza a la Prueba INICIA.

La END comparte con su antecesora el carácter centralizado y estandarizado, y su sustento en los estándares nacionales para las carreras de pedagogía<sup>6</sup> (CPEIP, 2021), pero difiere de la prueba anterior en su carácter obligatorio, en su aplicación un año antes de finalizar la FID y en la exigencia de obligar a los programas formativos a implementar acciones de nivelación y utilizar los resultados para mejorar continuamente (CPEIP, 2017), mientras INICIA centralizaba apoyos al comienzo. La ley estipula que las carreras de pedagogía utilicen los resultados que arroja la evaluación de dos maneras distintas: por una parte, señala la necesidad de formular planes de apoyo y acompañamiento en las áreas detectadas como deficitarias en los resultados de los estudiantes que ya han rendido la evaluación y, por otra, indica a las carreras que, a partir del análisis general, los resultados obtenidos por los estudiantes deben incorporar mejoras a los planes y programas de formación, con una mirada a más largo plazo (MINEDUC, 2016).

La END se aplica desde 2016 y se compone de dos instrumentos, una Prueba de Conocimientos Pedagógicos Generales, correspondiente al nivel de enseñanza (Educación Parvularia, Educación Básica o Educación Media) y una Prueba de Conocimientos Disciplinarios y Didácticos, de acuerdo con cada disciplina. Adicionalmente se aplica a los estudiantes un Cuestionario Complementario, que recoge información para caracterizar la opinión de los estudiantes de pedagogía, su proceso

---

6 Los estándares para Carreras de Pedagogía determinan las habilidades, conocimientos y competencias que debe adquirir un/a profesional de la educación en el transcurso de su formación inicial. <https://estandaresdocentes.mineduc.cl/estandares-carreras-pedagogia/>

formativo, sus creencias respecto de la docencia y sus expectativas profesionales.

La información acerca de la aplicación de la END se canaliza a través del Centro de Perfeccionamiento e Investigaciones Pedagógicas (CPEIP) del Ministerio de Educación de Chile (MINEDUC); este envía los reportes a las universidades y éstas los hacen llegar a sus estudiantes. Los resultados se entregan mediante informe digital a las universidades y a través de un informe individual confidencial que cada estudiante puede consultar en la web.

En 2018 la END pasó a ser digital, pero mantuvo un formato prioritario de selección múltiple, con foco en conocimientos declarativos, replicando la debilidad de INICIA, al no medir multidimensionalmente los estándares ni el desempeño (Donoso y Ruffinelli, 2020).

Las decisiones sobre esta política de la END han sido tomadas solo por expertos e instituciones, sin considerar la perspectiva de los evaluados. Explorar la subjetividad de los estudiantes y las direcciones de carrera permitiría agregar una perspectiva importante y generar conocimiento sobre la política, lo que podría ser útil para discutirla en Chile y en otros países que la observan con atención.

## 2. La literatura sobre la perspectiva de los evaluados

La bibliografía internacional muestra que las políticas que regulan la FID mediante estándares y con foco en la rendición de cuentas tienen pocas implicancias positivas; sin embargo, algunos autores señalan que cuando éstas se implementan bajo ciertas condiciones, mejoran las capacidades docentes y el estatus de la profesión (Demarchi, 2020, Darling-Hammond y Lieberman, 2012). Mayoritariamente se reportan implicancias negativas, como la focalización errónea en habilidades básicas y/o técnicas, sobrevigilancia del trabajo docente (Lerman, 2014), reducción de la autonomía local y tecnificación de la enseñanza, al omitir debates éticos y morales del currículum (Conway, 2013; Conway y Murphy, 2013).

O'Neill (2013) propone una “rendición de cuentas inteligente”, basada en la confianza y consideración de los sujetos de la política, que permite percibir los procesos evaluativos como señales de buenas intenciones, con evaluadores íntegros y capaces.

En Chile, Donoso y Ruffinelli (2020) relevan que la END es muy similar a INICIA y que no existe evidencia pública acerca de la forma en que mide los estándares. Solo existe un temario de la prueba que los replica. Por su parte, Montecinos (2014) plantea que el uso de evaluaciones educacionales orientadas por estándares requiere evaluar el desempeño en tareas “auténticas” del trabajo docente, aspecto desatendido en ambas evaluaciones y que en otros países se incorpora a través de instrumentos como portafolios.

Investigaciones internacionales señalan como aspectos críticos de este tipo de evaluaciones la estandarización (pérdida de sellos y particularidades de la formación) y la performatividad de los desempeños; restringiendo la formación a lo que se evalúa, promoviendo el entrenamiento para la evaluación y el estrechamiento curricular (Bartlett, Otis-Wilborn & Peters, 2017; Berliner, 2011; Greenblatt, 2018, Kessler, 2019; Ledwell y Oyler, 2016; Schmier, 2019).

Los formadores alegan querer ayudar a sus estudiantes a obtener buenos resultados evaluativos, sacrificando tiempo para retroalimentar aprendizajes fundamentales para el ejercicio de la profesión, en pos de focalizarse en aspectos específicos que se solicitan en los portafolios, disociando así un desarrollo profesional sustantivo (Clayton, 2018; Greenblatt, 2018).

La literatura también reporta que estas evaluaciones generan problemas de salud mental, por los altos niveles de ansiedad y angustia que genera la evaluación (Hilton, Saunders & Mansfield, 2020; Wilson y Goff, 2019; Förster, Ruffinelli, Cisternas, y Donoso, en prensa).

Por su parte, indagaciones en la percepción de los evaluados acerca de evaluaciones de este tipo, valoran que exista una prueba

nacional para validar sus saberes (Wilson y Goff, 2019) y estiman que “asegura” la calidad de la formación (Hilton et al., 2020; Wilson y Goff, 2019). También valoran la evaluación de desempeños en lugar de pruebas “de lápiz y papel” (Greenblatt, 2018); sin embargo, el hecho de tener “altas consecuencias” es percibido como una limitante a la capacidad formativa de los instrumentos (Kessler, 2019), aunque para algunos programas de formación alinearse a la evaluación ha significado una mejora (Ledwell y Oyler, 2016).

Existe evidencia (desde evaluaciones en el sistema escolar) en relación con la baja disposición de los evaluados y la escasa valoración que dan a esa instancia cuando no encuentran sentido a la evaluación (Cohen et al., 2020), y, al contrario, se incrementa el uso de la información que emana de los procesos evaluativos cuando perciben que ayuda a la mejora (Ehren y Baxter, 2021).

En Chile, Fernández y Madrid (2020) refrendan como efectos la generación de incentivos a la preparación de los estudiantes para rendir la END y el estrechamiento curricular, y muestran dudas sobre la relevancia de la información entregada a las instituciones de educación superior. Adicionalmente, Cabezas et al. (2019) reportan la inviabilidad de implementar apoyos y remediales por la demora en la recepción de resultados.

La política de la END ha considerado exclusivamente voces expertas y se desconoce la percepción de los evaluados. Una revisión de la literatura (Förster et al., en prensa) muestra un muy bajo número de investigaciones desde la perspectiva de estos actores. Entonces, explorar sus subjetividades y las de los directivos permite incorporar una perspectiva imprescindible respecto del efecto de esta política y generar conocimiento acerca de ella y sus posibilidades de mejora.

Considerando estos antecedentes, esta investigación buscó conocer los conocimientos, valoraciones, expectativas y recomendaciones a la END que expresa el estudiantado de pedagogía antes y después de rendir la evaluación, y también lo que opinan las direcciones de carrera, en torno al instrumento y la política, información previa recibida, aplicación, informe y uso de resultados.



### 3. Metodología

El estudio es descriptivo, de diseño mixto consecutivo, con predominio cualitativo. Se levantaron datos cualitativos y luego cuantitativos (Creswell, 2008). Para la fase cualitativa se utilizaron entrevistas semiestructuradas a directoras/es de carrera y grupos focales con estudiantes que ya habían rendido la evaluación — que en este artículo denominamos “Post END”— y con quienes les correspondía hacerlo el año 2020 (que aquí denominamos “Pre END”)7. Las entrevistas y grupos focales abordaron la experiencia vivida y los conocimientos, valoraciones y recomendaciones o expectativas respecto de la evaluación.

Se seleccionaron siete programas de pedagogía impartidos por cuatro instituciones formativas, resguardando la presencia de una universidad tradicional de alta matrícula y privadas de menor matrícula (3), que reflejaran la diversidad de las instituciones formativas que imparten las dos carreras con mayor matrícula nacional: Educación Diferencial (4 programas) y Educación Básica (3 programas). Fueron criterios de inclusión pertenecer a la Región Metropolitana, donde se concentra la mayoría de estos programas formativos, y estar acreditado8 por al menos 4 años, resguardando un criterio de calidad.

Con esta muestra se efectuaron 8 grupos focales remotos con 49 estudiantes Pre END, y 6 con 20 estudiantes Post END, mayoritariamente mujeres. También, 7 entrevistas con 8 directivos/as de carrera, de 4 programas de Educación Básica y 3 de Educación Diferencial, de 4 universidades: una tradicional y pública selectiva de gran matrícula y 3 universidades privadas consolidadas, 3 selectivas y una de menor selectividad. Los datos fueron procesados mediante análisis de contenido inter casos (Díaz y Navarro, 1998), combinando

---

7 Estudiantes PostEND que rindieron la evaluación el año 2019, y estudiantes Pre-END que les correspondía hacerlo en 2020. Dicha aplicación posteriormente fue suspendida por el Ministerio de Educación.

8 En Chile, las universidades solo pueden impartir programas de pedagogía que hayan sido acreditados por la Comisión Nacional de Acreditación, CNA, y pueden serlo por un máximo de 7 años. Uno de los criterios de acreditación que no se encuentra definido actualmente es acerca de los resultados obtenidos en la END.

estrategias inductivas y deductivas, y resguardando triangulación en el proceso de codificación.

Para la etapa cuantitativa se desarrolló una encuesta, cuyas dimensiones y preguntas consideraron el levantamiento cualitativo anterior. Se invitó a todas las direcciones de estas carreras en el país a animar a responder a sus estudiantes. Respondieron a dicho instrumento 371 estudiantes de 22 universidades. De ellos, 210 eran Pre END (120 de Educación Básica y 90 de Educación Diferencial) y 161 Post END (93 de Educación Básica y 68 de Educación Diferencial). Asimismo, participaron 31 directivos de 22 universidades (17 de Educación Básica y 14 de Educación Diferencial), representando al 46,3%<sup>9</sup> de los directores de estas carreras en el país. Son principalmente mujeres, y la mayoría indicó haber vivido 3 o más aplicaciones de la END (tabla 1).

Los datos cuantitativos fueron analizados descriptivamente, realizando cruces de información entre variables de caracterización y respuestas de los estudiantes y directivos.

Los aspectos éticos se resguardaron mediante aprobación del proyecto por parte del comité de ética autorizado y el consentimiento informado de quienes participaron.

Tabla 1  
*Distribución de la muestra cuantitativa*

Carrera	Directores		Estudiantes Pre END		Estudiantes Post END	
	n	%	n	%	n	%
Educación Básica	17	54,8	120	57,1	93	57,8
Educación Diferencial	14	45,2	90	42,9	68	42,2
Total	31	100	210	100	161	100

Fuente: Elaboración propia.

<sup>9</sup> Considerando como referencia que estudiantes de Educación Básica de 40 universidades rindieron END 2018 y estudiantes de Educación Diferencial de 27 universidades hicieron lo propio.

## 4. Resultados

Los resultados se organizan según a) conocimientos (en el caso de estudiantes) o experiencia (en el caso de directivos), b) valoraciones y recomendaciones (en el caso de Post END y directivos) o expectativas en torno al instrumento y la política (en el caso de Pre END), c) información previa recibida, d) aplicación, e) informe y f) uso de resultados. Estos se presentan articulando hallazgos de la encuesta, entrevistas y grupos focales, e integrando la visión de estudiantes y direcciones de carrera acerca de la END.

### 4.1. Acerca de la política y el instrumento

Los estudiantes (Pre y Post END) reportan un nivel de conocimiento muy general sobre la evaluación, aunque naturalmente es mayor entre Post END. Hay desconocimiento y confusión respecto de sus contenidos y características. Más del 80% reconoce dos propósitos: diagnóstico de conocimientos para mejorar la FID y medir y comparar programas, pero menos del 40% de los Pre END sabe que es un instrumento mayormente de selección múltiple:

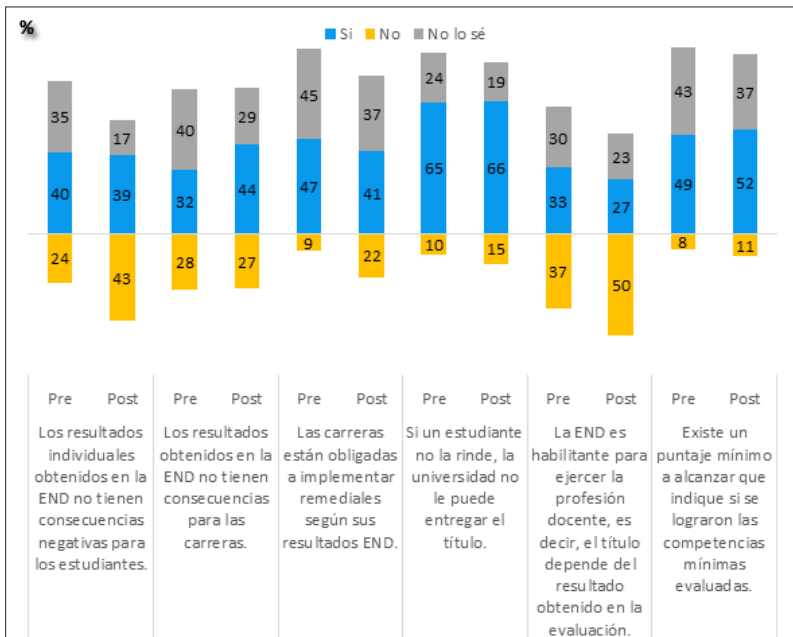
*...es una evaluación que reúne los conocimientos básicos que necesitamos para salir al mundo de nuestra labor docente, como lo básico, lo que nos tiene que entregar la universidad. También sé que el foco de esta prueba es cómo mejorar la formación docente de los profesores.* (Focus Pre END Estudiantes Ed. Diferencial, Universidad C)

En los grupos focales Pre END señalan no saber qué contenidos se evalúan y desconocer el temario de la prueba. Un pequeño grupo la asocia a los EO. En los grupos focales Post END dudan si la evaluación se basa en el currículum de la FID, el MBE o los EO. Sin embargo, en la encuesta, 86% reporta saber que se basa en los EO y haber conocido el temario.

*Yo desconozco mayores detalles de la prueba, solo sé que la tengo que rendir por mi último año de carrera y es un requisito para egresar (...) desconozco contenidos, qué evalúan, en qué saberes, algún contenido en específico. Eso es lo que no sé.* (Focus Pre END Estudiantes Ed. Básica, Universidad C)

Respecto de las consecuencias, más del 60% sabe que es obligatoria y requisito de titulación. Sin embargo, hay otras creencias erróneas: más del 60% de los Pre END y la mitad de los Post END cree que es habilitante para el ejercicio profesional o no tiene certeza al respecto, y alrededor de la mitad cree que existe un puntaje mínimo a lograr. Más del 40% de los Post END cree que tiene consecuencias individuales y más de la mitad no sabe o cree que no existe una exigencia a las instituciones de implementar remediales ante resultados descendidos (figura 1).

Figura 1  
 Conocimiento acerca de las consecuencias de la END. Estudiantes



Fuente: Elaboración propia.

Hay acuerdo con la idea de evaluar la FID (100% direcciones y más del 92% estudiantes) para asegurar un nivel mínimo formativo y por su potencial de ayuda a la mejora (68% directores, 88% Pre END y 87% Post END). Los Post END y directores/as agregan que significa un reconocimiento al valor de la profesión.

*Creo que es necesaria. Para mí sí o sí tiene que haber una evaluación de nuestra formación porque necesito saber qué es lo que espera el país que yo sepa como profesora (...) según lo que ellos determinan que se necesita, porque yo tengo mis calificaciones en la universidad, pero eso es mi universidad (...) hay otras universidades que por un trabajo les ponen un 7 y el mismo trabajo a mí me lo revisa una profe acá y me pone un 4. Las exigencias son distintas (...) tener la evaluación desde una visión más general en todo el país es mucho mejor. (Focus Post END Estudiantes Ed. Diferencial, Universidad A)*

Cerca del 70% de los Post END la considera relevante. Aunque algunos cuestionan la insistencia en la evaluación, en lugar de poner foco en la formación:

*No creo que sea necesario seguir creando tantas evaluaciones estandarizadas, por algo las universidades están acreditadas y deberían asegurar eso que dicen en su perfil de egreso (...) seguir generando evaluaciones estandarizadas para mí no tiene sentido, ¿dónde están los apoyos después cuando se egresa? (...) creo que pierde un poco el foco de la evaluación formativa. (Focus Post END Estudiantes Ed. Básica, Universidad B)*

Pese al consenso respecto de evaluar, en la sección cualitativa hay desacuerdo en hacerlo con la actual END. Los directores esgrimen desconocimiento del instrumento (más del 50%), lo que impide juzgarlo y utilizar los resultados para la mejora. También relevan como debilidad fundamental el que no mida desempeños:

*No me parece cómo está hecha la prueba (...) para medir ciertos conocimientos eran súper poquitas preguntas (...) tampoco estaba muy claro qué es lo que se iba a medir en esta parte de escritura. (Focus Post END Estudiantes Ed. Básica, Universidad A)*

Sin embargo, en la encuesta el 67% de los estudiantes Post END y 80% de los directores manifiesta que es una buena forma de evaluar la FID, pero identifican tres limitaciones: la sobre simplificación de la tarea docente al omitir los desempeños, la limitada capacidad de los instrumentos de selección múltiple para evaluar respuestas

complejas y la descontextualización respecto de enfoques o itinerarios formativos diversos.

Se observa cierto consenso respecto a la obligatoriedad de la END como requisito de titulación (63% Pre END, 66% Post END y 74% directores). Estiman que permite observar con fidelidad el panorama nacional y que incrementa su potencial como herramienta de mejora.

También otorgan una valoración positiva a la calidad de las preguntas, por su capacidad para abordar integralmente la formación, y es particularmente bien valorada la pregunta de “desarrollo o planificación”.

Asimismo, estudiantes y directores valoran la ausencia de consecuencias individuales por rendimiento y la existencia de consecuencias formativas para las carreras. No hay consenso sobre considerar los resultados en la acreditación de las carreras: 71% de los directores está en desacuerdo y una cifra similar de estudiantes está de acuerdo. Las mayores preocupaciones se refieren al uso de resultados para *rankear*, y que la estandarización invisibilice sellos formativos institucionales.

Los actores también hacen recomendaciones. Sobre el instrumento, plantean otras formas de evaluar. La principal sugerencia es complementar con evidencias de desempeños (94% directores, 83% Pre END y 91% Post END):

*Los conocimientos que se pueden evaluar en una prueba son súper importantes, pero te puedes sacar un 7 en una prueba y en el aula te puede costar muchísimo. Y al final, cuando nos titulemos, no vamos a estar haciendo pruebas, vamos a estar trabajando en aula, y ahí es donde se ve nuestra calidad de docente.* (Focus Pre END Estudiantes Ed. Diferencial, Universidad C)

Para incluir desempeños, proponen la participación de los programas formativos, observando esta dimensión en las líneas de prácticas:

*Nosotras tenemos mil prácticas (...) podría, así como lo hacen en la Carrera Docente (...) haber una observación de alguna experiencia de aprendizaje y también sacar de ahí una evaluación. Además que también sería mucho más cualitativa y nos podrían brindar información más enriquecida.* (Focus Post END Estudiantes Ed. Diferencial, Universidad A)

También sugieren incluir más preguntas de “aplicación” y análisis de casos, demandando alinear la evaluación con los principios pedagógicos y evaluativos con los que fueron formados, así como procurar contextualizar en los enfoques y sellos formativos de las universidades. 87% directores, 59% Pre END y 74% Post END sugieren incorporar más preguntas abiertas.

Más del 70% de los estudiantes y 49% de los directores propone que las carreras evalúen con instrumentos propios. Esta idea se vincula a la demanda por contextualización según enfoques formativos (92% Pre END y 84% Post END).

*Quizás sería una buena opción llegar a acuerdos en temas de contenidos y que cada universidad hiciera su prueba según lo que ellos mismos han estado trabajando durante años con sus estudiantes.* (Focus Post END Estudiantes Ed. Básica, Universidad C)

Consecuentemente, en la encuesta todos los directores proponen transparentar los indicadores de validez y confiabilidad del instrumento y contar con ejemplos de ítems. Casi todos (97%) demandan transparentar la forma en que la END evalúa los estándares formativos. Otras demandas casi unánimes son: transparentar el proceso de construcción del instrumento, indicadores de validez y confiabilidad, peso de cada estándar e indicador, y contar con niveles de logro por estándar.

Algunos sugieren acotar los propósitos de la END, prescindiendo de la exigencia de remediales a la cohorte y asumiendo un objetivo de mejora a más largo plazo:

*Los resultados nos llegan súper tarde y, dado que es su último año de carrera, no permite tener espacios para instalar esta mejora. Yo me*

*haría la pregunta si no vale la pena hacerla al final de la carrera y que el sentido sea retroalimentar a las generaciones que vienen, más que a esa generación en particular.* (Entrevista Directivo Ed. Básica, Universidad B)

Algunos estudiantes Post END sugieren considerar que sus resultados tengan efectos en la acreditación. Consecuentemente, estudiantes y direcciones de carrera recomiendan insistir en consecuencias formativas institucionales más que individuales, focalizando en apoyos y remediales promovidos desde el nivel central.

#### 4.2. Sobre la información previa recibida

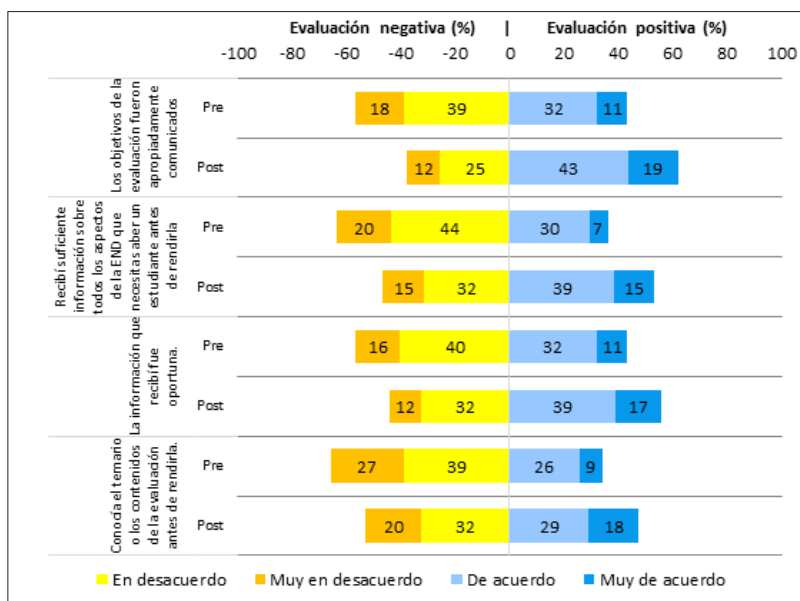
En un proceso evaluativo, la información disponible puede mediar la disposición y desempeño. De allí su importancia a la hora de indagar en la visión de los actores evaluados. En cuanto a la información recibida, hay una visión más positiva de los Post END que de los Pre END, que tienden a considerarla insuficiente e inoportuna (figura 2).

*La información que nos ha dado la universidad ha sido muy general y lo que hemos sabido lo hemos buscado en internet, en las páginas oficiales de la prueba.* (Focus Pre END Estudiantes Ed. Básica, Universidad B)



Figura 2.

Valoración de la información previa recibida. Estudiantes



Fuente: Elaboración propia.

Por su parte, las direcciones valoran positivamente la comunicación de los objetivos de la END (74%) y la información que llega para ser difundida a los estudiantes (58%), pero tienen una visión menos positiva de la utilidad del temario (49%) y de la suficiencia de información que entrega el CPEIP a los estudiantes (42%).

Consistentemente, más del 90% de los estudiantes recomienda proveer más y mejor información sobre el proceso, objetivos y consecuencias, y más del 80% sugiere adelantarla incluso al momento de postular o matricularse.

### 4.3. Sobre la aplicación

La experiencia vivida y las emociones asociadas podrían ayudar a explicar la valoración de la política. 51% Post END la considera una buena experiencia, aunque 61% sintió angustia, ansiedad y nerviosismo, sugerente del sentido de responsabilidad con la propia

formación. Los Pre END reportan que estas emociones se asocian a la incertidumbre con la que se enfrentan a la evaluación.

*Cuando a uno le dicen “la evaluación” tiene una carga emocional, pero esa carga disminuye cuando sabes qué es lo que te van a evaluar, en qué profundidad, de qué forma, si va a ser una prueba de alternativas, de desarrollo, etcétera, y ahora no hay nada de eso. (Focus Pre END Estudiantes Ed. Básica, Universidad A)*

Hay acuerdo respecto de que las condiciones de aplicación fueron inapropiadas (70% directores y 51% Post END). Consideran que es muy extensa y que hay rigidez en el uso del tiempo y aplicación de normas, espacios inadecuados y fallas técnicas. Los que la rindieron en sus universidades valoran positivamente el espacio. Se reporta que las condiciones de aplicación de 2019 fueron disímiles (64% de directores y 69% estudiantes Post END).

*A cada rato era más y más frustrante, y ya después, con los errores del sistema, que nos pasó a todas que se nos borró al menos una vez la pregunta abierta, yo también terminé respondiendo cualquier cosa. Así como que pasé de tener un párrafo gigante a un par de líneas y fue como “esto es lo que puedo dar en este momento. Adiós”. Fue peor de lo que pensaba la historia. (Focus Post END Estudiantes Ed. Diferencial, Universidad B)*

Consistentemente, los estudiantes recomiendan mejorar las condiciones físicas (86% Pre END y 76% Post END), rendirla en la propia universidad (85% Pre END y 83% Post END), realizarla en más días (75% Pre END y 56% Post END), flexibilizar el uso del tiempo y las normas al terminar de rendir el examen, y superar las fallas técnicas.

#### 4.4. Sobre el informe de resultados

Los Post END tienden a no recordar el contenido del informe individual de resultados. Coexisten opiniones que lo consideran completo y detallado: 71% por informar el nivel de logro, 66% por la posibilidad para la mejora y 64% por su claridad, junto con opiniones que estiman que es incompleto, al carecer de retroalimentación

descriptiva y de juicio sobre el significado de los datos, en ausencia de un parámetro de logro.

*No sabemos exactamente qué cosas respondimos mal. O sea, solamente sabemos que, en cierta dimensión, pero no sabemos qué respondimos mal.* (Focus Post END Estudiantes Ed. Diferencial, Universidad A)

Las direcciones, por su parte, reconocen que los informes institucionales han mejorado en calidad y oportunidad, pero 74% estima que siguen llegando tardíamente. Más del 40% indica que falta descripción en la comunicación de resultados por estándar y un tercio señala que no permite definir acciones remediales:

*Me preocupa cuando estoy por debajo del promedio nacional y es algo que hemos revisado y le hemos dedicado tiempo (...) yo digo “¿cómo habrán preguntado?”. Y ahí me falta el instrumento.* (Entrevista Directivo Ed. Básica, Universidad D)

También reportan disparidad en la recepción de resultados individuales por estudiante: solo un 36% dice recibirlos (figura 3).

Los directivos reconocen que no disponen de un parámetro para ponderar su rendimiento y que usan estos criterios de forma automática o inducida, sin que constituyan una referencia al tipo de rendimiento. En efecto, para atribuir valor al rendimiento los actores usan el promedio o porcentaje de logro, asumiendo algunos que sobre el 50% y otros que sobre 70% es un rendimiento satisfactorio. También esperan conocer los criterios o rúbricas con las que fueron evaluados.

*Qué es estar bien y qué es estar mal... no está en la discusión, que debiera ser una definición. ¿Dónde está el corte? ¿Dónde ponemos ese corte? No está la discusión.* (Entrevista Directivo Ed. Básica, Universidad A)

*Uno siempre hace la comparación a nivel nacional, entonces dice “bueno, lo mínimo esperado para nosotros es estar por sobre la media nacional...”. Pero ¿qué significa la media nacional? Vuelvo a decir, ¿cuál es la descripción del desempeño esperado por tramos y qué*

*significa que es capaz de hacer, no de saber?* (Entrevista Directivo Ed. Básica, Universidad B)

En esta misma línea, estudiantes (78% Pre END y 82% Post END) demandan incluir parámetros de logro, un dato de referencia para comparar su resultado.

*Haber puesto si pasas o no (...) o sea, yo tengo mis porcentajes y veo que unos son buenos y otros son malos, pero me faltó eso, como el ultimátum.* (Focus Post END estudiantes Ed. Básica, Universidad D)

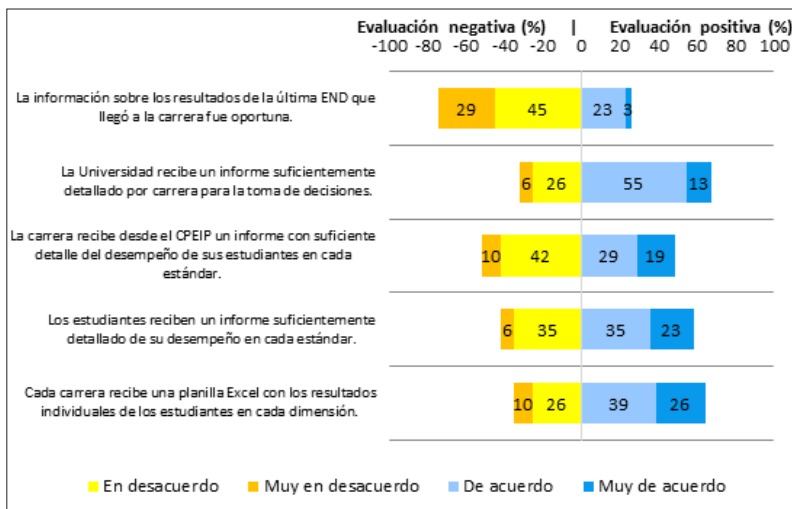
Los directores también enfatizan la necesidad de recibir retroalimentación oportuna para implementar las medidas exigidas por la ley. Asimismo, directores y estudiantes recomiendan hacer más descriptiva y cualitativa la comunicación de resultados de cada estándar, posibilitando usar los datos para la mejora:

*Requiere de entrega oportuna de los resultados, que permita hacer un proceso remedial. Eso pasa [también] por establecer mecanismos de entrega que reflejen efectivamente posibilidades de mejorar, no solo los datos generales.* (Entrevista Directivo Ed. Básica, Universidad A)

*Puede ser como una bajada cuali al puntaje, como, el 50% significa que han desarrollado tantas habilidades pero les faltan otras.* (Focus Post END Estudiantes Ed. Diferencial, Universidad B)

Figura 3.

Valoración del informe de resultados. Directores



Fuente: Elaboración propia.

#### 4.5. Sobre el uso de resultados

Tras conocer los resultados de la evaluación, la política exige a los programas: a) nivelación para estudiantes con bajos resultados y b) uso de la información para la mejora del programa formativo. Otro uso, no contemplado en la política, consiste en acciones preparatorias a la rendición: familiarización con el proceso e instrumento o de reforzamiento.

Respecto de las acciones preparatorias, la más recurrente es la charla informativa acerca de la evaluación: más del 40% de los estudiantes la señala y valora, mientras el 90% de los directores la señala.

Estudiantes y directores valoran de manera diferente las acciones de reforzamiento previas a la rendición de la evaluación: 83% de los Pre END y 60% Post END lo cree necesario, mientras solo un 29% de los directores. Sin embargo, 48% de los directores implementa talleres de reforzamiento de contenidos previos a la prueba, mientras solo 20% de los estudiantes sabe de estas instancias. Algunos

estudiantes reconocen que afecta la confiabilidad de la medición, pero valoran “reparar” contenidos para enfrentarla con más seguridad:

*Si es que nos empiezan a preparar sería como rendir algo un poco falso, porque han tenido todos estos años para ir preparándonos. Deberíamos familiarizarnos con el formato de la prueba y saber un poco de lo que se trata, qué objetivos tiene, lo que se mide, cómo se mide, pero más allá de eso siento que no deberíamos tener una preparación para tener un mejor puntaje, porque para algo tuvimos los años anteriores para prepararnos. (Focus Pre END Estudiantes Ed. Básica, Universidad D)*

Los estudiantes sugieren formas de preparación como actividades sobre la END en algún curso, así como charlas o talleres de reforzamiento de contenidos.

Respecto de acciones exigidas por la política —nivelación y mejora de programas—, como muestra la figura 4, solo las primeras son reconocidas por los estudiantes, y más del 80% de los Pre END y 70% de los Post END reporta no saber o señala que no hubo acciones remediales posteriores a los resultados evaluativos. Mencionan talleres grupales de reforzamiento disciplinar o didáctico en disciplinas principales. En contraste, 74% de los directores indica que implementó talleres grupales de reforzamiento, y 26% que implementó remediales individuales.

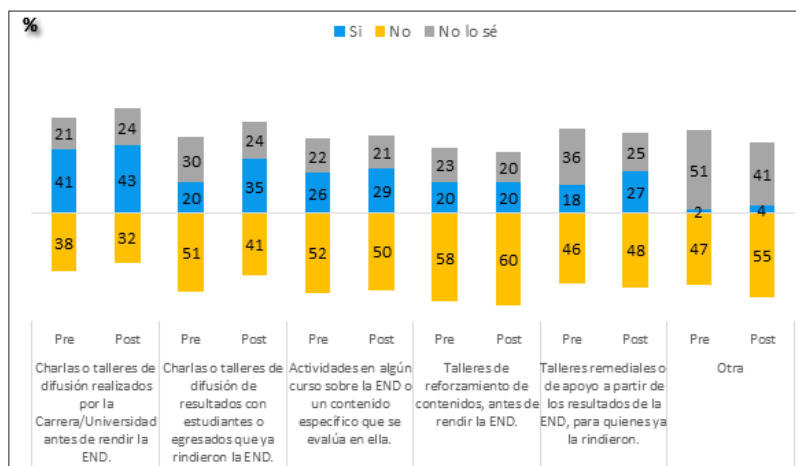
Los Post END estiman las medidas remediales insuficientes, poco oportunas, agobiantes, pero también beneficiosas. Por su parte, los directores creen que son valoradas por los estudiantes (66%), pero plantean que solo se alcanza a implementar medidas superficiales (78%), que no han podido evaluarlas y que los estudiantes no asisten (55%). En las entrevistas refieren condiciones inapropiadas para implementarlas: porque el tipo de resultados no permite diseñarlas y por el escaso tiempo para hacerlo, ya que se conocen muy tarde los resultados, en momentos críticos de la práctica profesional, lo que redundo en baja participación:

*Creo que la falencia de esa implementación es que no hemos evaluado ni siquiera la percepción de los profes que han dictado los talleres*

*respecto a lo que han visto y cuál es la percepción que los estudiantes tienen. (Entrevista Directivo Ed. Básica, Universidad A)*

Los estudiantes proponen remediales de análisis de resultados, búsqueda colaborativa de soluciones, así como modificación de cursos o redistribución de asignaturas. En cambio, las acciones más reportadas por directores/as son cambios curriculares que consideran los resultados END como un insumo más. Lo más recurrente es incorporar contenidos en programas de cursos (81%), elaborar/analizar matrices de tributación de los cursos a los estándares formativos (65%), enfatizar los estándares en todos los cursos (48%), incorporar cursos (42%) y, mucho menos, adelantar cursos (13%).

Figura 4.  
Acciones implementadas por los programas. Estudiantes



Fuente: Elaboración propia.

## 5. Discusión y conclusiones

¿Hay que evaluar la FID? Los actores consultados estiman que sí, pero asociando estos procesos a una cultura evaluativa para la mejora. Se argumenta que refuerza la trascendencia de la profesión y las implicancias del ejercicio docente sin una formación adecuada, refrendando lo que ocurre en otras latitudes con evaluaciones

similares (Darling-Hammond & Lieberman, 2012, Demarchi, 2020). Sin embargo, señalan no estar de acuerdo con el formato, que omite la demostración de desempeños, como advierten Donoso y Ruffinelli (2020), e indican que tiene las limitaciones de la estandarización, además de enfatizar que desconocen el instrumento y señalar que la forma de comunicar los resultados imposibilita comprender su alcance y usarlos para la mejora.

¿Qué saben acerca de la END? Los estudiantes manejan un paquete muy básico y tardío de información: obligatoria para titularse, propósito de diagnóstico, momento de aplicación y, menos, el tipo de instrumento. Desconocen aspectos relevantes, como contenidos, tipos de conocimientos evaluados, responsables de la elaboración, tipo de información que se obtiene y el tipo de consecuencias individuales.

Se reconocen también múltiples propósitos evaluativos, de los que se desprenden dilemas y cuestionamientos: se evalúa para medir, para comparar y para mejorar, una evidente mezcla de propósitos de la política en la que aparece una encrucijada que exige a un único y acotado instrumento resolver diferentes problemas: rendición de cuentas y mejora continua para la cohorte que rinde, que ya no tiene tiempo, y también para el programa formativo.

¿Qué valoran de la END? Valoran una evaluación nacional externa que valide saberes y asegure la calidad de la formación, en la línea de los hallazgos de Wilson y Goff, (2019), y de Hilton et al. (2020), pero no la actual. Valoran su obligatoriedad como garante de representatividad. También la pregunta abierta, y reconocen la calidad de las preguntas cerradas, aunque critican los límites de la estandarización y su descontextualización respecto de particularidades de su formación. Identifican ventajas y desventajas de la aplicación a un año del egreso: posibilidad de remediales, pero que afecta especialmente a los programas de menor duración el restar todavía tiempo de carrera.

¿Qué critican y demandan a la END? Estudiantes y direcciones critican la ausencia de desempeños prácticos, esto en la línea de los hallazgos de Greenblatt, (2018), de que en otras latitudes valoran



evaluaciones precisamente de desempeño por sobre evaluaciones declarativas. También relevan debilidades en las condiciones de aplicación, contenido de los informes de resultados y su uso en la implementación de remediales y mejora continua, en sintonía con los hallazgos de Fernández y Madrid (2020), así como falta de oportunidad en su entrega. Críticas similares se observan en los estudiantes que rinden LANTITE en Australia (Hilton et al., 2020) y refrenda también los hallazgos de Cabezas et al. (2019) en Chile. Los estudiantes agregan la estandarización que invisibiliza las particularidades formativas, tal como reportan Bartlett et al. (2017), Berliner (2011), Greenblatt (2018), Kessler (2019), Ledwell y Oyler (2016) y Schmier (2019); en tanto los directivos reclaman la falta de disponibilidad pública de información respecto del instrumento, lo que impide juzgarlo, pues se desconoce la forma en que da cuenta de los estándares y sus indicadores, e impide el uso de los resultados para mejorar, tal como han advertido antes Donoso y Ruffinelli (2020). Esto configura un problema de usabilidad de la información, o de validez consecuencial, al no lograr ser útil a la mejora, que es el propósito de la política. También hay preocupación por el riesgo de *rankear* programas.

¿Cómo se usan los resultados de la END? Los estudiantes informan de escasas acciones de nivelación individuales y colectivas, aunque las valoran, pese a que no se ha sistematizado evidencias. Se refrenda la inviabilidad de implementar medidas remediales a causa de la todavía demorada recepción de resultados (Cabezas et al., 2019), añadiéndose la imposibilidad de diseñarlas con el tipo de resultados entregados.

Respecto de los eventuales efectos performativos que la literatura ha vinculado con las evaluaciones nacionales estandarizadas (incentivo a la preparación de la evaluación y estrechamiento curricular —reportados por Bartlett et al., 2017; Berliner, 2011, Greenblatt, 2018, Kessler, 2019, Ledwell y Oyler, 2016, y Schmier, 2019—, los hallazgos muestran una tendencia poco clara a entrenar a los estudiantes y no hay evidencias del efecto de estrechamiento curricular. Proponemos cuatro posibles hipótesis: a) los programas ya efectuaron cambios curriculares para ajustarse a los estándares

formativos que mide la prueba; b) el único documento oficial disponible relativo a los contenidos evaluados es el temario, que considera la totalidad de los estándares, entonces, no existen antecedentes públicos que permitan focalizar la formación hacia tópicos evaluados más específicos; c) por los efectos de la contingencia social y sanitaria, y d) la falta de percepción de los resultados como valiosos para la mejora.

En suma, la mirada desde la perspectiva de los evaluados entrega un conjunto de elementos novedosos acerca de una política de alcance nacional que debe revisarse a la luz de la evidencia y sus propósitos.

## 6. Recomendaciones

Las principales demandas refieren tanto a elementos estructurales (qué evaluar) como a procesos. Recogemos estos aspectos para formularlos como recomendaciones. La disponibilidad pública de información acerca del instrumento (indicadores de validez y confiabilidad); la incorporación de desempeños y la diversificación del tipo de ítems, así como la necesaria participación de las instituciones formadoras en el proceso. Asimismo, se sugiere mejorar el tipo de información que comunican los informes de resultados, incorporando un juicio global que indique si se ha alcanzado el nivel de logro esperado y transitar a una lógica descriptiva de desempeños que permita interpretar los resultados y generar procesos de mejora.

En términos operativos, sugieren mejorar, adelantar y diversificar la información previa relativa al instrumento, proceso evaluativo y resultados. También se recomienda mejorar e igualar las condiciones de aplicación, extendiendo los días y flexibilizando normas de aplicación.

Por otra parte, trascendiendo la voz de los actores, caben algunas recomendaciones más estructurales: parece urgente tensionar el multipropósito de un solo instrumento que pretende servir a la rendición de cuentas y a la vez a la mejora, a remediar debilidades formativas de una cohorte y a la vez promover procesos

de mejora permanente en los programas. Parece prudente reducir las expectativas de la política a las posibilidades del instrumento, optando, por ejemplo, por la mejora estructural de largo plazo.

Otra alternativa es asumir la complejidad de la docencia y diseñar un conjunto de instrumentos con fines formativos, considerando evaluaciones con evidencias directas e indirectas, así como centralizadas y descentralizadas. Avanzar a una modalidad de evaluación diversificada y participativa en que la evaluación de desempeños puede abordarse en las líneas de prácticas, a partir de una construcción colaborativa que resguarde sellos formativos particulares.

Esto significa diseñar un sistema formativo que incluya procesos evaluativos capaces de generar confianza de las instituciones formadoras integrándolas al proceso, devolviendo información descriptiva útil para comprender resultados y promover la mejora, y acompañando dicha mejora al promover el fortalecimiento de capacidades institucionales.

Con todo, una condición para cualquier sistema evaluativo es la de contar con la confianza en la política y en la validez y confiabilidad de sus instrumentos. Esto demanda hacer públicos los elementos del proceso evaluativo, aunque esto apareja una paradoja: el riesgo de performatividad que deriva en estrechamiento curricular de conocerse las tablas de especificaciones del instrumento y sus rúbricas de evaluación. Mientras tengamos un sistema altamente competitivo entre instituciones de educación superior, este riesgo se incrementa. Si los estándares se asocian a la END, en cuyo sentido se avanza actualmente, se incrementa la probabilidad de incentivo a preparar para la prueba, cuestión que tensiona la conveniencia de optar por dicho camino.

Entonces, parecen obligadas algunas preguntas fuera de paradigmas evaluativos: ¿hay que evaluar? ¿Por qué? ¿La mejora de la formación inicial de docentes requiere una evaluación externa, centralizada y estandarizada? Existe experiencia de países que optaron por robustecer sus sistemas formativos y confiar en las capacidades

de las instituciones y formadores, apoyando al fortalecimiento de programas y equipos. Plantearse estas preguntas implica desnaturalizar dicha idea.

En este escenario, parece clave promover mayor investigación acerca de las implicancias de esta evaluación para la mejora, y robustecer el incipiente campo investigativo relativo a cómo se construye evidencia confiable acerca de los resultados de la formación, para informar mejor las decisiones acerca de la política en Chile y considerar los procesos de otros países que observan su implementación y resultados, evaluando su eventual replicabilidad.

## Referencias

- Ávalos, B. (2014). La formación inicial docente en Chile: Tensiones entre políticas de apoyo y control. *Estudios pedagógicos (Valdivia)*, 40(Especial), 11-28. <https://doi.org/10.4067/S0718-07052014000200002>
- Bartlett, M., Otis-Wilborn, A., & Peters, L. (2017). Bending or breaking: Appropriating edTPA policy in special education teacher education. *Teacher Education and Special Education*, 40(4), 287-298. <https://doi.org/10.1177/0888406417720188>
- Berliner, D. (2011). Rational responses to high stakes testing: the case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education*, 41(3), 287-302. <https://doi.org/10.1080/0305764X.2011.607151>
- Cabezas, V. y Claro, F. (2011). Valoración social del profesor en Chile: ¿cómo atraer a alumnos talentosos a estudiar pedagogía?. *Temas de la agenda pública, Centro de Políticas Públicas UC*, 6(42). <https://politicaspUBLICAS.uc.cl/wp-content/uploads/2015/02/valoracion-social-del-profesor-en-chile.pdf>
- Cabezas, V. Medina, L., Müller, M. y Figueroa, C. (2019). Desafíos y tensiones entre las nuevas políticas educativas y los programas de formación inicial de profesores en Chile. *Temas de la Agenda Pública*, 14(116), 1-30. Centro de Políticas Públicas UC. <https://politicaspUBLICAS.uc.cl/wp-content/uploads/2019/10/Paper-serie-Temas-de-la-Agenda-Pu%CC%81blica-N%C2%B0-116-2.pdf>
- Clayton, C. (2018). Policy meets practice in New York State: Understanding early edTPA implementation through preservice candidates' eyes. *Teacher Education Quarterly*, 45(3), 97-125. <https://files.eric.ed.gov/fulltext/EJ1185422.pdf>

- Cohen, J., Hutt, E., Berlin, R. L., Mathews, H. M., McGraw, J. P. & Gottlieb, J. (2020). Sense Making and Professional Identity in the Implementation of edTPA. *Journal of Teacher Education*, 71(1), 9-23. <https://doi.org/10.1177/0022487118783183>
- Conway, P. F. (2013). Cultural flashpoint: The politics of teacher education reform in Ireland. *The Educational Forum*, 77(1), 51-72. <https://doi.org/10.1080/00131725.2013.739021>
- Conway, P. F. & Murphy, R. (2013). A rising tide meets a perfect storm: new accountabilities in teaching and teacher education in Ireland. *Irish Educational Studies*, 32(1), 11-36. <https://doi.org/10.1080/03323315.2013.773227>
- Cox, C., Meckes, L. y Bascopé, M. (2010). La institucionalidad formadora de profesores en Chile en la década del 2000: Velocidad de mercado y parsimonia de las políticas. *Pensamiento Educativo*, (46-47), 205-245.
- Creswell, J. (2008). *Mixed Methods Research: State of the Art*. [Power Point Presentation]. University of Michigan. [sitemaker.umich.edu/creswell.workshop/files/creswell\\_lecture\\_slides.ppt](http://sitemaker.umich.edu/creswell.workshop/files/creswell_lecture_slides.ppt)
- CPEIP. (2017). *Informe Nacional Evaluación Nacional Diagnóstica de la FID 2016*. <https://www.cpeip.cl/wp-content/uploads/2019/03/Evaluaci%C3%B3n-Diagn%C3%B3stica-FID-2016.pdf>
- CPEIP. (2021). *Estándares para Carreras de Pedagogía*. Ministerio de Educación de Chile.
- Darling-Hammond, L. (2012). *Educación con calidad y equidad: los dilemas del siglo XXI*. Fundación Chile.
- Darling-Hammond, L. y Lieberman, A. (2012). *Teacher education around the World: Changing policies and practices*. Routledge.
- Darling-Hammond, L., D. Burns, C. Campbell, A. L. Goodwin, K. Hamerness, E. L. et al., (2017). *Empowered Educators: How Leading Nations Design Systems for Teaching Quality*. San Francisco, CA: Jossey-Bass.
- Demarchi, G. (2020). La evaluación desde las pruebas estandarizadas en la educación en Latinoamérica. *Revista En-Contexto*, 8(13), 107-133. <https://doi.org/10.53995/23463279.716>
- Díaz, C. y Navarro, P. (1998). Análisis de contenido. En J. M. Delgado y J. Gutiérrez, *Métodos y Técnicas Cualitativas de investigación en Ciencias Sociales*. Eds. Síntesis.
- Donoso, F. y Ruffinelli, A. (2020). ¿Hacia una Formación Inicial Docente de calidad?: La Evaluación Nacional Diagnóstica en las voces de actores partícipes del proceso de redacción e implementación de la Ley

- 20.903. *Revista de Estudios y Experiencias en Educación (REXE)*, Vol. 19, Núm. 41, 125-147. <https://doi.org/10.21703/rexe.20201941donoso8>
- Ehren, M. & Baxter, J. (2021). Trust, capacity and accountability. A triptych for improving learning outcomes. In *Trust, Accountability and Capacity in Education System Reform: Global Perspectives in Comparative Education*. Routledge.
- Feldfeber, M. (2007). La regulación de la formación y el trabajo docente: un análisis crítico de la “agenda educativa” en América Latina. *Educação e Sociedade*, 28(99), 444-465.
- Fernández, B. y Madrid, R. (2020). Profesionalización desde una perspectiva democrática: nudos críticos y propuestas para formación y desarrollo docente. En: M. T. Covera y G. Muñoz (eds.), *Horizontes y Propuestas para Transformar el Sistema Educativo Chileno* (pp. 206-233). Ediciones Biblioteca del Congreso Nacional de Chile, Colección Senado.
- Förster, C., Ruffinelli, A., Cisternas, T. y Donoso, F. (en prensa). Assessment of Undergraduate Teacher Education from the Subjects’ Perspective: A Literature Review (Artículo académico inédito).
- Greenblatt, D. (2018). Neoliberalism and teacher certification. *Policy Futures in Education*, 16(6), 804-827. <https://doi.org/10.1177/1478210318771827>
- Hilton, A. L., Saunders, R. & Mansfield, C. (2020). “In LANTITE, no one can hear you scream!” student voices of high-stakes testing in teacher education. *Australian Journal of Teacher Education*, 45(12), 57-72. <https://doi.org/10.14221/ajte.202v45n12.4>
- Ingvarson, L. & Rowley, G. (2017). Quality Assurance in Teacher Education and Outcomes: A Study of 17 Countries. *Educational Researcher*, 46(4), 177-193. <https://doi.org/10.3102/0013189X17711900>
- Kessler, M. A. (2019). The regulatory effects of high-stakes accountability in preservice teacher evaluation. *Teaching Education*, 1-17. <https://doi.org/10.1080/10476210.2019.1685485>
- Ledwell, K. & Oyler, C. (2016). Unstandardized Responses to a “Standardized” Test: The edTPA as Gatekeeper and Curriculum Change Agent. *Journal of Teacher Education*, 67(2), 120-134. <https://doi.org/10.1177/0022487115624739>
- Lerman, S. (2014). Mapping the effects of policy on mathematics teacher education. *Educational Studies in Mathematics*, 87(2), 187-201. <https://doi.org/10.1007/s10649-012-9423-9>
- Montecinos, C. (2014). Análisis crítico de las medidas de presión propuestas para mejorar la formación inicial de docentes en Chile por el Panel

- de Expertos para una Educación de Calidad. *Estudios pedagógicos* (Valdivia), 40(Especial), 285-301. <https://doi.org/10.4067/S0718-07052014000200017>
- O'Neill, O. (2013). Intelligent accountability in education. *Oxford Review of Education*, 39(1), 4-16. <https://doi.org/10.1080/03054985.2013.764761>
- Ruffinelli, A. (2017). Gubernamentalidad, pedagogía neutra y (des) profesionalización docente. *Educ. Soc., Campinas*, v. 38, n°. 138, 191-206. <https://www.scielo.br/j/es/a/dFGHH8Xm68Rf6Q4VfwDsQpf/?format=pdf&lang=es>
- Schmier, S. A. (2019). Performing the performance assessment. *English Teaching*, 18(3), 382-396. <https://doi.org/10.1108/ETPC-11-2018-0104>
- Vaillant, D. (2018). *Estudio exploratorio sobre los modelos organizacionales y pedagógicos de instituciones dedicadas a la formación docente inicial: un análisis en clave comparada*. Informe final. Buenos Aires: INFOD e IPE-UNESCO.
- Wilson, A. & Goff, W. (2019). “Hopefully, I Will Gain Confidence”: Hope in Pre-Service Teachers’ Mathematics and Numeracy Testing. *Australian Journal of Teacher Education*, 44(10). <https://doi.org/10.14221/ajte.2019v44n10.4>
- Zeichner, K. & Hollar, J. (2016). Developing professional capital in teaching through initial teacher education: comparing strategies in Alberta Canada and the U.S. *Journal of Professional Capital and Community*, 1(2). <https://doi.org/10.1108/JPC-01-2016-00>

Recibido: 04/10/2023

Aceptado: 20/11/2023