

EVALUACIÓN ESCOLAR: EXPLORANDO LA REALIDAD DE LA PRÁCTICA EVALUATIVA EN UN COLEGIO CHILENO

SCHOOL ASSESSMENT: EXPLORING THE REALITY OF ASSESSMENT PRACTICE IN A CHILEAN SCHOOL

Tamara Fabiola Cepeda Acevedo

University of Bristol, UK.

Resumen

Este artículo explora las prácticas evaluativas en un colegio chileno, con el objetivo de conocer perspectivas docentes y estudiantiles al respecto, métodos utilizados y resultados obtenidos, en términos de notas y retroalimentación. El marco conceptual del estudio en el que se basa el artículo considera los aspectos claves de la evaluación (*School Based Assessment*) según literatura internacional, que luego son contextualizados en la realidad chilena, considerando como relevante el concepto de evaluación sociocultural (Elwood y Murphy, 2015) y responsiva (Brown et al., 2022). El estudio busca identificar las prácticas evaluativas usadas en la realidad, desde lo sugerido por el decreto de evaluación que rige al país. Se trata de un estudio cualitativo exploratorio que utiliza análisis de documentos y entrevistas semiestructuradas a 10 docentes y 10 estudiantes de un colegio en Santiago, Chile, que fueron analizadas usando análisis temático. Los resultados indican que las intenciones de la política evaluativa no se ven reflejadas en

la práctica, principalmente por condiciones históricas externas impuestas al proceso de enseñanza: falta de tiempo docente, sobrecarga laboral y la inequidad social reinante en el país. La fuerte cultura sumativa y el sistema de consecuencias (*high stakes*) dificultan la evaluación formativa. Finalmente, se sugieren algunos elementos a considerar para que el sistema evaluativo apunte a ser socioculturalmente responsivo, es decir, que responda a las necesidades de los/as docentes y los/as estudiantes.

Palabras clave: política en evaluación, evaluación escolar, evaluación sumativa, evaluación formativa, sistemas de consecuencias.

ABSTRACT

*This article explores assessment practices in a Chilean school with the aim of understanding teachers and students perspectives on these practices, methods used, and results obtained in terms of grades and feedback. The conceptual framework of the study on which the article is based considers the key aspects of evaluation (School Based Assessment) according to international literature, which are then contextualized in the Chilean reality, considering the concepts of sociocultural (Elwood and Murphy, 2015) and responsive (Brown et al., 2022) evaluation as relevant. The study seeks to identify the evaluation practices used in reality, based on the suggestions of the evaluation decree that governs the country. This is an exploratory qualitative study that uses document analysis and semi-structured interviews with 10 teachers and 10 students from a school in Santiago, Chile, which were analyzed using thematic analysis. The results indicate that the intentions of the evaluation policy are not reflected in practice, mainly due to external historical conditions imposed on the teaching process: lack of teacher time, work overload, and the prevailing social inequality in the country. The strong summative culture and the system of consequences (*high stakes*) hinder formative evaluation. Finally, some elements are suggested to consider so that the evaluation system aims to be socioculturally responsive, that is, that it responds to the needs of teachers and students.*

Keywords: policy in assessment, school based assessment, summative assessment, formative assessment, high-stakes systems.

School Based Assessment (SBA de ahora en adelante) es el término usado internacionalmente para referirse a los sistemas donde la evaluación es llevada a cabo por los colegios, por los/as docentes de cada asignatura (Hong Kong Examination and Assessment Authority,[HKEAA], 2025). Este tipo de evaluación se contrapone a la evaluación externa, ya que permite mayor variedad de procedimientos evaluativos (Yates y Johnston, 2018). El concepto SBA fue presentado por primera vez en Hong Kong en el año 2000, entendido como un proceso evaluativo llevado a cabo por los docentes en cada asignatura, caracterizado por ayudar a los y las estudiantes a entender sus fortalezas y debilidades a través de la retroalimentación (HKEAA, 2021). Como se puede apreciar, este tipo de evaluación entrega una perspectiva integrada, incorporando elementos formativos y sumativos, que es la misma perspectiva del decreto que rige los procesos evaluativos en Chile (Ministerio de Educación de Chile, [Mineduc], 2018) y es usada en lugares como Australia (Cumming y Maxwell, 2004), Escocia (Scottish Government, 2011), Finlandia (Mäkipää y Ouakrim-Soivio, 2019), Hong Kong (Cheng, 2019), Indonesia (Puad y Ashton, 2021), Irán (Rasooli et al., 2023), Suecia (Hirsh, 2020) y Nueva Zelanda (Yates y Johnston, 2018). SBA entrega información esencial respecto al aprendizaje de los y las estudiantes, en forma de notas obtenidas por distintas actividades diseñadas por los y las docentes en el ambiente natural de la sala de clases (Cheng, 2019; HKEAA, 2021; Murchan, 2018). Sin embargo, SBA también incluye retroalimentación detallada que motiva al/la estudiante a trabajar en áreas donde necesita mejorar (Cheng, 2019). Así, se entiende que este tipo de evaluación sirve a propósitos sumativos y formativos (Cheng et al., 2011; Cheng, 2019). La literatura sugiere que la evaluación formativa se da cuando el proceso de aprendizaje está ocurriendo (Black y William, 2009), mientras que la evaluación sumativa ocurre después de un cierto período de tiempo, por ejemplo, al final de una unidad (Mäkipää y Ouakrim-Sovio, 2019). La evaluación sumativa se suele relacionar con logros generales en términos de productos concretos y valorables para asignar notas, certificar el logro de los y las estudiantes y reportar (Briggs y Ellis, 2013; Meng et al., 2014; Shepard, 2019; Stiggins, 2002). Por otro lado, la evaluación formativa destaca la participación del estudiante en el proceso de evaluación a través de distintas actividades que son claves para monitorear el aprendizaje (Black y William, 1998; Cheng, 2019; Tong y Adamson, 2015). Los resultados de estas actividades son usados para mejorar el proceso de enseñanza y, por extensión, los resultados

de aprendizaje de los y las estudiantes (Black y William, 1998; Briggs y Ellis, 2013; DeLuca et al., 2018; Shepard, 2019). Precisamente esta característica destaca la política evaluativa chilena, invitando a los colegios a implementar decretos internos que potencien y utilicen ambos tipos de evaluación, a través de una variedad de métodos de evaluación (Mineduc, 2018).

Sin embargo, la literatura muestra dificultades y desafíos en la implementación y uso de evaluación formativa en sistemas de altas consecuencias (o *high-stakes*) (Mujica-Johnson, 2020; Tan, 2017). Según el decreto de evaluación, las notas que arroja SBA implican la posibilidad de repetir el año escolar (Mineduc, 2018). Además, las notas también implican un porcentaje del puntaje que es usado para ingresar a la universidad, a través del NEM (Notas de Enseñanza Media) y el ranking (www.demre.cl). Por lo tanto, la particularidad del sistema evaluativo de cada país implicará que las reacciones a las estrategias formativas difieran enormemente, según sus intenciones y acciones, las percepciones de la evaluación y también las salas de clases y escuelas (DeLuca et al., 2018; Kennedy et al., 2008). Es importante destacar que muchos autores consideran que la dicotomía formativa/sumativa es problemática, ya que los conceptos se separan y muestran casi como incompatibles (Taras, 2005; Kennedy et al., 2008). Esto, sumado a que el campo de la evaluación implica de por sí un complejo proceso que incluye la recolección, análisis y evaluación de la evidencia sobre el proceso de aprendizaje del estudiante (Remesal, 2011). Por ello, autoras como Elwood y Murphy (2015) señalan que la dicotomía es deficiente e improductiva y que, en cualquier caso, siempre debe ser considerada de acuerdo a los contextos culturales.

Un enfoque socioculturalmente responsivo a la evaluación

Hace más de 20 años, Broadfoot (1999) reconocía que la evaluación es un producto social impactado por valores y tradiciones de una cultura particular, donde los intereses de sus distintos actores se combinan de forma diferente para definir conceptos como calidad o mérito. Por lo tanto, este artículo reconoce la relevancia e impacto de la cultura al implementar la práctica evaluativa que, en teoría, esta propuesta como una combinación de estrategias formativas y sumativas. Elwood y Murphy (2015) también enfatizan que la práctica evaluativa necesita ser comprendida dentro de su contexto social,

histórico y cultural. Los procesos evaluativos dentro de la escuela están inmersos en una cultura y se espera que sean influenciados por distintos elementos de esa cultura, creando distintos significados o “guiones” respecto a qué es la evaluación y para qué sirve (Elwood y Murphy, 2015). Por lo tanto, es esperable y comprensible que una cultura tendrá determinadas preferencias respecto a la evaluación, creará ambientes de aprendizaje específicos y promoverá ciertas conductas entre docentes y estudiantes (Kennedy et al., 2008).

Estas ideas de respetar los ambientes socioculturales en los que se da la evaluación (y que la determinan) han sido complementadas recientemente por varias experiencias de evaluación culturalmente responsiva (Brown et al., 2022; Burns et al., 2019). En la experiencia de Brown en Irlanda, la evaluación responsiva es simplemente la que constituye buena, es decir, justa evaluación de aula para estudiantes que provienen de diversos contextos culturales. Uno de los puntos centrales de esta experiencia es el diseño de evaluaciones que realmente respondan a las necesidades de los y las estudiantes, mientras que, a la vez, les entrega información relevante sobre su proceso de aprendizaje (Brown et al., 2022). La idea de evaluación formativa es una idea que ha sido implementada vertiginosamente desde que Black y William (1998) establecieron que producía mejoras significativas en el aprendizaje. Sin embargo, es una idea occidental (Kennedy et al., 2008) y, por lo tanto, cabe preguntarse ¿cómo se adapta una idea occidental a la realidad evaluativa chilena, de acuerdo a los propósitos, métodos e información propia que SBA promueve en nuestro país? Esta pregunta se relaciona con la forma en que esa idea puede ser introducida en una cultura evaluativa donde predominantemente ha reinado la nota como único lenguaje de comunicación (Flórez Petour, 2014).

Este artículo muestra una mirada inquisitiva en la forma en que SBA y la idea de evaluación formativa pueden introducirse en una cultura donde el guion evaluativo (Elwood y Murphy, 2015) nos ha enseñado que la evaluación es y se usa para lo sumativo (o “poner notas”).

Evaluación en Latinoamérica y Chile

Cuando se estudia la evaluación en Latinoamérica y en Chile, el foco de la investigación no siempre es el mismo que el internacional. Hay pocos estudios en Chile que describan la realidad evaluativa dentro

de la sala (Castillo et al., 2023; Maldonado-Fuentes, 2021; Martínez Quiñones, 2019; Mineduc, 2016; Peters Rodríguez y Contreras, 2019; Ravela et al., 2014; Vera et al., 2017). Ravela et al (2014) comparan procesos evaluativos en Chile, Colombia, Perú y Uruguay. Prieto y Contreras (2008) analizan concepciones docentes en estudiantes de pedagogía, es decir, previo a la real experiencia de evaluar en la sala de clases. Recientemente, Castillo et al (2023) realizaron un estudio cuantitativo exploratorio buscando identificar concepciones y estrategias de evaluación formativa durante la pandemia por COVID 19. Por ejemplo, el estudio realizado por Martínez Quiñones (2019) en Santiago, Chile, sobre prácticas evaluativas en un colegio, identifica distintos propósitos evaluativos, como medición y clasificación, pero también mejora de la enseñanza y el aprendizaje, como parte de los cambios y las nuevas políticas educacionales del país. Sin embargo, no se usan los conceptos “formativo” o “sumativo”. Adicionalmente, hay evidencia, en países como Perú y México, de que la evaluación formativa no ha sido correctamente promovida para poder ser aplicada en la realidad de las escuelas (Falcón et al., 2021; Monzón-Troncoso, 2015). El tono de la discusión sobre evaluación no siempre está centrado en lo formativo y lo sumativo y exhibe complejidades. Sin embargo, el Mineduc promueve ambos conceptos y métodos asociados en el decreto N° 67 que rige los procesos evaluativos en los colegios (Mineduc, 2018).

Este artículo explorará propósitos, métodos, e información producida por la evaluación realizada en un colegio chileno, con el fin de conocer las características del proceso en la práctica real dentro del contexto cultural de nuestro país, considerando lo que es sugerido en la teoría por el decreto de evaluación y tendencias internacionales en evaluación.

Propósitos evaluativos, métodos, retroalimentación y notas

Usualmente la literatura reporta propósitos evaluativos que van mucho más allá de lo formativo y lo sumativo, y que son más complejos. Harris y Brown (2009) indican que los propósitos suelen ser multifacéticos e incluir, por ejemplo, motivar a los y las estudiantes o individualizar el aprendizaje. Los autores reportan fuertes tensiones acerca de lo que los/as profesores/as sienten que es mejor para sus estudiantes y lo que es necesario para los propósitos de rendición de cuentas

(*accountability*) del colegio. Zhao et al. (2017) también destacan la compleja naturaleza de los propósitos evaluativos. Por lo tanto, es interesante indagar qué propósitos identifican los y las docentes, y estudiantes chilenos para la evaluación (que será la primera pregunta de investigación del estudio).

Respecto a los métodos de evaluación, el decreto N°67 promueve varios (Mineduc, 2018), al igual que SBA que plantea recolectar múltiples recursos de evidencia sobre el aprendizaje de los y las estudiantes, dentro el ambiente natural de la sala de clases, durante el año escolar (Davidson, 2009). Sin embargo, la literatura muestra, consistentemente que, aunque los docentes declaran usar métodos de evaluación variados, en la práctica usan métodos tradicionales como pruebas de selección, papel/lápiz (Bonner et al., 2018; Bota y Tulbure, 2015; Mäkipää y Ouakrim-Sovio, 2019; Puad y Ashton, 2021; Smyth y Banks, 2012). Los métodos evaluativos serán el segundo foco de interés de este artículo.

Los instrumentos de evaluación finalmente entregan información valiosa sobre el aprendizaje de los y las estudiantes. En el caso de SBA, se obtienen (o deberían obtener) notas y retroalimentación, lo que constituye el tercer foco de interés y pregunta de este artículo. En ese sentido, se cuestiona que condicionen a los y las estudiantes a estudiar para “pasar”, más que para mejorar su aprendizaje (Harlen, 2004). Por lo tanto, es esperable que estas notas causen efectos en los y las estudiantes como, por ejemplo, motivar a los de mejor rendimiento y desmotivar a los de bajo rendimiento (Harris y Brown, 2009; Sivenbring, 2019). Por otro lado, las notas también se asocian a problemas de objetividad y aumento de la carga laboral docente (Harlen, 2004; Murchan, 2018). Shepard (2019) ha planteado el problema de asignar puntos por esfuerzo, hábitos de trabajo, colaboración y desempeño, lo que ha normalizado conductas que impactan en las notas, incluyendo en ellas dimensiones que no son necesarias ni defendibles. Estas prácticas, que usualmente se conocen como “*score pollution*” consisten en diferentes aspectos como esfuerzo, conducta o trabajo (Remesal, 2011). Por ejemplo, el caso de Indonesia es particularmente interesante, ya que se promovió una nueva política evaluativa en el año 2020 (mismo año del decreto N°67 de evaluación en Chile), y muestra un enfoque sumativo donde ciertas conductas y actitudes reciben notas, evidenciando control de conducta de los y las estudiantes a través de la evaluación (Puad y Ashton, 2021). En Chile, por ejemplo, la práctica se describe

a través de dar décimas que posteriormente incrementan las notas, en trabajos como el de Martínez Quiñones (2019) y Peters Rodríguez y Contreras (2019).

Finalmente, la retroalimentación ha sido descrita como un elemento clave e importante de la evaluación formativa, ya que la nota es insuficiente para mejorar el aprendizaje por sí misma (Sivenbring, 2019). Se describe impactando significativamente en el aprendizaje de los y las estudiantes, ayudándolos/as a avanzar (Black y William, 1998; Cheng, 2019; Stiggins, 2002; Tong y Adamson, 2015). Zhao et al. (2017) analizaron 266 artículos de evaluación en el aula con especial foco en la retroalimentación entregada por los docentes. La caracterizaron como relacionada con la tarea/proceso/persona, con su naturaleza (positiva/negativa y lo que necesita mejorarse), su forma (verbal/escrita), su tiempo (inmediata o tardía) y la adaptación de la enseñanza promovida por la retroalimentación. Estas categorías resultan interesantes para analizar las características de la retroalimentación en otros contextos educativos.

Diseño del estudio

Enfocándose en las características de SBA en un colegio chileno, este estudio adoptó una metodología cualitativa exploratoria. El paradigma del estudio es interpretativo e implica una ontología donde la realidad es socialmente construida por la gente que experimenta el mundo dentro de contextos específicos (Prasad, 2017). Epistemológicamente, el conocimiento sobre el mundo no es factual sino co-construido por la gente y especialmente influenciado por sus perspectivas culturales compartidas (Creswell, 2013). Autores como Raczynski y Salinas (2008) han destacado la urgencia de investigar cualitativamente el proceso educacional, especialmente con los actores involucrados dentro de la sala de clases. Por lo tanto, el estudio se fundamenta en el sujeto como ser que crea realidad y conocimiento a través de su experiencia viviendo directamente el proceso de evaluación, a través de entrevistas semi-estructuradas con 10 docentes y 10 estudiantes de educación media (n=20). Adicionalmente, para analizar y comprender los propósitos y métodos de evaluación usados, se analizaron documentos, creados en el colegio, es decir, sin intervención de la investigadora (Bowen, 2009). Los documentos analizados fueron el decreto interno de evaluación del colegio y 8 libros de clases de educación media (dos cursos

por nivel desde I° a IV° Medio). Los libros fueron seleccionados por contener registros hechos a mano por los docentes, ya que en el país ellos son libres de planificar sus clases siguiendo el currículum nacional y creando evaluaciones de acuerdo a sus propios criterios y objetivos (Mujica-Johnson, 2020). La información de los libros se registró a mano y luego fue organizada digitalmente en 8 archivos Word Office, uno por curso.

De esta forma, los objetivos del estudio fueron:

1. Analizar y describir cómo y por qué los profesores utilizan prácticas específicas de SBA y las perspectivas de ellos y sus estudiantes respecto a esas prácticas.
2. Conocer las perspectivas de profesores y estudiantes respecto a las notas y la retroalimentación que resulta del proceso de SBA.

La forma de hacer SBA es considerada en el contexto sociocultural en el que ocurre y que la sustenta. En este sentido, las perspectivas docentes en SBA son una parte fundamental de su labor y están fuertemente ligadas a su práctica (Pizarro Gamero y Gómez Muskus, 2019; Remesal, 2011). Remesal (2011), por ejemplo, notó que importantes estudios en evaluación, como los de Stiggins (1999) o McMillan (2001), no incluyen las creencias y concepciones docentes que guían su práctica evaluativa. Por otro lado, las voces de los y las estudiantes, quienes tienen una larga historia de presencia y colaboración en investigación educacional en educación básica y media, reflejan sus visiones respecto a sus experiencias de aprendizaje (Cook-Sather, 2018). Sin embargo, en el campo de la evaluación, un pequeño número de estudios han considerado sus perspectivas y concepciones de evaluación escolar (DeLuca et al., 2018; Sivenbring, 2019).

El estudio se focalizó en educación media ya que son los y las estudiantes expuestos de mayor forma a las consecuencias de las notas por el NEM y en ranking de ingreso a la universidad. Se eligió un colegio particular subvencionado, ya que es donde se encuentra concentrada la matrícula en el país (Acción Educar, 2021). En este artículo se presentan los resultados de tres preguntas de investigación:

1. ¿Cuáles son los propósitos evaluativos, según documentos del colegio y las perspectivas de docentes y de estudiantes?
2. ¿Cuáles son los métodos de evaluación usados, según documentos del colegio y las perspectivas de docentes y de estudiantes?

3. ¿Cuáles son las perspectivas de docentes y estudiantes respecto a la información entregada por la evaluación, en términos de notas y retroalimentación?

Las preguntas de la entrevista se basaron en las características de SBA según la literatura en el tema, las que permiten comprender los significados subjetivos que las personas dan a sus experiencias (Denzin y Lincoln, 2008; Hennink et al., 2011).

El análisis realizado fue análisis temático (Braun y Clarke, 2012), con utilización de NVivo *software* tras una primera fase de lectura interpretativa del set de datos. Este tipo de análisis permite identificar, analizar y reportar patrones de temas dentro de un set de datos (Bowen, 2009; Braun y Clarke, 2006; Braun y Clarke, 2012; Clarke y Braun, 2013). El set de datos de transcripciones de entrevistas (n=20) fue analizado en su conjunto, es decir, perspectivas de docentes (D) y estudiantes (E), en tanto experimentan el mismo fenómeno de evaluación en el mismo espacio, la sala de clases. Los archivos digitales con la información de los libros y el decreto de evaluación fueron analizados por separado. En ambos casos, el *software* NVivo permitió agrupar citas desde los documentos y transcripciones a través de patrones transversales que, posteriormente, permitieron triangular la información, al combinar métodos para indagar el mismo fenómeno (Denzin, 1970). Los patrones fueron identificados por su recurrencia en el set de datos, a través de una codificación flexible inductiva, y luego fueron nombrados como temas. Los temas fueron asociados a los datos mismos/*data-driven* (Braun y Clarke, 2006; Kiger y Varpio, 2020) ya que la estructura de análisis se derivó de ellos. Por ejemplo, NVivo permitió codificar 629 menciones en los documentos y 271 en las entrevistas para métodos de evaluación, 181 menciones para complejidades de poner nota y 46 para el impacto del currículum en la evaluación, tema que no estaba directamente relacionado con ninguna pregunta de entrevista. Estos temas fueron finalmente relacionados con las características de SBA, para presentar los resultados del estudio en términos de propósitos, métodos y resultados.

Resultados

Propósitos de la evaluación

Según el reglamento de evaluación del colegio, la evaluación tiene propósitos sumativos y formativos y sigue los lineamientos del decreto del Mineduc. Sin embargo, en las entrevistas se constata que el propósito de la evaluación para los docentes es, principalmente, cubrir el currículo y pasar contenidos de acuerdo a los planes y programas del Mineduc para cada asignatura (46 referencias). De los 10 profesores entrevistados, 6 identificaron inmediatamente los contenidos como objetivo de la evaluación: “En el caso de [nombre de la asignatura] siempre las actitudes y las habilidades van a ir asociadas a un contenido en particular” (D03). A la vez, cubrir los contenidos propuestos por el currículo se siente como una gran sobrecarga para estos docentes: “Jamás logramos la cobertura curricular, o sea, de cuatro unidades, a veces, con suerte, logramos dos” (D01).

Los y las estudiantes también identifican los contenidos como un elemento importante en el propósito de la evaluación: “Controlar que todos los estudiantes alcancen un nivel... pueden notarlo en las pruebas, cuando tú haces un ejercicio o puedes entender el contenido” (E02). También resultaban críticos al identificar que el objetivo de la evaluación debería ser monitorear aprendizaje, pero no lo es:

Por ejemplo, en [nombre de la asignatura] era pasar materia y hacer una prueba para que nos fuera bien a todos y no era preocuparse de que hubiéramos aprendido, en verdad, la materia, sino que era para tener buenas notas, entonces pasaba un contenido y la prueba era al tiro, para que el contenido estuviera fresco. Así lo hacían. (E03)

En general, aunque tanto docentes como estudiantes son capaces de hablar de aprendizaje y habilidades, el contenido o la “materia” va primero, y todo queda supeditado a ello.

Métodos evaluativos

Los métodos evaluativos, siguiendo el decreto de evaluación del colegio, son sugeridos como variados tanto para la evaluación sumativa como para la formativa. La descripción de ambos métodos de evaluación está contenida en la sección de “notas”.

En los métodos evaluativos, según los libros de clases, prima una cultura sumativa y de calificación (Tema 1, 293 menciones), con baja variedad (mayoritariamente pruebas de selección múltiple) y uso de formatos estandarizados (SIMCE y PAES). En segundo lugar, (Tema 2, 222 menciones) aparece el uso sumativo de lo formativo, es decir, actividades formativas tipo guías en clase, que reciben apoyo/ayuda del docente y compañeros/as de clase, pero que siempre son evaluados con una nota. Finalmente, el menor patrón temático es para las rúbricas (Tema 3, 110 menciones), que constituyen trabajos que evalúan desempeño bajo criterios, asociados a puntajes, que de todas formas siempre llevan una nota.

Cuando se le pregunta a docentes y estudiantes sobre los métodos evaluativos utilizados, los docentes reconocen que usualmente usan formatos clásicos de pruebas de alternativas, simplemente porque son formatos más rápidos y fáciles de revisar:

También hago prueba de selección múltiple y de ítem de desarrollo, pero no es lo más común porque, siendo súper honesta, para facilitarme la revisión, trato de no poner tantas preguntas de desarrollo. Cuando pongo preguntas de desarrollo son muy puntuales... En verdad uno en el tiempo va aprendiendo; yo, cuando empecé a ser profe, hacía tantas preguntas de desarrollo y en el fondo es contraproducente, te quedas pegado con las evaluaciones revisándolas mucho tiempo, entonces es difícil... (D03)

Los y las estudiantes también reconocen que estos formatos tienen que ver con el tiempo docente y la carga laboral, más que con razones pedagógicas: “Es más rápido para el profe, es más fácil para el profe, tener una prueba con todas las repuestas y poner si está buena o está mala” (E01).

Por otro lado, son capaces de cuestionar la alta prevalencia de uso de la prueba como método de evaluación: “Porque prueba, prueba, prueba, prueba, como que te estresa..., uno anda tenso y todo eso, entonces, si hubiera otros métodos para medir las capacidades que no fuera una prueba, creo que subirían las notas” (E10).

La revisión de los libros de clase (8 libros) en educación media, arrojó como segundo patrón el trabajo formativo con nota, de tipo guías en clase. Usualmente este trabajo es del mismo formato de la prueba, con la diferencia de que se hace en la clase, con apoyo (docente, pares, libros/cuaderno). Los y las estudiantes valoran este

tipo de actividades: “Me gustan porque me ayudan mucho y me hacen entender mejor porque en la guía puedo preguntar si no entiendo algo” (E08).

Respecto a las rúbricas como método de evaluación, pese a que usualmente son consideradas una herramienta formativa que permite guiar el trabajo de los y las estudiantes, estas se utilizan con puntajes asociados y, finalmente, una nota, y fueron relacionados con actividades de “desempeño”. Fue el tema con menos referencias en el análisis temático, e involucró actividades del tipo debate, laboratorio, creación de videos, exposiciones orales en ferias/stands, modelos, actuación y debates. Los y las estudiantes conocen previamente los criterios y puntajes asociados a cada criterio en la rúbrica y valoran estas oportunidades, pero siempre para “subir notas” y asociadas a algo distinto a los “contenidos”, que ya fueron mencionados como un elemento importante: “Es una forma de subir el promedio, es como trabajo, es fácil, pero la prueba es diferente porque es puro contenido” (E06). Por otro lado, para los docentes, la alta presencia de métodos de evaluación basados en pruebas impacta en los y las estudiantes, generando resistencia a estos ejercicios evaluativos con rúbricas:

De repente siento que hay más rechazo a las evaluaciones flexibles que a las pruebas, porque están muy acostumbrados a las pruebas, muy acostumbrados. Por ejemplo, me pasó en [nombre curso] el año pasado que me decían los chiquillos: “¡ay tía!, a mí me iba excelente en [nombre de asignatura]” y yo les decía: pero, ¿por qué? ‘no sé, tía, porque yo daba las pruebas y me iba bien’, entonces, pasa mucho que hay una resistencia a los trabajos, a que ‘me voy a tener que esforzar más’, a que ‘voy a tener que crear algo’, ‘qué lata’, en verdad, una prueba es más fácil. (D05)

En general, los métodos evaluativos están marcados por la importancia de los contenidos, la falta de tiempo que hace que los/as docentes elijan formatos rápidos de revisión a los que los y las estudiantes están acostumbrados/as, y la gran relevancia que tienen las notas dentro de la cultura escolar.

Información producida por la evaluación

El estudio monitoreó la información que entrega el proceso evaluativo que, según la intención del decreto del Mineduc, debería presentarse

en formato notas (aspecto sumativo) y retroalimentación (aspecto formativo).

En primer lugar, respecto a las notas, se constata una fuerte cultura sumativa donde todo debe llevar nota, de otra forma los y las estudiantes no trabajan: “Lamentablemente en el contexto en el que trabajamos tiene que ser todo evaluado porque si no, el estudiante no le ve un fin a hacerlo de manera personal” (D01). Otro docente comenta la misma idea: “Ellos estudian, o lo que sea, hacen el trabajo, solo por la nota” (D08). A su vez, las notas reflejan varias consecuencias negativas en las entrevistas, tanto para los docentes como para los y las estudiantes. Para los profesores, son un vicio del sistema y una amenaza, una forma de control y presión. Los y las estudiantes reconocen, por otro lado, que son estereotipados de acuerdo a ellas. El proceso de evaluación sumativa genera estrés en los y las estudiantes e impacta en su autoestima: “había días en que me sentía súper vulnerable en ese sentido, me sentía tonta, me sentía... no estaba motivada a mejorar en mis estudios” (E04); “Me hubiera encantado tener mejores notas... sí, para sentirme mejor conmigo misma” (E10).

Adicionalmente, los participantes opinaron respecto al impacto de las notas en la repitencia y el ingreso a la universidad. Algunos profesores creen que los y las estudiantes “realmente merecen o necesitan” (D04) repetir. Las consecuencias para ellos dicen relación con ser estereotipados y descomprometerse del sistema educacional:

Yo siento que la repitencia es un vicio del sistema porque no necesariamente va a generar un cambio de actitud en el estudiante. Muchas veces genera lo contrario, un rechazo a la escuela. Es sacarlo de su contexto, separarlo de sus amistades, se ve súper perjudicado, genera rechazo en los compañeros del curso a donde llega el alumno, todos saben que es repitente, entonces... ‘es que debe ser flojo’, nadie quiere trabajar con él, ‘es que debe ser medio tonto’, y así se genera una serie de prejuicios. Hay alumnos que no logran encajar en otros cursos. Siento que es una falla del sistema porque eso significa que no hubo un monitoreo adecuado, no solo del colegio o de los profes, no le estoy echando la culpa al profe jefe que no se dio cuenta porque uno suele saber para donde va el asunto, pero también hay un tema familiar de por medio que, lo más probable, no supimos leer con claridad. (D05)

Por otro lado, los profesores reconocen que los métodos de evaluación poco variados podrían tener un impacto en la repitencia por las pocas posibilidades de diversificación para lograr que el/la estudiante logre un nivel adecuado para ser promovido. Los y las estudiantes también son capaces de reconocer este factor, identificando que con otra evaluación podrían tener otros resultados: “La verdad es que ellos son gente que puede avanzar y que puede aprender, pero en otras formas, no en las formas en que ellos nos evalúan” (E01).

Sin embargo, el estudio constató que los profesores suelen usar sistemas de décimas para ayudar a sus estudiantes a subir sus notas, con actividades que les permiten obtenerlas y utilizarlas para mejorar, especialmente, las notas de las pruebas, lo que invita a reflexionar sobre la objetividad del sistema.

Respecto a la retroalimentación, el decreto de evaluación contiene escasa información respecto a sus características, solo se limita a mencionar que se debe entregar. Sin embargo, debido a falta de tiempo, los profesores reconocen que no es una práctica que puedan realizar y que generalmente se limitan a entregar la nota, sin mayores comentarios que permitan al estudiante mejorar o avanzar: “Con las pruebas escritas, debo reconocer que a mí me cuesta un poco más hacer retroalimentación, siento que es instrumento más útil cuando uno está apurado, entonces, la verdad es que ahí me caigo” (D05).

Por las pocas experiencias relatadas, la retroalimentación se limita a revisar las respuestas de las pruebas, identificando errores, o haciendo comentarios del tipo “felicitaciones” en los que obtienen altas notas, pero que no son significativos para avanzar en el aprendizaje. También se constata retroalimentación verbal durante los trabajos del tipo guía (uso sumativo de lo formativo) que los alumnos declararon preferir por recibir ayuda de docentes y de sus pares. Sin embargo, no identifican la estrategia como retroalimentación ni como evaluación formativa y no hay una sistematización de la estrategia intencionada desde el decreto, que permita que se use para mejorar el aprendizaje.

Discusión

Si bien el objetivo de este artículo es conocer la práctica evaluativa en un colegio chileno según el modelo que promueve SBA, hay que hacer algunas precisiones respecto a las condiciones estructurales y

de política evaluativa. En primer lugar, tanto para las características de la evaluación en términos de propósitos y métodos, este estudio constata una desconexión entre lo que establece la política evaluativa (Mineduc, 2018) y la práctica en la sala de clases. Pese a que el foco del estudio no era directamente la política educacional, es importante destacar que no se ha prestado suficiente atención a la realidad de los colegios. Autores como Maroy y Pons (2021) destacan la importancia de traducir y adaptar las políticas a escala, de acuerdo a los contextos en donde serán implementadas y a los actores que las implementarán. Esta idea tiene sentido en el caso chileno, donde el Mineduc ha propuesto, desde hace algunos años, la evaluación formativa en un contexto educacional fuertemente marcado por la evaluación sumativa, es decir, en una cultura donde el guion, recordando a Elwood y Murphy (2015), ha estado dictado por números, que además están enmarcados por otras condiciones, como la falta de tiempo docente y la sobrecarga laboral. Esta realidad no se ha considerado de forma tal que la evaluación formativa tenga un espacio donde pueda volverse una realidad y donde genere oportunidades de mejora en el aprendizaje de los y las estudiantes.

La principal filosofía de trabajo y propósito para los/as docentes es cubrir el currículum, lo que se entiende porque, volviendo al contexto sociocultural de la evaluación sugerido por Elwood y Murphy (2015), saber en Chile es saber contenidos, por lo tanto, el propósito de la evaluación está centrado en el objeto (el currículum) y no en el sujeto (el/la estudiante y sus habilidades). En este ambiente, la interacción formativa y sumativa es compleja, ya que las condiciones no están dadas para que lo formativo tenga reales oportunidades de ocurrir de forma significativa para el aprendizaje de los y las estudiantes.

Respecto a los métodos evaluativos usados por los docentes, también se constata una desconexión entre lo que sugiere la política y la práctica. Se usan métodos tradicionales y poco variados (Bonner et al., 2018; Bota y Tulbure, 2015; Mäkipää y Ouakrim-Sovio, 2019; Puad y Ashton, 2021; Smyth y Banks, 2012). Las instancias formativas, como guías de trabajo en clase o rúbricas, al estar asociadas a notas, promueven que el estudiante se enfoque solo en el logro de la nota (y aprobar) y no en la mejora del aprendizaje (Harlen, 2004).

La información entregada por la evaluación está fuertemente marcada por las notas, es decir, por el lenguaje numérico que ha sido el predominante, históricamente, en el contexto escolar (Flórez Petour, 2014). El guion sociocultural respecto a qué es la evaluación

y para qué sirve (Elwood y Murphy, 2015) le ha enseñado a esta comunidad que lo que importa son las notas. Y esto es comprensible dentro de un sistema de altas consecuencias: si las notas implican consecuencias, en ellas nos enfocaremos. Sería injusto pedirles a estos/as docentes o a los y las estudiantes que fuera de otra forma, si se juegan pasar de curso, el NEM, el ingreso a la universidad y la fama/prestigio del colegio.

El estudio también constata consecuencias negativas asociadas a las notas, como han corroborado otros estudios previamente (Harris y Brown, 2009; Sivenbring, 2019). Además, se constatan estrategias de contaminación de notas (*score pollution*) a través de sistemas de otorgamiento de décimas (Martínez Quiñones, 2019; Peters Rodríguez y Contreras, 2019), que permiten cuestionar la objetividad del sistema. Sin embargo, estas estrategias creativas son vistas como una forma que los/as profesores/as tienen de ayudar a sus estudiantes a navegar las reglas de este sistema de altas consecuencias. En este sentido, la estrategia desafía lo que es considerado “incorrecto” o “correcto” y también muestra posturas éticas docentes al momento de evaluar a sus estudiantes, pese a no haber sido preparados éticamente para tomar decisiones evaluativas (Rasooli et al., 2023), lo que es un interesante aspecto de la evaluación para futura investigación en el área.

Finalmente, respecto a la retroalimentación, el sistema no tiene condiciones ni estructurales ni sistémicas que protejan su ejecución dentro del proceso evaluativo. Además, la retroalimentación es una estrategia formativa que está basada en el lenguaje y el sistema evaluativo chileno tiene un fuerte componente cultural sumativo, es decir, está basado en números. Cambiar la evaluación socioculturalmente desde números a palabras, o abrir espacios para la existencia de ambos, requiere esfuerzos desde la política educativa, protegiendo las condiciones necesarias para que ese lenguaje ocurra y permita potenciar, mejorar y apoyar el proceso de aprendizaje de los y las estudiantes.

Conclusión

Sabemos cómo funciona la evaluación escolar, pero ¿cómo debería ser una evaluación socioculturalmente responsiva en el país, es decir, pensada en lo que los y las estudiantes y docentes necesitan? (Brown et al., 2022; Burns et al., 2019; Elwood y Murphy, 2015). Este artículo

muestra cómo es el panorama real evaluativo en los colegios, así como también las dificultades que emanan del uso de SBA, que muestra una separación abismal entre lo sugerido en la política evaluativa y la práctica en los colegios. ¿Cómo subsanar esta separación?

En primer lugar, la política evaluativa debe reconocer la realidad de los colegios y las necesidades de los y las estudiantes. Para esto, es clave el concepto de “traducción de política” (Maroy y Pons, 2021) que apunta a que la política evaluativa no sea copiada de países extranjeros sin un proceso de adaptación a nuestra realidad. Sobre todo, para conceptos occidentales como la evaluación formativa (Kennedy et al., 2008) que necesitan ciertas condiciones para poder ser ejecutados en los sistemas educacionales. Esto implica reconocer también las condiciones históricas en las que se ha desarrollado la labor docente en el país, por ejemplo, la falta de tiempo de los y las docentes. Implementar nuevas políticas (o lo formativo) en esas condiciones resulta simplemente imposible.

En segundo lugar, del punto anterior se desprende que los y las docentes necesitan tiempo para poder adaptar sus prácticas y métodos evaluativos de acuerdo a las necesidades de sus estudiantes. Esto implica poder diversificar la evaluación pensando en lo que es mejor para la realidad de cada curso y cada estudiante de acuerdo a la comunidad en la que el colegio está inserto. Este trabajo es posible solo contando con el recurso esencial que lo permite: tiempo no lectivo.

En tercer lugar, la evaluación necesita volver a tener características humanas y formativas, más que de proceso técnico y numérico y de medición. Esto implica volver a poner al estudiante, el sujeto, y no un objeto (el currículum o “la materia”) al centro del proceso educativo, significa trabajar en modificar el lenguaje para que las habilidades de los y las estudiantes sean el foco del proceso de enseñanza-aprendizaje. Estos aspectos también pueden ser protegidos y promovidos a través de la política educacional.

Finalmente, este artículo genera reflexiones sobre SBA, considerando las voces de sus actores, que no siempre son consideradas. Urge considerarlas para repensar a quién sirve la evaluación y comprender que las tendencias internacionales no siempre resultan en efectos iguales. Por esa razón, este artículo invita a considerar el impacto del ambiente sociocultural en el que la evaluación se desarrolla, así como la necesidad de diseñar políticas públicas que consideren

las necesidades de docentes y estudiantes según esos contextos, es decir, políticas públicas socioculturalmente responsivas.

Referencias

- Acción Educar (2021). Aporte de la educación particular subvencionada en Chile. <https://accioneducar.cl/wp-content/uploads/2021/09/Aporte-de-la-educacion-particular-subvencionada-en-Chile-Accion-Educar.pdf>
- BLACK, P. Y WILLIAM, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, 5(1), 7-74.
- BLACK, P. Y WILLIAM, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, 21(1), 5-31.
- BONNER, S. M., TORRES RIVERA, C. Y CHEN, P. P. (2018). Standards and assessment: coherence from the teacher's perspective. *Educational Assessment, Evaluation and Accountability*, (30), 71-92.
- BOTA, O. A. Y TULBURE, C. (2015). Aspects regarding the relationship between teaching styles and school results. *Procedia-Social and Behavioral Sciences*, 203, 285-290.
- BOWEN, G. A. (2009). Document analysis as a qualitative research method. *Qualitative research journal*, 9(2), 27-40.
- BRAUN, V. Y CLARKE, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101.
- BRAUN, V. Y CLARKE, V. (2012). *Thematic analysis*. American Psychological Association.
- BRIGGS, J. Y ELLIS, V. (2013) *Assessment for learning, on learning and teaching in secondary schools*. Learning matters. SAGE.
- BROADFOOT, P. (1999) Preface. En A. Filer (Ed.), *Assessment: Social practice and social product* (pp.ix-xii). Routledge Falmer.
- BROWN, M., BRUNS, D., MCNAMARA, G. Y O'HARA, J. (2022). Culturally responsive classroom-based assessment: a case study of five secondary schools in Ireland. *Revista de Investigación Educativa*, 40(1), 15-32.
- BURNS, D., BROWN, M., MCNAMARA, G., O'HARA, J., ALTRICHTER, H., NAYIR, F., NORTVEDT, G., FELLNER, M., HELM, C., PUNZENBERGER, B., SKEDSMO, G. Y WIESE, E.F. (2019). *Aiding Culturally Responsive Assessment. A User-Friendly Handbook for Culturally Responsive Assessment in Educational Settings*. DOI: 10.13140/RG.2.2.31798.75844/1
- CASTILLO RETAMAL, F., MEJÍAS RODRÍGUEZ, M., & VÁSQUEZ-GÓMEZ, J. (2023). Conceptos, estrategias y aplicación de la evaluación formativa en tiempos de pandemia: un caso de la realidad chilena. *Meta: Avaliação*, 15(48).

- CHENG, A. (2019) School-based Assessment and Assessment for Learning: An Exploratory Study of Hong Kong English Teachers' Practices, *The Journal of Asia*, 16(4), 1368-1375.
- CHENG, L., ANDREWS, S. Y YU, Y. (2011). Impact and consequences of school-based assessment (SBA): Students' and parents' views of SBA in Hong Kong. *Language Testing*, 28(2), 221-249.
- CLARKE, V. BRAUN, V. (2013). Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning. *The psychologist*, 26(2), 120-123.
- COOK-SATHER, A. (2018). Tracing the evolution of student voice in educational research. En R. Bourke y J. Loveridge (Eds.), *Radical collegiality through student voice. Educational experience, policy and practice* (pp.17-38). Springer.
- CRESWELL, J. W. (2013). Steps in conducting a scholarly mixed methods study. *DBER Speaker Series 48*.
- CUMMING, J. J. Y MAXWELL, G. S. (2004). Profiles of educational assessment systems worldwide. Assessment in Australian schools: current practice and trends. *Assessment in Education: Principles, Policy & Practice*, 11(1), 89-108.
- DAVIDSON, J. (2009). Exhibitions: Connecting classroom assessment with culminating demonstrations of mastery. *Theory into practice*, 48(1), 36-43.
- DELUCA, C., CHAPMAN-CHIN, A. E., LAPOINTE-McEWAN, D. Y KLINGER, D. A. (2018). Student perspectives on assessment for learning. *The Curriculum Journal*, 29(1), 77-94.
- DENZIN, N. K. (1970). *The research act: A theoretical introduction to sociological methods*. Aldine.
- DENZIN, N. Y LINCOLN, Y. (2008). *The Landscape of Qualitative Research*. SAGE.
- ELLWOOD, J. MURPHY, P. (2015). Assessment systems as cultural scripts: A sociocultural theoretical lens on assessment practice and products. *Assessment in Education: Principles, Policy & Practice*, 22(2), 182-192.
- FALCÓN, Y., AGUILAR, J., LUY, C. A., & MORILLO, J. (2021). Formative assessment, reality or good intentions? Case study in primary level teachers. *Propósitos y Representaciones*. 9, <http://dx.doi.org/10.20511/pyr2021.v9n1.1041>
- FLÓREZ PETOUR, M. T. (2014). Systems, ideologies and history: A three-dimensional absence in the study of assessment reform processes. *Assessment in Education: Principles, Policy & Practice*, 22(1), 3-26.
- HARLEN, W. (2004). *A systematic review of the evidence of reliability and validity of assessment by teachers used for summative purposes*. EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- HARRIS, L. R. Y BROWN, G. T. (2009). The complexity of teachers' conceptions of assessment: Tensions between the needs of schools and students. *Assessment in Education: Principles, Policy & Practice*, 16(3), 365-381.
- Hennink, M., Hutter, I. y Bailey, A. (2011). *Qualitative research methods*. SAGE.

- HIRSH, Å. (2020). When assessment is a constant companion: Students' experiences of instruction in an era of intensified assessment focus. *Nordic Journal of Studies in Educational Policy*, 6(2), 89-102.
- Hong Kong Examination and Assessment Authority. (2 de Mayo 2025). *School Based Assessment-Introduction*. <https://www.hkeaa.edu.hk/en/sba/introduction/>
- KENNEDY, K. J., CHAN, J. K. S., FOK, P. K. Y YU, W. M. (2008). Forms of assessment and their potential for enhancing learning: Conceptual and cultural issues. *Educational Research for Policy and Practice*, 7, 197-207.
- KIGER, M. E. Y VARPIO, L. (2020). Thematic analysis of qualitative data: AMEE Guide 131. *Medical Teacher*, 42(8), 846-854.
- MÄKIPÄÄ, T. Y OUAKRIM-SOIVIO, N. (2019). Finnish Upper Secondary School Students' Perceptions of Their Teachers' Assessment Practices. *Journal of Teaching and Learning*, 13(2), 23-42.
- MALDONADO-FUENTES, A. C. (2021). Representación escolarizada de la evaluación: Un aprendizaje social profesional. *ALTERIDAD. Revista de Educación*, 16(2), 184-197.
- MARROY, C. Y PONS, X. (2021). Re-conceptualising education policy trajectories in a globalised world: lessons from a multi-level comparison of accountability in France and Quebec. *Comparative Education*, 57(4), 560-578.
- MARTÍNEZ QUIÑONES, G. (2019). *Percepción de docentes y estudiantes sobre implementación de estrategias diversificadas de evaluación. Una mirada a un establecimiento educativo secundario de la comuna de San Miguel*. Universidad de Chile.
- McMILLAN, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice*, 20(1), 20-32.
- MENG, C. C., IDRIS, N. Y EU, L. K. (2014). Secondary Students' Perceptions of Assessments in Science, Technology, Engineering, and Mathematics (STEM). *Eurasia Journal of Mathematics, Science & Technology Education*, 10(3), 219-227.
- Ministerio de Educación de Chile. (2016). "Percepciones, significados y necesidades sentidas en relación a la evaluación de aula en profesores, directivos y estudiantes de educación escolar de tres regiones del país". https://www.curriculumnacional.cl/614/articles-70901_archivo_01.pdf
- Ministerio de Educación de Chile. (2018). Decreto de calificación y promoción escolar N°67/2018. https://www.curriculumnacional.cl/614/articles-70983_archivo_01.pdf
- MONZÓN-TRONCOSO, M. Y. (2015). Evaluación del aprendizaje: un recorrido histórico y epistemológico. *Revista de Ciencias de la Educación Academicus*, 6(1), 12-21.

- MUJICA-JOHNSON, F. N. (2020). Análisis crítico del currículo escolar en Chile en función de la justicia social Critical Analysis of the School Curriculum in Chile Based on Social Justice. *Revista Electrónica Educare* 24(1), 1-14.
- MURCHAN, D. (2018). Introducing school-based assessment as part of junior cycle reform in Ireland: a bridge too far? *Educational Assessment, Evaluation and Accountability*, 30(2), 97-131.
- PETERS RODRÍGUEZ, B. Y CONTRERAS, G. (2019). Criterios de evaluación interpretados por estudiantes de segundo y tercer año medio en la asignatura de Matemáticas en Chile. *Perspectiva Educacional*, 58(3), 46-72.
- PIZARRO GAMERO, E. A. Y GÓMEZ MUSKUS, S. S. (2019). Concepciones docentes sobre evaluación: de los lineamientos, el discurso y la práctica. *Praxis & Saber*, 10(22), 71-88.
- PRASAD, P. (2017). *Crafting qualitative research: Beyond positivist traditions*. Taylor & Francis.
- PRIETO, M. & CONTRERAS, G. (2008). Concepciones de los docentes que informan sobre sus prácticas de evaluación: un problema por revelar. *Estudios pedagógicos (Valdivia)*, 34(2), 245-262.
- PUAD, L. M. A. Z. Y ASHTON, K. (2021). Teachers' views on classroom-based assessment: an exploratory study at an Islamic boarding school in Indonesia. *Asia Pacific Journal of Education*, 41(2), 253-265.
- RACZYNSKI, D. Y SALINAS, D. (2008). Aportes y desafíos en la evaluación de políticas y programas de mejoramiento educativo: reflexiones desde nuestra práctica de evaluación en Chile. *RIEE. Revista Iberoamericana de Evaluación Educativa*, 1(1). <https://doi.org/10.15366/riee2008.1.1.007>
- RASOOLI, A., RASEGH, A., ZANDI, H. Y FIROOZI, T. (2023). Teachers' Conceptions of Fairness in Classroom Assessment: An Empirical Study. *Journal of Teacher Education*, 74(3), 260-273.
- RAVELA, P., LEYMONIÉ, J., VIÑAS, J., & HARETCHE, C. (2014). La evaluación en las aulas de secundaria básica en cuatro países de América Latina. *Propuesta Educativa*, (41), 20-45.
- REMESAL, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education*, 27(2), 472-482.
- SCOTTISH GOVERNMENT (2011). *Curriculum for excellence. Building the curriculum 5. A framework for assessment: reporting*. The Scottish Government. <https://education.gov.scot/media/zexgg4ky/btc5-reporting.pdf>
- SHEPARD, L. A. (2019). Classroom assessment to support teaching and learning. *The Annals of the American Academy of Political and Social Science*, 683(1), 183-200.
- SIVENBRING, J. (2019). Making Sense and Use of Assessments. *Scandinavian Journal of Educational Research*, 63(5), 759-770.

- SMYTH, E. Y BANKS, J. (2012). High stakes testing and student perspectives on teaching and learning in the Republic of Ireland. *Educational Assessment, Evaluation and Accountability*, 24(4), 283-306.
- STIGGINS, R. J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational measurement: Issues and practice*, 18(1), 23-27.
- STIGGINS, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10), 758-765.
- TAN, K. H. K. (2017). Asking questions of (what) assessment (should do) for learning: the case of bite-sized assessment for learning in Singapore. *Educational Research for Policy and Practice*, 16(2), 189-202.
- TARAS, M. (2005). Assessment—summative and formative—some theoretical reflections. *British Journal of Educational Studies*, 53(4), 466-478.
- TONG, S. Y. A. Y ADAMSON, B. (2015). Student voices in school-based assessment. *Australian Journal of Teacher Education*, 40(2), 15-28.
- VERA SAGREDO, A., POBLETE CORREA, S., & DÍAZ LARENAS, C. (2017). Percepción de los docentes chilenos sobre sus perspectivas, habilidades y prácticas evaluativas en el aula. *Estudios pedagógicos (Valdivia)*, 43(3), 361-372.
- YATES, A. Y JOHNSTON, M. (2018). The impact of school-based assessment for qualifications on teachers' conceptions of assessment. *Assessment in Education: Principles, Policy & Practice*, 25(6), 638-654.
- ZHAO, X., VAN DEN HEUVEL-PANHUIZEN, M. Y VELDHIJS, M. (2017). Classroom assessment in the eyes of Chinese primary mathematics teachers: A review of teacher-written papers. *Studies in Educational Evaluation*, 52, 42-54.