

EVALUACIÓN Y ENSEÑANZA: CONTRIBUCIONES PARA LA PRÁCTICA EDUCATIVA

APORTES Y REFLEXIONES
DEL PRIMER SEMINARIO DE
INVESTIGACIÓN DE LA UEICEE

COMPILADORAS

Tamara Vinacur y Lorena Landeo



EVALUACIÓN Y ENSEÑANZA: CONTRIBUCIONES PARA LA PRÁCTICA EDUCATIVA

APORTES Y REFLEXIONES
DEL PRIMER SEMINARIO DE
INVESTIGACIÓN DE LA UEICEE

Tamara Vinacur y Lorena Landeo (compiladoras)

OEI



Vinacur, Tamara

Evaluación y enseñanza: contribuciones para la práctica educativa: aportes y reflexiones del primer seminario de investigación de la UEICEE / Tamara Vinacur; Lorena Landeo. - 1a ed. - Ciudad Autónoma de Buenos Aires: Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura, 2022.

Libro digital, PDF

Archivo Digital: descarga y online

ISBN 978-987-3753-82-4

1. Educación. I. Landeo, Lorena. II. Título.

CDD 371.2607

Diseño

Biblioteca del Congreso de la Nación

Maquetador:

Gabriel Martín Gil

© Ministerio de Educación e Innovación del Gobierno de la Ciudad Autónoma de Buenos Aires y Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura, 2019

ISBN 978-987-3753-82-4

INDICE

INTRODUCCIÓN

Tamara Vinacur y Lorena Landeo 5

PRÓLOGO

Jorge Apel 11

“ENSEÑANZA Y EVALUACIÓN”

Alicia W. de Camilloni 19

“JUSTICIA EN LAS EVALUACIONES”

Tamara Vinacur 39

“LA RETROALIMENTACIÓN FORMATIVA: UNA OPORTUNIDAD PARA MEJORAR LOS APRENDIZAJES Y LA ENSEÑANZA”

Rebeca Anijovich y Graciela Cappelletti 77

“ALGUNOS PUENTES NECESARIOS”

Lilia V. Toranzos 95

“REFLEXIONES Y PROPUESTAS EN TORNO A LA ARTICULACIÓN EXTERNA Y LA EVALUACIÓN EN EL AULA”

Pedro Ravela 109

INTRODUCCIÓN

Tamara Vinacur y Lorena Landeo¹

5

Esta publicación reúne ponencias realizadas en las Jornadas de Evaluación y Enseñanza organizadas conjuntamente por la Unidad de Evaluación Integral de la Calidad y Equidad Educativa (UEICEE) y la Organización de Estados Iberoamericanos (OEI) en los años 2017 y 2018. Estos encuentros se han constituido como espacios de intercambio y reflexión, congregando a investigadores y especialistas con amplia trayectoria con la finalidad de realizar contribuciones valiosas para las prácticas de enseñanza en las escuelas primarias y secundarias de la Ciudad de Buenos Aires. Asimismo, el libro incorpora otros artículos de referentes en el campo para profundizar en algunos aspectos de la evaluación que constituyen desafíos no sólo a nivel escolar, sino también para las políticas de evaluación en la Ciudad.

Ejerciendo su rol de generar información y conocimiento sobre el sistema educativo, la UEICEE concibe como parte fundamental de su quehacer, reflexionar sobre los propios procesos que lleva adelante y -en este marco- la construcción de espacios de diálogo e intercambio plural constituyen instancias privilegiadas para tal fin.

A partir de esta motivación, en este libro se exploran –a través de las voces de diversos autores– algunos temas centrales en el debate actual sobre la evaluación escolar y de sistema, entre ellos, las características y condiciones que favorecen y obstaculizan la evaluación formativa en el aula y los procesos de retroalimentación orientados a la mejora del aprendizaje y de las prácticas docentes. Se abordan, asimismo, los

1. Tamara Vinacur es Directora Ejecutiva de la UEICEE y Lorena Landeo es Coordinadora General de Evaluación dentro de la misma unidad.

alcances y las limitaciones de las evaluaciones externas; el desafío de ponerlas en diálogo con la evaluación de aula; la aparente pugna entre la idea de una medición objetiva y la evaluación que busca abarcar la subjetividad y la heterogeneidad. Finalmente, se echa luz sobre la necesaria convivencia y reconciliación entre ambas perspectivas, sobre la importancia de aproximarse al conocimiento del complejo proceso de enseñanza y evaluación de manera colegiada, y a partir de múltiples fuentes y vías que permitan asegurar las oportunidades de aprendizaje que todos los estudiantes requieren.

A continuación, se realiza una breve presentación del contenido del libro.

El primer artículo retoma la Conferencia de apertura a cargo de Alicia Camilloni en la Jornada “Evaluación y enseñanza: contribuciones para la práctica educativa”, organizada por la Unidad de Evaluación Integral de la Calidad y Equidad Educativa de la CABA y la Organización de Estados Iberoamericanos en agosto de 2018. En su ponencia, Camilloni presenta algunas de las dificultades y posibles soluciones para trabajar la evaluación de los aprendizajes en vínculo con la enseñanza. En este sentido, desarrolla lo que considera uno de los grandes desafíos: conciliar el logro de mejores aportes para el aprendizaje de los estudiantes con la necesidad de llegar a todos y sostener el funcionamiento de un sistema educativo.

La autora despliega un cuidadoso trabajo analítico entre dos modelos puestos en juego que involucran distintas concepciones de la enseñanza, el aprendizaje y la evaluación. Por un lado, el modelo empírico-analítico que se viene desarrollando desde comienzos del siglo XX, y pone en acción una evaluación uniforme, objetiva, explícita, cuantitativa, sumativa, con base en la evidencia y destinada a calificar. En este modelo, la evaluación de los aprendizajes se apoya en un supuesto: que existe la posibilidad de saber lo que el otro sabe y que con ese fin podemos construir instrumentos que midan los saberes. Por otro lado, el modelo no-técnico, ubica su eje en la comprensión del significado de las actividades educativas y no en la medición. Sus rasgos residen en el involucramiento de los docentes y estudiantes en la generación de información y en la reflexión sobre los procesos educativos. Su abor-

daje cualitativo reconoce y pondera la heterogeneidad a la que le debe corresponder una pluralidad de objetivos de enseñanza y de instancias de evaluación. Se trata, en otras palabras, de ofrecer una oportunidad de expresión de las diferencias entre los procesos que cada alumno atraviesa. A la preocupación permanente del enfoque empírico-analítico por mejorar la definición y la observación de lo que el estudiante aprendió y, en suma, construir juicios más justos, se contraponen el reconocimiento de sus límites: alcance parcial, relativo a los criterios fijados y con una tendencia uniforme, sin llegar a ser nunca una lectura completa del desarrollo de la diversidad presente en el aula. Por otro lado, el enfoque no-técnico puede resultar complejo a la hora de ponerse en práctica.

En el segundo artículo, Tamara Vinacur repone el marco conceptual en el que se desarrolla la justicia evaluativa, explicitando sus principios, dimensiones y evidencias a partir de los aportes teóricos de diversos autores. Seguidamente, la autora identifica algunas iniciativas de justicia en las evaluaciones de los aprendizajes desplegadas a nivel internacional, nacional y jurisdiccional en distintos sistemas educativos. Se destacan en este marco: I. la línea de trabajo vinculada con la atención a la diversidad cultural en las evaluaciones a gran escala que ha desarrollado el Instituto Nacional para la Evaluación de la Educación (INEE) de México; II. las prácticas de atención a la diversidad puestas en marcha desde la Agencia de Calidad de la Educación (ACE) de Chile para la consideración del contexto socioeconómico, el género y la ruralidad de los estudiantes, y para la acomodación de las pruebas para alumnos con discapacidad sensorial; III. las líneas de trabajo referidas a los alumnos con discapacidad y a los estudiantes con una competencia limitada en inglés, como prácticas sistemáticas de atención a la diversidad en tanto garantía de la participación en las evaluaciones nacionales en Estados Unidos; IV. el marco regulatorio de las evaluaciones estandarizadas explícitamente orientado a garantizar el acceso de todos los estudiantes a las pruebas en igualdad de condiciones en Inglaterra. Para el caso jurisdiccional, se presentan tres experiencias que la Unidad de Evaluación Integral de la Calidad y Equidad Educativa (UEICEE) del Ministerio de Educación e Innovación de la Ciudad Autónoma de Buenos Aires se en-

cuentra desarrollando con vistas a la atención y el abordaje de la justicia en las evaluaciones: el desarrollo de un sistema integral de evaluación de los aprendizajes, la conformación de una Comisión Asesora sobre justicia en las evaluaciones y las intervenciones realizadas en el diseño de los instrumentos de evaluación.

En el tercer artículo, Rebeca Anijovich y Graciela Cappelletti señalan que históricamente la retroalimentación se impuso como una práctica unidireccional ofrecida por el docente y recibida por el estudiante, con bajo impacto en la mejora de sus aprendizajes y con un interés asociado exclusivamente con la calificación. Actualmente, convivimos con este paradigma tradicional y con la emergencia de una concepción diferente del proceso, que se propone como principal objetivo un involucramiento activo de los estudiantes con foco en el desarrollo de habilidades meta-cognitivas acerca del modo en que se aprende. En este sentido, las autoras identifican los desafíos actuales de la retroalimentación formativa como impulsora de la mejora de los aprendizajes, realizando un recorrido por diversos aportes teóricos y advirtiendo que cuanto mayor es la presión y el predominio de la cultura focalizada en resultados educativos, más difícil se hace alcanzar progresos en la retroalimentación. El análisis de las autoras hace hincapié en la naturaleza dialógica y relacional del proceso de retroalimentación formativa, cuya efectividad radicaría en dar impulso a una práctica reflexiva por parte de los docentes, y a un involucramiento de los estudiantes en un ejercicio de autorregulación, que involucraría razonamientos y emociones, para alcanzar así el logro de metas y objetivos académicos. Finalmente, Anijovich y Cappelletti presentan una serie de condiciones y obstáculos de la retroalimentación formativa, con la intención de orientar la estrategia de los docentes y posibilitar su puesta en práctica en el aula.

El cuarto artículo toma como punto de partida la conferencia de cierre de la jornada 2017 “Aportes para la enseñanza desde la investigación y la evaluación educativas” a cargo de Lilia Toranzos, quien recupera una de las grandes preocupaciones en materia de evaluación en tanto acción de carácter político: ¿qué es lo que se puede proponer con la acumulación de evidencias que genera la práctica evaluativa? y ¿qué caminos transitar entre espacios de trabajo diferentes?

Por un lado, Toranzos refiere al plano de los abordajes macro y micro, esto es, del sistema y el del aula. Por otro, la autora pone el foco en las posibilidades del uso de la evidencia producida y en la capacidad de su “traducción” en una herramienta más práctica que permita la toma de decisiones en diversas instancias, entre ellas: de enseñanza en el ámbito del aula, de orientaciones institucionales para el conjunto de los docentes de un ciclo de una escuela, y de intervención política para la formación docente. Toranzos distingue la noción de indicio de la de evidencia, como dos conceptos que -lejos de ser contrapuestos- plantean facetas diferentes a tener en cuenta en la práctica evaluativa. El concepto de indicio hace énfasis en la cautela con la que se debe mirar, trabajar y tratar la información generada a partir de los procesos de evaluación a nivel del sistema. Finalmente, la autora sistematiza los desafíos para la construcción de un sistema de evaluación de los aprendizajes que proporcione información valiosa, para distintos propósitos y destinada a una amplia gama de actores.

En el artículo quinto, se repone la conferencia “Reflexiones y propuestas en torno a la articulación externa y la evaluación en el aula” a cargo de Pedro Ravela en la Jornada del año 2018: “Evaluación y Enseñanza: contribuciones para la práctica educativa”. El autor presenta su investigación iniciada en 2008 en torno a la evaluación dentro del aula, en cuyo marco se entrevistaron docentes de sexto grado de primaria y de tercer año de secundaria en diez países de América Latina. El puntapié de su trabajo fue la preocupación en torno a una tendencia latente en la política educativa coyuntural y en particular en el ámbito de la evaluación: “estar poniendo mucha cabeza, inversión y tiempo en la evaluación externa sin mirar lo que ocurre dentro del aula”. Advierte entonces sobre la necesidad de articular estos dos enfoques, identificando qué puentes se pueden tender entre la evaluación de sistema y la evaluación en el aula. Al momento de presentar los resultados de su investigación, Ravela destaca que la evaluación externa ofrece una serie de contribuciones al docente y a la evaluación en el aula: 1) aporta un cambio de mirada en las expectativas de aprendizaje y en los modos de comunicación con los estudiantes; 2) ayuda a los docentes a conocer ejemplos de actividades que evalúan procesos más complejos o situa-

ciones auténticas; 3) permite al docente contrastar el modo en que está calificando el desempeño del estudiante con los resultados de la prueba externa. En este sentido, Ravela realiza un recorrido por iniciativas de diversos países con especial foco en el involucramiento docente en los procesos evaluativos. Entre éstas, se reponen: la experiencia de evaluación en línea en Uruguay; las evaluaciones aplicadas a principio de año en Australia para brindar información útil y oportuna a los docentes; la evaluación progresiva y voluntaria en Chile; y la elaboración de actividades de evaluación y de rúbricas para la corrección en los Estados Unidos. Sostiene Ravela que la clave del diálogo entre la evaluación de sistema y la evaluación en el aula reside en el compromiso de “abrir el aula” con vistas a que la profesión y la práctica docente estén sujetas a la observación y colaboración colectiva.

PRÓLOGO

Jorge Apel

11

Escribir estas palabras a modo de prólogo es una gran oportunidad. Todos los autores fueron mis maestros y no siempre se tiene la posibilidad de hacer público el agradecimiento para aquellos que con sus enseñanzas, ideas o gestos nos señalan los obstáculos, desbrozan confusiones, o expresan conceptos que luego haremos propios y transitaremos con cierta comodidad.

Y si el profesor Ravela comienza su artículo diciendo “A cierta edad parece que uno que ya tuvo varias vidas...”, la extensión de la mía -que por cierto hace pensar en cierta filiación gatuna- me da la perspectiva de observar los hilos que se tienden entre miradas, la aparición de nuevos mojones, el borramiento de diferencias entre ideas, conceptos o prácticas que se consideraban alejadas, distantes o directamente opuestas. Son tensores anclados en las orillas de un problema complejo, que luego se transformarán en senderos por donde circularán enseñantes y aprendices con la seguridad que da el camino ya socialmente habilitado. Los bordes donde se planta bandera no son accidentes “naturales”, sino construcciones sociales establecidas en intentos anteriores y se transformaron en acciones, que por habituales nos parecen instaladas desde siempre y que ocultan discusiones y reconstrucciones anteriores. Y así avanzamos, de maestro en maestro...

La profesora Camilloni en una de sus clases a las que asistí como alumno, preguntó: “*La didáctica... ¿progresas?*”. Tal vez no haya respuesta sencilla para esta pregunta; tal vez estaba destinada a poner entre paréntesis mi noción de tiempo como algo lineal; tal vez me ofrecía la posibilidad de comprender grandes ciclos en los modos de abordar los problemas de la enseñanza y el aprendizaje, sin considerar un antes y un después como avance o retroceso, sino como modos de entender, enfrentar, construir la realidad. En su conferencia distingue, a cientos

de horas de haber formulado la pregunta mencionada y con total claridad, dos modos de pensar la evaluación para obtener la información buscada de manera válida y con algún grado de confiabilidad. Por un lado, el modelo empírico analítico y, por el otro, el que denomina “no técnico”. Desde mi mirada, lo notable es la diferencia en el modo de *construcción* de esa información, pues allí se revela la idea de “sujeto de conocimiento” que, una y otra perspectiva teórica, tienen al momento de producir o recoger las evidencias o indicios necesarios para el proceso evaluativo. En ambos la influencia del maestro -autoridad de conocimiento- está presente de una manera diferente, ya que en el segundo se otorga voz al alumno que interviene activamente en el descubrimiento. Se confía en su apreciación y se le dan oportunidades de mejorarla. Para aquellos que con anterioridad estudiamos de sus escritos, la profesora ya había señalado “tendríamos que comprender en primer lugar, que el propósito principal de la evaluación no se puede lograr si la evaluación no se convierte en autoevaluación tanto para el docente cuanto para el alumno²”. Sin embargo, en la descripción del modo “no técnico” no se trata solamente de conciencia -tema notable de por sí-, ni de los procesos metacognitivos que la habilitan sino también de autorregulación. Si todo va bien, docente y alumno, se influyen mutuamente cada uno desde su rol, enseñar y aprender. Nos advierte, sin embargo, sobre un aspecto distintivo de la evaluación: la formulación de un juicio evaluador calificativo, un juicio performativo que crea una realidad. Tanto en clave “empírico analítica” como en modo “no técnico”, la búsqueda de la información sobre el saber se diferencia al momento de emitir un juicio sobre ese saber.

Cuando leí el texto del profesor Ravela recordé inmediatamente sus “fichas”³. A mediados de los 90 todavía coordinaba una escuela prima-

2. Camilloni, A. (1998). La calidad de los programas de evaluación y los instrumentos que los integran. En A. Camilloni, S. Celman, M. Paolu de Maté, & E. Litwin, La evaluación de los aprendizajes en el debate didáctico contemporáneo (pág. 74). México: Paidós.

3. Ravela, P. (2006). Fichas didácticas para comprender las evaluaciones educativas.

ria privada en Capital Federal y no estaba interesado especialmente en la evaluación. El primer ONE (Operativo Nacional de Evaluación)^{4,5} se constituyó para mí en un estímulo formidable ya que lo entendía como una modalidad diferente que asumía el Estado en la supervisión de las escuelas. Tal vez al llamarlas “las fichas de Ravela”, no hago honor ni a la organización donde se desempeñaba ni a sus colaboradores, pero la facilidad, familiaridad y sencillez de plantear problemas complejos, así como la frecuencia de su utilización, llevaba a esa denominación casi familiar.

Al escribir estas líneas, supongo que por el hecho de la nacionalidad o la similitud del nombre de su autor, recordé el libro “Apreciación de los resultados de la acción educativa” del profesor Alfredo Ravera, que me acompañó cuando comencé a trabajar a mediados de la década del 60. Hay varias frases que podríamos hacer nuestras por la actualidad de sus planteos. Por ejemplo: “*La escuela debe enseñar maneras de aprender*” y “*al colocar al niño en situación de que por sí mida el resultado de su trabajo, le estamos dando normas de autoconducción que irán afirmando su independencia y multiplicando el desenvolvimiento de su personalidad*”⁶. El texto que estamos prologando aquí, presenta en forma resumida las conclusiones de la investigación que realizara, desde el año 2008, el profesor Pedro Ravela junto con Beatriz Picaroni y Graciela

PREAL. Recuperado en Enero de 2019, de https://www.mineduccion.gov.co/cvn/1665/articles-125590_archivo_pdf.pdf

4. Perassi, Z. (Junio de 2017). Contextos de Educación. El sistema nacional de evaluación en la República Argentina. Políticas y concepciones., ISSN 2314-3932(22), 1-9. Recuperado en Enero de 2019, de <http://www.hum.unrc.edu.ar/ojs/index.php/contextos/article/view/495>

5. Argentina. Secretaría de Programación y Evaluación. Ministerio de Cultura y Educación. Educativa. Calidad de la Educación. (1993). Sistema Nacional de Evaluación Primer Operativo Nacional 1993: resultados nacionales, primer informe. Ciudad Autónoma de Buenos Aires: Buenos Aires. Recuperado en Enero de 2019, de <http://repositorio.educacion.gov.ar:8080/dspace/handle/123456789/90459>

6. Ravera, A. (1953). *Apreciación de los resultados de la acción educativa* (pág. 25). Buenos Aires: Kapelusz.

Loureiro⁷ para estudiar lo que está pasando dentro del aula, la mirada de los docentes y lo que pueden hacer los funcionarios y organismos al momento de enfrentar decisiones sobre los modos de obtener información cuantitativa de grandes cantidades de alumnos para facilitar el diálogo entre sus conclusiones y la puesta en acción dentro del aula.

Pero ¿acaso el trabajo del profesor Ravera no estaba destinado a lo mismo? ¿Las conclusiones y esfuerzos del ONE y de algunos de los operativos que le siguieron, no pretendían influir en las decisiones prácticas, concretas, habituales de los maestros, más allá de la declarada “información necesaria para la toma de decisiones políticas” a nivel macro y de alcance nacional?

El profesor Ravela toma en consideración otra mirada, “centrada en los agentes”, que parte de preguntarse “quién quiero que use esto” y “cómo me imagino que lo puedo usar”. El maestro tiene voz, es escuchado. Se confía en él y se le da las opciones de mejorar su mirada. Para Ravera, en cambio, el maestro era un agente que debía seguir las indicaciones de los expertos. Dice: “*El maestro **deberá** saber poner al alcance del niño*” (...) “*el maestro **debe** recurrir a diversas pruebas y exámenes adecuados*” y en el apéndice señala además otras consideraciones sobre el “*deber ser*”, sobre una normativa aplicable de arriba para abajo⁸.

Ravela, por su parte, analiza las creencias de los docentes para apoyarse en ellas y comprender qué está pasando con la evaluación en el aula. Uso la palabra “creencias” en el sentido que le da Vinciane Despret: “*una creencia es aquello que pone a las entidades ‘a disposición’ de los acontecimientos*”⁹. Lo importante no es lo que dicen los docentes sobre sus supuestos sino qué hacen con ellos. Es decir, la escucha se

7. Ravela, P., Picaroni, B. y Loureiro, G. (2017). ¿Cómo mejorar la evaluación en el aula?; Magro; Montevideo.

8. *Ibidem* (pág. 171).

9. Despret, V. (2008). El cuerpo de nuestros desvelos. En T. Sánchez-Criado, *Tecnogénesis: la construcción técnica de las ecologías humanas* (2 Vols.) (pág. 281). Madrid: AIBR, Antropólogos Iberoamericanos en Red.

centra en la práctica habitual, en establecer un diálogo de aprendizaje mutuo entre docentes y pruebas estandarizadas. Munido de los conocimientos que su indagación le presta, el profesor Pedro Ravela introduce en la parte final de la exposición la idea de evaluación auténtica, como oportunidad de vinculación entre la mirada de sistema, propia de las acciones de evaluación a nivel macro, y el trabajo cotidiano en el aula, que podría inspirar en los docentes.

Las profesoras Rebeca Anijovich y Graciela Cappelletti trabajan sobre la retroalimentación formativa y nos dejan una serie de consideraciones para su implementación, destacando los obstáculos más frecuentes y los aspectos nodales para la instalación de esta perspectiva en las aulas. Quiero destacar cuestiones que allí aparecen y asumo como esenciales, ideas que integran aspectos disociados de la enseñanza y el aprendizaje, no solo en la consideración del proceso evaluativo. En primer lugar, las emociones. Este tema es por lo general dejado por fuera del aprendizaje, e incluirlo desde algún aspecto recupera al alumno como persona. La aparición de los pares señala la posible emergencia de una entidad mayormente ignorada: el grupo, los subgrupos existentes en el curso. No tendremos frente a nosotros una suma de individualidades a las que llamamos “grado”, todos igualados por la edad, la obligación curricular y la evaluación. Por supuesto, si aparece el grupo encontraremos el “clima” -tal como ellas lo mencionan- como un aspecto influyente en la enseñanza y el aprendizaje. Es cierto que la diferenciación de los individuos como parte de una clase estaba presente en el trabajo de la profesora Camilloni. Son dos aspectos de lo mismo: aparece el grupo y el individuo. No solo tiene voz el alumno, tiene participación el conjunto de alumnos que forma parte del clima emocional a considerar. Se introduce una tensión apasionante: por un lado, la complejidad de la concepción de un aula que se espera que sea homogénea en cada clase como una parte significativa de la gramática escolar¹⁰ y por otro lado, el reconocimiento de las diferencias entre grupos, subgrupos e indivi-

10. Fernández Enguita (2018): Más escuela y menos aula; Morata Madrid. Pag 30 y capítulo: “Es el aula, estúpido”.

duo, tensión que tiene relevancia concreta, práctica diaria y profunda significatividad teórica. Como individuo, me reconozco perteneciente a un grupo-grado o clase particular dentro de la escuela, me incluyo en algunos subgrupos, pero también soy único y me consideran como tal. Por último, Anijovich y Cappelletti nos invitan a pensar sobre el proceso y no solo sobre los resultados. Este cambio de enfoque está asociado a una serie de creencias pedagógicas y a los constructos teóricos que operan en los profesores a la hora de tomar decisiones, permitiendo iluminar nuevos aspectos.

Al final menciono en forma conjunta dos escritos: el presentado por la profesora Toranzos, y el capítulo de Tamara Vinacur, pues ambos hablan de temas complementarios entre sí, desde la mirada de la UEICEE. Toranzos advierte sobre el peligro de considerar los resultados del trabajo de evaluación como una totalidad comprensiva, enfatizando que es solo parte de la compleja realidad de la enseñanza y el aprendizaje. La cuestión de la justicia en las evaluaciones pone el foco sobre otra de las aristas del cambio de paradigma, introduciendo un tema apasionante que se vincula con el conjunto de los cambios que estuve señalando.

La relación que se establece entre “todos, algunos y uno” varía a lo largo del tiempo. Por ejemplo, la cultura occidental y cristiana se creyó que encarnaba el “todos” anulando la condición humana de millones de personas, aniquilando tribus y culturas. En el absolutismo político, era el derecho divino lo que justificaba a esa persona -ese “uno”- como totalmente diferente y especial con respecto a cualquier otro “uno”¹¹.

La cuestión de la justicia encierra muchos otros interrogantes que nos interpelan ¿Podemos partir de la idea de que todos los alumnos no son iguales en su condición pero iguales en lo que pretendemos que aprendan? ¿Debemos asumir que la determinación del mismo punto de llegada para todos es una meta justa? ¿o partir y admitir las diferencias

11. Para decirlo de otra manera; será necesario que los “occidentales” estemos presentes de una manera por completo diferente, primero para nosotros mismos y luego para los demás. Latour, B. (2013) Investigación sobre los modos de existencia (pág. 32). Buenos Aires: Paidós.

a lo largo de toda la escolaridad con diferentes niveles de logro? ¿Las distintas finalidades de las evaluaciones de aula y de sistema plantean “injusticias” entre sí o deben pensarse con una mirada diferente sobre la justicia?¹²

El tema de la justicia en la evaluación tiene un costado político e ideológico, manifestaciones prácticas y advertencias técnicas. No necesariamente estas instancias son congruentes entre sí. En todos los casos nos expone a examinar nuestra coherencia entre el pensar, el sentir y el hacer en la docencia.

Las características que subrayamos anteriormente en la lectura de cada uno de los artículos -la voz del maestro y del alumno, la aparición de los pares y del grupo, la inclusión de las emociones, el empeño en señalar el trabajo evaluativo como palanca para la enseñanza y el aprendizaje, el énfasis en el compromiso del alumno para aprender-, se incluyen como problemáticas propias de la justicia en la evaluación.

Recuerdo haberme encontrado una y otra vez con docentes que mencionaban su disposición para modificar estrategias de enseñanza, innovar en nuevas modalidades de aprendizaje, incorporar tecnologías y recursos, pero que sin embargo sentían que todo se congelaba al momento de la evaluación. Ellos tenían una sola manera de entender y poner en práctica la evaluación. Hoy podemos decir que los conceptos y las prácticas evaluativas están ligadas a los posibles avances en el lento proceso de cambio en la educación. No entorpecen sino que facilitan una educación que supone un encuentro entre alumnos, maestros y saberes, en donde se produzca y robustezca *“alguna forma de apropiación subjetiva y singular del saber y una decisión de hacer”*¹³.

Y ya que hablamos de escuelas, maestros y alumnos concluyo citando a otra maestra a la que tuve la posibilidad de agradecer personalmente: Edith Litwin, quien nos enseña que:

12. Evaluación Fepba. Informe 2017 en <http://bde.operativos-ueicee.com.ar/documentos/507-evaluacion-fepba-informe-2017>. Recuperado Enero 2019.

13. Cerletti, A. (2008). Repetición, novedad y sujeto en la educación (pág. 188). Buenos Aires: Del Estante.

“En las escuelas no se debería actuar pensando que hay una sola manera posible de evaluar. Una buena evaluación debe otorgar confianza y generar, contrariamente al clima habitual que se instala con las evaluaciones, un espacio para permitir que los aprendizajes fluyan y se expresen con naturalidad. Posibilita reconocer nuestros límites, reírnos de nuestras propias exigencias y recuperar el sentido de lo humano en el acto de aprender. Desde esta perspectiva, la evaluación es proveedora de seguridad, protege a los estudiantes y por tanto, se imbrica en la buena enseñanza”¹⁴.

14. Litwin, E. (2016). El oficio de enseñar (pág. 178). Buenos Aires: Paidós.

“ENSEÑANZA Y EVALUACIÓN”¹⁵

Alicia W. de Camilloni

19

INTRODUCCIÓN

La evaluación es un género de acción que acompaña a muchas de las tareas que están implicadas en los procesos educativos. Cuando nos referimos a ella podemos aludir a la evaluación de los sistemas educativos, a la evaluación institucional, a la evaluación del currículo, a la evaluación de un proyecto de un curso, a la evaluación de la calidad de la enseñanza y, también, a la evaluación de los aprendizajes. Aquí nos ocuparemos exclusivamente de la evaluación de los aprendizajes, tema no menor porque constituye un espacio en sí mismo y porque tiene, por otra parte, un impacto muy grande sobre la evaluación de todos los otros objetos que encontramos situados en el campo de la educación. No se puede evaluar un sistema educativo, en verdad, sin tener en cuenta los aprendizajes de los alumnos. Tampoco es posible evaluar un currículo, ni una institución ni una programación didáctica, ni la calidad de la enseñanza sin atender a los resultados de aprendizaje. Y como, generalmente, la información que proviene de la evaluación de los aprendizajes es tan importante y significativa para otros tipos de evaluación y es, al mismo tiempo, muy compleja y difícil de resolver,

15. Conferencia de Apertura de la Jornada: “Evaluación y enseñanza: contribuciones para la práctica educativa” organizada por la Unidad de Evaluación Integral de la Calidad y Equidad Educativa de la CABA y la Organización de Estados Iberoamericanos, Buenos Aires, 10 de agosto de 2018.

porque en ella confluyen un sinnúmero de variables diferentes que establecen entre sí, por lo demás, muy diversas relaciones, observamos que se constituye en el eslabón más débil de la cadena. Así, pues, si no está bien resuelta la evaluación de los aprendizajes tampoco va a estar bien fundada la evaluación de la enseñanza, la del currículo, la de una institución o la de un sistema educativo. Sin embargo, no es lo mismo evaluar los aprendizajes para evaluar un sistema o una institución, que evaluar los aprendizajes cuando nuestro principal interés está centrado en el alumno que aprende y el profesor que enseña. Por esta razón, y porque es necesario diferenciar la evaluación según el uso que se haga de ella, precisamos que la evaluación de los aprendizajes a la que acá nos referiremos es la evaluación que docentes y alumnos realizan en calidad de acción concurrente con sus procesos de enseñanza y aprendizaje. Diagnosticar con estrictez la causa de un problema de aprendizaje cuando estamos al frente de un curso no es lo mismo que diagnosticar la causa de un problema en un sistema educativo. En atención a los planos y los actores, las decisiones y, en consecuencia, las informaciones necesarias, son disímiles.

Los administradores de la educación encuentran valiosa la información estadística en la que se agrupa información cuantitativa sobre grandes grupos de alumnos e instituciones. Los profesores, por el contrario, requieren contar con una información muy minuciosa y cuidada sobre grupos e individuos, tal que permita diferenciar niveles de calidad en el aprendizaje, identificar errores y sus posibles causas y, en particular, discriminar entre etapas en los procesos de aprendizaje. En sus evaluaciones se trata de recabar información que, en última instancia, resulte de utilidad para tomar decisiones respecto de qué es lo que se ha de enseñar y cómo y cuándo se debe enseñarlo, cuánto tiempo requiere para la enseñanza y el aprendizaje, cuánta ejercitación necesitan los alumnos y, cuando los aprendizajes presentan algún problema en particular, cómo definir la dificultad y cómo hallar sus causas con el propósito de ayudarles a resolverla.

Toda acción de evaluación se asienta sobre tres bases: primeramente, alguna información que se ha recogido en relación con un atributo, una característica, un conocimiento, una habilidad aprendida; en se-

gundo lugar, alguna forma de medir que consideramos apropiada para interpretar esa información de manera válida y, si es posible, con algún grado de confiabilidad y, por último, algún tipo de juicio de valor que se construye sobre el objeto evaluado en atención a la información que se ha recogido. Se trata de un proceso en el que se adoptan, indudablemente, múltiples decisiones, y en el que los diferentes pasos no están claramente delimitados a la hora de tomar las variadas determinaciones que están enhebradas en él. Hay distintas operaciones implicadas en la acción de evaluación y, por ende, cada una de ellas es difícil de definir en razón de la naturaleza de la cuestión y de la diversidad de opciones que se presentan para su resolución.

La naturaleza de la información necesaria cambia según sea el profesor o el alumno quien debe resolver cuáles son las acciones que, en consecuencia, conviene que realice. Si el docente debe dar por finalizada la enseñanza de un tema, o debe re-enseñarlo a todos o solamente a algunos alumnos, empleando la misma u otra estrategia de enseñanza, con el propósito de reemplazar la que hasta el momento venía empleando habitualmente, requiere cierta información que, a esos fines, resulte específicamente relevante. Si es el estudiante quien se enfrenta, por caso, a la decisión de resolver qué hacer cuando sus calificaciones son bajas, le es imprescindible recibir noticias detalladas acerca de qué contenidos está aprendiendo bien o mal o si es conveniente que reemplace el material de estudio con el que está trabajando por otro material o si necesita dedicarle más tiempo y concentración al estudio. La información proveniente de la evaluación debiera facilitarle al alumno descubrir, por ejemplo, que hay ciertos conceptos que en realidad no comprende aunque creía haberlos entendido, de modo de poder aplicarlos correctamente en la resolución de un problema o en la respuesta a una pregunta de elaboración. En todas esas situaciones, el profesor y el estudiante requieren contar con una información muy bien seleccionada y pertinente, no solamente cuantitativa sino también cualitativa porque las distinciones que es preciso establecer se refieren, siempre, a cuestiones delicadas y, en ocasiones, privadas, propias del mundo interior del alumno y del docente.

Nos referimos a modalidades de información que aluden, entre otros, a los progresos que están realizando los alumnos de su clase en conjunto y cada alumno en particular, en relación con lo que el docente espera que sus alumnos aprendan, las calidades diferentes que pueden aparecer en los desempeños que los docentes estamos tratando de desarrollar, los avances o retrocesos que se producen en los aprendizajes. Este tipo de información requiere gran robustez en los significados asignados por los profesores a los aprendizajes diferenciales en los que están trabajando en un mismo programa de formación y un acuerdo entre ellos en torno de la definición de criterios transparentes respecto de cuáles son los rasgos de lo que denominan “un buen trabajo” o “un mal trabajo”, así como la determinación respecto de cuáles son los distintos niveles y calidades en el desempeño, cuáles serían aceptables y cuáles no e, igualmente, cuáles serían los límites entre los distintos grados de calidad y cómo se habrían de apreciar esas diferencias. Definiciones y conceptos que resultan útiles en la medida en que revierten sobre los procesos de enseñanza y las actividades de estudio y aprendizaje que realiza el alumno. En esos casos también serán útiles para las autoridades que tienen que tomar decisiones en el nivel institucional, aunque las informaciones que estas necesitan para tomar este tipo de decisiones son relativamente más gruesas que las muy finas que, comparativamente, deben manejar un profesor o un alumno. Es decir, estamos hablando de una evaluación que, como la ha definido en algún momento Daniel Stufflebeam, consiste en recoger información útil para tomar decisiones. Debemos preguntarnos, en consecuencia, qué información necesitamos, quién va a recibir la información, qué utilidad le queremos dar, qué nivel de decisiones tenemos que tomar y, finalmente, agregamos, que sobre esa información útil tenemos que tomar decisiones emitiendo un juicio de valor. El problema es, entonces, cómo se recoge la información, con qué propósitos, qué es lo que se hará con ella, y cómo se procederá para construir juicios de valor.

El tema de la evaluación es, actualmente, uno de los temas más complejos en el terreno de la didáctica en todos los niveles. Es complejo en la educación inicial, es complejo en la primaria y la secundaria,

y es complejo, asimismo, en la educación superior y en el nivel de posgrado.

Debemos destacar, sin embargo, que la evaluación, hoy, está asentada sobre un conjunto de teorías que se han ido construyendo sobre la base de más de cien años de investigación. Hay mucho conocimiento acumulado pero, como ocurre en las ciencias sociales, no hay una única teoría, no hay un único conjunto de supuestos, ni hay una sola línea de acción que se derive de ellos. Evaluar y calificar bien es, indudablemente, una tarea muy difícil. Los estudios empíricos demuestran que teoría y técnica en evaluación no siempre son fáciles de conciliar y la advertencia que debemos tener presente permanentemente, es que los peligros de las evaluaciones defectuosas son considerables. Las evaluaciones y los juicios de valor que emanan de ellas tienen consecuencias para las personas, para los grupos sociales y para la sociedad en su conjunto. En los extremos, los fracasos escolares marcan a los individuos ante sí mismos y ante los demás; los éxitos, en cambio, los enaltecen. La autoimagen y las aspiraciones se definen con estas etiquetas que institucionalizan, con el aval explícito de una autoridad que se presume objetiva, el papel de los individuos en su grupo y en la sociedad. Los juicios evaluativos calificadores tienen carácter categórico, aunque dependan, lo que ocurre habitualmente, de un proceso en el que las dudas lo hayan impregnado en la mayor parte de las decisiones que, en cada etapa, han ido tomando los docentes al construir los instrumentos de evaluación y al calificar los trabajos de los alumnos.

Sabemos que, con frecuencia, esos juicios son inapelables por los alumnos. No obstante, es imprescindible tomar conciencia de que evaluar es una de las tareas más laboriosas y serias que debe realizar un docente. En este sentido, no se trata de que el juicio evaluativo constituya solo una reseña o un retrato de uno o varios de los caracteres de lo evaluado porque, dado el marco institucional en el que ocurre, va más allá. En realidad, el juicio evaluativo calificador es un juicio performativo, esto es, un acto de habla en el que al enunciarse el juicio se realiza la acción que está contenida en su enunciado, y como depende del contexto en el que el juicio se pronuncia, este carácter performativo se instaura porque se establece en un contexto escolar. Es un enunciado

que, instituye una realidad. Convierte al alumno en un “aprobado” o un “desaprobado”, en un “graduado” o un fracasado. Gravísimo sería, por ende, que esto ocurriera sin contar con el fundamento justo, honesto, equitativo y apropiado para la circunstancia en que ese estado se asigna y adquiere reconocimiento social y crea, por consiguiente, una línea en la configuración de la identidad de las personas. Es un poder extraordinario el que tiene el docente que puede emitir esta clase de juicios calificadores, y debe ser asumida con gran responsabilidad y compromiso ético.

LA RELACIÓN DE LA ENSEÑANZA Y LA EVALUACIÓN DE LOS APRENDIZAJES

Afirmamos antes que la evaluación es una acción concurrente con los procesos de enseñanza y aprendizaje, lo cual se puede interpretar normativa o descriptivamente. La afirmación es normativa cuando sostenemos que ambas, enseñanza y evaluación, deben estar alineadas. No corresponde evaluar el aprendizaje de lo que no hemos enseñado y, particularmente, si no hemos dado oportunidad real a los alumnos de lograr los aprendizajes que vamos a evaluar. Es este un principio básico de justicia que es garantía de respeto al derecho del alumno al ser evaluado. Este principio se encuentra entre los ejes principales de la ética del ejercicio de la docencia y no puede, y no debe, ser vulnerado.

Desde un punto de vista descriptivo, la relación entre enseñanza y aprendizaje no responde a una secuencia causal simple en la que el profesor enseña y el alumno aprende los contenidos enseñados que serán luego objeto de evaluación. Esta relación lineal se complica cuando se advierte que el alumno estudia más, o solamente, aquello que él supone que será evaluado. Con frecuencia, su propósito se centra en la aprobación de la evaluación y su esfuerzo de aprendizaje responde a esta lectura y a la demanda o el registro de claves acerca de lo que es importante para el docente. Estudia, y entonces aprende, en función de la evaluación, más que en respuesta a la enseñanza. La evaluación es el mensaje más claro para él sobre lo que debe aprender y demostrar que

sabe en el momento de la evaluación. Se ha observado, a menudo, que esta centralidad que los alumnos otorgan a la evaluación, producto de actitudes estudiantiles pero, asimismo, de actitudes de los propios docentes, estrecha el campo de lo aprendido ya que es obvio para todos que la totalidad de los contenidos que prescribe el diseño curricular, aunque fueran efectivamente enseñados, no podrían ser objeto de evaluación y que, además, para los alumnos puede ser suficiente cumplir con un aprendizaje superficial o meramente memorístico para aprobar la “prueba de evaluación”. Por esta razón, es relevante que en la información recogida sobre cuáles son los eventos que se presentan en la enseñanza que los alumnos consideran como instancias de evaluación, se compruebe que se refieren exclusivamente a aquellos en los que han recibido calificaciones. La evaluación, entonces, es importante para ellos porque culminarán con una calificación. Esta tiende a convertirse en el centro del interés de la actividad escolar. El alumno estudia, principalmente, para ser calificado y por la calificación. Es esta una de las cuestiones que debemos encarar muy pronta y enérgicamente pues es apremiante estimular, en cambio, la curiosidad de los alumnos, recuperar en ellos el interés por los saberes que les entrega la escuela e impulsarlos a emprender el camino de su autorrealización, determinaciones que implican recuperar para todos los alumnos el sentido que tiene la educación escolar.

Las instancias de evaluación integran la programación didáctica de los cursos escolares. Su cuidada determinación en lo referido a contenidos, naturaleza, frecuencia y formato de los instrumentos de evaluación a emplear debe responder a la estructura integral de la enseñanza proyectada.

En este campo, la programación de la evaluación que, como anteriormente afirmamos, presenta ante los docentes variadas opciones, puede sintetizarse en la disyuntiva de elección entre dos grandes modelos. Estos son el modelo empírico- analítico y el modelo no técnico¹⁶.

16. Leonhard Coomer, Donna (1986) “Reformulating the Evaluation Process,” in Kenneth A. Sirotnik, and Jeannie Oakes, eds., *Critical Perspectives on the Organization and Improvement of Schooling*. Boston: Kluwer/Nijhoff, 163-204.

El modelo “empírico-analítico”, que se inició y se viene desarrollando desde los comienzos del siglo XX, se caracteriza por preconizar y poner en acción una evaluación uniforme, objetiva, explícita, cuantitativa, sumativa, con base en la evidencia, destinada a calificar y diferenciar éxitos y fracasos. Se imbrica con una teoría del aprendizaje, una teoría curricular y una modalidad de programación didáctica que parten de la definición de fines y objetivos expuestos de manera específica y analítica en términos de las conductas que los alumnos deben demostrar como producto de sus logros al término del aprendizaje. Un rasgo central, admitida la especificidad de los objetivos sobre los que se estructura la enseñanza es, por ende, la fragmentación de los aprendizajes y de las evaluaciones que son un espejo de la segmentación del currículo, de los horarios escolares y de la organización de la actividad docente. Las evidencias sobre las que se construye la evaluación en este modelo, son conocimientos proposicionales (el alumno lo dice, lo escribe), conductas observables (acciones) y productos que son también evaluables por medio de la observación. Las evaluaciones son objetivas, independientes de quien evalúe, y son uniformes, iguales en lo que respecta a los criterios aplicados para todos los alumnos. Por lo tanto, se espera que siempre sean clara y definitivamente comparables con base en las mediciones cuantitativas de las evidencias recogidas. Los objetivos están prescriptos previamente y constituyen el criterio básico de la programación de la enseñanza y de la evaluación. En la denominada “evaluación referida a normas”, se impone la comparación entre los aprendizajes de los alumnos como fundamento para la calificación. En la llamada “evaluación referida a criterios”, la calificación surge de un análisis de discrepancia entre la respuesta del alumno y la respuesta preestablecida de acuerdo con el objetivo dispuesto. Este enfoque, posteriormente, pasó a reemplazar un régimen de calificaciones usualmente empleado en las escuelas, constituido por puntajes derivados definidos con base en la comparación de los alumnos entre sí, y lo substituyó por las formulaciones de los criterios en términos de objetivos operacionales específicos y conductuales.

En esta última modalidad del enfoque empírico-analítico, también se evalúan preferentemente los logros terminales en el aprendizaje

de los alumnos. Algunos ejemplos de autores que han desarrollado sus propuestas partiendo de los principios de este modelo se asocian con el movimiento de la eficiencia social iniciado por Franklin Bobbitt (1918:41 y sig.)¹⁷, ubicado en la línea iniciada por Edward Thorndike y continuada luego por Ralph Tyler¹⁸, Robert Mager¹⁹ y W. James Popham²⁰ y, más recientemente, por Daniel Stufflebean quien se ubica ya en la pauta de toma de decisiones y construye un nuevo marco, el modelo CIPP (Contexto, Insumo, Proceso y Producto)²¹. En este, se monitorean también los procesos o eventos diacrónicos o sincrónicos que van ocurriendo, en atención a los estados intermedios y a los logros finales. Es una evaluación que, además de considerar los aspectos cuantificables, atiende a dimensiones cualitativas. Asigna a la evaluación las funciones de las evaluaciones inicial, formativa y sumativa.

En el modelo empírico-analítico, la evaluación de los aprendizajes se apoya sobre un supuesto que es su regla básica de juego, y es que tenemos la posibilidad cierta de saber lo que otra persona sabe y que, con ese fin, podemos construir instrumentos que midan los saberes. Estos instrumentos serán los estímulos a los cuales los alumnos deben dar respuesta, lo cual sólo sería aceptable si se tiene en cuenta que esas demostraciones de saber tienen un alcance parcial porque son respues-

17. Bobbitt, Franklin (1918) *The Curriculum*, Cambridge, Mass. Hoghton Mifflin Co.

18. Madaus, George F. and Stufflebeam Daniel L. (eds), (1989), *Evaluation: Classic Works of Ralph W. Tyler*, Boston: Kluwer Academic Publishers, Una postura crítica respecto de algunos aspectos de este enfoque fue asumida posteriormente por algunos de sus fundadores, por ejemplo, Ralph Tyler en "On educational evaluation: a conversation with Ralph Tyler", Jeri Ridings Nowakowski (1983) *Educational Leadership*. ASCD p. 24-29.

19. Mager, R. & Pipe, P. (1984) *Analyzing Performance Problems, or You Really Oughta Wanna* (2nd Edition). Belmont, CA: Lake Publishing Co.; Mager, R. (1988). *Making Instruction Work*. Belmont, CA: Lake Publishing Co.

20. W. James Popham (1983) *Evaluación basada en criterios*. Madrid. Magisterio Español.

21. Stufflebeam, Daniel y Shinkfield, Anthony (1987) *Evaluación sistemática*. Paidós, Barcelona.

tas que dependen del instrumento, que adquieren sentido en tanto son respuestas relativas al instrumento empleado. Partiendo de este fundamento, se establece un principio en el cual se asienta la evaluación. En el modelo empírico-analítico la preocupación es cómo mejorar la definición de lo que hay que aprender, cómo mejorar la observación de lo que el alumno aprendió, cómo mejorar la medición en las observaciones y cómo interpretarlas mejor, en suma, cómo construir juicios justos con base en la evidencia²². Implícita se encuentra la idea de que se deben enseñar saberes que puedan ser medidos de acuerdo con estos cánones, a lo cual algunos agregan que, en consecuencia, y como son los que presentan características que los hacen pasibles de ser medidos con exactitud, ellos son los únicos saberes que deben ser enseñados. En apoyo a la implementación exitosa de los fines atribuidos a la educación escolar y a la programación y puesta en escena de la enseñanza y la evaluación, el enfoque empírico-analítico promueve la determinación de estándares genéricos y específicos para el aprendizaje.

Pero si algo hay que no podemos olvidar es que todo esto siempre va a constituir una evaluación parcializada, relativa a los criterios fijados y a los instrumentos construidos y que no va a ser una lectura completa y realista de todo lo que el alumno sabe y no sabe, ni de cómo lo sabe, qué aciertos podría demostrar y qué errores podría cometer si otros hubieran sido los instrumentos empleados en la medición.

El segundo modelo, denominado “no técnico”, coloca su eje en la comprensión del significado de las actividades educativas y no en la medición, como lo hace el modelo empírico-analítico. Presenta una multiplicidad de rasgos, entre ellos, la iluminación de los procesos de enseñanza y de evaluación, la implicación de los participantes, no solamente de los docentes sino también de los alumnos, en la producción de información y en la reflexión sobre los procesos de enseñanza, de aprendizaje y de evaluación, y con ese fin, el establecimiento de co-

22. Van Der Vleuten, C. P. M. M. Dolmans, D. H. J. Scherpbier, A. J. J. A.(2000) “The need for evidence in education”, *Medical Teacher*, 22:3, 246-250, DOI: 10.1080/01421590050006205.

municaciones entre todos los participantes en las situaciones de enseñanza, la construcción intersubjetiva de significados, el desarrollo de significados y sentidos cambiantes acerca de la realidad y, por último, la inclusión en la evaluación de conceptos, creencias, deseos y conocimientos que no se expresan verbalmente, esto es, de conocimientos tácitos y aspectos correspondientes a dimensiones no cognitivas del aprendizaje.

El modelo no técnico se define por la denominación negativa que porta. Se enfrenta al modelo empírico-analítico al que considera inadecuado tanto en su propuesta de enseñanza y de currículo cuanto al enfoque que plantea para la evaluación de los aprendizajes. Quizá la expresión más terminante de ese rechazo sea la comparación que Elliot W. Eisner hiciera de los dos enfoques. Sostuvo que en el modelo “no técnico”, los estudiantes no se limitan a aprender lo que los profesores les enseñan, por el contrario de lo que el modelo empírico-analítico manifiesta que es:

“la visión de un ejército uniformado de jóvenes adolescentes marchando todos al compás de un mismo tambor, al mismo ritmo, hacia un mismo objetivo, (esta) puede ser una visión que alegra los corazones de los tecnócratas, es una visión que tiene poco o nada que ver con esos deliciosos resultados que constituyen las sorpresas de la experiencia educativa. Los estudiantes aprenden, de hecho, mucho más y mucho menos que lo que se proponen los profesores. Los educadores deben estar agradecidos por esto” (2000:344)²³.

Los dos modelos de evaluación que hemos reseñado, en efecto, se oponen. Varios temas marcan diferencias fundamentales. En primer lugar, los logros de aprendizaje que uno y otro persiguen. En un caso se trata de objetivos puntuales, específicos, en los que la calidad y efectividad del aprendizaje se miden fragmentariamente como producto de

23. Eisner Elliot W. (2000) “Those who ignore the past...: 12, ‘easy’ lessons for the next millennium”. Curriculum Studies, Vol. 32, No. 2, 343-357.

la subdivisión de conductas y pensamientos complejos, casi pulverizados, que se pueden aprobar o desaprobado aisladamente, segmentados e independientes unos de otros. Traducido en la forma de pruebas de evaluación está claro el tipo de trabajo que se solicita a los alumnos y, lo que es más grave, la forma de estudiar que suscitan en los alumnos. El modelo no técnico se cuestiona, en cambio, la modalidad de la enunciación de los llamados “logros de aprendizaje”. Son ya muchas las críticas que esta ha recibido no sólo debido a la especificidad abusiva que se confiere a los objetivos de aprendizaje sino, también, y paradójicamente, por causa de la ambigüedad y la falta de definición de aspectos de valor pedagógico²⁴. Si bien se pretende que sean de clara lectura, el empleo de términos como habilidad, capacidad o destreza siempre referidos a una conducta particularizada, no señalan ni magnitudes de efectividad ni niveles en las etapas de desarrollo para cada uno. Se pretende que admiten una única lectura y que no generan dudas acerca de lo que está bien o mal y que, además, se trata de logros que se obtienen o no se obtienen. Son, sin embargo, términos polisémicos que docentes y alumnos pueden interpretar de diversos modos. Se ha demandado y proclamado que la precisión es indispensable, pero lejos está de haberse plasmado en la práctica.

¿Es innata la capacidad o la habilidad, o es un producto del esfuerzo del alumno? ¿Alcanza un nivel máximo de desarrollo que es propio de cada alumno y que, cuando llega a su techo, éste no puede superar? ¿Las condiciones que generan fracaso son naturales o son producto de la falta de esfuerzo del alumno? ¿O dependen de la enseñanza que ha recibido, la que no ha sido efectiva en general para el conjunto de la clase o no ha sido apropiada para ese alumno en particular? ¿Se trata de un alumno que como aprendiz requiere más tiempo para apropiarse del conocimiento, para pensar y para responder a sus propias preguntas?

24. Davies, Philip (1999) “What is evidence-based Education?” *British Journal of Educational Studies*. Vol. 47 N° 2; Trevor Hussey and Patrick Smith (2008) “Learning Outcomes; a conceptual analysis”. *Teaching in Higher Education*, 13:1,107-115 Taylor & Francis. DOI: 10.1080/13562510701794159

Las respuestas a estas cuestiones tienen una importancia decisiva no solamente a la hora de formular juicios evaluativos sino, mucho más aún, cuando se elabora la programación didáctica de un curso y de sus unidades didácticas.

Otro aspecto crítico del enfoque empírico-analítico fue señalado por Eisner: la uniformidad de propuestas en la enseñanza y de logros esperados. La falta de atención a la diversidad que, a veces, no se manifiesta tan claramente en la enseñanza, sí resulta determinante en el momento de la evaluación. En cada objetivo o logro de aprendizaje hay un producto o un desempeño que representa el punto de llegada deseado, el nivel de excelencia. Más abajo quedan los errores y los fallos o vacíos, que son castigados con disminución de puntos. Sólo los muy buenos se elevan hasta la cima de la mejor calificación. Advirtiendo los aspectos negativos de este rasgo del enfoque que critican, Duncan-Andrade y Morrell (2008) expresan que *“El sistema de creencias darwinianas está incorporado en la mayoría de las escuelas a través de la existencia, en gran medida indiscutida, de un sistema pedagógico de las calificaciones y las pruebas de evaluación que, por las características de su diseño, garantiza el fracaso de algunos”* (cit. por Gillian Parekh (2017:2)²⁵.

El modelo no técnico, a diferencia del empírico-analítico, es interpretativo. Trata de comprender los significados compartidos con y entre los alumnos y se preocupa, asimismo, por los pensamientos, razonamientos, niveles de comprensión, intereses, aspiraciones y preocupaciones, temores y ansiedades de los alumnos. Sostiene que es preciso vencer muchas dificultades y tener en cuenta un conjunto de variables relevantes para formular objetivos de aprendizaje. No niega que estos son parte de la programación didáctica pero considera que deben ser formulados de manera más general, abierta y flexible para adecuarse

25. Duncan-Andrade, J. M. R., & Morrell, E.(2008). The art of critical pedagogy: Possibilities for moving from theory to practice in urban schools. New York: Peter Lang [Crossref], [Google Scholar], p. 2) cit por Gillian Parekh (2017) “The tyranny of “ability”, Curriculum Inquiry, 47:4, 337-343, DOI: 10.1080/03626784.2017.1383755 <https://doi.org/10.1080/03626784.2017.1383755> Published online: 21 Nov 2017.

a la variedad de intereses y de experiencias de los alumnos y que, a la reconocida heterogeneidad de los estudiantes, le debe corresponder una cierta heterogeneidad de objetivos. Advierte que, si se han formulado objetivos en algún modo diferenciados, sea por los temas, los tipos de conocimiento, los niveles de dificultad, las tareas, las formas de expresión, los ritmos de aprendizaje, la progresión en los niveles de aprendizaje o los saberes y las experiencias previas, tanto la enseñanza cuanto la evaluación deben procesar esas variables, atenderlas en clase y ofrecer múltiples oportunidades de aprendizaje. En consecuencia, las instancias de evaluación deben relacionarse con la variedad de los procesos que cada alumno ha transitado, está recorriendo y continuará atravesando. El modelo no técnico procura que enseñanza y evaluación estén alineadas y que contribuyan, en conjunto, al desarrollo personal de los alumnos. Se propone desarrollar en ellos capacidades de metacognición, de autoevaluación y de autorregulación.

Incorporar en la evaluación estas ideas es una tarea creativa y por esto mismo, muy atrayente. Algunas de sus propuestas se han ido desarrollando desde hace ya casi tres décadas. Entre ellas, se le ha dado un lugar importante al error en el aprendizaje. No se trata de penalizarlo sino de resolverlo porque se acepta que no hay proceso de aprendizaje sin que se comentan errores. El problema siempre es qué se hace con el error. La evaluación formativa, que se convirtió en la estrella del firmamento de la evaluación, es considerada el dispositivo más eficaz para la mejora del aprendizaje y la enseñanza. Multitud de trabajos se han realizado para determinar las mejores formas de brindar retro y pro-alimentación útil a los alumnos a partir de la presentación progresiva de sus trabajos. El uso de rúbricas contribuye a hacer más transparentes los criterios de calificación, más creíbles las calificaciones recibidas, los alumnos las pueden emplear como apoyo en el estudio y, lo que es más importante, contribuyen a desarrollar la capacidad de autoevaluación cuando las rúbricas se presentan y se discuten con los alumnos antes de usarlas en evaluaciones formativas o sumativas. La evaluación, en la medida en que cumpla esta misión, va adquiriendo una tesitura diferente, porque de ser un instrumento de control destinado a calificar, pasa a convertirse funcionalmente, sirviendo ahora de acompañamiento de un

aprendizaje que, siempre que sea posible, se propone responder intrínsecamente al interés del alumno. Su motivación rotará, entonces, de la búsqueda de la pura calificación al saber que satisface una necesidad. Es así cómo, a los instrumentos convencionales de evaluación, sean subjetivos u objetivos, se ha incorporado un abanico extenso y variado de nuevas instancias de evaluación. Son las evaluaciones, llamadas, por esta razón, evaluaciones alternativas, con una fuerte presencia entre ellas de evaluaciones auténticas en las que se exponen situaciones, problemas, casos y proyectos de la vida real. La evaluación abandona la centralidad del conocimiento proposicional que se expresa con la palabra escrita o hablada para asumir una gran diversidad de formas de expresión del saber. La evaluación ha ampliado sus horizontes y, sin perder las formas instaladas en la escuela desde hace mucho tiempo, se abre a nuevos formatos. La vida entra al aula y hace un lugar al desarrollo de la creatividad. Dos modalidades son útiles en estos procesos de crecimiento: las evaluaciones multinivel y las evaluaciones de proceso.

Las evaluaciones multinivel se corresponden con la atención a la diversidad de los alumnos. Las evaluaciones empírico-analíticas son uniformes y tienen una respuesta única final correcta. En las evaluaciones multinivel se presentan diferentes formatos que los estudiantes pueden adoptar. Uno de ellos consiste en la inclusión de preguntas, problemas, casos o proyectos abiertos, que admiten distintas respuestas correctas producidas en diferentes niveles de complejidad o dificultad e, incluso, eventualmente, en sucesivas etapas de desarrollo en el aprendizaje (introdutorio, intermedio, superior). Una misma consigna, por ejemplo, puede ser respondida en estos distintos niveles.

En otra modalidad de la evaluación multinivel, se ofrece un conjunto de consignas de diverso tema o tarea y la consigna señala que de ese listado, cada alumno, individualmente o en grupo, según se establezca, debe desarrollar los trabajos que según el profesor indica, representan el nivel básico indispensable para la aprobación. Empero, si lo prefieren, pueden sumar otros trabajos que encuentren entre las consignas y que resulten, también, de su interés, lo cual les permitirá incrementar su puntaje por sobre el mínimo necesario establecido para la aprobación.

Finalmente, otra posible modalidad consiste en ofrecer a los alumnos un menú de tareas diferentes en las que se trabajan alguno o algunos de los temas estudiados en el curso. Las consignas representan distintos niveles de posible dificultad, atendiendo a las estrategias de conocimiento que los alumnos deberían emplear para desarrollarlos. Cada uno de esos trabajos tiene señalado un valor en puntos que el alumno irá adicionando a su calificación al realizar cada tarea. Se establece una escala según la suma de puntos que cada alumno obtenga y se comunica la suma mínima de puntos que es necesario alcanzar para lograr la aprobación. En todos estos casos, el alumno tiene oportunidad de obtener el mejor puntaje que corresponde al trabajo elegido. Es preciso señalar que la calificación no se asigna restando puntos del máximo posible. La intención es que todos los alumnos encuentren uno o más trabajos que puedan realizar bien.

A las evaluaciones multinivel se añaden las evaluaciones de “seguimiento de proceso”. Estas evaluaciones se diferencian de las corrientemente empleadas “evaluaciones instantáneas” o “de estado” en las que los alumnos rinden una prueba de evaluación respondiendo a las consignas entregadas por el docente en un momento y lapso determinado, como ocurre, por ejemplo, con los exámenes. El profesor recoge las respuestas y califica según los puntos obtenidos por el alumno en esa instancia de una o más horas, un día establecido. Se las llama evaluaciones instantáneas porque su recopilación es semejante a la de un álbum de fotografías. Son instantáneas de momentos separados e independientes unos de otros. En las evaluaciones de proceso, por el contrario, se solicita a los alumnos la producción de un trabajo en un tiempo no inferior a dos o tres semanas. Se establecen etapas en la elaboración y se fijan fechas para sus sucesivas presentaciones. En cada una, el docente efectúa con el alumno una evaluación formativa indicando los aspectos bien logrados y solamente los errores más importantes y, orientando hacia el trabajo futuro del alumno, expresa los señalamientos que le serán útiles para continuar y mejorar su trabajo.

Estas presentaciones no se califican²⁶. Solamente se calificará el trabajo final cuando el estudiante lo presente en la instancia que es fijada con ese carácter. La elaboración del producto se convierte, de esta forma, en un proceso de aprendizaje monitoreado periódicamente por el profesor. En los tiempos intermedios de trabajo durante las distintas etapas, los alumnos, en grupo, pueden compartir sus ideas y ayudarse colaborativamente para perfeccionar la tarea y sus resultados. Los trabajos y el registro de las evaluaciones formativas permiten que el álbum de fotografías sea reemplazado por la construcción de una película que muestra qué ocurre en el transcurso del tiempo.

En la evaluación de los aprendizajes los dos enfoques, empírico-analítico y no técnico se encuentran en uso. Ambos proveen información a docentes y alumnos. La enseñanza también responde a uno u otro modelo. Como hemos dicho, enseñanza y evaluación deben estar alineadas. Pensar en la evaluación nos obliga a pensar en la enseñanza. Las concepciones de ambas responden a diferentes modos de concebir la enseñanza, el aprendizaje y la evaluación. Condiciones de practicidad imperan en algunas circunstancias y conducen a elegir, prevalentemente, uno de estos modelos. Sin embargo, y a pesar de que representan enfoques pedagógicos de distinta naturaleza, no es imposible pensar en una combinación de ambos.

La misión del docente, no obstante, en el modelo no técnico está mejor desarrollada en los términos de las acciones que podemos considerar propias de una buena enseñanza y, lejos de poner a la evaluación al servicio de la selección de los mejores, cumple con el propósito de ayudar a todos y de hacer florecer sus talentos en cada uno al brindarle oportunidades y apoyos para ampliar y engrandecer su yo ante sí mismo y ante los demás.

26. Gibbs, Graham and Simpson, Claire "Does your assessment support your students' learning?" (s/d) Centre for Higher Education Practice, Open University 1-30; Gibbs, Graham, (2010) Using assessment to support student learning. Metropolitan University of Leeds, 1-71.

Cuando limitamos la evaluación al uso de los formatos de evaluaciones empírico-analíticas y empleamos exclusivamente evaluaciones instantáneas y uniformes mononivel, por ejemplo, estamos estrechando el campo de la enseñanza y el aprendizaje en un doble sentido. En primer lugar, separamos evaluación y enseñanza, distanciándolas y perdemos, en consecuencia, la riqueza que, a través de las orientaciones proporcionadas durante el proceso de aprendizaje, podríamos ofrecer al alumno para mejorar su nivel de comprensión y su capacidad de materialización de las ideas en el trabajo que realiza. Y, en segundo lugar, la limitación de la información que nos brinda el enfoque empírico-analítico cuando se emplea con exclusividad, nos plantea una cuestión que caracteriza la debilidad de ese enfoque para apoyar positivamente lo que denominamos “buena enseñanza”.

Esta requiere que los profesores fundamenten sus juicios acerca de los resultados de aprendizaje de los estudiantes en una variedad de tipos de actividad adecuados para los objetivos, los contenidos, el nivel de escolaridad, la edad y las características personales de los estudiantes. Una enseñanza rica, interesante y sugerente ofrece a los alumnos oportunidades para demostrar lo que pueden hacer a través de tareas y actividades que abordan no solamente la matriz de objetivos de aprendizaje formulados en el diseño curricular, sino que va más allá de estos, permitiendo su ampliación en un abanico de contenidos temáticos y habilidades que responden a opciones abiertas a diferentes saberes.

Con ese propósito, en la evaluación que responde a un enfoque no técnico se emplea una variedad de instrumentos que permiten a los alumnos manifestar el avance que han logrado en una pluralidad de aprendizajes diversos. Las ocasiones para recoger evidencias de aprendizajes no se limitan al uso de instrumentos formales en circunstancias formales de evaluación. Numerosas situaciones de un sin número de especies distintas permiten, por consiguiente, que los profesores reúnan información sobre sus alumnos a través de la observación, las preguntas y respuestas en la clase, escuchando las conversaciones informales, interesándose por las actividades extra-clase de sus alumnos, y revisando trabajos escritos, videos, audios, individuales y grupales. En la evaluación formativa esta información se puede utilizar inmediatamente para

ayudar a los alumnos o puede ser almacenada y utilizada para planificar las oportunidades de aprendizajes futuros.

La evaluación hoy se enfrenta a los problemas de conciliar el empeño por proporcionar las mejores aportaciones con el fin de apoyar el aprendizaje de los alumnos, con la necesidad de abarcar a todos, sin exclusiones y contribuir, por estas vías, al mejor funcionamiento del sistema educativo de acuerdo con principios de justicia y equidad. Con este propósito, la evaluación de los aprendizajes de los alumnos debe estar acompañada por la evaluación de la enseñanza, ambas desarrolladas con base en el enfoque que interpreta a la evaluación como un proceso que nos proporciona una ayuda a partir de brindar información muy fina, muy delicada, relevante y personalizada, necesaria para mejorar la calidad de la enseñanza y la calidad del aprendizaje. La condición para que este proceso evaluativo sirva efectivamente de soporte de las decisiones que toman los alumnos para perfeccionar su aprendizaje y los profesores para mejorar la enseñanza, es que todos se autoevalúen, que construyan esta evaluación no solamente con la condición de ser fundadamente justa sino que, además, visualicen claramente que son imprescindibles para decidir qué acciones les conviene emprender en adelante.

Si bien estas evaluaciones no coinciden con el tipo de información que requieren las autoridades para poder tomar decisiones en el nivel de los sistemas educativos y que provienen en parte de evaluaciones en gran escala, cuando se emprende la tarea de realizar una evaluación institucional, del currículo o de la calidad de la enseñanza, es menester advertir que la evaluación de los aprendizajes de los alumnos que nos brindará aspectos sustanciales de la información que precisamos, es la que nos podrá suministrar esa información fina, necesariamente cualitativa, propia del enfoque no técnico porque es particularmente relevante para el tipo de decisiones que en todos estos casos hay que adoptar.

Sin duda, en todo proceso de evaluación, contribuye a su buena realización, el contar con información cuantitativa, pero la base de la información que todos, docentes y alumnos, demandan para ponerla al servicio de un mejor aprendizaje y una mejor enseñanza, incluye la que proporciona el enfoque no técnico porque es interpretativo y se preo-

cupa por los significados que el docente y el alumno construyen acerca de los conocimientos que se han categorizado como saberes necesarios en el aprendizaje escolar y sobre los procesos mismos de aprendizaje. No se puede resolver la evaluación, por tanto, de una única forma. Sabemos que hay clases de información que no pueden omitirse. En este sentido, si es una condición establecida que nos manejemos necesariamente con enfoques empírico-analíticos, lo cual puede ser una decisión que se adopta en el nivel del sistema, de la escuela y del propio docente, debieran incluirse, igualmente, prácticas evaluativas orientadas por un enfoque no técnico.

Hasta acá, he tratado de presentar algunos de los problemas y de las dificultades así como algunas de las posibles soluciones y caminos para examinar el estado actual de la cuestión de la evaluación de los aprendizajes en su relación con la enseñanza.

JUSTICIA EN LAS EVALUACIONES

Tamara Vinacur

39

INTRODUCCIÓN

Desde fines del siglo pasado y hasta la actualidad, en Argentina, se identifica un incremento de la segregación escolar, lo que constituye un obstáculo para el desarrollo igualitario de la educación para la construcción de una sociedad más justa, al mismo tiempo que favorece la reproducción de las desigualdades sociales (Tenti Fanfani, 2019). Ante este contexto, hablar de evaluaciones más justas pareciera ser una preocupación menor, sin embargo, la evaluación en tanto actividad social, solo puede ser entendida considerando los contextos culturales, económicos y políticos en los que opera, reconociendo su potencial político, técnico y su contribución a visibilizar el grado en que los sistemas educativos consideran la diversidad y heterogeneidad de los estudiantes. De este modo, el proceso de construcción de una educación justa y la justicia social, implica entre otros factores, una evaluación que promueva la equidad, la inclusión, la participación y la democracia (Gipps y Stobart, 2010; Murillo, Román y Hernández Castilla, 2011).

A partir de esta perspectiva, el concepto de justicia social fue analizado desde múltiples miradas (Rawls, 2006; Sen, 1996; Fraser, 2005; Walzer, 2001). Nancy Fraser (2008) reconoce la naturaleza multidimensional y compleja de la justicia social y propone un abordaje conformado por tres esferas que requieren ser abordadas conjuntamente: a) la primera hace referencia a la redistribución o justicia económica, b) la segunda, al reconocimiento o justicia cultural, y c) la tercera, a la representación o justicia política. La perspectiva de la redistribución, se basa

en las reivindicaciones que pretenden una distribución más justa de los recursos y de la riqueza, por su parte, la esfera del reconocimiento, hace referencia a la valoración social y cultural de todos los individuos y considera los distintos modos de ser, hacer y pensar. Finalmente, la dimensión de la representación remite a la existencia de las condiciones adecuadas para la plena participación en la vida social (Murillo e Hidalgo, 2015).

Por su parte, Dubet (2017) reconoce la tensión entre los conceptos de igualdad de posiciones y la igualdad de oportunidades, que son los dos pilares sobre los que se asienta la justicia social. Define la igualdad de posiciones como aquella que posibilita que todos tengan un piso aceptable de condiciones de vida y acceso a la educación, a los servicios y a la seguridad. Por otra parte, la igualdad de oportunidades hace referencia a que cada uno pueda acceder a estos bienes y servicios de acuerdo a su esfuerzo y talento personal, en un contexto de competencia equitativa. Dubet enfatiza la necesidad de garantizar desde las políticas públicas el “modelo de las posiciones”, considerando que permite atenuar las brechas entre diferentes estratos sociales.

La profundización sobre el concepto de justicia social de acuerdo con las distintas perspectivas teóricas excede el alcance de este capítulo. Sin embargo, reconociendo la centralidad que adquiere la educación en la construcción de una sociedad más justa (Tedesco, 2012; Dubet, 2017; Connell, 1997) se considera relevante su abordaje, ya que brinda el marco conceptual a partir del cual se explicitan los principios que subyacen a los modos de diseñar, implementar y a utilizar los resultados de las evaluaciones de los aprendizajes a gran escala.

Tomando en consideración las perspectivas analíticas esbozadas más arriba, el presente capítulo presenta el marco conceptual en el que se desarrolla la justicia evaluativa. En él, se identifican algunas experiencias de justicia en las evaluaciones de los aprendizajes en distintos sistemas educativos tanto a nivel internacional, nacional y jurisdiccional. Para este último caso, se realiza una caracterización del modo en que institucionalmente se está avanzando con este desafío desde la Unidad de Evaluación Integral de la Calidad y Equidad Educativa (UEI-

CEE) del Ministerio de Educación e Innovación de la Ciudad Autónoma de Buenos Aires.

I. EL CONCEPTO DE JUSTICIA EN LA EVALUACIÓN EDUCATIVA

Hace más de 30 años que hay acuerdo en que la justicia (o “*fairness*”) es uno de los principales criterios que los sistemas de evaluación deben contemplar, aunque su significado fue variando a lo largo del tiempo (AERA, APA, NCME, 1999; Kunnan, 2000; 2013). De acuerdo con la definición tradicional, una evaluación justa es aquella en donde no hay sesgos, hay un tratamiento equitativo de todos los evaluados, y hay equidad en la oportunidad de aprendizaje del material que será objeto de evaluación (Xi, 2010: 147).

En este sentido, garantizar un tratamiento equitativo a todos los evaluados en el proceso de evaluación supone un conjunto de procedimientos vinculados con su diseño, su administración, el modo en que se proporcionan oportunidades de aprendizaje equiparables, como son las condiciones de accesibilidad a la evaluación, entre muchos otros factores que afectan el contexto en que se realizan las evaluaciones de aprendizajes. Sin embargo, si bien existe un consenso en la literatura revisada sobre la necesidad de atender a la justicia en las evaluaciones, la perspectiva desde la cual se la conceptualiza permite poner en foco distintos aspectos y diseñar intervenciones diferenciadas.

Warmington & Macintosh (2003) reconocen cuatro perspectivas en las que suele abordarse la discusión sobre justicia en la evaluación o “*fairness*”²⁷, que a los fines de este trabajo se los consideran como si-

27. Hay discusiones acerca si el concepto de justicia es homologable al de “*fairness*”, e incluso algunos autores (Kunnan, 2013) diferencian “*fairness*” para remitirse al plano de las personas (asumiendo que todas las personas debieran ser tratadas con el mismo respeto), del concepto de “justicia”, utilizado para dar cuenta del nivel institucional (pensando en las instituciones responsables del diseño, administración de evaluaciones y comunicación de resultados).

nónimos: justicia en tanto ausencia de sesgos; justicia en tanto un tipo de intervención equitativa en el proceso de evaluación; justicia en tanto igualdad en los resultados de la evaluación; y justicia en las oportunidades de aprendizaje.

Habitualmente la discusión técnica sobre la justicia en la evaluación se centra en los dos primeros, considerando que las intervenciones que se requieren para propiciar la igualdad en los resultados y garantizar oportunidades equitativas de aprendizaje para todos los estudiantes superan las competencias específicas de las áreas responsables de la evaluación de los aprendizajes a nivel del sistema educativo.

Desde ya, esto no supone que las otras perspectivas no sean relevantes ni requieran de consideraciones específicas como insumo para el diseño de evaluaciones alineadas a lo que los estudiantes hayan tenido oportunidad de aprender, y de manera similar, que los resultados sean analizados considerando las características de subgrupos específicos, tales como género o discapacidad.

En uno de los textos clásicos sobre evaluación educativa, Alicia Camilloni identifica las cuatro características generales que los instrumentos de evaluación deben presentar, en tanto criterios o requisitos a tener en cuenta al momento de tomar decisiones acerca del programa de evaluación de un curso y/o sobre cada uno de los instrumentos de evaluación que lo integran. Por ejemplo: la validez, confiabilidad, practicidad y utilidad. Si bien no hace referencia explícita a la justicia como un requerimiento técnico, la autora menciona que, en relación con la validez de contenido, “...la justicia de la evaluación se pone en juego, entonces, cuando se efectúa el muestreo de los contenidos que se incluirán en el instrumento a través de las consignas de trabajo (por ejemplo, las preguntas) que éste presenta” (Camilloni, 1998:75). Este fragmento evidencia que la autora concibe a la justicia como uno de los aspectos de la validez de constructo, perspectiva que también tiene su correlato en la literatura internacional (Messick, 1989; Kane, 2006).

En esta línea, la discusión acerca del vínculo entre los conceptos de validez y justicia ha sido abordada por varios autores que cuestionan si la validez incluye a la justicia o si se trata de dos conceptos independientes.

Si se presenta una definición acotada de la validez (Popham, 1997) y la justicia como más amplia (Kunnan, 2000), entonces la validez puede ser concebida como un componente de la justicia. De otro modo, si definimos justicia como un requerimiento técnico y a la validez en sentido más amplio (Messick, 1989; Kane, 2006), entonces la justicia se concibe como un componente de la validez. Desde esta última perspectiva, la validez se define como “*un juicio evaluativo integrado del grado en el cual la evidencia empírica y teórica soporta la adecuación y pertinencia de las inferencias y acciones basadas en los puntajes de un test o en otros modos de evaluación*” (Messick 1989: 13), que reconoce su vinculación con la idea de justicia en tanto que, sin haberse establecido la validez, difícilmente sea posible justificar cualquier interpretación de los puntajes o decisiones que se tomen a partir de los resultados (Chapelle, 2011; Púrpura, Brown & Schoonen, 2015).

Por otra parte, una serie de autores han cuestionado la efectividad o poca practicidad de la noción de validez para poder identificar dificultades específicas en la justicia de una evaluación (Kane, 2012: 8), lo cual dio lugar a distintos modos de conceptualizar la justicia en las evaluaciones: como parte de un marco conceptual de discusiones sobre la validez. Un referente de esta línea es Anthony John Kunnan (2004), quien diseñó un modelo conceptual para estudiar la justicia en las evaluaciones, a partir del campo de evaluación de idiomas. Este autor considera que “*fairness*” es un concepto que integra las tres dimensiones: validez, justicia y acceso, operando como marco que, en versiones posteriores, amplía a cinco dimensiones: validez, ausencia de sesgo en las evaluaciones, acceso equitativo al aprendizaje y a la evaluación, administración de la evaluación, y consecuencias o implicancias sociales. Entonces, para que una evaluación sea justa debieran proporcionarse evidencias de los siguientes aspectos:

Tabla 1

44

Dimensiones de la justicia en las evaluaciones	Evidencias
Validez	Representatividad del contenido a ser evaluado o evidencia que demuestre que aquello que es evaluado representa el objeto de evaluación previsto. Validez de constructo, que hace referencia a que los ítems incluidos en la evaluación representan el constructo o rasgos subyacentes a los que se pretende medir a partir de un instrumento determinado. Validez de criterio, en qué medida los puntajes de una prueba reflejan un cierto conjunto de habilidades o no.
Ausencia de sesgos	Hace referencia a la ausencia de sesgos tales como lenguaje o contenido ofensivo (McNamara and Roever, 2006), penalización injusta basada en el contexto de quien evalúa, resultados desiguales según el subgrupo al que pertenezcan los evaluados, y/o diferencias en el establecimiento de puntajes de corte.
Acceso equitativo al aprendizaje y a la evaluación	En términos de acceso educativo, financiero, geográfico, de equipamiento o por situación de discapacidad. En otros términos, todos los estudiantes debieran contar con las mismas oportunidades para aprender el contenido sobre el que serán evaluados y familiarizarse con el tipo de tareas y estrategias que se requieren en la producción de la respuesta.
Condiciones de administración de la evaluación	Hace referencia a las condiciones físicas en que se realiza la prueba, tales como la iluminación, la temperatura, así como también a las condiciones similares en que se realiza la evaluación en distintos espacios, y la equivalencia en las formas.
Consecuencias o implicancias sociales	Es un aspecto a profundizar que hace referencia al modo en que una evaluación podría afectar prácticas institucionales o el tipo de intervenciones que se proponen a partir de los resultados
Fuente: Elaboración propia a partir de Kunnan, A. J. (2013). Fairness and justice in language assessment. <i>The companion to language assessment</i> , 3, 1098-1114.	

En la práctica, los Estándares para la Evaluación Educativa y Psicológica (“Standards for Educational and Psychological Testing”, en inglés) elaborados en 1974 en Estados Unidos constituyeron la primera referencia a nivel internacional sobre cómo abordar la justicia en la evaluación.

Actualmente, los estándares definidos (AERA et al., 2014) en el año 2014, procuran orientar el desarrollo de procesos evaluativos más justos y libres de sesgo y se organizan en cuatro grupos que hacen referencia a:

1. Las diferentes fases de una evaluación: el diseño, la administración, el tratamiento de la información y los procedimientos de calificación.
2. La validez de los puntajes obtenidos como resultados y la validez de las interpretaciones de dichos resultados para los diferentes grupos de estudiantes.
3. Las acomodaciones o adecuaciones propuestas para remover las barreras que pudieran existir en el acceso a las evaluaciones y promover interpretaciones válidas de los puntajes y los usos de los resultados.
4. Las garantías que se ofrecen con el fin de evitar interpretaciones inadecuadas de los puntajes de las pruebas.
5. En esta edición de los estándares (2014), se presenta a la justicia y validez como aspectos inseparables, dado que la justicia hace referencia tanto a la validez de una evaluación en particular, como medida de la habilidad de un estudiante o grupo de estudiantes, como al proceso de evaluación en sentido amplio, en tanto contribuye o refleja la equidad social.

En este sentido, se define a la justicia como un derecho de todos los individuos y grupos que constituyen la población a ser evaluada, por lo que debe garantizarse que todos los estudiantes evaluados tengan igual oportunidad de contar con puntajes interpretados en forma válida para el uso que se pretenda (AERA, 2014: 52). Es decir, se entiende que una prueba justa es aquella que refleja el mismo constructo para todos los estudiantes evaluados, y que los puntajes obtenidos tienen el mismo

significado para todos los individuos en la población objetivo (AERA et al., 2014: 50).

46

El interés por la justicia en las evaluaciones está presente en la mayor parte de los sistemas de evaluación a nivel internacional, los cuáles expresan las preocupaciones por contar con evaluaciones de los aprendizajes más justas, tanto en relación con el diseño, con la aplicación como también con el análisis de los resultados. Esto se observa en la producción de resguardos en el diseño de los instrumentos de evaluación desde los principios de diseño universal (en Estados Unidos), o acceso justo por diseño (en Reino Unido). Estas prácticas benefician no solo a los alumnos de grupos minoritarios sino al conjunto de los estudiantes evaluados, ya que por ejemplo en el caso de diseño universal, garantizar mejoras en la accesibilidad para personas en situación de discapacidad implica introducir instrucciones simples, claras de comprender, procedimientos intuitivos, máxima claridad en la tipografía, por mencionar algunos ejemplos (Thompson, Johnstone, Christopher y Thurlow, 2002).

Un grupo de especialistas en evaluación de la Universidad de Stanford en Estados Unidos -*Stanford Center for Opportunity Policy in Education*- identificó cinco criterios que son necesarios para contar con evaluaciones de aprendizajes de alta calidad. Entre ellos se incluye el requerimiento de que las evaluaciones sean válidas, confiables y justas para evaluar las habilidades de los estudiantes pero que también debieran proporcionar consecuencias positivas para la calidad de la enseñanza y para las oportunidades de aprendizaje de los estudiantes.

Algunos de los indicadores que el equipo de técnicos y especialistas seleccionó para medir el nivel de justicia de una evaluación son los siguientes:

- Evidencia de que los contenidos y las habilidades que se pretenden evaluar están bien medidos.
- Evidencia de que los puntajes se vinculan con las habilidades sobre las que se intenta realizar inferencias.
- Evidencia de que las evaluaciones están bien diseñadas y son válidas para cada uno de los usos previstos y que dichos usos

son apropiados a los propósitos de la evaluación y pueden presentar evidencias de validez.

- Evidencia de que las evaluaciones están libres de sesgo y miden los conocimientos y habilidades de los estudiantes de diferentes contextos culturales, socioeconómicos, como así también de los alumnos en situación de discapacidad.
- Evidencias de que los resultados de las evaluaciones se ubican en un *continuum* de aprendizaje, coherente con los propósitos a los que la evaluación pretende contribuir.

Por otra parte, ETS -por sus siglas en inglés “Educational Testing Service”- (2016) establece un conjunto de principios para la justicia en las evaluaciones en donde define a la justicia en el contexto de la evaluación como la medida en que las inferencias y acciones basadas en los puntajes de una evaluación resultan válidos para diferentes grupos de personas evaluadas. A continuación se menciona el conjunto de principios definidos:

- Medir los aspectos más importantes del contenido priorizado como objeto de evaluación.
- Eliminar obstáculos cognitivos irrelevantes que pudieran afectar a los estudiantes evaluados.
- Eliminar obstáculos emocionales irrelevantes que pudieran afectar el desempeño de los estudiantes evaluados (por ejemplo contenido ofensivo).
- Eliminar obstáculos físicos irrelevantes que pudieran afectar el desempeño de los estudiantes evaluados (particularmente en el caso de los alumnos en situación de discapacidad).

En este contexto, se consideran algunas variables clave para la identificación de algunos aspectos que debieran recibir especial atención en los grupos a ser evaluados: edad, características atípicas, status de ciudadanía, etnicidad, género (incluyendo identidad de género o representación de género), discapacidad mental o física, origen nacional o regional, etnia, religión, orientación sexual y nivel socioeconómico.

También se identifican obstáculos cognitivos que debieran tratar de evitarse en el diseño de los instrumentos de evaluación, por ejemplo, aquellos vinculados al lenguaje (incluir vocabulario específico que no es pertinente en relación con el objeto de evaluación priorizado), temas que no se vinculan directamente con lo que se pretende evaluar y para lo cual se requiere un conocimiento específico (por ejemplo, la cantidad de jugadores que intervienen en un equipo de fútbol, si lo que se evalúa no fuera educación física), traducciones que podrían reflejar obstáculos culturales (por ejemplo, en el nombre de marcas de productos, de las festividades locales, de los aspectos propios de la geografía u organización política de un país, o incluso de las unidades de medida), contextos ajenos a los estudiantes (por ejemplo, si se presenta un texto estímulo sobre la nieve y la evaluación se aplica a estudiantes que no la conocen, o en una prueba de matemática hay un texto introductorio muy largo, que perjudica a los alumnos con dificultades en la comprensión lectora) y religión (hay ciertos tipos de textos, o de presentación de los ítems que descuidan elementos que afectan la sensibilidad de algunos grupos).

Otro tipo de obstáculos a considerar son los emocionales en tanto, los contenidos de una evaluación, no debieran producir malestar en los estudiantes, ni distraerlos de la atención que requiere la comprensión de la consigna solicitada. Se debe intentar que los ítems y el material de estímulo sea lo más neutral y equilibrado posible. Es por ello que se sugiere que no se utilice la evaluación para promover ciertos temas, que no se aborden temáticas de sensibilidad social (aborto, eutanasia, por señalar algunos ejemplos), que se eviten estereotipos en el lenguaje y en imágenes, a menos que sean parte de lo que se intenta evaluar; por ejemplo, en consignas que van acompañadas de ilustraciones que caracterizan a distintos subgrupos. Finalmente, se hace referencia a utilizar la terminología apropiada para los grupos étnicos, de género o raciales con la que los propios segmentos se identifican; e incluir a todos los miembros de distintos grupos cuando se menciona o hace referencia a personas, de modo que todos los estudiantes puedan sentirse representados.

La consideración de los alumnos en situación de discapacidad revisa otro tipo de observaciones, teniendo en cuenta que pueden existir obstáculos físicos que requieren de otra clase o tipo de intervenciones. Este es el caso de las recomendaciones para alumnos con discapacidad visual, que contemplan: la utilización del menor número de estímulos visuales (cuando no sean necesarios para la validez), diferencias sutiles en el sombreado, líneas de texto que no contrasten fuertemente con el fondo, letras o sonidos que resulten similares o símbolos especiales (tales como notación científica, cuando no forme parte del objeto de evaluación). Es aquí que adquiere centralidad el principio de accesibilidad, entendido como la habilidad de acceder a la funcionalidad y a los beneficios de un dispositivo o sistema de evaluación. El punto de partida para el diseño de evaluaciones que respondan a este principio radica en la comprensión de las necesidades de los alumnos y en el reconocimiento de que las discapacidades no siempre son visibles, por ejemplo alguna dificultad de aprendizaje (Nesbit y Shaw, 2019).

En términos generales, y para intentar sistematizar lo anteriormente analizado, podrían identificarse dos grandes cuestiones que intentan abordar las conceptualizaciones sobre el tema de la justicia en las evaluaciones, con niveles de desarrollo disímiles: a) la atención a las características de todas y cada una de las etapas involucradas en el proceso evaluativo y su orientación, de modo de garantizar la accesibilidad a todos los grupos de estudiantes; b) la vinculación entre la justicia en las evaluaciones, las oportunidades de aprendizaje y la oferta curricular.

Asimismo, la noción de accesibilidad hace referencia al mejoramiento de las condiciones de evaluación, las mismas que se procuran maximizar a través de diferentes soportes: a) aquellas llamadas universales, esto es, que hacen más accesibles las evaluaciones para todos los estudiantes; b) apoyos especialmente previstos para estudiantes con necesidades específicas; c) las denominadas acomodaciones, que son adecuaciones o cambios en los procedimientos y/o medios utilizados en la evaluación con el fin de incrementar la accesibilidad para todos los estudiantes involucrados.

Por último, la vinculación entre la justicia en las evaluaciones, las oportunidades de aprendizaje y la oferta curricular, resulta otro espacio

fértil para la reflexión sobre los modos de construcción de evaluaciones más justas. Si bien este es un área menos desarrollada que los temas relativos a la accesibilidad, es posible identificar algunos elementos que pueden resultar contribuciones en este terreno. Tal es el caso de la caracterización de las oportunidades de aprendizaje para los diferentes grupos que participan de una evaluación, de modo de incrementar el conocimiento sobre los contextos de aprendizaje para que retroalimenten las decisiones de política educativa.

II. EXPERIENCIAS DE JUSTICIA EN LAS EVALUACIONES DE LOS APRENDIZAJES EN DISTINTOS SISTEMAS EDUCATIVOS

MÉXICO

El Instituto Nacional para la Evaluación de la Educación (INEE) ha desarrollado una línea de trabajo específica vinculada con la atención a la diversidad en las evaluaciones a gran escala. En ese marco, ha profundizado su mirada en la atención a la diversidad cultural procurando promover un enfoque intercultural en el diseño e implementación de las evaluaciones y evitar sesgos que perjudiquen el desempeño de los estudiantes que pertenecen a culturas indígenas.

En ese sentido, uno de los aportes significativos del Plan Nacional de Evaluación de los Aprendizajes (PLANEA) es la incorporación de un protocolo de validez cultural que consiste en:

- I. Participación de lingüistas, docentes indígenas y antropólogos tanto en la elaboración de la tabla de especificaciones como de los ítems y posterior proceso de validación.
- II. Definición de las poblaciones participantes tomando en cuenta la diversidad lingüística, cultural y económica.
- III. Revisión de la información gráfica y contextual incluida en los ítems.

- IV. Revisión, por parte de jueces específicos, de la presencia de sesgo cultural en textos e ítems.
- V. Aplicación de análisis estadístico de sesgo y desarrollo de estudios de validez y confiabilidad para distintos grupos de poblaciones.

51

Si bien se afirma que es difícil garantizar que las pruebas estandarizadas hoy tengan un enfoque intercultural, sí se plantea como un desafío y esta búsqueda condujo al INEE a realizar una Consulta Previa, Libre e Informada a Pueblos y Comunidades Indígenas sobre la Evaluación Educativa en el 2014. Este proceso amplio de consulta trajo como resultado una serie de consideraciones sobre la evaluación (y el sistema educativo en general) muy valiosas para revisar las prácticas vigentes y plantear algunas orientaciones para el logro del enfoque intercultural en las evaluaciones:

- Es necesario asegurar que las evaluaciones carezcan de sesgo cultural en todas las etapas de diseño e implementación.
- Las organizaciones indígenas deben ejercer plenamente el derecho que les otorga la legislación vigente a una educación en su propia cultura y lengua, y así evaluar sus logros.
- Los docentes que trabajan con poblaciones indígenas deben contar con herramientas para llevar a cabo evaluaciones formativas y participativas con el fin de generar evidencia sobre los logros de aprendizaje comunes y de los logros diferenciados en términos de conocimientos, habilidades y destrezas.
- Es necesario evaluar, por medio de pruebas estandarizadas, la medida en que la diversidad cultural y lingüística del país es conocida, respetada y valorada por toda la población.

CHILE

Dentro de la gama de evaluaciones que realiza la Agencia de Calidad de la Educación (ACE), existen dos prácticas de atención a la diversidad

que es relevante destacar. Por un lado, la consideración del contexto socioeconómico, el género y la ruralidad de los estudiantes y, por otro, la acomodación de las pruebas para alumnos con discapacidad sensorial.

CONSIDERACIÓN DEL CONTEXTO SOCIOECONÓMICO, EL GÉNERO Y LA RURALIDAD

En el marco del Sistema de Medición de la Calidad de la Educación (SIMCE) que supone la aplicación de pruebas estandarizadas a nivel nacional, se han desarrollado en los últimos años algunos mecanismos específicos de atención a la diversidad, entre los que se destacan particularmente aquellos vinculados a la contextualización de los resultados a partir de la consideración de las condiciones socioeconómicas de la población evaluada.

La identificación de las características del contexto permite hacer comparaciones más justas de los resultados y aporta elementos efectivos para que las escuelas establezcan metas y plazos viables en términos de mejora de los logros de los estudiantes. Estas acciones tienen lugar en diferentes etapas del proceso de evaluación: el diseño de los instrumentos, análisis de resultados y su difusión. A continuación se detallan brevemente cada uno de ellos.

En el diseño de los instrumentos de evaluación

Como parte del diseño de los instrumentos de evaluación, se desarrolla un proceso de juicio de expertos que revisan los mismos a partir de un conjunto de criterios específicos. Los criterios previstos aluden a la consideración de que todos los estudiantes que rinden las pruebas tengan acceso igualitario a la comprensión de los estímulos y contextos, independientemente de su género, raza, religión, condición social, económica, cultural, física u otra; y que se tenga especial cuidado con que

los estímulos y contextos no sean emocionalmente perturbadores, evadiendo temas o situaciones traumáticas o violentas, o temas controversiales. También se considera que los estímulos propuestos como parte de las pruebas, deben ser adecuados al nivel escolar de los estudiantes. Esto implica que deben evitarse tanto estímulos y contextos complejos o distantes en términos temáticos, lingüísticos o de otra índole.

En una segunda instancia, y a partir del proceso de pilotaje, se analizan desde el punto de vista psicométrico el comportamiento de las consignas de los instrumentos, procurando identificar aquellos cuyo funcionamiento diferencial pueda ser atribuido a algún tipo de sesgo que eventualmente suponga un perjuicio para algún grupo de estudiantes.

En la aplicación de los instrumentos

Durante la aplicación, la evaluación SIMCE está acompañada por Cuestionarios de Calidad y Contexto Escolar que son respondidos por estudiantes, docentes y padres o apoderados. Estos cuadernillos relevan información sobre distintas dimensiones de la realidad educativa de los estudiantes y las instituciones a través de un conjunto de indicadores que denominaron Indicadores de Desarrollo Profesional y Social (IDPS).

A continuación se presenta una tabla que sintetiza los indicadores construidos, las dimensiones a las que atienden y los actores que responden a cada uno de ellos.

Tabla 2

Indicador	Dimensiones	Actores
Autoestima y motivación escolar	Autopercepción y autovaloración académica	Estudiantes
	Motivación escolar	Estudiantes
Clima de convivencia escolar	Ambiente de respeto	Padres, docentes y estudiantes
	Ambiente organizado	Padres, docentes y estudiantes
	Ambiente seguro	Padres, docentes y estudiantes
Participación y formación ciudadana	Participación	Padres y estudiantes
	Vida democrática	Estudiantes
	Sentido de pertenencia	Estudiantes
Hábitos de vida saludable	Hábitos alimentarios	Estudiantes
	Hábitos de autocuidado	Estudiantes
	Hábitos de vida activa	Estudiantes

Fuente: Indicadores de Desarrollo Profesional y Social (IDPS) de la Agencia de la Calidad de la Educación, Chile.

En el análisis de los resultados

Para el análisis de la información de resultados se organiza a la población de estudiantes en cinco grupos según sus condiciones socioeconómicas y, posteriormente, se identifican los parámetros de dificultad que presentan las pruebas para cada uno de estos grupos; esto permite hacer comparaciones más justas de los resultados.

En lo que respecta a género, la ACE ha construido el Indicador Equidad de Género que evalúa el logro equitativo de resultados de aprendizaje obtenidos por hombres y mujeres en establecimientos mixtos.

55

CONSIDERACIÓN DE LA DISCAPACIDAD SENSORIAL

Las pautas previstas para la aplicación de las pruebas SIMCE prevén un tratamiento especial a los estudiantes con discapacidad sensorial que se encuentran integrados a la educación común.

Los alumnos con necesidades educativas especiales (NEE), hasta el 2008, eran excluidos de los procesos de evaluación nacionales. A partir de una modificación en la normativa, se produjeron una serie de cambios tendientes a generar procesos evaluativos progresivamente más inclusivos.

Los cambios introducidos resultan acomodaciones²⁸ destinadas a garantizar la participación autónoma de estudiantes con discapacidad sensorial, eliminando o minimizando los obstáculos existentes a causa de su discapacidad y garantizando los contenidos evaluados. De este modo, es posible que los estudiantes con discapacidad sensorial realicen las pruebas en condiciones equivalentes a las de los estudiantes que rinden las pruebas regulares.

Diseño de las pruebas

Se desarrollaron una serie de estudios diagnósticos sobre diferentes características de las pruebas y a partir de los resultados de los mismos

28. Las acomodaciones son consideradas cambios en las pruebas que no alteran el constructo medido y es efectivo solo para los estudiantes que requieren de este cambio.

se generó una propuesta de acomodaciones para los textos y las preguntas.

56

Aplicación de las pruebas

Inicialmente se realizó un relevamiento para identificar la cantidad de alumnos en situación de discapacidad sensorial que asistían a escuelas regulares o especiales. Como parte del proceso de aplicación, se capacitó a los aplicadores informando sobre las acomodaciones permitidas durante la evaluación tanto las relativas a la presentación de las pruebas (uso de dispositivos de aumento), a la respuesta (uso del ábaco, responder en el propio cuadernillo de prueba) y al contexto de aplicación (ubicación del estudiante dentro de la sala).

Difusión de resultados

Los resultados se dan a conocer a partir de un informe específico dirigido a docentes y directivos. En el mismo se desagregan los resultados en los diferentes dominios y ejes evaluados y se especifican el tipo de discapacidad, el tipo de establecimiento educativo y el conjunto de acomodaciones utilizadas, y se incluyen ejemplos de tareas específicas que se solicitaron a los alumnos.

En el mismo informe se advierten el tipo de comparaciones que son posibles realizar y cuáles no dado que, por ejemplo, se emplea una escala distinta a la utilizada en las pruebas regulares y se destaca que las puntuaciones conseguidas por alumnos con diferente tipo de discapacidad no son equivalentes.

ESTADOS UNIDOS

La ley conocida como *No Child Left Behind* NCLB del 2001 enfatiza la responsabilidad del Estado de asegurar que todos los niños tengan oportunidades justas e igualitarias de recibir educación de calidad y alcanzar el dominio básico de los niveles de logro académico en las evaluaciones.

57

En este marco, se manifiesta una práctica sistemática de atención a la diversidad en tanto garantía de la participación en las evaluaciones nacionales en dos líneas de trabajo: la referida a los alumnos con discapacidad y a los estudiantes con una competencia limitada en inglés.

LA PARTICIPACIÓN EN LAS EVALUACIONES DE ALUMNOS CON DISCAPACIDAD

De acuerdo al marco legal, los estudiantes con discapacidad tienen derecho a contar con las acomodaciones necesarias para participar de las pruebas nacionales. La forma de determinar las acomodaciones requeridas es a través del Programa Individualizado de Educación. Se trata de un instrumento en el que participa un equipo especializado junto con los padres de los estudiantes y en él se definen las condiciones necesarias para facilitar el acceso y la plena participación en las evaluaciones provinciales y del distrito escolar. Este programa individualizado se revisa periódicamente al menos cada tres años.

En la definición de las acomodaciones requeridas se diferencian aquellas consideradas estandarizadas y las no estandarizadas según modifiquen o no las habilidades evaluadas. Por su parte, las políticas relacionadas con acomodaciones de todo tipo varían entre los Estados.

La selección de las acomodaciones tanto para la enseñanza como para la evaluación es responsabilidad del equipo del Programa Individualizado de Educación de cada estado y cada vez que se decide la implementación de un nuevo tipo de acomodación, se prevé la consi-

deración del tiempo requerido por el estudiante para la utilización de la misma.

58

En este marco se ha desarrollado una guía específica para seleccionar la acomodación más adecuada de acuerdo con la discapacidad, considerando tanto las discapacidades físicas, como las sensoriales y/o dificultades de aprendizaje. Entre los tipos de acomodaciones previstos se encuentran los relativos a la presentación de los materiales, los relativos a las formas de brindar respuesta, los que suponen extensión de tiempo y los relativos a los espacios.

LA INCLUSIÓN DE ESTUDIANTES CON COMPETENCIA LIMITADA EN INGLÉS

En el marco de la Evaluación Nacional de Progreso Educativo (NAEP) se elaboraron una serie de lineamientos tendientes a incluir a los estudiantes con una competencia limitada de inglés. Las modificaciones previstas incluyen el desarrollo de los instrumentos de evaluación en la lengua materna, la utilización de inglés estándar en los instrumentos con varios tipos de apoyo, y modificaciones de la versión en inglés estándar. Algunas de las estrategias alternativas propuestas son modificaciones de los ítems e instrucciones que tienden a disminuir la carga de la lengua y otras son la provisión de apoyos durante la administración de las pruebas sin modificar los ítems.

Si bien se destaca la importancia del desarrollo de mecanismos que aseguren la inclusión de estos estudiantes con competencia limitada de inglés, se advierte reiteradamente que son acciones que deben ser monitoreadas ya que, por ejemplo, hay muchas variaciones entre los estados en la forma en la que se identifican a estos estudiantes. En el mismo sentido, se advierte la relevancia de reportar resultados diferenciando el tipo de acomodación o de ajuste que se lleva a cabo.

INGLATERRA

La Agencia de Estándares y Pruebas (STA) es la responsable del desarrollo e implementación de las evaluaciones estandarizadas conocidos como SAT (Scholastic Aptitude Test). El marco regulatorio de las pruebas está explícitamente orientado a garantizar el acceso de todos los estudiantes a las pruebas en igualdad de condiciones sin importar el género, lengua, religión, origen étnico o social, o discapacidad alguna.

Los tipos de diversidad que se consideran son, por una parte los niños y niñas con necesidades educativas especiales y, dentro de este grupo, se consideran las limitaciones auditivas o visuales, lesiones o enfermedades, dificultades para concentrarse y dificultades para entender diagramas bidimensionales. El segundo grupo son aquellos niños y niñas con dominio limitado del idioma en que están desarrolladas las pruebas, en este caso el inglés.

El diseño de las evaluaciones por parte de la STA sigue los principios de diseño universal y se especifican una serie de condiciones que deben cumplir las mismas, entre las que se destacan: a) usar medios adecuados para que los estudiantes demuestren sus conocimientos; b) proveer desafíos adecuados y generar oportunidades para que los estudiantes puedan alcanzar desempeños altos, sin importar su género, necesidades educativas especiales, discapacidad, contexto cultural, social y lingüístico; c) no afectar la autoestima y confianza de los niños; d) estar libre de estereotipos y todo tipo de discriminación y; e) usar materiales familiares para los niños y garantizar que estén adecuadamente preparados para usarlos.

Adicionalmente a las orientaciones sobre los procesos de desarrollo de las evaluaciones, se promueve el monitoreo de los propios procesos de evaluación. Un organismo especializado – la Oficina Regulatoria de Calificaciones y Examinaciones – recupera evidencia que permite valorar la calidad de los procesos de evaluación implementados y que hace posible la mejora y ajuste de los mismos.

III. LA EXPERIENCIA EN CABA. AVANCES Y DESAFÍOS.

60

En la Ciudad Autónoma de Buenos Aires (CABA), la Unidad de Evaluación Integral de la Calidad y Equidad Educativa (UEICEE), creada bajo la Ley Nº 5049 en el año 2014, tiene las atribuciones de definir los lineamientos para la evaluación de la educación y del sistema educativo de la Ciudad a fin de garantizar el derecho personal y social a la educación y la mejora continua de su calidad y la del sistema (Artículo 1).

Por consiguiente, la consideración de las características individuales de los estudiantes y de los contextos escolares en los que tienen lugar las evaluaciones es ineludible para generar intervenciones específicas que contribuyan a generar un sistema educativo cada vez más justo y equitativo.

A continuación, se presentan tres experiencias que la UEICEE se encuentra desarrollando en vistas a la atención y el abordaje de la justicia en las evaluaciones: el desarrollo de un sistema integral de evaluación de los aprendizajes, la conformación de una Comisión Asesora sobre justicia en las evaluaciones y las acciones realizadas en el diseño de los instrumentos de evaluación.

A. Sistema integral de evaluación de los aprendizajes

Son varios los autores que identifican la necesidad de ampliar la mirada respecto de los aprendizajes logrados por los estudiantes, reconociendo que muchos otros saberes, competencias y habilidades que son objeto de enseñanza y de aprendizaje, no son capturados por medio de las estrategias de evaluación tradicionalmente utilizadas. En ese sentido, Murillo e Hidalgo (2015) advierten sobre la necesidad de considerar el desarrollo integral de los estudiantes, en lugar de reducir los resultados en los aprendizajes a aquellas áreas curriculares que son habitualmente privilegiadas en las evaluaciones de sistema (Lengua/ Matemática). Esto supone reconocer que

“...para mejorar la calidad de la oferta educativa y distribuirla de manera justa, no basta con la aplicación de buenas evaluaciones aisladas. Se requiere que estas dialoguen entre sí y formen parte de un mismo proceso de análisis, reflexión y proyección de lo hecho y necesario por hacer, desafío que, centrándose en el campo educativo, se inserta en una visión de cambio de la sociedad en su conjunto” (Murillo e Hidalgo, 2015: 45).

61

Al respecto, también se identifica que son diversas las instancias de evaluación en donde participan los alumnos de la CABA, desde las evaluaciones que realizan los docentes cotidianamente en el aula, algunas iniciativas impulsadas desde las supervisiones o redes de escuelas, las evaluaciones a gran escala: Finalización de Estudios Primarios en la Ciudad de Buenos Aires (FEPBA), Tercer año de Estudios Secundarios en la Ciudad de Buenos Aires (TESBA), Pausas Evaluativas (solo en el sector estatal), Aprender (a nivel nacional) e incluso en algunos casos la participación en estudios internacionales. Independientemente que se trate de grados/años distintos, de áreas curriculares u objetos de evaluación diferentes, es necesario poner en diálogo todas estas evidencias. En ese sentido, la UEICEE está trabajando en el desarrollo de un sistema de evaluación de los aprendizajes, que intenta articular las distintas evidencias sobre los aprendizajes de los estudiantes realizados en establecimientos educativos de la CABA. Esto incluye tanto las evaluaciones de los aprendizajes realizadas a nivel del sistema educativo -jurisdiccionales, nacionales e internacionales- como otras que serán de participación voluntaria -desarrolladas en el marco de una Plataforma de Evaluación Formativa de los Aprendizajes- y el modo en que dialogan con las prácticas de evaluación que habitualmente realizan los docentes en las aulas. Este desarrollo se encuentra en una etapa inicial, partiendo de la sistematización de estas iniciativas en un primer momento, y en una segunda instancia, de la articulación de las evidencias correspondientes a aquellos contenidos que son abordados de manera recurrente en algunos de los estudios anteriormente mencionados, y para los que ya fue desarrollada una progresión de aprendizajes correspondiente a las áreas de Prácticas del Lenguaje/Lengua y Literatura y Matemática.

B. Comisión Asesora sobre justicia en las evaluaciones

A partir de la revisión de experiencias internacionales y de las acciones de la UEICEE, surgió la necesidad de conformar un espacio de consulta y trabajo entre expertos que permita seguir avanzando hacia el logro de la construcción de evaluaciones libres de sesgo y progresivamente más justas. En este marco, se creó una Comisión Asesora sobre Justicia en las evaluaciones.

El objetivo central de esta Comisión es generar un espacio de diálogo e intercambio que nutra la formulación de criterios para el desarrollo de instrumentos y estrategias de evaluación desde la perspectiva de la diversidad. Se trata de una instancia de diálogo, intercambio, orientación y seguimiento de las acciones referidas al logro de evaluaciones que brinden a los estudiantes las mismas oportunidades para evidenciar sus aprendizajes a través de las evaluaciones desarrolladas. Para ello, se convocó a un conjunto de especialistas a integrar esta Comisión e iniciar así un espacio de trabajo provechoso, que identificó como posibles temáticas:

1. Atención a la diversidad en la evaluación: conceptualización sobre la diversidad y tipos de diversidad considerados.
2. Identificación de posibles grupos de interés.
3. Normas o criterios de consideración de la diversidad y procedimientos aplicados:
 - a. En el diseño de los instrumentos
 - b. En la implementación de la evaluación
 - c. En el procesamiento y análisis de la información
 - d. En la interpretación de los resultados
 - e. En la presentación de resultados
 - f. En la difusión de resultados
 - g. En el uso de los resultados
4. Evaluación de las condiciones de las escuelas con alta matrícula en algunos de los grupos de interés.
5. Evaluación de la perspectiva de género.

6. Evaluación de la atención educativa a la discapacidad.

Dicha Comisión se reunió en dos oportunidades para discutir algunas experiencias reconocidas a nivel internacional en relación a la temática, poner en conocimiento las acciones que en la Ciudad se habían estado desarrollando hasta el momento e identificar acciones de cara a los principales desafíos del sistema educativo de CABA.

63

C. Acciones realizadas en el diseño de los instrumentos de evaluación

La atención puesta en el diseño de los instrumentos y en la implementación de las pruebas parte de dos grandes conceptualizaciones en función de las características de los estudiantes, abriéndose así distintos caminos de acción. En la literatura revisada (AERA et al., 2014; Gipps y Stobart, 2010; Solano-Flores, 2011) se distinguen:

- las dimensiones culturales, que comprenden la diversidad lingüística, étnica, socioeconómica, de género, racial y religiosa, entre otras;
- las denominadas necesidades educativas especiales de los alumnos, que aluden por lo general a situaciones de aprendizaje (ritmos de aprendizaje significativamente distintos a los del grupo de edad del alumno). Esto comprende un abanico amplio de situaciones de discapacidad, como motrices, auditivas, intelectuales, entre otras.

El abordaje de las dimensiones culturales requiere el trabajo de especialistas que revisen el contenido de las pruebas junto con un trabajo focalizado en la formulación de los ítems, para evitar sesgos desde el diseño de las evaluaciones.

Por otro lado, respecto a las necesidades educativas especiales, la principal acción desarrollada consiste en la implementación de acom-

daciones, como las mencionadas previamente, para garantizar la accesibilidad de todos los estudiantes. Desde la aprobación de la Resolución Nº 3773 de 2014 hasta la actualidad, se observa un crecimiento sostenido de los acompañantes pedagógicos no docentes (APND) en las escuelas de la modalidad Común. Esta situación, junto con las formas que está adoptando el trabajo de los profesionales del área de salud en las escuelas, así como la relación entre los centros de los que éstos dependen y las distintas áreas de educación, requiere de un análisis que contemple las implicancias de la actual forma de incorporar los APND en los procesos de inclusión educativa y su participación en las evaluaciones.

La Resolución Nº 3773/MEGC/11 aprueba el Reglamento para el Desempeño de Acompañantes Personales no Docentes para alumnos diagnosticados con Trastorno Generalizado del Desarrollo en los términos definidos en el Manual Diagnóstico y Estadístico de los Trastornos Mentales, Masson 2002 (DSM IV TR), en todo establecimiento educativo de Gestión Estatal de la Ciudad Autónoma de Buenos Aires. El mismo es antecedente de la Resolución Nº 3034/13 que regula el desempeño de los APND que atienden a los niños con cualquier tipo de discapacidad en la escuela Común. Los centros de integraciones son instituciones privadas que tercerizan los servicios de los APND, funcionando como intermediarios entre estos agentes, las Obras Sociales, la escuela y la familia del niño. El centro es la institución responsable del desempeño del APND dentro de la escuela.

A partir de lo mencionado, se presentan a continuación distintas experiencias desarrolladas a nivel jurisdiccional y nacional.

EXPERIENCIAS A NIVEL LOCAL: ACOMODACIONES EN FEPBA Y TESBA

65

Anualmente desde el año 2012, la UEICEE implementa evaluaciones de aprendizaje en todos los establecimientos de gestión estatal y privada que ofrecen educación primaria y secundaria común en la Ciudad de Buenos Aires. Se evalúa a los estudiantes de 7º grado (FEPBA) y de 3º año (TESBA)²⁹, y el objetivo de dichas evaluaciones es identificar el nivel de concreción de algunas de las metas de aprendizaje en las áreas de Matemática, Prácticas del Lenguaje y Lengua y Literatura, planteadas en los marcos curriculares de la jurisdicción.

Dimensiones culturales

El análisis realizado comprende aspectos vinculados a la nacionalidad de los estudiantes, sexo/género y religión.

Nacionalidad

En relación a la nacionalidad, en el caso de FEPBA, del total de alumnos que declararon haber nacido en el exterior, la mayor proporción proviene de países limítrofes (10.2%), de los cuales el 20.3% proviene de Bolivia, 1.8% de Paraguay y 1.6% de Perú. En el caso de TESBA, donde el porcentaje de alumnos migrantes asciende a 10.9%, se observa un patrón similar: Bolivia (3.5%), Paraguay (1.9%) y Perú (1.6%).

29. Hasta 2016 la evaluación de secundaria se aplicó en 5º/6º año y a partir de 2017 se realiza en 3º año. Este cambio obedece, por un lado, a que el dispositivo nacional APRENDER evalúa a la misma población de estudiantes y por otro lado, a que evaluar los logros de aprendizaje en tercero permite contar con mayor tiempo para la toma de decisiones que orienten la mejora de los estudiantes antes de finalizar la secundaria.

A partir de lo relevado y en función de los análisis multinivel realizados, se observa que la migración no ha emergido como una variable estadísticamente significativa asociada a los resultados. Se seguirá profundizando con información de los resultados del 2019.

Sexo/Género

Respecto de las preguntas sobre sexo y género, en el 2017 se observó que 11.9% de los estudiantes de FEPBA omitían la pregunta o marcaban más de una opción. Se ha observado que incluso algunos estudiantes proponen una tercera opción “otros”.

Dadas las dificultades encontradas en la formulación de la pregunta referida al género/sexo, se realizaron consultas en organismos oficiales (Dirección General de Estadísticas y Censos de CABA, considerando la experiencia de reclasificación para la Encuesta Anual de Hogares) y organismos de defensa de los derechos de los niños, jóvenes y adolescentes, como así también revisión de las normativas referidas a ESI e igualdad de género para mejorar los cuestionarios complementarios. Obtener un mejor dato permitirá avanzar en el análisis para identificar posibles sesgos en el diseño de las pruebas.

Religión

Desde 2016 la UEICEE implementa acciones para abordar sesgos religiosos. Con este objetivo, se ha trabajado con la Dirección General de Educación de Gestión Privada, para revisar las pruebas y desarrollar criterios de elaboración e implementación de evaluaciones libres de sesgos, iniciándose la estrategia con escuelas observantes de la colectividad judía.

Dimensiones de inclusión de alumnos en situación de discapacidad

La mejora en el desarrollo y la implementación de las evaluaciones constituye una tarea permanente que asume la UEICEE para la generación de datos válidos y confiables y, al mismo tiempo, de instancias de evaluación que sean pertinentes para el conjunto de estudiantes de la jurisdicción. En este sentido, la UEICEE se encuentra desarrollando una serie de estrategias de acción concretas que buscan contemplar las necesidades de los estudiantes en situación de discapacidad.

67

a) *Registro de estudiantes en situación de discapacidad presentes en las pruebas FEPBA y TESBA 2017*

Desde el 2016, durante la visita previa que los veedores encargados de supervisar el desarrollo de la evaluación realizan a las escuelas, se solicita completar una planilla para consignar la cantidad de estudiantes en situación de discapacidad y el tipo de apoyos habitualmente utilizados en el aula, que se desplegarán al momento de la evaluación.

En función de dichos relevamientos, se observó que hay un 9% de estudiantes con algún tipo de discapacidad que participaron de FEPBA y 0.2% de TESBA en 2017. Respecto a la distribución porcentual de los estudiantes según su situación de discapacidad, en FEPBA se observa que el 68% ha sido identificado con algún tipo de discapacidad intelectual, y un 48% en el caso de TESBA; seguido por un Trastorno del Espectro Autista (14% en FEPBA y 13% en TESBA), discapacidad motora (3% y 9% respectivamente), discapacidad auditiva (2% y 11%) y disminución visual (2% y 7%).

Respecto del grado de participación de los estudiantes en situación de discapacidad en las evaluaciones anteriormente mencionadas, se observa que en el caso de FEPBA, la cobertura en matemática fue del 75.2% y en Prácticas del Lenguaje del 75.6%. En el caso de TESBA, la cobertura fue de 82.6% en Matemática y del 78% en Lengua y Literatura.

b) ***Plan para la inclusión de estudiantes en las pruebas FEP-BA y TESBA***

68

El relevamiento realizado evidencia la necesidad de introducir acomodaciones que posibiliten atender las necesidades singulares de los estudiantes. Ello ha implicado la flexibilización de ciertas condiciones de aplicación, como por ejemplo, incrementar el tiempo disponible para desarrollar las pruebas en estos casos, incorporar las configuraciones cotidianas de apoyo como la presencia de maestros integradores u otras figuras de acompañamiento al momento de realización de las pruebas y habilitar el uso de espacios específicos para que los estudiantes que lo requieran puedan desarrollar la evaluación. Dado que las pruebas de sistema buscan mantener la mirada del conjunto de los estudiantes y utilizan como referente ciertas metas comunes de aprendizaje, este instrumento no resulta idóneo para evaluar contenidos específicos referidos a las diversas situaciones de discapacidad, por ello el trabajo se ha focalizado en mejorar las condiciones de evaluación.

En todos los casos se ofrecen las siguientes acomodaciones para realizar las pruebas, que son acordadas con el equipo directivo, de acuerdo con la configuración de apoyo que habitualmente utiliza cada estudiante en clase:

- 1) La realización de la prueba en un área únicamente (Matemática o Prácticas del Lenguaje/Lengua y Literatura).
- 2) El ofrecimiento de mayor tiempo para realizar las pruebas (durante el mismo día).
- 3) La realización de las pruebas en más de un día.
- 4) La realización de las pruebas en un espacio separado.
- 5) La realización de las pruebas junto al docente integrador.

De acuerdo a la conceptualización de las adecuaciones realizada por M. Thurlow, los cambios en la modalidad de aplicación de las pruebas se pueden agrupar en diferentes dimensiones. Se señalan en negritas las utilizadas actualmente en CABA:

Tabla 3

AMBIENTACIÓN	DURACIÓN	PLANIFICACIÓN
Se aplica la prueba a un grupo pequeño en un salón separado	Se extiende el tiempo para completar la prueba	Se permite que los temas de la prueba puedan tomarse en un orden diferente. (Si la maestra integradora acompaña al estudiante, puede definir el recorrido en la resolución de la prueba)
Se aplica la prueba de modo individual en un salón separado	Se realizan cortes de descanso en algunas instancias de la prueba	Se administra la prueba en varias sesiones, especificando la duración de cada sesión
Se proporciona iluminación/ mobiliario/ acústica especial	Se permite un cronograma flexible	Se administra la prueba durante varios días, especificando la duración de la sesión en cada día
Se aplica la prueba en un lugar determinado para evitar la máxima cantidad de distracciones		Se administra la prueba por la mañana o por la tarde

69

c) ***Uso de la tecnología para la mejora la accesibilidad***

Adicionalmente, la UEICEE ha desarrollado experiencias piloto de uso de evaluaciones en línea para mejorar el acceso y las condiciones de evaluación para estudiantes en situación de discapacidad motora, visual y/o auditiva.

III. DESAFÍOS PENDIENTES

70 Muchas de las acciones anteriormente descritas se encuentran en proceso, por lo que, en gran medida, los desafíos previstos están orientados al acompañamiento y profundización de las acciones mencionadas en relación a la construcción de un sistema de evaluación de los aprendizajes. Este proceso contempla tanto aspectos conceptuales como procedimientos estadísticos que posibilitan el alineamiento de las distintas evidencias obtenidas. Ello, con la intención de contar con información más robusta acerca de los aprendizajes logrados por los estudiantes permitiendo que se tomen decisiones de política educativa e intervenciones institucionales que contribuyan a acompañar mejor a los actores escolares.

Resulta necesario advertir que la evaluación tiene el potencial de contribuir a la mejora de la enseñanza y los aprendizajes, solo en la medida en que es comunicada y trabajada con los equipos de supervisión, los equipos docentes y directivos de modo adecuado. En este sentido, no resulta suficiente con diseñar pruebas más justas, alineadas con el diseño curricular, que dialoguen con otras fuentes de información disponible, sino que es preciso construir conversaciones comprometidas y profesionales con distintas audiencias en relación a los resultados obtenidos, contribuyendo al desarrollo de interpretaciones contextualizadas que recuperen las características específicas de cada establecimiento educativo, y sean complementarios de otras fuentes de información (internas o externas a la propia escuela). “En este marco, la autonomía reflexiva, por un lado, y el apoyo técnico y la evaluación, por el otro, no deben tratarse como opciones antagónicas y excluyentes, sino como alternativas complementarias que se fortalecen entre sí. Entre las alternativas polarizadas del docente como mero implementador de políticas y el docente crítico protagonista de su propia transformación, es posible plantear matices, secuencias en el tiempo y modelos superadores” (CI-PPEC, 2011: 119-120).

La referencia a lo contextualizado de la evaluación pone también de relevancia la necesidad de entender los resultados de las evaluaciones de aprendizaje en relación con otros indicadores, tanto en términos del

sistema educativo en su conjunto (para lo cual la UEICEE ha logrado avanzar en el desarrollo de Panorama Educativo³⁰) como a nivel institucional. Es en este aspecto, donde se abren nuevos desafíos y oportunidades para profundizar en la autoevaluación institucional como estrategia que apoye la reflexión institucional y el diseño de planes de mejora institucional (como puede ser el Proyecto Escuela). Al respecto, se reconoce que “...para que la información generada sea útil y oportuna para la toma de decisiones y optimice la calidad de la educación de los sistemas, debe ser analizada y considerada como parte de un todo y no de manera aislada” (Román y Murillo, 2010: 102).

La preocupación por un sistema de evaluación más justo no se reduce al diseño, implementación utilización de las pruebas de evaluación de aprendizajes, sino que comprende otros posibles ámbitos de evaluación (alumnos, instituciones, programas y políticas educativas), reconociendo desde ya los objetivos, estrategias metodológicas y alcances diferenciados.

Finalmente, y considerando que se trata de una problemática recurrente en varios sistemas de evaluación de la región, resulta necesario propiciar instancias de encuentro y/o intercambio a fin de relevar acciones desarrolladas por dichas unidades para el desarrollo de evaluaciones más justas. El contacto con otras unidades de evaluación de la región permitiría relevar los modos en que se aborda la diversidad en la evaluación educativa (sea de aprendizajes, institucional o de programas y proyectos), de modo de identificar desafíos y soluciones comunes.

30. Se trata de un proyecto cuyo objetivo es desarrollar y poner en funcionamiento una plataforma (con accesos abiertos y otros diferenciados para funcionarios del MEI-GCBA) que brinde información integrada y periódica sobre el estado de la situación educativa de la Ciudad de Buenos Aires en clave de dimensiones del derecho a la educación, para poder identificar avances, alcances y desafíos persistentes en la Ciudad de Buenos Aires.

REFERENCIAS BIBLIOGRÁFICAS

- 72 American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: Author.
- CAMILLONI, A. (1998). La calidad de los programas de evaluación y de los instrumentos que los componen. *Camilloni, A., Celman, S., Litwin, E. y Palou de Maté, M., La evaluación de los aprendizajes en el debate didáctico contemporáneo, Buenos Aires, Paidós.*
- CIPPEC (2011). *La construcción de la justicia educativa. Criterios de redistribución y reconocimiento para la educación argentina /*. Cecilia Veleda, Axel Rivas y Florencia Mezzadra. - 1a ed. - Buenos Aires, octubre de 2011. Disponible en: <https://www.cippec.org/wp-content/uploads/2017/03/2536.pdf>
- CHAPELLE, C. (2011). Validation in language assessment. En E. Hinkel (Ed.), *Handbook of research in second language teaching and learning*, Vol.II, 717-730. Nueva York, NY: Routledge.
- CONNELL, R. W. (1997). *Escuelas y justicia social*. Ediciones Morata.
- DUBET, F. (2017). *Repensar la justicia social. Contra el mito de la igualdad de oportunidades*. 4a ed. Ed. Siglo XXI.
- FRASER, N. (2005). Reframing justice in a globalizing world. *New Left Review*, 36, 69-88.
- FRASER, N. (2008). La justicia social en la era de la política de identidad: redistribución, reconocimiento y participación. *Revista de trabajo*, 4(6).

- GIPPS, C., Y STOBART, G. (2010). Fairness. En *International Encyclopedia of Education* (3.a ed., Vol. 4, pp. 56-60). Oxford, GB.: Elsevier.
- KANE, M. T. (2006). Validation. En R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17-64). Washington, DC: American Council on Education/Praeger.
- KANE, M. T. (2012). Validating score interpretations and uses: Messick Lecture, *Language Testing Research Colloquium*. Cambridge, April 2010. Language Testing, 3-17.
- KUNNAN, A. J. (Ed.). (2000). *Fairness and validation in language assessment: Selected papers from the 19th Language Testing Research Colloquium, Orlando, Florida* (Vol. 9). Cambridge University Press.
- KUNNAN, A. J. (2004). Test fairness. *European language testing in a global context*, 27-48.
- KUNNAN, A. J. (2013). Fairness and justice in language assessment. *The companion to language assessment*, 3, 1098-1114.
- POPHAM, W. J. (1997). Consequential validity: Right concern-wrong concept. *Educational measurement: Issues and practice*, 16(2), 9-13.
- PURPURA, J. E., BROWN, J. D., SCHOONEN, R. (2015). Improving the Validity of Quantitative Measures in Applied Linguistics Research. *Language Learning*, 65(Supplement 1), 36-73.
- MESSICK, S. (1989). Validity. In R. Linn (Ed.). *Educational Measurement* (3rd ed., pp. 13-100). Washington, DC: American Council on Education.
- MURILLO, F. J., & ROMÁN, M. (2010). Retos en la evaluación de la calidad de la educación en América Latina. *Revista Iberoamericana de educación*, 53(53), 97-120.
- MURILLO, F. J., ROMÁN, M. Y HERNÁNDEZ CASTILLA, R. (2011). Evaluación educativa para la justicia social. *Revista Iberoamericana de evaluación educativa*.

MURILLO, F.J. E HIDALGO, N. (2015). Enfoques fundamentantes de la evaluación de estudiantes para la justicia social. *RIEE. Revista Iberoamericana de Evaluación Educativa*.

74 NESBIT, I & SHAW, S.D. (2019). Fair assessment viewed through the lenses of measurement theory, *Assessment in Education: Principles, Policy & Practice*, DOI: 10.1080/0969594X.2019.1586643

RAWLS, J. (2006). *Teoría de la justicia* (6ª reimp.). EUA: Cambridge.

RUIZ, G., PÉREZ, M., LANGFORD, P., Y GARCÍA, A. (2015). *Atención a la diversidad en evaluaciones educativas externas. Muestra de prácticas internacionales*. México: INEE.

SCHMELKES DEL VALLE, S. (2016). Hacia una evaluación de logro escolar con enfoque intercultural en: *Gaceta de la Política Nacional de Evaluación en México*. Año 2 N°5. México INEE.

SEN, A. (1996), "Capacidad y bienestar" en: *Nussbaum, Martha y Sen, Amartya (comps.), La calidad de vida*, Buenos Aires, FCE.

SOLANO-FLORES, G. (2011). Assessing the cultural validity of assessment practices. En *Cultural validity in assessment* (pp. 3-21). Nueva York: Routledge.

STOBART, G. (2005). Fairness in Multicultural Assessment Systems. *Assessment in Education: Principles, Policy & Practice*, 12(3), 275-287. <http://doi.org/10.1080/09695940500337249>

TEDESCO, J. C. (2012). *Educación y justicia social en América Latina*. Buenos Aires: Fondo de Cultura Económica. Primera parte.

TENTI FANFANI, E. (2019) ¿La escuela soluciona las desigualdades sociales? en *Le Monde Diplomatique (Edición Cono Sur), Suplemento "La Educación en debate"*, N° 73.

THOMPSON, S. J., JOHNSTONE, C. J., & THURLOW, M. L. (2002). *Universal Design Applied to Large Scale Assessments. Synthesis Report*.

TORRECILLA, M., JAVIER, F., & HIDALGO FARRAN, N. (2015). Enfoques fundamentantes de la evaluación de estudiantes para la justicia social. RIEE. *Revista Iberoamericana de Evaluación Educativa*.

WALZER, M. (2001). *Las esferas de la justicia*. Ed. Ciudad de México: Fondo de Cultura Económica.

75

WARMINGTON, P., & MACINTOSH, H. (2003). *Fair assessment and equal opportunities*. Discussion paper for QCA.

XI, X. (2010). How do we go about investigating test fairness? *Language Testing*, Sage Journals, 27, 147-170.

LA RETROALIMENTACIÓN FORMATIVA: UNA OPORTUNIDAD PARA MEJORAR LOS APRENDIZAJES Y LA ENSEÑANZA

77

Rebeca Anijovich y Graciela Cappelletti

Este trabajo presenta los desafíos que ofrece la retroalimentación formativa como promotora de aprendizajes valiosos. Para ello, presentaremos consideraciones conceptuales que dan cuenta de las características centrales de este proceso, atendiendo a aportes teóricos. Además, dado que el propósito es poner a disposición sugerencias para su utilización en las aulas, definiremos presentar las condiciones y los obstáculos en los que se apoya la retroalimentación formativa, para promover comprensiones profundas y complejas (no mecánicas ni en formato “receta”) de las estrategias para poder ofrecerla.

ACERCA DE LA RETROALIMENTACIÓN FORMATIVA

La evaluación es un proceso indisoluble de la enseñanza y el aprendizaje. Cuando evaluamos los aprendizajes de nuestros estudiantes, la mirada se vuelve sobre cómo están aprendiendo y cómo se encaró la propuesta de enseñanza.

Dylan Wiliam (2011) propone comprender la evaluación como un puente entre la enseñanza y el aprendizaje para lo cual define cinco estrategias:

- Visibilizar y compartir con los estudiantes las metas de aprendizaje y los criterios de logro.
- Diseñar y llevar a la práctica actividades variadas que evidencien lo que los estudiantes están aprendiendo.

- Proponer estrategias para que los estudiantes sean responsables de su proceso de aprendizaje.
- Favorecer las interacciones entre pares como fuente de aprendizaje.
- Ofrecer retroalimentaciones formativas que favorezcan avances en los procesos de aprendizaje.

La idea principal alrededor de la cual se desarrolla este trabajo es que la retroalimentación formativa es valiosa cuando contribuye a mejorar los aprendizajes de los estudiantes. Una variedad de investigaciones sostienen esta afirmación: Sadler (1989, 2010), Perrenoud (2008), Taras (2002), Shepard (2000), Wiliam y Black (2002), Camilloni (2004), Allal y López Mottier (2005), Brookhart (2001, 2008), Stiggins (2005), Hattie & Timperley (2007), Shute (2008), Brooks (2017), Costa & Garmston (2017), Anijovich y Cappelletti (2017), entre otros.

Los distintos investigadores coinciden en definir a la retroalimentación como formativa haciendo foco en dos aspectos: 1) en su contribución para modificar los procesos de pensamiento y las acciones de quienes la reciben y 2) en la información que ofrece para contribuir a reducir la distancia entre donde se encuentra el estudiante, al inicio de un aprendizaje y adónde tendría que llegar en función de los propósitos de la enseñanza.

Tal como lo define Sadler (2010) la retroalimentación tiene que ayudar a los estudiantes a reconocer, a ser conscientes de la brecha existente entre el nivel en el que se encuentran y el que tienen que alcanzar con respecto a un aprendizaje. En este sentido, identifica tres condiciones para que un estudiante se beneficie de la retroalimentación:

- que tenga claridad acerca del nivel de aprendizaje que requiere alcanzar,
- que pueda comparar el nivel actual en el que se encuentra con el que requiere alcanzar,
- que esté comprometido con el aprendizaje y ponga en juego estrategias orientadas para reducir la brecha entre su nivel actual y el que tiene que alcanzar.

La retroalimentación formativa tiene el potencial de influir en el aprendizaje, pero el simple comentario o la entrega de un resultado no conduce necesariamente a una mejora.

Al mismo tiempo, se reconoce el impacto de la retroalimentación formativa sobre la autoestima de los estudiantes tanto en lo referido al contenido (tareas, producciones, persona), como al modo de comunicación (oral, escrito, individual, grupal) y al sentido de oportunidad (cuándo, dónde).

Históricamente la retroalimentación era unidireccional, ofrecida por el docente y recibida por el estudiante con bajo impacto en la mejora de sus aprendizajes y con un interés asociado solo a la calificación, dando cuenta que los comentarios referían a lo correcto o incorrecto de la tarea y/o a los errores cometidos. De este modo las calificaciones se convertían en los únicos indicadores de aprendizaje. Ha dominado la interpretación de la retroalimentación como transmisión de información o “decir”. Reconocer la retroalimentación como algo que se le da al estudiante para corregir sus errores, es una visión estrecha del aprendizaje.

En relación con el impacto de las calificaciones, una investigación desarrollada por Sadler (2010) recoge evidencias acerca de que los estudiantes ponen en juego diferentes enfoques para interpretar las calificaciones junto con los comentarios que reciben de sus profesores. Los datos muestran diferentes impactos: a) la retroalimentación “tomada a bordo” (se recibieron comentarios de retroalimentación junto con la calificación), b) la calificación más importante que la retroalimentación y c) efecto negativo de calificaciones bajas. En la primera, algunos estudiantes pudieron desasociarse de las reacciones emocionales relacionadas con el resultado de la calificación y actuar positivamente a partir de la retroalimentación.

Hoy nos encontramos ante una concepción diferente en tanto se propone involucrar activamente a los estudiantes en el proceso de retroalimentación, focalizando en el desarrollo de habilidades metacognitivas y de conciencia acerca de cómo se aprende; ofreciendo comentarios sobre hacia dónde ir y cómo hacer para llegar hasta ahí. La

retroalimentación formativa focaliza en la mejora de los aprendizajes y no tiene como finalidad justificar una calificación.

Este enfoque centra el interés en favorecer la retroalimentación como conversación, en la que se ponen en juego las dimensiones cognitivas (por ejemplo, formular preguntas); emocionales, sociales (por ejemplo, expresar empatía) y estructurales (por ejemplo, ofrecer suficiente tiempo para dialogar).

Una condición para comprender los efectos de la retroalimentación formativa es tener claridad acerca de cómo la definen profesores y estudiantes. Carless et al. (2011:397) definen la retroalimentación como *“procesos y actividades dialógicas que pueden apoyar e informar al estudiante sobre la tarea actual y así desarrollar la capacidad de autorregular el desempeño en tareas futuras”*. Reconocen la retroalimentación dialógica como parte de una evaluación sostenible en términos de que promueve el compromiso de los estudiantes con la autorregulación del aprendizaje. Para que esto ocurra no alcanza con compartir los objetivos de aprendizaje y los criterios de evaluación, es necesario un estudiante involucrado tanto en sus aprendizajes como en los procesos de evaluación.

Una característica relevante de la retroalimentación formativa refiere al tiempo pasado o futuro. Al decir de Wiliam (2009), las retroalimentaciones focalizan habitualmente su atención más retrospectivamente que prospectivamente. La pregunta relevante no es ¿qué hizo bien o mal el estudiante?, sino ¿qué hacer ahora? ¿Cómo avanzar? ¿Qué le falta? ¿Qué grado de conciencia tiene el estudiante acerca de su recorrido, avances, logros y dificultades?

También resulta importante indagar acerca de la calidad del contenido de la retroalimentación formativa, cómo se presenta a los estudiantes, cómo es aceptada o qué hacen los estudiantes con lo que reciben y cómo se integra al futuro aprendizaje y a la enseñanza. Particularmente en lo referido a su contenido encontramos en distintas prácticas de retroalimentación formativa, referencias a:

Valoraciones sobre la persona: cuyo propósito es impactar sobre la autoestima del estudiante y provocar mejoras en sus aprendizajes, reco-

nociendo que este impacto puede ser negativo en muchas ocasiones. En ese sentido existe controversia acerca de ofrecer valoraciones sobre la persona.

Las producciones y desempeños: el propósito aquí es impactar sobre la calidad y la profundidad de las tareas y los logros.

81

El proceso de aprendizaje: El propósito es focalizar en las estrategias que utilizan los estudiantes, en la identificación de sus fortalezas, obstáculos y dificultades para reconocerse como aprendices.

Consideremos algunos aportes de investigaciones, como por ejemplo la que han llevado adelante Anton Havnes, Kari Smith, Olga Dys-the y Kristine Ludvigsen (2012) acerca de las visiones de estudiantes y profesores de escuela secundaria en Noruega, sobre el valor de la retroalimentación. Entre sus hallazgos aparece que la retroalimentación se ofrece casi siempre junto con la calificación, que es más útil cuando se ofrece durante el desarrollo de una tarea y que la retroalimentación entre pares es poco frecuente. También describen las situaciones de aula que encontraron con mayor oportunidad para desarrollar una retroalimentación formativa: 1) las devoluciones después de un examen o trabajo práctico; 2) en la presentación de proyectos por parte de los estudiantes; 3) en los trabajos en equipo; 4) en los diálogos entre estudiantes y docentes en tanto contribuyen a reducir mal entendidos sobre la evaluación.

Desde otra perspectiva, Hattie (2012) plantea que la retroalimentación es poderosa cuando el estudiante se la ofrece al docente, haciendo visible lo que sabe, lo que entiende, sus ideas previas y de este modo podrá impactar sobre la enseñanza.

En una investigación Handley & William (2009) encontraron que los profesores que usaron la retroalimentación para mejorar su enseñanza tomaron un enfoque reflexivo a su enseñanza con preguntas tales como:

- ¿hasta qué punto aprendieron mis estudiantes?
- ¿qué ha beneficiado o bloqueado su aprendizaje?

- ¿hasta qué punto propicié su aprendizaje?
- ¿cómo puedo mejorar los resultados del aprendizaje de mis estudiantes?
- ¿cómo puedo ofrecer una mejor experiencia de aprendizaje?
- ¿cómo puedo mejorar mi enseñanza?
- ¿qué otro feedback necesito?

Estas preguntas deberían contrastarse con las preguntas que se haría un docente con una “actitud sumativa”: ¿soy lo suficientemente bueno como docente? ¿Qué tan bien enseñé? Resulta valioso considerar esta perspectiva para ampliar la función formativa de la retroalimentación.

Desde la perspectiva del estudiante, la retroalimentación remite también al aprendizaje autorregulado que comprende distintas acciones, pensamientos y emociones que permiten, mediante el control y la revisión, el logro de metas y objetivos académicos (Zimmerman, 2000). Uno de los aspectos centrales en el monitoreo del desempeño es la retroalimentación, generada por el propio sujeto, los pares o profesores y favorece procesos de reflexión, revisión y mejora de los aprendizajes. En los desarrollos teóricos actuales acerca de la noción de autorregulación se incluyen los aspectos sociales y contextuales. El aprendizaje autorregulado “ocurre cuando los estudiantes están motivados a involucrarse reflexiva y estratégicamente en actividades de aprendizaje dentro de ambientes que estimulen la autorregulación” (Butler, 2002: 60).

Por último, es importante considerar los aportes acerca de la retroalimentación entre pares. Topping (1998, 2009), investigó esta perspectiva y lo explicita a través de un ciclo de procesos cognitivos y emocionales. Identifica factores que intervienen en la retroalimentación entre pares y contribuyen a los aprendizajes tanto de quienes reciben como de quienes ofrecen.

FACTORES	MODOS DE INTERVENCIÓN
Aspectos organizativos y de compromiso	Definir objetivos y planes, organizar los tiempos y los espacios para que estudiantes puedan interactuar.
Conflictos cognitivos	Tratar de desmitificar creencias anteriores, errores conceptuales.
Andamiaje	Actividades como modelar, monitorear, diagnosticar, identificar errores, ofrecer información.
Habilidades de comunicación	Requeridas para ambos, quien recibe y quien ofrece retroalimentación ya que se trata de saber describir, explicar, sintetizar, argumentar, preguntar, escuchar.
Aspectos afectivos	Por ejemplo, la confianza mutua, la responsabilidad, la capacidad de influencia de uno sobre el otro.

El autor reconoce niveles de profundidad en que la retroalimentación formativa entre pares puede avanzar desde un nivel inicial que implica agregar, ajustar y reestructurar habilidades procedimentales, nuevos conocimientos y/o dificultades de comprensión hasta un nivel profundo de metacognición: explícita, intencional y estratégica y de autoestima (Anijovich, 2010: 141).

Entre las ventajas de la retroalimentación entre pares, consideramos las oportunidades de diálogo y la menor dependencia de los profesores, es decir, mayor autonomía (Sadler, 2010).

LA RETROALIMENTACIÓN FORMATIVA Y SUS CONDICIONES

84

En este apartado nos referirnos a las condiciones en las que se llevan adelante los procesos de retroalimentación formativa. Estas condiciones pueden facilitar o por el contrario, obstaculizar el proceso.

A las condiciones que aquí presentaremos y problematizaremos, no podemos adjudicarle una única definición para juzgar su eficacia. Pero en tanto condiciones, nos permiten reflexionar y establecer criterios para orientar las decisiones que tomamos al ofrecer retroalimentaciones a los estudiantes.

Cabe la pregunta ¿cuáles condiciones? Canabal y Margalef (2017) consideran: a) el tiempo en el que se realiza, b) el medio o formato para hacerlo, y c) el contexto o clima en el que se ofrece.

Respecto del tiempo en el que se realiza, si bien no es posible establecer un “plazo temporal único” para ofrecer retroalimentaciones, es necesario considerar que sea recibida en una temporalidad valiosa para los estudiantes y que, además, les permita utilizarla en aprendizajes posteriores. Una retroalimentación de excelente calidad, pero a des-tiempo, no cumple su función formativa.

El medio por el que se ofrece la retroalimentación también la condiciona: aunque el contenido de la retroalimentación sea el mismo (si es que pudiéramos afirmarlo), no es equivalente para quien la recibe que sea oral- una conversación-, o que se trate de un formato escrito. Algunos pueden considerar las facilidades de la retroalimentación escrita porque permite ajustarse a ritmos, tiempos, a ir y volver sobre lo escrito por el lector para poder mejorar a futuro. Sin embargo, este condicio-nante también es variable en función de quien recibe la retroalimen-tación (que puede ser una persona más afín a expresarse en formatos orales que escritos, o viceversa). Además, puede variar el potencial formativo según la facilidad y o la empatía de quién lo ofrece con el tipo de formato: algunos podrán ofrecer mejores retroalimentaciones escritas y otros, orales.

También, tal como mencionamos en trabajos anteriores, el clima en el que se realice la retroalimentación es el que, en definitiva, mayor impacto tiene en la función formativa (Anijovich y Cappelletti, 2017).

Sin un escenario de confianza, más allá del tiempo en que se ofrezca la retroalimentación o de su formato, la posibilidad de contribuir a la mejora de los aprendizajes disminuye.

Además de las condiciones, focalizaremos ahora en los factores que pueden obstaculizar los procesos de retroalimentación formativa. Estos factores se relacionan con las situaciones reales y los contextos en los cuales se ofrece la retroalimentación.

85

- Una tendencia de los profesores y maestros a evaluar el “volumen” de trabajo y la presentación en lugar de la calidad del aprendizaje alcanzado por los estudiantes. En el marco de las tradiciones de la escolaridad este obstáculo se hace presente con la fuerza de un paradigma tradicional, aunque desde un análisis reflexivo nadie dudaría en priorizar el valor de la calidad.
- Una mayor atención a la calificación, como lo hemos mencionado al inicio de este trabajo, que puede tender a disminuir la autoestima de los estudiantes en lugar de proporcionar orientaciones para mejorar sus aprendizajes futuros.
- Un fuerte énfasis en la comparación entre estudiantes que desmoraliza a los menos exitosos. Si bien esta comparación puede no atribuirse a los profesores, su presencia en el sistema educativo tiene las consecuencias mencionadas.
- La información que se ofrece en la retroalimentación no se relaciona con los propósitos de aprendizaje y los criterios de logro de la tarea.
- La retroalimentación de los profesores en ocasiones direcciona la tarea en lugar de ayudarles a los estudiantes a aprender más y mejor.
- La percepción de estudiantes y profesores acerca de que ofrecer retroalimentaciones genera una carga de trabajo excesiva dado que se utilizan variedad de instrumentos y se recogen evidencias múltiples a lo largo del proceso.
- El poco uso que los estudiantes hacen de la retroalimentación recibida de sus pares y profesores.

- Los hábitos muy arraigados de una cultura tradicional de evaluación como las pruebas de lápiz y papel como instrumento privilegiado.
- Cuando las retroalimentaciones son extensas pueden abrumar a quienes la reciben y desencadenar una respuesta frustrada o que no haya respuesta.
- Dificultad para comprender los criterios de evaluación de sus producciones.
- Dificultad para involucrarse con la retroalimentación porque no ven la conexión de ella con los objetivos de aprendizaje o con los criterios de evaluación.

Cuanto mayor es la presión y el predominio de culturas de resultados, más difícil es hacer progresos en la retroalimentación formativa. Este es un aspecto central para considerar en el marco de los sistemas educativos.

Para proponer estrategias que posibiliten una retroalimentación valiosa para estudiantes y profesores es necesario primero reconocer condiciones y obstáculos, tal como lo mencionamos en este apartado, considerando que no es posible pensar estrategias “a prueba de situaciones reales”. Ofrecer retroalimentaciones formativas requiere de una práctica reflexiva profunda de quienes la ponen en juego. Sus condiciones, en qué focaliza, en qué temporalidad, si es oral u escrita, si se ofrece o no a través de protocolos son decisiones profundas y diversas, en las que los profesores trabajan con grados de incertidumbre, presentes en todas las situaciones de enseñanza. Es en este encuadre que presentamos el apartado siguiente.

ESTRATEGIAS PARA DESARROLLAR RETROALIMENTACIÓN FORMATIVA

Nos vamos a referir a la retroalimentación formativa escrita entendiendo que es posible planificarla y revisarla antes de ponerla en acción.

Los profesores planifican la retroalimentación formativa para identificar, explorar y desafiar las ideas previas de sus estudiantes y ofrecer múltiples oportunidades y andamios para establecer vínculos con información nueva.

Nicol (2010) sostiene que la retroalimentación formativa debería verse en un contexto más amplio, como diálogo entre estudiantes y docentes. Dicho diálogo puede tomar un formato escrito y tiene que estar integrado en “*contextos dialógicos en los cuales se comparten las actividades de retroalimentación entre docentes y estudiantes, siendo éstas adaptativas, discursivas, interactivas y reflexivas*” (p.503).

Dysthe (1996) ofrece un marco de trabajo sobre cómo los profesores pueden usar diferentes estrategias para lograr que los estudiantes reflexionen y por ende construyan significado. Uno de los elementos claves en este marco de trabajo es la pregunta auténtica con final abierto sin respuestas preestablecidas. Tratar la respuesta del estudiante como un dispositivo de pensamiento en lugar del efecto de un “mecanismo de transmisión” en el que se muestra lo aprendido. Sin embargo, es necesario considerar que la pregunta que el docente intenta que sea auténtica puede no ser percibida como tal por el estudiante. Este aspecto no resulta menor: formular preguntas auténticas para los estudiantes requiere de consideraciones específicas a tener en cuenta por los profesores (Anijovich, 2014). Formular preguntas es una estrategia privilegiada para ofrecer retroalimentaciones. Y es posible, en función de lo que los profesores prioricen como contenido de la retroalimentación, que se trate de preguntas abiertas o cerradas, generales o direccionadas, en tanto y en cuanto promuevan la reflexión de los estudiantes y contribuyan a la autorregulación de sus aprendizajes.

Una de las estrategias que proponemos es “cambiar el tiempo de la retroalimentación” ¿Por qué exclusivamente al cerrar el trabajo? Sugérimos:

- Retroalimentación “en borrador”: Se ofrece sobre los borradores de las tareas de los estudiantes para promover que se hagan mejoras para la entrega final.

- Retroalimentación a través de ejemplos: Esta retroalimentación se basa en la comparación. Se ofrece ejemplos a los estudiantes para que ellos los comparen con los criterios establecidos para la elaboración de la tarea, y que puedan, en función de ellos, mejorar sus producciones. Estos ejemplos pueden obtenerse de tareas previas de otros grupos de estudiantes que ya la hayan realizado. Cabe señalar que, como se trata de “cambiar el tiempo de la retroalimentación” se propone hacerlo en el transcurso de la elaboración, y no necesariamente al finalizar la tarea.

Identificamos en distintas investigaciones, de las cuales seleccionamos la de Mulliner & Tucker (2017), estrategias que contribuyen a una retroalimentación formativa:

- Centrar la retroalimentación en el trabajo realizado y los logros alcanzados de acuerdo con criterios claros, preestablecidos, así como sobre aspectos de autorregulación de los aprendizajes.
- Reconocer las fortalezas y debilidades del estudiante y abordar los obstáculos y las dificultades del aprendizaje.
- Comunicar al estudiante qué debe hacer para mejorar.
- Crear oportunidades para dialogar sobre la evaluación y la retroalimentación.
- Involucrar a los estudiantes desde el comienzo de sus cursos, explicando los propósitos de la retroalimentación, cuándo, dónde y cómo se generarán las oportunidades de retroalimentación.
- Incluir espacios y tiempos para asegurar una comprensión completa de la retroalimentación antes de que los estudiantes utilicen los comentarios.
- Compartir los criterios de evaluación y ofrecer ejemplos, haciendo visible lo que se espera que los estudiantes aprendan. Lo que es visible para un profesor o tutor, puede no serlo para un estudiante.

Concluimos que algunas formas de ofrecer retroalimentaciones son más productivas y tienen más chances que otras para promover los

aprendizajes. Cuando los docentes ofrecen retroalimentación trabajan activamente para reconocer e integrar lo que los estudiantes están diciendo, para poder seguir y desarrollar el pensamiento que, según Dolin, Black, et al. (2018) es central para la evaluación formativa.

89

Entonces:

- diferentes docentes tienen diferentes maneras de ofrecer retroalimentaciones formativas,
- cada docente individual también ofrece diferentes tipos de comentarios,
- diferentes grupos de estudiantes pueden responder de maneras diferentes al mismo profesor.

Este análisis enfatiza la naturaleza dialógica y relacional de la retroalimentación formativa y hace foco en la manera en la que se formula, atendiendo a la posibilidad de ser recibido y usado de forma constructiva. La importancia de ser percibido como una invitación en la formulación es central para este diálogo. Como mencionamos anteriormente, una manera específica de ofrecer retroalimentaciones no garantiza un tipo de respuesta o de reflexión específica de parte de quien lo recibe.

La retroalimentación formativa es central en la enseñanza y para promover los aprendizajes de los estudiantes. Para que sea efectiva, los estudiantes deben involucrarse. Esto no es evidente, dado que en las actuales configuraciones de los sistemas educativos formales y sus tradiciones de “resultados”. Como ejemplo de estas limitaciones mencionamos que muchos estudiantes perciben que las retroalimentaciones que les ofrecen son irrelevantes para las próximas tareas.

Además, el involucramiento no es sólo la suma de las acciones que los profesores observan dentro y fuera del aula, con compañeros o tutores. Existen acciones opacas a la observación directa, requiere “interacción mental con el contenido” (Moore, 1989). Este proceso que podemos denominar involucramiento cognitivo, puede darse de modo superficial, o en un nivel más profundo.

Las investigaciones referidas sugieren que los estudiantes que se involucraron con las retroalimentaciones ofrecidas las consideraron útiles para orientar sus producciones.

90

En síntesis, la retroalimentación para ser formativa tiene que ser oportuna, constructiva, alentadora, focalizada y vinculada a criterios.

BIBLIOGRAFIA

ALLAL, L. & LÓPEZ MOTTIER, L. (2005). Formative assessment of learning. A review of publication in French, en J.Looney (ed), *Formative assessment: Improving learning in secondary classrooms*, París, OCDE, pp 241-264.

91

ANIJOVICH, R. (2010). La retroalimentación en la evaluación, en Anijovich, R. (Comp.) (2010) *La Evaluación significativa*. Buenos Aires. Paidós.

ANIJOVICH, R. (2014) *Gestionar una escuela con aulas heterogéneas*. Buenos Aires. Paidós.

ANIJOVICH, R. & CAPPELLETTI, G. (2017). *La evaluación como oportunidad*. Buenos Aires. Paidós.

BROOKHART, S. M. (2001). Successful students' formative and summative uses of assessment information, *Assessment in Education*, 8, (2), pp. 153-169.

BROOKHART, S. M. (2008). *How to give effective feedback to your students*. Virginia. ASCD publications.

BROOKS, C. (2017). *Coaching teachers in the power of feedback*. University of Queensland, Australia.

BUTLER, D. (2002). Qualitative approaches to investigating self-regulated learning: Contributions and challenges. *Educational Psychologist*, 37(1), 59-63.

CAMILLONI, A. (2004). Sobre la evaluación formativa de los aprendizajes, *Quehacer educativo*, año XIV, N 68, Diciembre, pp.6-12. Uruguay.

CANABAL, C. & MARGALEF, L. (2017). *La retroalimentación: La clave para una evaluación orientada al aprendizaje Profesorado*, 21(2): 149-170 (2017). [<http://hdl.handle.net/10481/47669>]

- CARLESS, D., D. SALTER, M. YANG, & J. LAM. (2011). Developing Sustainable Feedback Practices. *Studies in Higher Education* 36 (4): 395–407. doi:10.1080/03075071003642449.
- 92 COSTA, A. & GARMSTON, R. (2017). A feedback perspective in Wallace, I & Kirkman (eds). *Best of the best: Feedback*. Carmarthen, Wales: Crown House Publishing.
- DYSTHE, O. (1996). “The Multivoiced Classroom: Interactions of Writing and Classroom Discourse.” *Written Communication* 13 (3): 385–425. doi:10.1177/0741088396013003004.
- DOLIN, J., P. BLACK, W. HARLEN, & A. TIBERGHIE. (2018). “Exploring Relations between Summative and Formative Assessment.” In *Transforming Assessment: Through Interplay between Practice, Research and Policy*, edited by Jens Dolin and Robert Evans. *Contributions from Science Education Research*, vol. 4. 53–80. Cham: Springer. doi:10.1007/978-3-319-63248-3_3.
- HANDLEY, K. & WILLIAMS, L. (2009). *From copying to learning: using exemplars to engage students with assessment criteria and feedback*. <http://dx.doi.org/10.1080/02602930903201669>.
- HATTIE, J. & TIMPERLEY, H. (2007). *The power of feedback*. *Review of Educational Research*, 77 (1): 81-112.
- HATTIE, J. (2012). *Visible learning for teachers*. Routledge. New York.
- HAVNES, A., SMITH, K., DYSTHE, O. & LUDVIGSEN, K. (2012). Formative assessment and feedback: Making learning visible. *Studies in Educational Evaluation*. March 2012. <http://dx.doi.org/10.1016/j.stueduc.2012.04.001>.
- MOORE, M. (1989). Three types of interaction. *American Journal of Distance Education* 2: 1–6.
- MULLINER, E. & TUCKER, M. (2017). Feedback on feedback practice: perceptions of students and academics, *Assessment & Evaluation in Higher Education*, 42:2, 266-288. doi:10.1080/02602938.2015.1103365.

- NICOL, D. (2010). "From Monologue to Dialogue: Improving Written Feedback Processes in Mass Higher Education." *Assessment & Evaluation in Higher Education* 35. (5): 501–517.
- PERRENOUD, P. (2008). *La evaluación de los alumnos*. Buenos Aires. Colihue.
- SADLER, R. (1989). Formative assessment and the design of instructional assessment. *Instructional Science*, 18, 119-144.
- SADLER, R. (2010). Beyond feedback: developing students capability in complex appraisal. *Assessment and Evaluation in Higher education*, 35 (5): 535-550.
- SHEPARD, L. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29 (7), 4-14.
- SHUTE, V. (2008). Focus on formative feedback, *Review of Educational Research*, 78 (1). 153-189.
- STIGGINS, R. (2005). Student- involved assessment for learning. *Upper Saddle River*. Prentice Hall.
- TARAS, M. (2002). Using assessment for learning and learning from assessment. *Assessment & Evaluation in Higher Education*, 27(6), 501–510.
- TOPPING, K. (1998). Peer assessment between students in college and university. *Review of Educational Research*, 68, 249-276.
- TOPPING, K. (2009). Peer assessment. *Theory into Practice*, 48 (1), 20-27.
- WILIAM, D. & BLACK, P. (2002). Feedback is the best nourishment, *Times Educational Supplement*, special supplement: 'Mind Measuring'. October 4. 8–9.
- WILIAM, D. (2009). "Una síntesis integradora de la investigación e implicancias para una nueva teoría de la evaluación formativa. *Archivos de Ciencias de la educación*. Universidad de La Plata, año 3, N 3, 4 época, pp 15-44.

WILLIAM, D. (2011). *Embedded formative assessment*. Bloomington, Solution Tree.

94 ZIMMERMAN, B. (2000). Attaining self-regulation: A social cognitive perspective. En M. Boekaerts, P. R. Pintrich & M. Zeidner. *Handbook of self-regulation*. Academic Press. San Diego.

ALGUNOS PUENTES NECESARIOS

Lilia V. Toranzos

95

Por definición la evaluación es siempre una acción de carácter político en el sentido de un quehacer cuyo objetivo es intervenir y modificar la realidad sobre la cual se asienta. Por lo tanto, es una actividad que de por sí moviliza, divide aguas, genera discusiones, y conversaciones intensas. Todo ello nos permite disponer de mucho material de trabajo variado y rico para quienes nos ocupamos de la evaluación, pero a la vez nos desafía a sortear el principal reto: cómo transitar el camino que conduce desde las controversias que se generan alrededor de la evaluación hacia propuestas que permitan orientar (en sentido de mejora) las intervenciones que se derivan de sus resultados.

Es por ello, por su naturaleza política, que toda acción de evaluación conlleva una responsabilidad primordial respecto a que todos los esfuerzos estén orientados con una intención clara y explícita, que sean provechosos, constructivos y potentes, ya que una intervención inútil o vacua –en este sentido– es doblemente costosa y perjudicial.

Frente a esta situación es posible distinguir, al menos, dos planos diferentes en los que tiene lugar el desafío de acortar las brechas existentes y que, por tanto, proponen una interesante agenda de trabajo en el área de la evaluación educativa. El primero, la vinculación entre los distintos niveles de análisis que pueden proporcionar información sobre los aprendizajes de los estudiantes. El segundo, en relación a la preocupación por el uso que se hace de la información sobre los aprendizajes para orientar acciones de mejora, tanto a nivel del diseño de políticas educativas, como al nivel institucional y/o del aula.

ACERCA DE LA RELACIÓN ENTRE LOS DISTINTOS PLANOS

96

En primer lugar, el plano más evidente es el de la relación entre la perspectiva panorámica o macro sobre lo que sucede a nivel del sistema educativo en sentido amplio, con el nivel meso, correspondiente a las instituciones educativas y más específicamente en el aula, y la perspectiva más micro que está directamente vinculado con el quehacer cotidiano de docentes y alumnos en esas aulas. Cada nivel de análisis procura mirar y aportar desde distintas ópticas que, lejos de ser antagónicas, debieran ser complementarias. ¿Cómo tender puentes o ensayar senderos que aproximen estas perspectivas y permitan transitar (entre una y otra) en ambos sentidos? ¿Qué se necesita para construir esos caminos?

Bajo esta primera consideración cabe señalar que es un hecho evidente la coexistencia de diferentes dispositivos y prácticas de evaluación que se producen en diferentes niveles del sistema educativo y que inciden en tanto en las intervenciones pedagógico didácticas, como en la percepción social sobre la educación.

Hay un conjunto de prácticas de evaluación a nivel macro que ya se encuentran instaladas: evaluaciones jurisdiccionales/locales, nacionales, e incluso internacionales que proporcionan información de distinto tipo acerca de los aprendizajes de los estudiantes, como así también otro conjunto de información complementaria que da cuenta de las circunstancias y condiciones -personales, educativas y sociales- bajo las cuales estos estudiantes aprenden. Por otra parte, existe un amplio repertorio de prácticas de evaluación que regularmente los docentes llevan a cabo en cada escuelas, y dentro de cada aula, que en algunos casos se encuentra más sistematizada, en otros menos y a su vez podríamos reconocer los esfuerzos- individuales y colectivos- que cotidianamente se realizan para la mejora de dichas prácticas de evaluación como parte de las prácticas de enseñanza.

Entre estas instancias es posible identificar actualmente, en la Ciudad de Buenos Aires, una serie de dispositivos e iniciativas que intentan consolidarse como instancias intermedias, estaciones en ese recorrido y que comienzan a instalarse como parte de ese puente necesario entre

la mirada macroscópica del sistema educativo y la consideración más cercana a lo que acontece a diario en el aula. En este sentido se identifican, entre otros dispositivos:

- FEPBA y TESBA: las evaluaciones de aprendizaje que se implementan con carácter censal y periodicidad anual, relevando algunos aprendizajes logrados por los estudiantes en Prácticas del Lenguaje/ Lengua y Literatura y Matemática en 7° grado del nivel primario y 3° año del nivel secundario respectivamente.
- Pausas Evaluativas: se trata de una iniciativa orientada a alumnos de 3° y 6° grado del nivel primario y gestión estatal para relevar aprendizajes correspondientes a las áreas de Prácticas del Lenguaje y Matemática. En este caso, se trata de una propuesta en donde los supervisores, directivos y docentes intervienen en el diseño y corrección de las pruebas.
- Progresiones de Aprendizaje³¹ se definen como

“... un complemento del marco curricular en el que se basan y dialogan con los materiales de desarrollo curricular que apuntan a fortalecer la intervención didáctica (...), a partir de la descripción de recorridos posibles y pertinentes para la enseñanza y el aprendizaje de contenidos fundamentales de la trayectoria escolar. (...) Proporcionan criterios claros y compartidos para el monitoreo de los aprendizajes con una perspectiva formativa, orientada a la toma de decisiones para la enseñanza.”

En todos los casos, se trata de propuestas que se instalan como parte de ese recorrido y procuran aportar herramientas, puntos de apoyo y espacios de reflexión que fortalezcan las estrategias de trabajo de los docentes y acorten las distancias percibidas entre unos y otros espacios evaluativos.

31. Para mayor información al respecto, es posible consultar en: <https://www.buenosaires.gob.ar/calidadyequidadeducativa/evaluacion/progresiones>

RESULTADOS: ¿PARA QUÉ Y PARA QUIÉN?

98

La otra preocupación que resulta de interés es aquella que vuelve la mirada sobre la información que se produce a partir de los procesos de evaluación. La producción de información es el corazón de los procesos de evaluación sobre el cual invertimos mucho esfuerzo técnico; requerimos que esta información se constituya en una evidencia sólida y consistente, metodológicamente muy cuidada, entre otros criterios técnico- metodológicos que es preciso atender. Sin embargo, el verdadero propósito y lo que le otorga sentido a la evaluación es que esa evidencia se transforme en conocimiento, que se convierta en herramienta en algún sentido más práctica, más tangible, más maleable y que permita tomar decisiones en términos de enseñanza en el ámbito del aula, decisiones en términos de orientaciones institucionales para el conjunto de los docentes de una escuela y también decisiones de intervención en materia de política educativa.

Es por ello que, no resulta suficiente que sólo se mejoren los procedimientos metodológicos para generar evidencia, si nadie la usa. O que los resultados posibiliten la construcción de noticias, para su utilización política partidaria. El desafío consiste en encontrar formas de transitar (y acortar) las distancias existentes entre la producción de la información consistente y la transformación de la misma en conocimiento e insumos de trabajo para diferentes actores del sistema educativo (en los distintos niveles) y otros posibles interesados, fuera del sistema educativo.

Si se ubica el centro de las preocupaciones en el uso y el aprovechamiento de la información que producen las evaluaciones, lo que le otorga verdadero sentido a las mismas, surge la necesidad de construir una mirada integral sobre los diferentes dispositivos de evaluación y por ende la posibilidad de generar un sistema de evaluación. El desafío consiste en promover mayores niveles de articulación entre las diferentes iniciativas a la vez que se identifican necesidades y requerimientos diferenciados para los diferentes actores del sistema: docentes, directivos, supervisores, familias, alumnos, niveles de gestión intermedia y central.

Simultáneamente, se trata de que se comprenda y reconozca que los resultados de las evaluaciones pueden responder a diversos propósitos y, en consecuencia, su aprovechamiento dependerá de fortalecer las capacidades de uso de los sujetos y las instituciones.

El diseño de un sistema de evaluación de los aprendizajes implica una decisión estratégica para ofrecer información con diferentes propósitos y a variados actores permitiendo establecer una mayor coherencia entre la enseñanza y los aprendizajes de los estudiantes.

La definición de sistema comprende de manera necesaria, la coexistencia de múltiples evaluaciones, debiendo ser cada una de ellas válida y confiable para el propósito y destinatarios establecidos. Como tal, el sistema organiza información sobre los procesos y el contexto de aprendizaje lo que permite disponer de insumos que contribuyen a la toma de decisiones tendientes a mejorar la enseñanza a través de diferentes tipos de intervención pedagógica en los niveles de gestión central, intermedio y local del sistema educativo.

¿Qué tipo de información que requiere cada actor destinatario para promover y mejorar los aprendizajes de los alumnos? ¿De qué manera, la producción de información sobre los aprendizajes de los estudiantes se constituye en un factor de mejora y no en un instrumento de control o sanción explícita o implícita?

Esta orientación implica una transformación en la mirada de la evaluación que no es nueva pero que aún no logra permear de manera efectiva y consistente sobre todo el “mundo educativo y social”.

CONFIANZA Y COMPROMISO COMO CONDICIONES CENTRALES

Todavía persisten ideas respecto que es normal y aceptable que no todos los estudiantes alcancen los objetivos de aprendizaje y, en consecuencia, otros no lo logren. Bajo este supuesto la evaluación se transforma en un “parteaguas” cuyo sentido es la distinción entre unos y otros. Sin embargo, bajo la regulación normativa que establece la obligatoriedad de la educación hasta el nivel secundario el desafío consiste

en que todos los alumnos logren alcanzar los objetivos de aprendizaje establecidos curricularmente y con ello, dispongan de las competencias y habilidades necesarias para vivir en una sociedad en la cual el conocimiento constituye una pieza clave para el ejercicio de una ciudadanía plena.

En este sentido, el marco de derecho en el que se desenvuelve el sistema educativo, así como el contexto social interpela fuertemente y reclama formas más adecuadas para evaluar los aprendizajes y por ende, la revisión papel de la evaluación como un elemento constructivo que aporta a la mejora del quehacer pedagógico en tanto acompaña y dota a los docentes de más y mejores herramientas para intervenir.

¿Cómo avanzar en la tarea de tender puentes entre los diferentes componentes de un sistema de evaluación? ¿Y entre los resultados y la construcción de herramientas para la mejora de los mismos?

Tal como lo expresa Huebner, (2009) se presentan dos desafíos centrales; por una parte **construir confianza** con los actores de sistema, ofreciendo instancias de capacitación y formación específicas (para docentes, directivos, supervisores) y proponiendo apoyo y acompañamiento a partir de procesos de retroalimentación que colaboren con el diseño de propuestas de políticas educativas para los niveles centrales y propuestas de gestión pedagógica y enseñanza, para los docentes y directivos.

El segundo desafío consiste en generar **compromiso** en todos los niveles de gestión para construir una cultura que se guíe por la evidencia y que contribuya a enriquecer las prácticas y las intervenciones en todos los ámbitos de gestión del sistema educativo (Huebner, 2009: 85-87).

Estos dos desafíos también se presentan a nivel institucional, e incluso al nivel del aula, cuando reflexionamos en el modo en que el tipo de comunicación que se hace de los resultados de la evaluación puede afectar a los estudiantes.

Cuando los alumnos obtienen bajos puntajes en sus evaluaciones en forma reiterada desde los primeros grados construyen una autopercepción que les ratifica la falsa idea de que no van a ser capaces de aprender y por ello, pierden confianza en sí mismos. Por esta razón, el

uso de la información sobre los resultados de las evaluaciones de los alumnos se debe inscribir en un espacio de retroalimentación cuyo propósito esencial es la mejora y la identificación de los logros, es decir de aquello de lo que los alumnos son capaces de hacer y resolver.

Si la sociedad quiere que todos los alumnos alcancen buenos niveles de desempeño, las escuelas tienen que trabajar para que todos los estudiantes construyan la confianza necesaria para creer que pueden conseguirlo y los organismos encargados de la producción y administración de pruebas a gran escala deben hacer los esfuerzos necesarios para formular los argumentos indispensables, en función de los resultados obtenidos, para construir confianza y la disposición a enfrentarse a situaciones evaluativas en mejores condiciones.

Si los estudiantes son los usuarios privilegiados de los resultados de las evaluaciones es necesario apoyarlos para que sus decisiones sobre los aprendizajes jueguen a su favor. En cambio, si se consideran a los adultos como los usuarios se perderá de vista que un alumno puede considerar que determinados aprendizajes se encuentren fuera de su alcance y terminen inmovilizando las decisiones de enseñanza de maestros y profesores. Los estudiantes cuentan con el poder de tornar ineficientes a los educadores, aun cuando no se lo propongan.

Quando *“un alumno decide que un aprendizaje está fuera de su alcance o que el riesgo de fracaso es demasiado grande o amenazador, entonces, hagamos lo que hagamos los adultos, el aprendizaje termina”* (Stiggins, 2008:8).

¿EVIDENCIAS O INDICIOS?

La idea de un accionar basado en la evidencia es algo bastante generalizado en el discurso pero no tan desarrollado en la práctica efectiva, el valor de la evidencia suele ponerse en cuestión con frecuencia en el ámbito educativo y eso constituye un elemento invalidante en el marco de los procesos de evaluación. Tanto la noción de evidencia como la de indicio son dos conceptos que lejos de ser contrapuestos plantean facetas diferentes a tener en cuenta en la práctica evaluativa.

Por una parte, la idea de evidencia hace referencia a una cualidad abstracta y a la vez a manifestaciones concretas, lo cierto es que en nuestro ámbito frente a un conjunto de información como son los resultados de las evaluaciones cabe preguntarnos a qué refieren esas evidencias, es decir ¿qué es lo que evidencian? ¿Qué es lo que manifiestan en forma clara y certera? ¿Resultados de qué proceso se hacen evidentes?

En este punto es que la idea de indicios se presenta como otro aspecto a considerar. En términos generales, la realidad sobre la que procuramos intervenir es sumamente compleja, y la información que se genera a partir de los procesos de evaluación nos permite acercarnos a ella siempre parcialmente. Por muy bien diseñado e implementado que sea el proceso de evaluación de los aprendizajes, o todo el conjunto de dispositivos de evaluación, en el mejor de los casos lograremos captar algunas señales, algunos indicios, pistas de aquellos procesos más complejos que se despliegan en la realidad de nuestras instituciones escolares.

Esto implica que si se ha generado un amplio espacio de conversación sobre qué es lo que interesa y cuál es el foco; si se han tomado todos los recaudos y resguardos metodológicos producto de discusiones intensas y fundadas; si se han generado acuerdos sobre el propósito de los procesos de evaluación en marcha y su necesaria relación con la mejora, estaremos en mejores condiciones para construir dispositivos que generen indicios sobre aquello que queremos capturar.

Frente a la complejidad que presentan los procesos sobre los cuales intentamos intervenir la idea de indicios nos obliga a adoptar una posición de relativa modestia y cautela frente a la información derivada de la evaluación y a la vez nos desafía a revisar constantemente el conjunto de los procesos de producción de información educativa con el fin de producir más y mejores evidencias sobre aquellos aspectos centrales del quehacer educativo. En igual sentido, se vuelve necesario explorar otras fuentes que generen información complementaria que permita enriquecer el análisis de los resultados de las evaluaciones promoviendo una mirada más comprehensiva de la realidad educativa.

Respecto a la producción de información educativa, la investigación aporta un modo de producir conocimiento muy interesante, y que en

algunos casos no se instala de modo habitual como una conversación sólida en el ámbito de la evaluación. La investigación nos proporciona sus resultados, su enfoque, un modo de tratar lo que temáticamente se reconoce como central, hay mucho que aprender en este proceso de acumulación de conocimiento que genera la práctica investigativa. En el ámbito de la evaluación no estamos tan acostumbrados a estas formas de acumulación sino a generar acciones de evaluación que generan resultados y que a su vez permiten iniciar un nuevo ciclo de acciones. Esta práctica cíclica está relativamente pautada y estandarizada, suele variar muy poco y cuesta mucho esfuerzo generar alguna movilización en ese sentido.

En este sentido, la forma de producción del conocimiento en el ámbito de la investigación puede promover otro modo complementario de producción de información. Hoy más que nunca, parece necesario generar procesos de acumulación de conocimiento evaluativo derivado de las prácticas realizadas, y de las reflexiones sobre la información producida y sobre los propios procesos de evaluación. La sistematización de ciertas acciones de evaluación no debe naturalizar estos procesos ni convertirlos solo en una rutina sin variantes dentro del sistema, deben convertirse una y otra vez en una ocasión para producir más y mejores aprendizajes en todo sentido y orientar las decisiones hacia la mejora continua. En la misma línea, la idea de construir procesos y generar procesos de producción de conocimiento desde una perspectiva de igualdad en el peso relativo de los aportes de evaluadores, investigadores y docentes configura una propuesta de construcción colectiva sumamente potente.

NUEVOS (Y NO TAN NUEVOS) DESAFÍOS PARA LOS SISTEMAS DE EVALUACIÓN

En síntesis, qué requerimos hoy para integrar un sistema de evaluación de los aprendizajes que proporcione información valiosa, para distintos propósitos y destinada a una amplia gama de actores. En primera instancia se tratará de un sistema que incluya múltiples dispositivos de

evaluación, válidas y confiables y que como expresan Sigman y Mancuso “...comprende múltiples evaluaciones-siendo cada una de ellas válida y confiable para el propósito y los destinatarios específicos-que organiza información sobre los procesos y el contexto de aprendizaje para generar insumos que contribuyan a la toma de decisiones orientadas a mejorar la enseñanza en los distintos niveles de gestión del sistema educativo” (2017: 8). Para ello, se requiere la construcción de una mirada compartida, por parte del conjunto de los actores educativos involucrados, sobre los ejes o focos centrales, los propósitos de cada propuesta de evaluación y los usos de la información derivada de misma. Esto demanda altos niveles de coordinación y articulación entre las diferentes áreas involucradas y la complementariedad de la información que se produce a partir de diferentes fuentes.

Por último, cabe señalar un conjunto de desafíos que aporta la experiencia internacional en evaluación educativa respecto a la construcción y mejora de los sistemas de evaluación de los aprendizajes en distintos aspectos de su desarrollo: la política de evaluación, el diseño y los procedimientos, las capacidades técnicas requeridas, y el uso de los resultados.

DESAFÍOS PARA LA CONSTRUCCIÓN DE UN SISTEMA DE EVALUACIÓN DE LOS APRENDIZAJES³²

- Política de evaluación

- Elaborar un marco integrado de evaluación

32. Extracto de OCDE (2013) Sinergias para un mejor aprendizaje: una perspectiva internacional sobre evaluación. París, OCDE en Pérez Centeno, C, Asprella, G y Toranzos, I. (2015) Los estudios internacionales de evaluación de la calidad en América Latina: estado de situación y de debate en el ámbito del Mercosur. En Monarca, H. (2015) Evaluaciones externas. Mecanismos para la configuración de representaciones y prácticas en educación. Buenos Aires, Ed Miño y Dávila.

- Alinear el marco de evaluación con los objetivos educativos y los objetivos de aprendizaje de los estudiantes.
- Reforzar la capacidad del Estado para definir y sostener una política de evaluación.

- *Diseño y procedimientos técnicos*

- Asegurar que los componentes principales estén suficientemente desarrollados dentro del marco de evaluación.
- Lograr convergencia técnica entre iniciativas de evaluación en diferentes niveles de gestión del sistema educativo.
- Establecer articulaciones entre los componentes y el marco de evaluación.
- Colocar a los estudiantes en el centro de la evaluación.

- *Capacidad técnica*

- Incrementar las capacidades técnicas y la formación de equipos solventes en materia de evaluación educativa.
- Mejorar la articulación entre los niveles gestión dentro del sistema educativo vinculado con los ejes de la evaluación.
- Comprometer a los interesados y a los profesionales en el diseño e implementación de las políticas de evaluación.

- *Uso de los resultados*

- Mantener una adecuada gestión del conocimiento dentro del marco de evaluación.
- Comprometerse con el uso de evidencia para el desarrollo de intervenciones en el marco del sistema educativo.

En síntesis, la agenda de trabajo en materia de evaluación de los aprendizajes es extensa y desafiante, es necesario generar consensos claros sobre los propósitos de cada iniciativa de evaluación y lo modos de implementarla así como promover la participación más intensa y simétrica de evaluadores, docentes/directivos escolares e investigadores en los procesos de diseño e implementación y mejora de las acciones de evaluación. Pero todo requiere no perder de vista que el objetivo central es mejorar las trayectorias escolares, consecuentemente sus

aprendizajes y las condiciones en que estos se producen para el conjunto de niñas, niños y adolescentes que habitan hoy nuestras escuelas.

REFERENCIAS

HUEBNER, T. A. (2009) *What research says about.../Balanced Assessment*. 67(3), pp.85-86.

107

PÉREZ CENTENO, C. ASPRELLA, G. Y TORANZOS, L. (2015). Los estudios internacionales de evaluación de la calidad en América Latina: estado de situación y de debate en el ámbito del Mercosur. En Monarca, H. *Evaluaciones externas. Mecanismos para la configuración de representaciones y prácticas en educación*. Buenos Aires, Ed Miño y Dávila.

SIGMAN, H. Y MANCUSO M. (2017). *Designing a Comprehensive Assessment System*. San Francisco, CA; WestEd.

STIGGINS, R. (2008). *Assessment Manifesto: A call for the development of balanced assessment systems*, ETS Assessment Training Institute, Portland.

CONFERENCIA REFLEXIONES Y PROPUESTAS EN TORNO A LA ARTICULACIÓN EXTERNA Y LA EVALUACIÓN EN EL AULA

Jornada “Evaluación y enseñanza: contribuciones para la práctica educativa”.

109

Mg. Pedro Ravela³³

A cierta edad parece que uno que ya tuvo varias vidas. Yo empecé siendo profesor de aula en Filosofía, después en una segunda vida profesional me dediqué a la evaluación y trabajé en Uruguay en evaluaciones estandarizadas -siempre con la preocupación de que sirvieran en alguna medida para transformar algo en la práctica de aula-. Estuve a cargo de PISA y trabajé en varios países en torno al uso de los resultados de dichas evaluaciones. Desde el 2008 para acá, como resultado de una historia larga en la que no me detengo, nos pusimos a trabajar con un equipo de gente en torno a la evaluación dentro del aula, la que hacemos los docentes. Fruto de esta última vida es este libro que co-

33. Pedro Ravela es Profesor de Filosofía (IPA) y Master en Ciencias Sociales y Educación (FLACSO Buenos Aires). A lo largo de su carrera se especializó en temas de evaluación educativa, sobre los que ha sido investigador, docente y asesor en varios países de la región. En Uruguay fue Director de Evaluación en la Administración Nacional de Educación Pública (1996-2002), Coordinador Nacional del Estudio PISA (2002-2006), Director del Instituto de Evaluación Educativa en la Universidad Católica (2007-2012) y Director Ejecutivo del Instituto Nacional de Evaluación Educativa (INEED) (2012-2014). Actualmente integra el Consejo Técnico del Instituto Nacional para la Evaluación de la Educación (INEE) de México. En los últimos años ha enfocado su trabajo en la investigación sobre evaluación en el aula y en el trabajo con docentes en torno a esta práctica. Entre sus publicaciones sobre evaluación educativa se destacan el libro *Fichas Didácticas. Para comprender las evaluaciones educativas* (2006) y *¿Cómo mejorar la evaluación en el aula?* (2017).

mentó recién Tamara [“¿Cómo mejorar la evaluación en el aula?” (2017)], que escribimos con dos colegas de Uruguay, Beatriz Picaroni y Graciela Loureiro. Como parte de los trabajos de investigación que hicimos a partir de 2008, tuvimos la oportunidad de entrevistar docentes de sexto de primaria y tercero de secundaria en diez países de América Latina, y de pedirles que compartiesen con nosotros lo que ellos quisiesen de sus propuestas de evaluación, las que usaban en sus aulas.

El puntapié inicial de estos trabajos fue la siguiente preocupación: estamos poniendo mucha cabeza, inversión, tiempo y dinero en la evaluación externa y nadie mira qué está pasando dentro del aula, qué hacen y cómo evalúan los docentes. A partir de entonces he trabajado en torno a la relación entre los dos temas [evaluación estandarizada y evaluación en el aula] y, en los últimos años, me estoy dedicando más a trabajar con maestros o profesores sobre la evaluación en el aula.

Una introducción necesaria: enfoques de política de evaluación a gran escala

1. Sistemas de información al público y monitoreo para la toma de decisiones y la gestión
2. Difusión de resultados para responsabilidad compartida y rendición de cuentas - derecho a la educación
3. Sistemas de mercado y competencia entre escuelas
4. Incentivos económicos directos a partir de los resultados
5. Devolución a docentes y escuelas con fines de desarrollo profesional e intervención pedagógica
6. Pruebas con consecuencias para los estudiantes / Exámenes - incidencia en las calificaciones
7. Información para toma de decisiones de política educativa

Imagen obtenida de la presentación de Pedro Ravela en la Jornada “Evaluación y enseñanza: contribuciones para la práctica educativa”, que se desarrolló el día 10 de agosto de 2018.

Esta tarde quiero hacer una reflexión sobre cómo articular estas dos cosas, qué puentes se pueden establecer entre la evaluación estandarizada y la evaluación en el aula. Lo primero que es indispensable, como marco general, es recordar que la evaluación estandarizada a gran escala se puede utilizar de muchas maneras. Es posible distinguir al menos siete grandes lógicas de política educativa para su uso. En primer término, podemos pensar la evaluación externa como un sistema de información pública. Del mismo modo que tenemos datos periódicos de desempleo, pobreza, crecimiento económico e inflación, necesitamos tener información sobre educación. Creo que lo que aporta un sistema de evaluación externa estandarizada, desde esta perspectiva, es incorporar a la agenda pública el tema de los aprendizajes, porque muchas veces queda afuera del escenario. Más allá de que muchas veces se haga mal, de que los debates mediáticos a partir de los resultados de evaluaciones estandarizadas suelen ser poco enriquecedores de la discusión pública, es una ganancia en la región que la información de las evaluaciones estandarizadas esté disponible. Este es un primer rol básico, que no garantiza ningún cambio, del mismo modo que las estadísticas sobre empleo no van a generar empleo, eso lo hacen las políticas.

Un segundo enfoque consiste en afirmar que “la sociedad tiene que saber sobre lo que está pasando en las escuelas”. Es interesante notar que en América Latina el país que avanzó en esa línea fue el Brasil de Lula. Brasil hasta entonces tenía una prueba muestral y anual, el Sistema Nacional de Evaluación de la Educación Básica (SAEB), que se aplicaba cada dos años y entregaba resultados muestrales. Cuando Lula llega a la Presidencia dice “no”, la prueba debe ser censal, todas las escuelas del país deben ser evaluados y los datos de cada escuela deben ser públicos. Todo esto desde un enfoque de transparencia y a asegurar el derecho a la educación. Así se creó la llamada “Prova Brasil”, que tuvo impactos bastante positivos en la última década. No hicieron *rankings* [como habían hecho en Chile] sino que simplemente se buscó que las familias y la sociedad pudiesen saber qué había pasado en cada escuela con las pruebas. Junto con esto se hicieron muchas más cosas, no es tan simple como eso. Entre otras, crearon un el Índice de Desarrollo de la Educación Básica (IDEB), que combina los resultados de apren-

dizaje de los estudiantes con datos de retención de los alumnos y de finalización en tiempo de los ciclos educativos. Esto permitió combinar la evaluación de la capacidad de la institución para que los estudiantes aprendan, con la de asegurarse de que los estudiantes terminen sus estudios, evitando que simplemente se saquen de encima a los estudiantes con problemas.

Este paso fue algo inesperado porque el debate en América Latina estaba más bien marcado por la experiencia chilena, que era la del tercer enfoque de la diapositiva, usar la evaluación externa como un mecanismo de mercado y de generación de competencia entre las escuelas, con publicación de rankings y demás (que es algo que Chile ahora ha abandonado, cambiando el rumbo en los últimos años).

El cuarto enfoque que ha tenido peso en la región, especialmente en México (aunque también en Chile) es el de usar los resultados de pruebas externas para otorgar incentivos económicos a los docentes. México lo hizo de distintos modos a partir del año 92, a través de acuerdos de la Secretaría de Educación Pública (SEP) y el sindicato de maestros (SNTE). La experiencia más reciente fueron las Pruebas ENLACE, que terminaron siendo canceladas en el 2013. A partir de ese año hubo un vuelco de política, porque se constató que esa forma de uso generaba una cantidad de pequeñas corruptelas. Por otra parte, la política de valorar el trabajo de los docentes a partir de los resultados de sus estudiantes no se sostiene teóricamente. Un docente puede ser muy bueno y hacer las cosas muy bien, pero no necesariamente garantizar que sus estudiantes aprendan si estos últimos no ponen su parte. El aprendizaje “lo hacen” los estudiantes, no los docentes.

Luego ha habido una quinta lógica, que fue desarrollada en varios países de la región, en especial en Uruguay y Perú. Esta consistió en pensar la evaluación externa principalmente como una herramienta que le diera insumos de trabajo a los docentes. Tanto Perú como Uruguay, han mantenido durante dos décadas la política de que los resultados por escuela son confidenciales, y que la información se les entrega a los centros educativos para que sea utilizada, buscando que sea de valor para los docentes.

La sexta lógica existe en muchos países del mundo. Tal vez los casos más conocidos son los exámenes en China y el bachillerato francés. Pero hay experiencias en una gran cantidad de países, que tienen alguna forma de examen nacional mediante el cual las pruebas externas estandarizadas tienen consecuencias directas para el estudiante. Por ejemplo, definiendo la acreditación de un ciclo. Este es otro modo de usar las pruebas externas.

Por último, un enfoque que suele ser el caballito de batalla de los discursos políticos pero muy difícil de concretar, es el que sostiene el uso de la evaluación como información de base para la toma de decisiones de política educativa. Es una linda idea, pero por lo general las políticas educativas se rigen por otras lógicas, no es tan fácil. Gilbert Valverde, un colega que trabaja en Albany NY y conoce mucho América Latina, tiene una frase muy linda que dice *“toda política educativa es una hipótesis que uno debería someter a validación”*. La implementación de cualquier acción o intervención en el sistema educativo debería tener un buen dispositivo de evaluación, justamente para constatar si lo planeado realmente se implementa o no y, en caso afirmativo, cómo (porque los programas siempre son transformados por los agentes durante su implementación). Recién después viene la pregunta por el impacto: qué resultados tiene la política implementada. Enfatizo esto porque muchas veces se realizan “evaluaciones de impacto” de cosas que nunca llegaron realmente a implementarse, por lo que difícilmente puedan tener algún impacto.

Dicho lo anterior como marco general, y teniendo en cuenta que la pregunta que motiva esta presentación es “cómo puede la evaluación externa aportar algo a la evaluación en el aula” o “dialogar con la evaluación en el aula”, una segunda reflexión que quiero plantearles es la tensión entre una lógica centrada en los datos y una lógica centrada en los agentes.

¿A qué me refiero? Habiendo trabajado dentro de las áreas de evaluación, soy consciente de que muchas veces uno tiene una preocupación especial, e incluso una cierta fascinación, con los datos. Quiere generar datos confiables, precisos, pertinentes. Pero después viene la queja: la gente no usa los datos o no los entiende. Es una lógica un

poco egocéntrica. “Estoy especializado en evaluación, me encanta la psicometría y quiero que todos entiendan y sepan usar los datos que generamos...”. Pero, en realidad, si uno quiere que otros usen los datos, debería adoptar una postura distinta, que denomino “lógica centrada en los agentes”. Esto quiere decir que la Unidad de Evaluación debería partir de preguntarse “¿quién queremos que use esta información?” y “¿cómo y para qué la podrían utilizar?”. Para responder estas preguntas es necesario conocer cómo es el mundo y la lógica de trabajo de nuestros destinatarios. No deberíamos enfocarnos tanto en cómo querríamos que usen “nuestros” datos, sino en aportar a los destinatarios información que les sea relevante a ellos, en función de sus formas de trabajo y tipos de decisiones.

Por estas razones, una Unidad de evaluación debería conocer muy bien las lógicas de toma de decisiones de política educativa, en los distintos ámbitos y niveles del sistema educativo. Y conocer muy bien cómo es el mundo, la cultura y la lógica de trabajo de los docentes, si se pretende que los datos derivados de una evaluación les sirvan para algo. Es necesario además tener en cuenta que los seres humanos trabajamos y tomamos decisiones con la información de que disponemos, que nunca es completa. Es lo que se llama “racionalidad limitada”. El supuesto de que todos deberían poder entender datos estadísticos complejos y saber cómo usarlos, es poco realista.

A partir de lo anterior, es importante llamar la atención acerca de la complejidad que tiene modificar las prácticas o incidir sobre lo que hacen los docentes en el aula. A veces tenemos cierta ingenuidad, en cuanto a que es posible generar cambios muy rápidamente. La enseñanza es una práctica que se construye muy lentamente. A mí me gusta usar la imagen del vitral. Un curso es como un vitral que uno va armando de a poco a lo largo de los años: una piecita de acá, una piecita de ahí, esto no funcionó, lo corrijo, comparto cosas con colegas, agrego una lectura, desecho o modifico la consigna de una actividad... un curso es una obra de arte que uno no puede desmontar y volver a armar de un año para otro.

Por eso, si se quiere ayudar a los docentes a trabajar de manera diferente o a mirar el aula de un modo distinto, creo que necesariamente

se debe partir de lo que ellos ya están haciendo. Y empezar a pensar con ellos, “esto se podría hacer de otro modo”, con qué concepciones, con qué finalidades, etc. Es muy importante mirar lo que hacen los docentes. Pero en general sabemos muy poco de lo que acontece realmente dentro de las aulas. Hace un instante la última colega hacía énfasis en la importancia de entrar al aula y observar lo que allí sucede. En nuestros trabajos de investigación sobre evaluación no entramos al aula, pero descubrimos que es posible analizar las evaluaciones de los docentes. Fue como abrir una ventana que nos permitió mirar hacia dentro del aula. Porque la evaluación dice mucho acerca de qué está valorando el docente y qué le está comunicando a los estudiantes que es el aprendizaje.

A través de dos investigaciones en diez países de la región pudimos conformar una base de 8.000 registros fotográficos de propuestas de evaluación utilizadas por docentes, que son muy elocuentes acerca de lo que acontece en las aulas. Pongo aquí algunos ejemplos, simplemente para traer a mi exposición la realidad de las aulas, el tipo de tareas que los estudiantes realizan a diario.

Esta tarea es de un sexto de primaria. Traducir del lenguaje coloquial al lenguaje simbólico: *La cuarta parte del cuadrado de doce aumentada en el doble de dieciséis*. Y a la inversa, traducir del lenguaje simbólico al lenguaje coloquial.

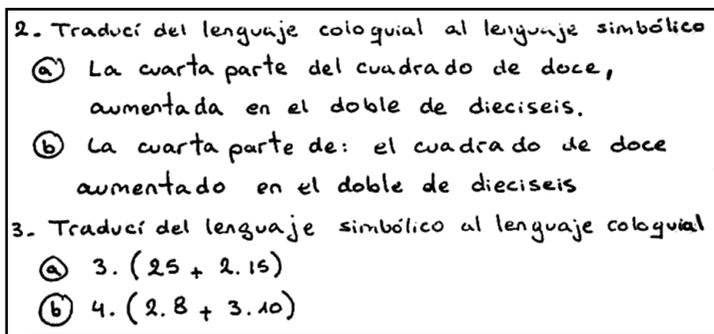


Imagen obtenida de la presentación de Pedro Ravela en la Jornada “Evaluación y enseñanza: contribuciones para la práctica educativa”, que se desarrolló el día 10 de agosto de 2018.

El que sigue es otro típico problema de sexto grado. *Un camino de 36 kilómetros, se asfaltó en tres etapas, en la primera se hizo un tercio, en la segunda un cuarto, qué parte falta asfaltar, cuánto va a costar el total de la obra.*

PROBLEMAS 6TO. GRADO

- 1) Un camino de 36 km se asfaltó en tres etapas, en la primera se concretó $\frac{1}{3}$ del total, en la segunda, $\frac{1}{4}$ del total.
 - a) ¿Qué parte del total falta asfaltar?
 - b) Si 10 km asfaltados cuestan \$ 1.000.000 (cuánto cuesta te 3ra. etapa?
 - c) ¿Cuál es el monto total de la obra?

Imagen obtenida de la presentación de Pedro Ravela en la Jornada “Evaluación y enseñanza: contribuciones para la práctica educativa”, que se desarrolló el día 10 de agosto de 2018.

Esta otra es muy divertida, uno de los típicos “problemas” que se plantean en las escuelas. *Laura comió un quinto de un chocolate y le dio al hermano la mitad de lo que quedaba, qué fracción comió el hermano de Laura.*

Leé atentamente y resolvé

1- Laura comió $\frac{5}{9}$ de un chocolate y le convidó a su hermano la mitad de lo que le quedaba.

a- ¿Qué fracción del chocolate comió el hermano de Laura?

b- ¿Qué fracción del chocolate comieron entre los dos?

2- Para repartir en partes iguales $3\frac{1}{2}$ litro de perfume en frasquitos de $\frac{1}{2}$ litro, de $\frac{1}{4}$ litro y de $\frac{1}{8}$ de litro: ¿Cuántos frasquitos se necesitan en cada caso?

Imagen obtenida de la presentación de Pedro Ravela en la Jornada “Evaluación y enseñanza: contribuciones para la práctica educativa”, que se desarrolló el día 10 de agosto de 2018.

Este tipo de propuestas son intentos, desde nuestra perspectiva fallidos, de vincular el contenido con situaciones cotidianas. Digo fallidos porque en la vida real nadie usa fracciones para repartir un chocolate.

En tercero de secundaria trabajamos con profesores de ciencias naturales. La mayoría de las tareas requieren básicamente recordar y reproducir. En este caso se le da a los estudiantes una figura del sistema nervioso y hay que identificar las partes del mismo.

2. Observa la imagen a continuación:

a) ¿Qué observas? ¿A qué nivel de organización pertenecen?

b) Identifica sus partes más importantes

1.....

2.....

3.....

4.....

5.....

6.....

c) Anota brevemente la función de las partes identificadas.

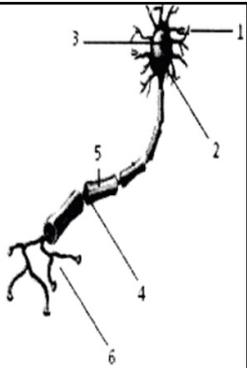


Imagen obtenida de la presentación de Pedro Ravela en la Jornada “Evaluación y enseñanza: contribuciones para la práctica educativa”, que se desarrolló el día 10 de agosto de 2018.

Esta que viene a continuación es muy ilustrativa (pido perdón si hay profesores de física en la sala). *Un gato persigue a un ratón en una trayectoria rectilínea. Empieza a perseguir al ratón desde el reposo y alcanza una velocidad de 8 m/s en cinco segundos.* Entonces hay que calcular la aceleración del gato, la velocidad del gato y la pregunta clave: *“Si el ratón estaba a tres metros del gato cuando lo empezó a perseguir a qué velocidad constante debió correr para que el gato no lo atrapara”.* La actividad es ingeniosa y entretenida, pero ni los gatos ni los ratones corren en línea recta ni a velocidad constante, ni la física clásica se usa para analizar este tipo de situaciones. Allí hay todo un tema sobre el que hay que pensar, los tipos de situaciones en que contextualizamos los contenidos escolares. La mayor parte de ellas son artificiales y mecánicas: tienen una única respuesta correcta y se resuelven de una sola manera.

- 2) Un gato se mueve persiguiendo a un ratón sobre una trayectoria rectilínea.
- El gato comienza a perseguir al ratón desde el reposo y alcanza una velocidad de 8,0 m/s en 5,0 s. Luego continúa moviéndose con velocidad constante durante 4,0 s más.
- A partir de ese instante comienza a moverse con una aceleración de -0,5 m/s/s, hasta detenerse.
- En estas condiciones determina:
- La aceleración del gato en los primeros 5,0 s
 - La velocidad del gato a los 12 s
 - ¿A qué distancia del punto de partida se detuvo el gato?
 - Si el ratón se encontraba a 3,0 m del gato cuando lo comenzó a perseguir. ¿Con que velocidad constante mínima debió correr para que el gato no lo atrapara en este movimiento?

Imagen obtenida de la presentación de Pedro Ravela en la Jornada “Evaluación y enseñanza: contribuciones para la práctica educativa”, que se desarrolló el día 10 de agosto de 2018.

Luego están este tipo de pruebas, que requieren recordar fórmulas o denominaciones de dudosa relevancia para la formación ciudadana, sin descalificar la necesidad o el interés de este tipo de tareas para

aprender química. Si notan el título, dice “Prueba Especial de Química”. Cuando leímos “Especial” pensábamos: “más de un estudiante se jugaba la vida en esta prueba”. Tal vez varios reprobaron el curso. Sin embargo, estoy seguro que en esta sala nadie que no sea de química puede responder esta tarea. Pensemos que una prueba como esta puede generar la reprobación de un curso de un estudiante y, eventualmente, el abandono de la enseñanza media, que sabemos que es un problema importante hoy en día.

PRUEBA ESPECIAL DE QUÍMICA: NOMENCLATURA

1- Formula:

- a) Trióxido de di-cloro
- b) Hidróxido férrico
- c) Óxido estánnico
- d) Anhídrido carbonoso
- e) Ácido clórico
- f) Óxido de cobre (II)

2 - Nombra los siguientes compuestos químicos:

- a) I_2O
- b) FeO
- c) $Cu(OH)_2$
- d) HBr
- e) H_2SO_4
- f) Li_2O

Imagen obtenida de la presentación de Pedro Ravela en la Jornada “Evaluación y enseñanza: contribuciones para la práctica educativa”, que se desarrolló el día 10 de agosto de 2018.

Para resumir lo que encontramos a partir del relevamiento de actividades, construimos cinco grandes categorías: 1) Actividades que simplemente requerían recordar y reproducir contenido; 2) actividades que requerían aplicar lo aprendido de algún modo pero en ejercicios mecánicos, con una sola respuesta posible y un solo camino de resolución; 3) actividades que requerían algún nivel de comprensión y de

explicación por parte del estudiante. Un ejemplo de esta categoría es una actividad muy simpática que decía que hay un charco de agua en el patio de hormigón. Después de una fuerte lluvia los chicos salen y quieren jugar al fútbol, pero está el charco. Entonces uno de los chicos dice pidamos un lampazo para extender el agua y que se seque más rápido. La consigna era *“estás de acuerdo con la propuesta, fundamentá tu respuesta utilizando los conceptos de cambio de estado, temperatura y calor que vimos en clase”*. Este planteo implica un buen aterrizaje del contenido a una situación cotidiana, en la que el estudiante tiene que poner en juego ciertos conocimientos aprendidos en ciencias, pero no de un modo memorístico. Tenía que haber logrado cierto nivel de comprensión para explicar esa situación. 4) Luego viene una categoría relativa a situaciones que requieren algún tipo de juicio de valor, toma de decisión o argumentación por parte del estudiante. Y por último, 5) aquellas que suponen la creación de modelos o propuesta de soluciones nuevas. En general, a grosso modo, en nuestra base de 8.000 registros la distribución fue 85/15. Es decir, 85% estaban en las dos primeras categorías -recordar, reproducir o aplicar mecánicamente sin entender- y apenas 15% implicaban comprender, valorar o crear.

Casi no encontramos situaciones auténticas, muy poquitas. Las situaciones auténticas no necesariamente deben corresponder a la vida cotidiana. Había algunas pruebas interesantes en situaciones sociales, algunas vinculadas a la salud reproductiva, a cuestiones de medio ambiente, etc. Y encontramos muy pocas situaciones auténticas en el sentido disciplinar, que reflejasen el tipo de preguntas o el modo en que trabaja un investigador o un profesional de la física, de la química, o de las distintas disciplinas.

Otro hallazgo importante fue la presencia muy fuerte de la práctica de la calificación como forma de devolución. Cuando el profesor quiere darle una pista al estudiante o le devuelve un trabajo, básicamente lo que hace es darle una nota. Lo vimos en las producciones que analizamos. No entramos a ver el aula pero hicimos un cuestionario a los estudiantes sobre cómo eran evaluados y básicamente el modo predominante de comunicar por parte de los docentes es, o bien marcas de correcto o incorrecto, o bien una nota directa, o puntos. Hubo muy

pocos casos de devoluciones que hicieran reflexionar o pensar al estudiante.

Los profesores tienen ahí un lío con la evaluación formativa. Algunos nos decían que la evaluación formativa es tener en cuenta todo: lo actitudinal, si viene limpio, si levanta la mano, si se porta bien, etc. Eso es evaluación formativa. Para otros la evaluación formativa es ponerle nota a todo lo que hacen los estudiantes. Significa que para la nota final no cuente solo la prueba, sino que tengan nota por oral, nota por presentar una tarea, nota por esto, nota por lo otro.

Luego está la cuestión de la homogeneidad, la dificultad para reconocer y asumir la diversidad dentro del aula. Las propuestas de enseñanza y de evaluación suelen ser las mismas para todos. Se supone que hay un alumno promedio y que todos deben “saber” todo.

También hay una confusión fuerte entre los denominados “procesos” y la cuestión del esfuerzo. Nos decían los profesores: *“Se valora el esfuerzo que hacen por llegar a un mínimo aceptable”*; *“El grado de cumplimiento de las tareas es evaluado más allá de si está bien o mal la tarea”*. *“Si intentó hacer algo, la información, si hizo una búsqueda, tiene que haber un trabajo continuo, el que hizo el trabajo de hormiga que se mantuvo todo el año, con altibajos pero todo el año trabajó, ese pasa”*.

Esta es una postura delicada: valorar el esfuerzo que no conduce a nada. Muchos docentes parecían creer que evaluar el proceso desde una actitud formativa tiene que ver con tener en cuenta el esfuerzo, pero no captaban la importancia que tiene dar buenas devoluciones para que el estudiante destrabe las dificultades que tiene. Creo que también tiene algo de bajar los brazos y decir “bueno, por lo menos se preocupó”.

Esta contextualización es muy breve pero nos permite traer al escenario qué es lo que pasa mayoritariamente dentro de las aulas. En este contexto es que quiero abordar la pregunta: ¿qué le puede aportar la evaluación estandarizada a los docentes?

Esto fue parte de lo que les preguntábamos a los docentes en las entrevistas. Hablábamos con ellos sobre cómo evaluaban, cómo calificaban, cómo hacían devoluciones, qué idea tenían de la evaluación formativa y, además, cuál era su visión sobre las evaluaciones externas,

si las usaban, si las habían mirado, si habían recibido informes, qué habían hecho, etc.

122

Algunas de las respuestas rápidas a la pregunta acerca de qué les interesa a los docentes de las evaluaciones externas son las siguientes. Primero, los datos no importan mucho. Es la verdad, la información estadística no es la que va a mover al docente a hacer otras cosas. Si bien puede ser útil conocerla por curiosidad o como información general, como cualquier otro ciudadano, para tener una idea general de qué es lo que está pasando con los aprendizajes a gran escala. Tampoco les preocupa o interesa mucho compararse, ver cómo están sus estudiantes en relación con otros. Salvo casos excepcionales de países como Chile, donde estaba muy metido todavía el tema de los rankings del Sistema de Medición de la Calidad de la Educación (Simce).

Entonces ¿qué es lo que les interesa a los maestros? Primero, conocer las pruebas. Conocer las actividades que se aplicaron, más que los datos. Les interesa ver ejemplos de actividades, ver qué y cómo se evaluó. Lo segundo, que es bastante obvio aunque no siempre es posible: les interesa saber cómo le fue a sus alumnos. Si mis alumnos no participaron en la prueba, los resultados me interesan poco.

Tercero y muy importante: las evaluaciones externas aportan una visión de cómo puede progresar el aprendizaje. Lo enfatizo a propósito porque sé que hoy presentaron un trabajo nuevo sobre progresiones de aprendizajes. La noción de progresión es útil para el docente, porque le ayuda a pensar en el largo plazo y, sobre todo, a tener una mirada de dónde está hoy el alumno y hacia dónde podría ir avanzando. Le ayuda a alinear el currículum. Esto es importante sobre todo en el Río de la Plata, donde tanto en Argentina como en Uruguay (voy a decir algo que no es políticamente correcto) tenemos una extraña noción de libertad de cátedra en la educación obligatoria. La libertad de cátedra es una noción de la enseñanza universitaria. En la educación obligatoria habría que discutirla. Porque está muy bien la libertad del docente, pero también y sobre todo, está el derecho del estudiante al aprendizaje. Digo esto porque en Uruguay muchas veces los profesores no quieren coordinar las evaluaciones con sus colegas, porque eso los “obliga a enseñar lo que enseña el otro”. Cada uno quiere enseñar lo que le pare-

ce. Hablamos mucho de trabajo colaborativo pero trabajamos en forma muy individualista y compartimentada.

Me desvió un momento sobre este aspecto. Acompañé una tesis de Maestría sobre cómo se evalúa y cómo se enseña a evaluar en la Formación Docente. Se hacían entrevistas a los docentes formadores. Uno de los temas era el trabajo en los departamentos - departamento de Matemática, de Lengua y otros-. Cuando la tesista les preguntaba, “¿ustedes qué grado de acuerdo tienen para las evaluaciones?”, una respuesta paradigmática fue “*nuestro único acuerdo es que estamos en desacuerdo*”. Entonces -pregunta-: “¿cómo hacen cuando tienen que tomar examen?”. Respuesta: “*cada uno pone una propuesta de acuerdo a lo que enseñó y después califica a sus estudiantes*” - aunque hay tribunales de tres miembros-. Tenemos una identidad profesional muy individualista. A pesar de que tenemos un discurso solidario, el trabajo colectivo es muy difícil, también por las condiciones de trabajo, porque no hay tiempo, porque el trabajo fuera del aula no está reconocido en la remuneración. Se supone que el docente trabaja cuando está en el aula. Pero en fin, este es otro tema.

Vuelvo a la pregunta que nos ocupa: ¿qué cosas aporta la evaluación externa a los docentes y a la evaluación en el aula? Otra respuesta es que aportan a un cambio de mirada. El docente, cuando piensa en la enseñanza, naturalmente piensa en los temas que tiene que dar. Cuando le comunica a los estudiantes qué será el curso o una unidad nueva que va a empezar, les habla de los temas que va a dar. Y cuando los estudiantes preguntan “¿qué va para la prueba?” o “¿cómo nos va a evaluar, profe?”, la respuesta son los temas que van para la prueba... “*esto va para el escrito*”, “esto no va para el escrito”... Siempre nos enfocamos en los temas. Nos cuesta mucho salir de ese lugar y expresar qué esperamos que sean capaces de hacer ellos, comunicar expectativas de aprendizaje. La evaluación externa de algún modo ayuda a empezar a hacer ese cambio de perspectiva, a poner el atención en los aprendizajes más que en los temas, algo que tiene también está vinculado a la concepción de la educación como derecho. Qué tienen derecho a aprender los estudiantes, en lugar de qué temas de mi disciplina me gusta enseñar.

Muchas veces, cuando hay buenas pruebas estandarizadas -esto depende de la calidad de las pruebas externas-, ayudan a los docentes a ver ejemplos de actividades que evalúan procesos cognitivos más complejos o situaciones auténticas. Por ejemplo, las pruebas de Ciencias de PISA son sumamente interesantes. Es notorio el contraste entre ciertas formas de enseñanza como, por ejemplo, estudiantes que aprenden de memoria “los pasos del método científico”, y las situaciones de PISA en las que entran en juego, por ejemplo, cuestiones relacionadas con el medio ambiente, en las que el estudiante tiene que desarrollar argumentos o construir hipótesis. O situaciones en las que tiene que resolver cuál sería un diseño experimental para probar una determinada hipótesis. Los profesores que tienen contacto con ese tipo de actividades de evaluación, en general, sacan muchas buenas ideas para sus propias clases. Esto no quiere decir que simplemente las reproduzcan, sino que sacan ideas que luego adaptan.

Sin embargo, existe un riesgo en países donde hay una presión muy fuerte con las pruebas externas, como son los casos de Chile y Colombia. Y por supuesto, en el caso de México, donde además había plata de por medio. En estos casos, además, las pruebas estandarizadas externas son de múltiple opción. En estos países muchas veces los docentes terminan adoptando ellos mismos las pruebas de múltiple opción como modo de evaluación dentro del aula. Pero cuando no existe esa presión, cuando el uso es menos amenazador, los docentes suelen sacar buenas ideas de las pruebas.

Mencioné antes la importancia que tienen los mapas de progreso y los niveles de desempeño. Otro aporte interesante de la evaluación externa a la evaluación en el aula se produce cuando el docente puede ver los resultados de sus propios estudiantes, porque puede contrastar el modo en que él está calificando el desempeño de cada estudiante con su nivel de desempeño en la prueba externa. Por lo menos para preguntarse cómo está evaluando, qué está teniendo en cuenta, etc.

Para enfatizar la importancia de las progresiones de aprendizaje, quiero detenerme ahora a analizar un problema importante. Mi visión es que en el fondo, lo que está en cuestión y lo que deberíamos cues-

tionar, es el modelo de escuela graduada que tenemos, y que viene del paradigma de la producción en serie en la sociedad industrial.

El supuesto central de la enseñanza graduada

125

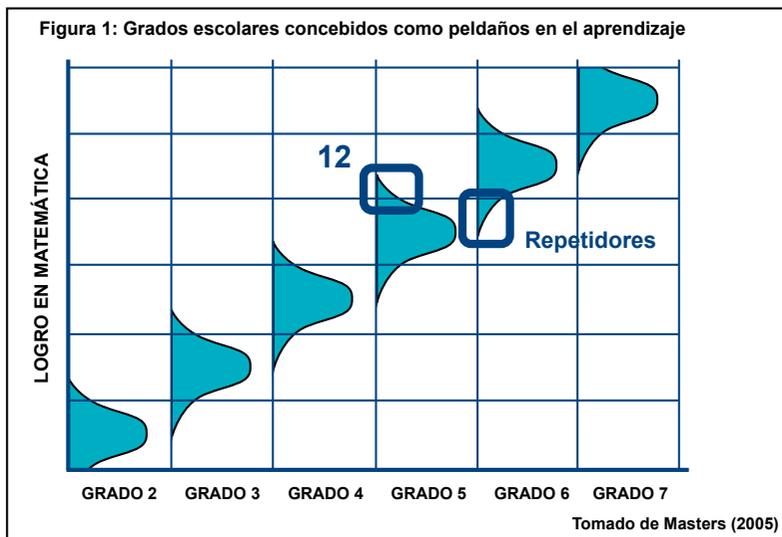


Imagen obtenida de la presentación de Pedro Ravela en la Jornada “Evaluación y enseñanza: contribuciones para la práctica educativa”, que se desarrolló el día 10 de agosto de 2018.

El supuesto del régimen académico es que los saberes se distribuyen del modo que está representado en esta gráfica. Los niños están agrupados por grado en función de su edad y en cada grado hay una pequeña curva normal de aprendizaje. La mayoría está en un determinado nivel promedio de conocimiento. Hay unos pocos de la cola de arriba de la distribución, que son aquellos a los que le ponemos la nota más alta. Y hay unos pocos en la cola de abajo, que son los que reprueban. Pero suponemos que la mayoría está en la zona media y que hay un escalón de un grado a otro.

Pero la realidad está representada en el segundo gráfico, que está basado en datos empíricos. Es de Geoff Masters, un australiano. Él

muestra que en realidad la distribución de los estudiantes no está escalonada por grados. Tenemos en todos los grados estudiantes que están en distintos niveles de desempeño. En 7º grado hay todavía algunos estudiantes que tienen un nivel de desempeño, en este caso en matemática, similar al de un alumno de 2º grado. Pero, sobre todo, que en cada curso tenemos una gran diversidad de estudiantes.

La distribución real de los logros educativos

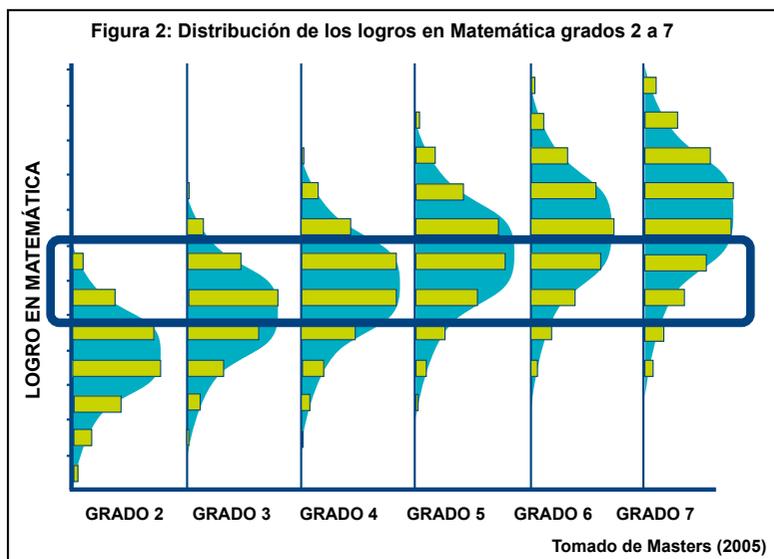


Imagen obtenida de la presentación de Pedro Ravela en la Jornada “Evaluación y enseñanza: contribuciones para la práctica educativa”, que se desarrolló el día 10 de agosto de 2018.

Entonces es necesario asumir la realidad de que uno va a tener diversidad de estudiantes en cada grado y que, por lo tanto, tiene que ver de qué modo diversifica su propuesta, tanto de enseñanza como de evaluación -en realidad deberíamos reemplazar la organización por grados y edades-.

En Uruguay hay un tipo muy interesante, ingeniero, se llama Juan Grompone, que tiene un librito titulado “El paradigma del laberinto”. Él

pone en cuestión la escuela graduada y la idea de que todos los estudiantes aprenden lo mismo en los mismos tiempos y van avanzando al mismo ritmo. Afirma que este mecanismo de producción en serie se va a transformar en un laberinto; que los jóvenes de la era de la informática van a recorrer, recorren el conocimiento, recorren los libros sin una secuencia establecida, ni siquiera cuando leen un libro de matemática. La manera como aprenden jóvenes es más parecida a lo que hacen en internet: navegan en un laberinto, lo van recorriendo en un orden aparentemente caótico -para uno que tiene otra cabeza armada de otro modo-, pero que tiene su lógica propia.

En este marco las progresiones ayudan a establecer un mapa. No deberían ser tanto un itinerario a seguir escalón por escalón, sino una hoja de ruta, un mapa que me permite ver dónde está cada estudiante, asumiendo que cada uno va a ir avanzando por un recorrido diferente.

Creo además que esto de las progresiones contribuye a un tema crucial: la educación es un derecho y una obligación hasta los 18 años. Esta definición modifica todas las reglas de juego. La lógica con la que nació la secundaria, por lo menos en el Uruguay, es que era la enseñanza previa a la universidad. En Uruguay la secundaria comenzó siendo la “Sección preparatoria” de la Universidad de la República. Luego se escindió en 1935. Es decir, la secundaria en Uruguay no nació como la educación posterior a la primaria, sino que nació como la educación previa a la universidad y se desprendió de ahí. Pero nunca se sacó de encima el peso de que su misión principal era preparar a los jóvenes para la universidad. Sin embargo, en el presente en Uruguay no más del 30% de los jóvenes llega a la universidad. Apenas un 40% termina la secundaria superior, en parte por esa carga, por ese peso histórico, que ha hecho de la secundaria un mecanismo de selección terrible.

Las encuestas de jóvenes en Uruguay muestran un dato muy fuerte: cuando se pregunta a los jóvenes que abandonaron la educación media por qué lo hicieron, la razón principal tiene que ver con la falta de relevancia. El 75% abandona por razones de este tipo: “no me interesaba, no me servía para nada”, ni por razones de embarazo adolescente ni por necesidad económica, ni por trabajo; sino por la percepción de

irrelevancia, de no entender de qué iba ese mundo de contenidos abstractos vinculados a “asignaturas”.

Si queremos una educación inclusiva, obligación y derecho hasta los 18 años, deberemos aceptar mucho más diversidad y dar más opcionalidad a los estudiantes. No necesariamente todos tienen que ser buenos en todo. Si un adolescente es muy bueno en letras y quiere hacer muchas letras y poca matemática ¡vamos arriba! Claro que algo de matemática te voy a enseñar, pero te sigo la corriente, te impulso a hacer algo que te entusiasme.

Hace poco me tocó un caso en una institución. Una chica con dificultades y malas notas, a la que le encantaba la fotografía. Pero le iba mal, tenía las notas bajas en el liceo. ¿Entonces los papás qué hicieron? Le prohibieron la fotografía hasta que subiera las notas ¿Cómo le van a sacar lo único que le interesa? ¡No!

Entonces, yo creo que es interesante la combinación de un mapa de progreso con un laberinto. Estas dos imágenes marcan un poco el sentido de un nuevo formato escolar. Esto no quiere decir que los estudiantes vayan a seguir secuencialmente una progresión. Y tampoco que cada uno va a estudiar solo una materia que le interesa. Tenemos que pensar en términos dialécticos entre estos dos polos. Un mapa de progreso nos permite saber por dónde anda cada estudiante, dejar que cada uno vaya haciendo su recorrido, su itinerario, y brindarle apoyo y acompañamiento apropiado.

Ahora quiero advertir sobre algunas dificultades o riesgos de cara a estos temas. La primera es trabajar excesivamente con ítemes de opción múltiple. Creo que esto es un problema para el diálogo con la evaluación en el aula y con las prácticas de enseñanza de los docentes. Yo creo que hoy, con el desarrollo tecnológico e informático, hay opciones técnicas para trabajar con formatos de prueba más abiertos. Con toda la experiencia que hay atrás de PISA, con toda la acumulación en el trabajo con rúbricas, hay mucho espacio para desarrollar pruebas abiertas, que puedan corregir y codificar los propios docentes.

Claro, aquí hay un dilema. Los especialistas en medición prefieren los ítemes de opción múltiples porque permiten mayor precisión. Al

respecto hay un autor norteamericano -no recuerdo el nombre- que dice que *“sacrificamos la validez en nombre de la confiabilidad”*.

¿Qué quiere decir esto? Que nos preocupa tanto la precisión de la medida psicométrica, que nos atamos a la prueba de múltiple opción, aunque perdamos la validez de uso de las pruebas y sus resultados por parte de los docentes. Volvemos al planteamiento del comienzo: qué esperamos de la evaluación estandarizada. Si mi finalidad es tener un buen sistema de información estadística, el primer modelo que planteé al principio, no hay problema. Pero sabemos que eso no nos servirá para dialogar con los docentes sobre cómo mejorar la enseñanza.

Ahora, si lo que quiero es justamente dialogar con los docentes y que la evaluación externa tenga valor para el trabajo en el aula, entonces probablemente tenga que sacrificar algo de confiabilidad y precisión en las mediciones, acercar más las pruebas al tipo de trabajo en el aula, involucrar a los docentes en la construcción y/o en la corrección de las actividades. Es muy fácil hoy subir las pruebas realizadas por los estudiantes a una plataforma, tenerlas en línea y hacerlas codificar por docentes con un buen manual. Y, a partir de pruebas de respuesta abierta, obtener un panorama de datos agregados de calidad razonable. En fin, creo que en las evaluaciones estandarizadas de la región tenemos todavía un exceso de actividades de opción múltiple y que no tienen demasiado contexto. Mi visión es que hay que aprender de PISA: siempre trabaja con un “ramillete” de preguntas que se derivan de una situación madre, que suele ser una situación relevante en términos sociales y de ciudadanía.

Una segunda dificultad o riesgo importante es el tema de los tiempos, de la oportunidad en la devolución de resultados en las pruebas nacionales. En Perú me habían pedido que revisara el informe que enviaban a las escuelas. Las pruebas eran confidenciales y hacían un informe para cada maestro, con los resultados de sus alumnos, en una prueba que era censal. La aplicaban al final de segundo de primaria. Entonces yo les pregunto, ¿y ustedes, cuándo mandan esos informes a las escuelas? Respuesta: *“en mayo más o menos”*. ¿Y a quién se los mandan? *“Al maestro de segundo”*. Pero los niños no están más con el maestro de segundo, están con el de tercero. El maestro de segundo ya

no puede hacer nada con los resultados. Este tipo de detalles a veces se escapan y son esenciales para una estrategia de uso de los resultados.

Para finalizar, quiero presentar cuatro pistas de experiencias interesantes de relación entre la evaluación externa y el trabajo en el aula.

130

La primera es lo que está haciendo Uruguay con la “evaluación en línea”. A partir de que cada niño tiene una computadora portátil (XO) personal, la Administración de la Nacional de Educación Pública lo que hace es colocar las pruebas estandarizadas en línea. Hay una ventana de tiempo de varias semanas, y cada maestro decide qué día quiere aplicarlas. Ese día los niños se conectan y responden en sus laptops. Esta aplicación es autónoma, o sea, no es controlada externamente, porque la evaluación no tiene como finalidad dar datos estadísticos precisos. El propósito es que sirva para fines pedagógicos, por eso se la denomina evaluación “formativa”. Se trata más bien de que sea un ejercicio de revisión por parte de maestros y alumnos.

Centro	Escuela N°																									
Grupo	2º B																									
Departamento	MONTEVIDEO																									
Asignatura:	Matemáticas																									
Prueba:	2º Matemáticas																									
Contenido:	Todos																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
RIVERA LÓPEZ	✓	✗	✓	✓	✓	✓	✓	✓	✓	✗	✗	✗	✗	✗	✗	✓	✓	✗	✗	✗	✗	✓	✓	✗	✗	13
LO ALFARO	✓	✗	✓	✓	✓	✓	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✗	✗	14
JISSPAUMER AGUI	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	10
ELLA GROSSO MELO	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	13
USON BURGUEZ	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
NA CORTI RODRIGUEZ	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	11
IGUEZ ABREU	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	15
I FAGOAGA	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
ELHO DE LOS S	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	17
NITA ARAUJO FERNÁNDEZ	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	13
COV SCLAVI IRIGOYEN	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	12
RA ITORBURO VIEYTE	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	8
JÉN VINCON CARTIER	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	14
JINEA MARTÍN	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
JALDI PICHEL	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	15
Z PEDEMONTÉ	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	8
JUEL MOURO SOSA	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	21
O LONGO	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	10
JNELLO VIERA	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	19
JALAVERRY	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	10
JTONI GONZÁLEZ	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	14
JSTI MONGELOS TECCO	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	9
JGAS TIZITZIOS CAMAÑO	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
JDÍAS PÉREZ FERRARI	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	10
	21	8	17	7	13	17	8	17	12	9	14	13	1	7	9	16	16	11	13	10	8	4	11	8	12	

Imagen obtenida de la presentación de Pedro Ravela en la Jornada “Evaluación y enseñanza: contribuciones para la práctica educativa”, que se desarrolló el día 10 de agosto de 2018.

Lo primero que ve el maestro, apenas terminan de hacer la prueba sus alumnos, es una tabla como esta. Es una grilla donde tiene todos sus estudiantes y todas las preguntas. Allí puede ver, estudiante por estudiante, quién respondió correctamente y quién no. Puede ver el total de respuestas correctas de cada estudiante y el total de estudiantes que respondieron correctamente cada pregunta. En ese mismo momento, enseguida de haber realizado la prueba o al día siguiente, puede discutir con sus alumnos ¿por qué respondieron? ¿Les parece que es correcto? ¿Por qué?

Además de lo anterior, y esto es muy interesante, el maestro puede ver cómo se distribuyen sus alumnos entre las distintas opciones de respuesta, qué opción seleccionó cada uno. Por ejemplo, la opción A la marcaron Camila, Antonella, Florencia, Jerónimo y Santiago. Y la

respuesta correcta, que era la C, Agustina, Juan, Maximiliano, Darío, Octavio, Joaquín y César.

132

Ejercicio 17 *Comprar en grupo*



Ana compra un cuarto kilogramo de jamón cocido de la oferta

Paga
 A) \$ 2
 B) \$ 16
 C) \$ 20
 D) \$ 32

Respuestas

Opción	Cantidad	Alumnos
A	5	CAMILA LASSERE, ANTONELLA TAFANI, FLORENCIA VIOLA, JERONIMO MORA, SANTIAGO TORRES,
B	2	Alumnos. LAURA FIGARES, IGNACIO ACOSTA,
C	7	AGUSTINA DEFAUT, JUAN BRUSCO, MAXIMILIANO CRAVINO, DARIO DATHAGUY, OCTAVIO LARROSA, JOAQUIN MARTINEZ, CESAR MARTINEZ,
D	7	AMALIA CONDON, CECILIA TERRA, SEBASTIAN DE LOS SANTOS, PABLO GONZALEZ, EMILIANO MORELLI, JUAN PIA, IGNACIO ROMERO,

Imagen obtenida de la presentación de Pedro Ravela en la Jornada “Evaluación y enseñanza: contribuciones para la práctica educativa”, que se desarrolló el día 10 de agosto de 2018.

Entonces el maestro puede mirar inmediatamente cómo respondieron los alumnos. Obviamente, detrás de cada distractor hay una hipótesis de error y demás. Entonces el maestro tiene ahí un material valioso para trabajar en el momento con los estudiantes.

Esta creo que es una linda experiencia. La crítica que yo hago siempre -ahora que lo miro de afuera- es por qué no usar más preguntas abiertas. Y el argumento es “bueno, pero se perdería la inmediatez de los resultados”. Pero yo creo que se ganaría mucho más si el maestro tiene que codificar las respuestas de sus alumnos con un buen manual e introducir esa codificación en la plataforma. Los involucrará más y te podrías acercar mucho más preguntas y actividades propias del aula.

Además tienen un banco de ítems donde el maestro puede elaborar una prueba por sí mismo cuando quiera, seleccionando ítems. Puede decir “*quiero armar una prueba sobre fracciones*”, elegir ítems y armarla... tiene sus pro y sus contras.

Una segunda experiencia, que si bien es complicada y costosa, tiene nociones muy interesantes, viene de la mano de lo que se hace en un país del primer mundo, Australia. Ellos hacen una evaluación censal

todos los años en tercero, quinto, séptimo y noveno grado. Todos los años.

Estas pruebas las aplican al principio del año, no al final. Cada docente recibe los datos de sus estudiantes en los primeros meses del año. Nosotros tenemos naturalizado que si es evaluación externa -y bueno, muchas veces con la evaluación en el aula también-, que la evaluación siempre se hace al final, después de la enseñanza. Como en Australia tienen como la finalidad que la información le sea útil al docente, la evaluación se hace a principio de cada año y los docentes reciben una suerte de mapa con la situación de sus estudiantes en los primeros meses.

Para organizar la información tienen una progresión de diez niveles (que denominan bandas) que abarcan esos grados de la escolaridad (de 3º a 9º). Vean la imagen. Aplican pruebas en Lectura, Numeración, Escritura persuasiva y Manejo de la gramática o de las convenciones del lenguaje. Para cada grado y para cada prueba, señalan cuál sería el nivel o banda esperable para cada grado escolar. En este caso, el que tienen el grado tres, la banda 2 dice “Estudiantes en el estándar mínimo nacional”, o sea, se espera que los estudiantes estén como mínimo allí, estudiantes por encima, estudiantes por debajo.

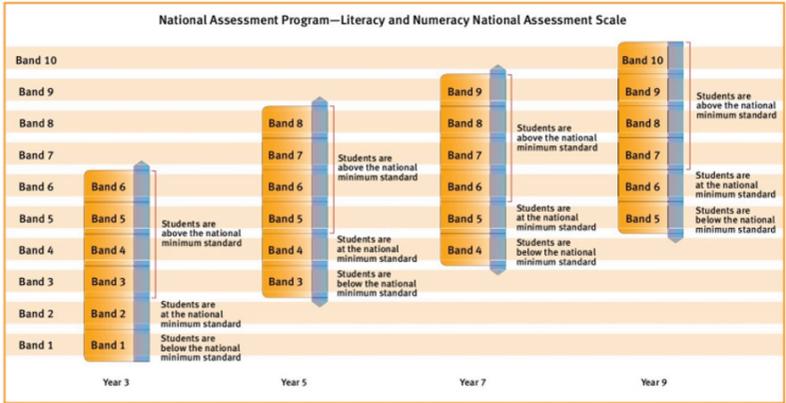
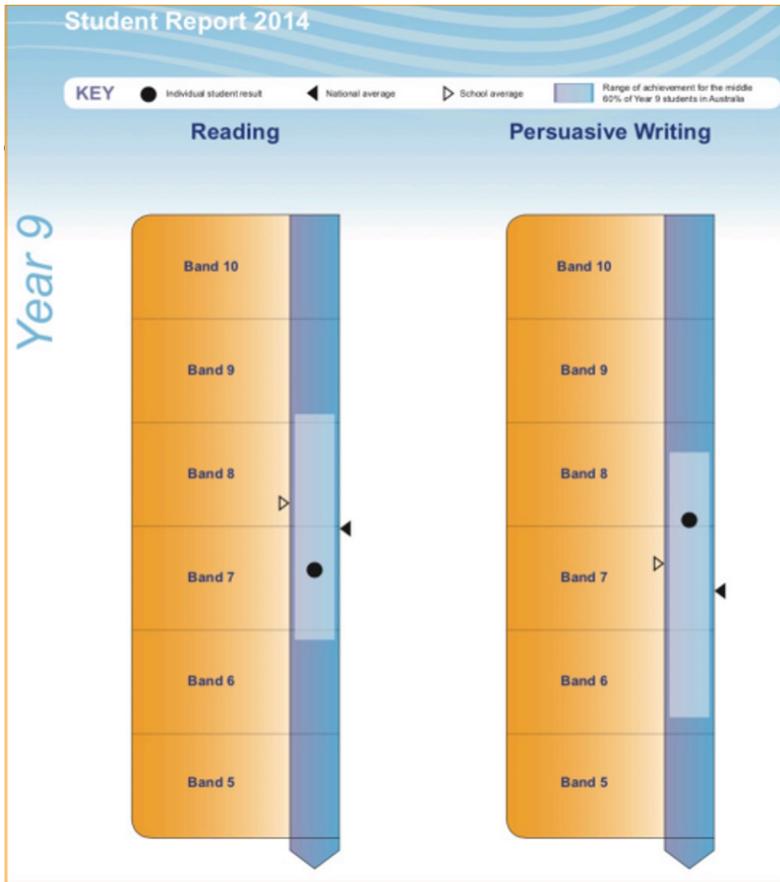


Imagen obtenida de la presentación de Pedro Ravela en la Jornada “Evaluación y enseñanza: contribuciones para la práctica educativa”, que se desarrolló el día 10 de agosto de 2018.

Eso lo traducen luego en lo que se ve en la siguiente figura, un reporte sobre cada estudiante. Si miran arriba dice Lectura y Escritura persuasiva. Utilizan estos símbolos que aparecen allí, un círculo negro y unos triangulitos que quieren lo siguiente: el círculo negro indica dónde está un estudiante en concreto. El triángulo negro señala el promedio nacional de los estudiantes de esta edad. El triángulo vacío muestra el promedio del colegio, de la escuela. El recuadro celeste clarito indica dónde se ubica el 60% de los estudiantes en el centro de la distribución, es decir, sacando las colas del 20% de peor desempeño y del 20% superior.



135

Imagen obtenida de la presentación de Pedro Ravela en la Jornada “Evaluación y enseñanza: contribuciones para la práctica educativa”, que se desarrolló el día 10 de agosto de 2018.

Luego, esto se transforma en un sistema de reporte a las familias, que incluye una descripción de qué es lo que es capaz de hacer el alumno en cada uno de los niveles o bandas, o sea, qué es capaz de hacer un estudiante en el nivel 7, en Lectura y en escritura, por ejemplo. Estos reportes le permiten a las familias ver dónde está el alumno, cómo está en relación al grado, y qué significa el hecho de estar en una u otra

banda. Esto es bien interesante como forma de establecer un puente con los sistemas de calificaciones, para que tengan un significado más claro.

Lo otro interesante del caso de Australia es que el docente puede ver cómo están los alumnos que tiene hoy, alumno por alumno. En el siguiente gráfico, cada línea es un alumno, un estudiante. Y lo que muestra es el grado de avance de cada estudiante respecto a dos años antes, cuánto avanzó en Lectura, en Escritura, con respecto a dos años antes.

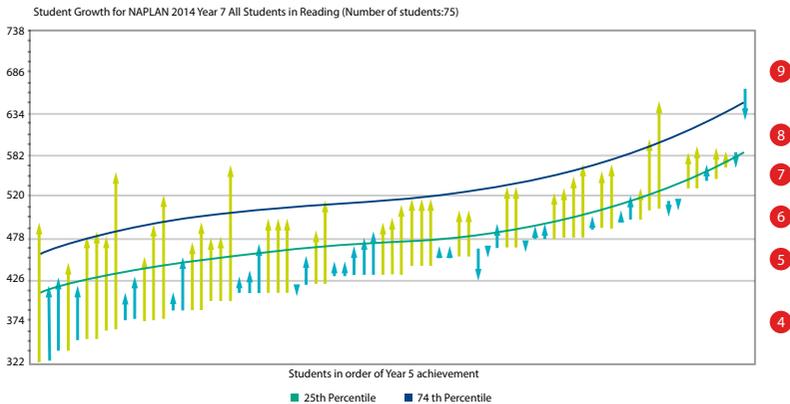


Imagen obtenida de la presentación de Pedro Ravela en la Jornada “Evaluación y enseñanza: contribuciones para la práctica educativa”, que se desarrolló el día 10 de agosto de 2018.

Imagino que un sistema de este tipo debe ser muy costoso.

Otro dato interesante es que la Agencia de evaluación a gran escala está dentro de lo que en Australia es la autoridad nacional de currículum. O sea, la evaluación estandarizada está integrada con el área de currículum, algo que en nuestros países suele ser bastante difícil de hacer.

Paso al tercer ejemplo. Los chilenos, que se han ido alejando de los rankings y demás, siguen haciendo sus evaluaciones nacionales. Una de las líneas que empezaron a desarrollar en los últimos años es una mo-

dalidad que se llama evaluación progresiva, que es voluntaria. O sea, las escuelas que quieren se inscriben, no están obligadas. A esas escuelas les ofrecen hacer evaluación y darle los datos de sus estudiantes, en tres momentos del año: al inicio, a la mitad, al final. La Agencia de Calidad de Chile les da un reporte inmediato sobre los resultados por curso y por estudiante, y orientaciones de trabajo a partir de los resultados. Les ayuda a ver la progresión de los estudiantes durante el año.

Además, la Agencia de Calidad construyó un portal sobre evaluación formativa, donde hay herramientas de evaluación para los docentes, espacios de intercambio de experiencias, etc.

Por último, la cuarta experiencia. En Estados Unidos hay varias movidas, en distintos lugares, para salirse del exceso de pruebas estandarizadas y empezar a avanzar en otras direcciones. Conley hace una reflexión interesante sobre cuál es el tipo de habilidades que necesita un estudiante para tener un buen desempeño en la vida universitaria, más allá de los contenidos. Él sostiene que

“necesita saber formular problemas e hipótesis; tener capacidad de identificar sintetizar información para investigar; tener capacidad de construir, organizar, comunicar información; tener capacidad de crear estrategias o diseños para resolver situaciones; y para analizar y evaluar perspectivas diversas. Obviamente necesita conocer los conceptos y los términos clave de una disciplina, y poder organizarlos y relacionarlos, así como un conjunto de actitudes hacia el aprendizaje.”

Partiendo de estas definiciones, le pidieron a los docentes que elaborasen actividades de evaluación (las escuelas se adhieren a esto voluntariamente): un análisis, un ensayo analítico literario, un paper en ciencias sociales, un experimento en ciencias naturales, problemas complejos en matemática. Los docentes elaboran las actividades y, además, las rúbricas para la corrección. Esto después es validado y se revisan las rúbricas. Luego, con los trabajos que van haciendo los estudiantes en distintos momentos del año, empiezan a montar un banco de “exhibiciones de desempeño” que está disponible en línea para otros profesores. Al mismo tiempo, tienen la visión de futuro de que están ar-

mando una especie de portafolio digital para cada estudiante, que luego le permitirá mostrar el tipo de proyectos en los que participó, el tipo de evaluaciones complejas que realizó, para que eso le sirva como una especie de currículum. En vez de tener una lista de números -cinco en matemática, siete en física, dos en literatura, promedio tanto-, el “certificado académico” del estudiante no será un conjunto de notas sino un perfil de tipos de tareas y proyectos. Por otra parte cada estudiantes tiene la posibilidad de elegir, cursos y evaluaciones. En Estados Unidos hay mucha opcionalidad y el estudiante va eligiendo cursos. Tiene ciertos cursos obligatorios, pero puede elegir hacer mucha matemática, o meterse mucho en otra área.

Reflexiones finales para cerrar. El que las evaluaciones externas estandarizadas puedan tener algún diálogo con la evaluación en el aula requiere de hacerse cargo, en algún momento, de cuestiones que tienen que ver con el trabajo docente. Un problema principal de la educación en América Latina, en mi opinión, es que concebimos el trabajo docente como las horas que está en el aula. La propia OCDE, en uno de los tantos informes que tiene, denominado TALIS, que es un estudio sobre profesión docente, muestra que en los países de la OCDE la norma es el contrato docente de 40 horas semanales en una institución. De esas 40 horas, entre un 50% y un 60% son horas de aula. El resto son horas que el docente está en la institución pero no está dando clase; está haciendo el resto del trabajo que tiene que hacer un docente para poder hacer buenas clases.

Creo que este es un tema clave, así como abrir el aula es un tema clave. El aula debe dejar de ser un espacio sagrado que se abre sólo cuando viene el inspector, porque no hay más remedio. Hay que romper la cultura de cada docente hace lo que le parece. Necesitamos empezar a pensar la profesión y la práctica en el aula como algo abierto a la mirada de otros, creo que este es un paso clave. Nosotros lo que estamos trabajando es una modalidad intermedia, tomando como puerta de entrada compartir y analizar las evaluaciones. Este un modo de analizar las prácticas colectivamente, sin el costo y la dificultad de organizar visitas a las aulas. El hecho de que cada docente abra sus evaluaciones

a los colegas y reciba sugerencias e ideas de los otros, ayuda mucho. Es un modo indirecto de abrir el aula.

Necesitamos pensar la profesión y la actividad docente como una cuestión más colectiva. Para eso se necesita liderazgo. De lo contrario puede no funcionar. En Uruguay se han creado espacios remunerados para reuniones de trabajo fuera del aula, espacios de coordinación y horas de departamento. Pero no se están usando para mejorar la práctica. En general son más bien sesiones de catarsis sobre todos los problemas que tenemos o para hablar de los alumnos. Pero no para revisar lo que estamos haciendo en el aula nosotros mismos.

Es difícil hacer fructíferos estos espacios si no hay alguien que conduzca, yo creo que es una necesidad que tenemos. En primaria en general el director, si es bueno, lo puede hacer. En secundaria se necesita una figura de orientador pedagógico. Hay algunas instituciones que los tienen. Yo acompaño a una que está trabajando con lo que ellos llaman “referentes académicos”, que son profesores que siguen dando clases en el aula, pero tienen diez horas a la semana liberadas que usan para acompañar a los colegas. Se sientan en grupo a planificar secuencias didácticas, organizan sesiones en las que van tres o cuatro profesores a ver una clase, después se sientan a discutirla, etc.

¿Qué quiero decir con todo esto? Que hacer una evaluación estandarizada y luego convocar a los docentes a una jornada para intentar explicarles toda la estadística que hay detrás de los datos, no es el mejor modo de entablar el diálogo entre la evaluación externa y la enseñanza. Más bien lo que se necesita es crear espacios para que la evaluación externa dialogue y tenga vínculo con lo que los docentes están haciendo dentro del aula.

