

CAMILLONI, Alicia R. W. et alli.: **La evaluación de los aprendizajes en el debate didáctico contemporáneo**. Buenos Aires, Paidós, 1998, 176 p.

*Luisa L. Lazzari**
*Pablo S. García***

El libro reúne cinco trabajos de diferentes autoras. En el primero, "La evaluación: campo de controversias y paradojas o un nuevo lugar para la buena enseñanza" (pp. 11 a 33), de Edith Litwin, se sostiene que la evaluación es parte del proceso didáctico e implica para los estudiantes una toma de conciencia de los aprendizajes adquiridos y, para los docentes, una interpretación de las implicancias de la enseñanza en esos aprendizajes. Evaluar el almacenamiento de información, esto es, la memoria a largo o corto plazo, en situaciones en donde el alumno fundamentalmente recuerda hechos y datos, ha sido una práctica constante en los diferentes niveles del sistema educativo. Desde una perspectiva cognitiva, la autora propone realizar actividades que cambien el lugar de la evaluación como reproducción de conocimientos por una evaluación como producción. Una buena evaluación, sostiene, requiere la formulación y explicitación de antemano de los criterios que se utilizarán para dar cuenta del nivel de la producción: reproducción de información (obtenida en clase o a partir de lecturas), resolución original de problemas, creatividad u originalidad en la respuesta, reconocimiento de niveles diferentes de análisis en lo que respecta a la profundización temática, etc.

El segundo trabajo, "¿Es posible mejorar la evaluación y transformarla en herramienta de conocimiento?" (pp. 35 a 66), de Susana Celman, sostiene que es posible transformar la evaluación en una herramienta de conocimiento, tanto para los profesores como para los alumnos. Según este criterio, se intenta sacar a la evaluación del lugar en el que comúnmente se la ubica: como un acto final desprendido de las acciones propias de la enseñanza y el aprendizaje. Las actividades evaluativas, en un contexto de educación en que se respete la constitución del sujeto de la enseñanza y del aprendizaje como un sujeto capaz de tomar decisiones fundadas, se constituyen y entrelazan en el interior mismo del proceso total. Para que esto sea posible se debe concebir la tarea educativa como una propuesta que se pone a consideración de sus actores, quienes la ejercen con autonomía responsable y transformadora. Las pruebas de evaluación de los aprendizajes serán valiosas si nos permiten conocer la manera y el grado

* Universidad de Buenos Aires, Facultad de Ciencias Económicas.

** Universidad de Buenos Aires, Facultad de Ciencias Económicas.

de apropiación que los estudiantes han hecho de un conocimiento que se considera importante y digno de ser conocido. No se debe considerar que sólo son valiosos los exámenes que pretendan evaluar ciertas temáticas, potencialmente valiosas, sino que tal cualidad depende también del tipo de conocimiento que hayan promovido, y de la calidad del sistema de evaluación para ponerlo de manifiesto. Al diseñar las actividades de evaluación tendientes a que los estudiantes pongan en juego procesos cognitivos (indagación, problematización, diferenciación, generalización, inducción y deducción de principios, aplicación y creación de procedimientos, resolución de problemas), también se propiciará que generen nuevos aprendizajes, como resultado de las nuevas relaciones conceptuales desencadenadas por esta situación.

En el tercer trabajo, "La calidad de los programas de evaluación y los instrumentos que los integran" pp. (67 a 92), Alicia Camilloni observa que existe una variedad de modelos de instrumentos de evaluación, pero cada uno tomado aisladamente resulta insuficiente para obtener información sobre el aprendizaje de los alumnos. De manera que el diseño de un programa de evaluación exige la combinación de diferentes instrumentos para obtener una cobertura adecuada: la eficacia de la evaluación dependerá de la pertinencia de tal combinación, de la oportunidad en que se emplea y de la interpretación de sus resultados.

En general, las características que los instrumentos de evaluación deben reunir son cuatro: validez, confiabilidad, practicidad y utilidad.

Un instrumento de evaluación es *válido* cuando evalúa lo que se intenta evaluar con él, esto es, cuando permite apreciar los logros del aprendizaje que se han determinado como aquellos que se deben evaluar. Sin embargo, la validez de un instrumento no puede determinarse de un modo absoluto sino, por el contrario, de un modo relativo a su adecuación o inadecuación respecto de los propósitos con que se lo aplica y la situación específica de su aplicación. Por otra parte, hay diversas clases de validez: (i) *validez de contenido*: un instrumento tiene validez de contenido cuando representa una muestra significativa del universo de contenido cubierto por el curso; (ii) *validez predictiva*: se refiere a que los instrumentos de evaluación deben contribuir a pronosticar éxitos y dificultades de los alumnos en los aprendizajes posteriores y en sus desempeños ocupacionales o su actuación social en general; (iii) *validez de construcción*: los instrumentos que integran el programa de evaluación deben estar contruidos según los principios de las teorías didácticas que sostienen el proyecto pedagógico; (iv) *validez de convergencia*: necesaria en los casos en que un programa o una prueba de evaluación es reemplazada por otros, en cuyo caso es preciso demostrar que los nuevos instrumentos son tan buenos o mejores que los anteriores; (v) *validez manifiesta*: se refiere a que los instrumentos de evaluación deben ser percibidos por los alumnos como instancias adecuadas

para la evaluación, no como un obstáculo para producir una selección más o menos azarosa e injusta entre los alumnos; (vi) *validez de significado*: se refiere al hecho de que las tareas de evaluación tengan significado para los alumnos y los motiven para lograr un mejor rendimiento; (vii) *validez de retroacción*: es la que impone la realización de acciones de corrección con la finalidad de reestablecer el verdadero sentido didáctico de los procesos de evaluación en caso de ser necesario.

Respecto de la *confiabilidad*, decimos que un instrumento de evaluación es confiable cuando presenta *exactitud* en la medición y *sensibilidad* para apreciar los rasgos que mide y las diferencias de magnitud en los mismos. La *confiabilidad*, al igual que la validez, no es absoluta sino relativa al uso que se quiera dar al instrumento. Los instrumentos de alta confiabilidad son aptos para evaluar cierto tipo de rasgos, pero no todos, y se caracterizan por eliminar el peso de la subjetividad. Pero en las así llamadas “pruebas subjetivas” (pruebas escritas de ensayo o de respuesta extensa) la validez aumenta a expensas de la confiabilidad, ya que están destinadas a evaluar la producción del alumno con libertad para escoger información, organizarla, expresarla y presentarla. Así, mientras que la confiabilidad es relativa a los instrumentos, la validez es relativa a los objetivos para los que se está recogiendo información. Una buena prueba debe ser válida y lo más confiable posible.

Respecto de la *practicidad*, depende de la confluencia de tres aspectos: (i) administrabilidad, (ii) facilidad de análisis e interpretación de los resultados, y (iii) economía de tiempo, esfuerzo y costo de su utilización.

Finalmente, la *utilidad* resulta de la capacidad de una prueba para satisfacer las necesidades específicas de los procesos de enseñanza y aprendizaje: los resultados de la evaluación deben ser útiles para orientar a los alumnos, a los docentes, a la institución y a los sectores interesados en la calidad de la educación.

En el cuarto artículo, “La evaluación de las prácticas docentes y la autoevaluación” (pp. 93 a 132), María del Carmen Polou de Maté aborda la tarea de dar cuenta de los diversos modos que asume la evaluación en el ámbito de la didáctica, advirtiendo que su conformación actual responde, por una parte, a la construcción histórica del concepto mismo, y por otra, a los entrecruzamientos contextuales que la afectan. De manera que es posible analizar la evaluación tanto desde una dimensión política como desde una perspectiva psicológica y desde un marco sociológico: tales dimensiones se entrecruzan en la práctica y determinan una configuración particular de la evaluación para cada caso.

Por otra parte, es posible tomar la tarea del docente como un proyecto dentro del cual la evaluación le permite analizar su propuesta didáctica en relación con los demás proyectos que constituyen su contexto. Desde esta perspectiva, la autoevaluación de la enseñanza ofrece genuinas posibilidades de revisar las

prácticas docentes, y al mismo tiempo, la autoevaluación del alumno se convierte en un modo de trabajo que facilita su vinculación con el conocimiento.

Finalmente, en el artículo "Sistemas de calificación y regímenes de promoción" (pp. 133 a 176), Alicia Camilloni observa que un problema al que se ve inevitablemente enfrentada la evaluación es el de los sistemas de calificación y regímenes de promoción. Estos constituyen un componente esencial de los sistemas de evaluación que se emplean en la evaluación formal. Aunque ambas, calificación y promoción, son producto de un conjunto de decisiones prácticas adoptadas en virtud de la adhesión a una u otra de las diversas corrientes teóricas de la didáctica, siempre ejercen, de hecho, un fuerte efecto normativo sobre las acciones de enseñanza y de aprendizaje, y suelen operar como determinantes de todas las prácticas pedagógicas.

Las escalas de medición pueden clasificarse en cuatro grupos: (i) *escalas nominales*: las categorías que las componen no se hallan en un orden de seriación que implique que un grado es mayor o mejor que el o los otros grados, de manera que todos los grados se encuentran al mismo nivel, y se emplean para medir la frecuencia con que los ejemplares evaluados se incluyen en cada una de las clases que la escala presenta; (ii) *escalas ordinales*: son aquellas que presentan una serie de categorías ordenadas de acuerdo con un criterio definido por un sentido u orientación determinada que no permite introducir alteraciones en la frecuencia; no son escalas universales, son específicas y se usan para medir aspectos determinados, como por ejemplo, para evaluar la participación en clase de los alumnos (participa muy activamente, activamente, poco, muy poco, nada); (iii) *escalas de intervalos*: cumplen los mismos requisitos que las ordinales, pero los intervalos entre pares de grados consecutivos deben ser iguales, lo cual permite operar con las mediciones obtenidas, como por ejemplo, hallar el promedio (media aritmética) de las mediciones; (iv) *escalas de razones o proporciones*: cumplen con todas las características de las escalas de intervalos y además poseen un cero absoluto; esto hace posible efectuar operaciones en las que se suman mediciones obtenidas independientemente o se comparan mediciones en las que, por ejemplo, el grado "2" de la escala se puede considerar equivalente a la mitad del grado "4".

Al construir una escala de calificación debe tenerse en cuenta que el número de grados debe ser elegido de acuerdo con la posibilidad real que tiene el evaluador de operar con una técnica apropiada, discriminando los diferentes grados de calidad de aquello que ha de ser evaluado con dicha escala. Por otra parte, hay dos maneras de calificar que responden a diferentes criterios, según se contrasten las realizaciones del estudiante con sus propias posibilidades o con un patrón ajeno a él, ya sea el rendimiento de sus compañeros o una pauta establecida por el docente como objetivo a alcanzar. Cada una de estas formas dará lugar a calificaciones diversas para un mismo trabajo y un mismo alumno.