



Avances y desafíos en la evaluación educativa

En las últimas décadas se han producido notables avances en el ámbito de la evaluación. Sin embargo, son todavía escasos los sistemas educativos que cuentan con modelos asentados de evaluación de escuelas y de la práctica docente. La evaluación de los aprendizajes en el aula necesita cambios profundos. La evaluación de las propias políticas educativas y de los programas de innovación que desde ellas se impulsan tiene también poca tradición en nuestro entorno. Este libro ofrece una panorámica de los diversos niveles y ámbitos de la evaluación, destacando los avances y los retos aún pendientes.

Series de la colección

Afabetización	Cultura escrita	Educación artística
Educación técnico-profesional	Evaluación	Infancia
Profesión docente	Reformas educativas	TIC

Evaluación

La evaluación de los sistemas educativos, de los programas, de las escuelas, de los docentes y de los alumnos es una condición necesaria para mejorar la calidad de la educación. La serie de libros sobre **evaluación** pretende ofrecer los principales modelos, indicadores, experiencias e innovaciones que se desarrollan en este campo.

Metas Educativas 2021

La conmemoración de los bicentenarios de las independencias debe favorecer una iniciativa capaz de generar un gran apoyo colectivo. Así lo entendieron los ministros de Educación iberoamericanos cuando respaldaron de forma unánime el proyecto **Metas Educativas 2021: la educación que queremos para la generación de los Bicentenarios**. Semejante tarea colectiva, articulada en torno a la educación, ha de contribuir al desarrollo económico y social de la región y a la formación de ciudadanos cultos y libres en sociedades justas y democráticas. La **Colección Metas Educativas 2021** pretende ampliar y compartir el conocimiento e impulsar el debate, la participación y el compromiso colectivo con este ambicioso proyecto.



Evaluación

Avances y desafíos en la evaluación educativa

Elena Martín
Felipe Martínez Rizo
Coordinadores

Metas Educativas 2021
La educación que queremos para la generación de los Bicentenarios



Fundación Santillana

Evaluación de los aprendizajes en el aula: una conceptualización renovada

Juan Manuel Esquivel

INTRODUCCIÓN

Este capítulo pretende principalmente presentar a los educadores una propuesta de evaluación de los aprendizajes congruente con una concepción del aprendizaje fundamentada en el socio-constructivismo, que reúne las teorías cognitiva y sociocultural. Shepard (2006) señala lo siguiente: “si bien hay muchos puntos discutibles no resueltos entre estas dos perspectivas y en el interior de ellas, las considero compatibles”. Esta propuesta se aleja del modelo de evaluación de los aprendizajes que ha predominado en las aulas de nuestras escuelas primarias y secundarias durante las últimas décadas. Tal modelo que podría caracterizarse, en términos generales, por el empleo de pruebas llamadas “objetivas” y por un uso frecuente de la evaluación con el fin de calificar a los estudiantes y, en muchas instancias, definir y decidir aspectos importantes en sus vidas.

Considera la propuesta de este capítulo que la enseñanza, el aprendizaje y la evaluación deben constituirse en una unidad indisoluble (Coll, Martín y Onrubia, 2001), de manera tal que la evaluación sirva para explicar tanto los aprendizajes de los estudiantes como la actividad instruccional del maestro. La evaluación entonces servirá para ofrecer información a los estudiantes acerca de los aprendizajes alcanzados, a fin de que puedan asumir la responsabilidad de controlar sus logros. Por otra parte, servirá al docente para obtener información que le permita mejorar su actividad didáctica al planear y ejecutar actividades de enseñanza.

La evaluación, desde la perspectiva señalada en el párrafo anterior, ofrecerá a los actores del hecho educativo información útil y relevante para mejorar cada uno su acción. A los alumnos, para que los conocimientos construidos sean más amplios, significativos y profundos, y a los profesores, para que su actividad didáctico-pedagógica sea más eficaz en cuanto a promover esos mejores aprendizajes. Como muy bien señalan Coll, Martín y Onrubia (2001), “la evaluación de los aprendizajes puede estar o ponerse al servicio de un proceso de toma de *decisiones de naturaleza pedagógica o didáctica*” (p. 4).

En los diferentes apartados de este capítulo se pretende ofrecer a los profesores, en primer lugar, una revisión de algunos conceptos teóricos básicos sobre evaluación de los aprendizajes. Seguidamente, se hace una crítica a la forma como se emplea, en términos generales, la evaluación de los aprendizajes en las aulas de la educación básica y secundaria, y se fundamenta, desde la perspectiva de una concepción socio-constructivista de los aprendizajes, una evaluación congruente con tal concepción. Finalmente, se ofrece una revisión general de los principios de la evaluación del desempeño y de la evaluación auténtica, que incluye información acerca de los principales medios e instrumentos que emplea, y se brinda algunos ejemplos de estos.

REVISIÓN DE CONCEPTOS FUNDAMENTALES

En su conceptualización más amplia, la evaluación de los aprendizajes es un proceso sistemático de acopio de información que permite al profesor emitir un juicio de valor sobre las adquisiciones o aprendizajes que alcanzan sus estudiantes como resultado de su participación en las actividades de enseñanza. La información que se recolecta está condicionada por la existencia de dos elementos fundamentales: 1) ¿Cuáles son los aprendizajes que se pretende que alcancen los estudiantes?, y 2) ¿Cuáles son las muestras observables en las tareas, realizaciones o ejecuciones de los estudiantes que se consideran como pruebas del logro de los aprendizajes? Comúnmente, los aprendizajes se formulan en términos de objetivos, mientras las pruebas se obtienen de la observación de las tareas, realizaciones y ejecuciones de los estudiantes y de la aplicación de escalas y rúbricas a estas.

Por otra parte, es importante analizar las funciones de la evaluación. La evaluación puede darse en diferentes momentos de los procesos de enseñanza y del aprendizaje, y también puede servir para la toma de diferentes clases de decisiones, según el propósito del empleo de la información que genera. De estas dos perspectivas, tiempo y propósito, se originan las tres funciones básicas de la evaluación: diagnóstica o inicial, formativa y sumativa.

La *evaluación diagnóstica* se da al inicio de los procesos de enseñanza y de aprendizaje. Sobre la base del juicio de valor emitido, se toman las decisiones pedagógicas que han servido tradicionalmente para fundamentar el planeamiento didáctico de una unidad o curso en la realidad de los conocimientos previos de los estudiantes. Los resultados de esta evaluación han favorecido además una enseñanza basada en los principios de la educación inclusiva, pues han permitido adaptar la enseñanza a las características y necesidades de los estudiantes (Ainscow, 2001; Arnaiz, 2003, y Esquivel, 2006).

La *evaluación formativa* es apropiada durante el desarrollo de los procesos de enseñanza y de aprendizaje. Permite ofrecer información a los estudiantes y a los profesores sobre los aprendizajes logrados en un momento determinado de estos procesos. Esta conceptualización tradicional se ha visto enriquecida con la propuesta que hace Sadler (1989), quien aportó el modelo más aceptado de la evaluación formativa. Este autor indicó que no es suficiente que los maestros simplemente señalen si las respuestas dadas en una prueba son correctas o incorrectas, o si la tarea ejecutada (o el producto) exhibido refleja aprendizajes significativos. Esta retroalimentación deberá ir necesariamente acompañada de criterios explícitos y claros de desempeño, así como de información a los estudiantes sobre estrategias para facilitar el aprendizaje. Este aporte de Sadler se acerca a la conceptualización aportada por Coll, Martín y Onrubia (2001), según la cual la información que ofrece el juicio de valor deberá ayudar al profesor a tomar decisiones que mejoren sus actividades de enseñanza, y a los estudiantes, a mejorar su aprendizaje. Esta función de regulación de los procesos de enseñanza y de aprendizaje hace que se lo denomine “evaluación reguladora”. Por otra parte, según señalan los autores citados arriba, “se ha subrayado su vertiente ‘formadora’, es decir, su utilidad para que los alumnos aprendan a regular sus procesos de aprendizaje” (p. 5).

Finalmente, la *evaluación sumativa*, también denominada “acumulativa” o de “resultados”, se realiza al terminar un período determinado de los procesos de enseñanza y de aprendizaje. Tiene como propósito principal calificar el grado del desempeño de los estudiantes en el logro de los objetivos curriculares. Esta evaluación se efectúa usualmente al terminar períodos escolares de

diferente duración, como los correspondientes a las unidades de aprendizaje, a las partes del curso señaladas en el calendario escolar, como los trimestrales o los semestrales, al final de los cuales deben entregarse a los padres de familia informes acerca del rendimiento académico estudiantil, y *a todo el curso lectivo*. También se da el caso de sistemas educativos en los cuales los profesores realizan evaluación sumativa en forma programada y sistemática. Con los resultados de la evaluación sumativa el profesor emite un juicio sobre el logro final de los aprendizajes y sobre esta base se asigna al estudiante la calificación numérica o alfabética correspondiente al período evaluado. Si se trata del final de un curso o de un período lectivo, esta calificación permite una certificación con fines de promoción.

Como se ha indicado anteriormente, la evaluación disociada de los procesos de enseñanza y de aprendizaje corresponde a un modelo de evaluación de los aprendizajes fundamentado en la tradición psicométrica, adoptada por la psicología educativa y por una gran mayoría de los evaluadores educativos. En este modelo (clásico o normativo) esencialmente se interpretan los resultados del aprendizaje de un individuo, evaluado mediante pruebas, desde una perspectiva comparativa. En otras palabras, el significado del resultado de un individuo adquiere sentido cuando se compara con los resultados del grupo al que pertenece o al grupo normativo. Sin embargo, existe otro modelo de elaboración y validación de pruebas (modelo referido a criterios) que, por sus fundamentos y procedimientos, difiere sustancialmente del descrito antes, aunque no tiene su solidez teórica y metodológica. En la interpretación de los resultados manifiesta un sentido más pedagógico, pues compara el resultado de un individuo con objetivos o intenciones de logro.

En esta revisión general de conceptos básicos sobre evaluación de los aprendizajes es necesario dedicar un espacio de análisis a algunos otros conceptos importantes. El primero se refiere a la conceptualización errónea del término “medición” como sinónimo de “evaluación”. Ya señalamos en un párrafo anterior que la evaluación implica la emisión de un juicio de valor, basado en información relevante. Esta información podría incluir el resultado de la medición del logro de algún objetivo, mediante la definición de indicadores observables en las tareas, exhibiciones o producciones escritas, artísticas o de otra índole de los estudiantes, que se comparan con los objetivos o intenciones.

Al análisis anterior hay que agregar la dimensión comunicativa que tiene la evaluación. La comunicación que emite el profesor de los juicios de valor sobre el resultado del proceso evaluativo siempre tiene un destinatario, sean estos los estudiantes, otros profesores, la colectividad de la institución educativa, los padres y madres y el mismo profesor, y también tiene un propósito. Sin embargo, muchas veces el propósito del profesor puede ser distorsionado por las decisiones que, con respecto al empleo de la información, tomen algunos de sus destinatarios.

LA EVALUACIÓN COMO PARTE INTEGRAL DE LA ENSEÑANZA Y EL APRENDIZAJE

La evaluación disociada de la enseñanza y el aprendizaje

Tradicionalmente, el maestro ha considerado la evaluación como un elemento importante de su quehacer de aula, pero la ha aplicado al final de una unidad o curso para comprobar mediante ella si los alumnos han logrado los aprendizajes esperados. Esa concepción disocia la evaluación de la enseñanza y del aprendizaje y la convierte en un ejercicio que influye muy poco en la modificación de su práctica de la enseñanza y en el mejoramiento de los aprendizajes de los estudiantes.

Como consecuencia de esta realidad, la evaluación, para la mayoría de los maestros, tiene un único propósito: justificar las calificaciones de sus estudiantes. Para estos, la evaluación se constituye en el filtro determinante del éxito o fracaso escolar, muchas veces lejos de la determinación de los verdaderos aprendizajes adquiridos.

Por otra parte, los maestros han recibido una formación en evaluación fundamentada teóricamente en el modelo psicométrico y con énfasis en la elaboración de pruebas según los dictados de la elaboración de los tests, en un modelaje de los tipos de preguntas o ítems que generalmente se emplean en las pruebas estandarizadas. La conceptualización de la evaluación como sinónimo de medición, que generalmente hacen los maestros, es una consecuencia de la importancia que tiene el modelo psicométrico en la formación inicial del educador. Esta conceptualización se refuerza con la presencia, en los cursos de evaluación de los aprendizajes, de una unidad de estadística descriptiva, que deriva en los análisis de calidad de los ítems y en el análisis de confiabilidad y validez dependientes de la comprensión de la estadística y de las premisas básicas del modelo psicométrico: la curva normal y la necesidad de una alta variabilidad de los puntajes. Son muy pocos los maestros que se cuestionan el hecho de que estas dos premisas son esencialmente contrarias a la finalidad última de su papel como docente: facilitar con sus actividades de enseñanza el mayor aprendizaje posible en el mayor número posible de sus estudiantes. Los puntajes de los alumnos que alcancen estos logros no conformarán una curva normal, ni lograrán una dispersión o variabilidad como se esperarían para considerar los parámetros de la prueba y sus ítems como adecuados.

La imitación del evaluador profesional por parte del maestro en la elaboración de preguntas, como se enfatiza en su formación, propugnarán la elaboración de preguntas que se caractericen por solicitar al estudiante que seleccione una respuesta verdadera entre varias opciones o que identifique una relación simple, que llene un espacio en blanco, que identifique la veracidad o falsedad de una frase o de preguntas que requieren para su respuesta escrita únicamente la memoria o que son de muy poca elaboración de parte del estudiante. Además se da el caso de que la mayoría de los libros de evaluación de los aprendizajes que se encuentran a la disposición de los maestros¹ hacen énfasis en la elaboración de preguntas como las antes citadas. Todas ellas presentan una característica común: la respuesta se califica como buena o mala, lo cual implica que la calificación del test compuesto por estas se realiza sin la participación del criterio del profesor. De allí que se encuentre que los maestros emplean preguntas de los tipos citados anteriormente, que solicitan del alumno solamente la memorización o la comprensión de relaciones simples. Cada vez es más difícil encontrar en las pruebas confeccionadas por el maestro preguntas de ensayo en las que se solicita al estudiante la construcción de una respuesta.

Rápidamente, el estudiante se convence de que hay unas reglas de juego que conviene seguir, que indican que hay que lograr buenas calificaciones y que, para hacerlo, debe contestar bien las preguntas de una prueba. Esta prueba exige memorizar hechos, relaciones, fórmulas, procedimientos y comprender relaciones simples entre estos contenidos. Como resultado, pocos aprendizajes son significativos. El alumno estudia para responder en la prueba, haciendo énfasis en

¹ He examinado los 18 libros que aparecen más comúnmente en la bibliografía de los cursos de evaluación de los aprendizajes en escuelas formadoras de docentes de Costa Rica, y solamente uno no hace énfasis en la elaboración de preguntas objetivas o caracterizadas por una respuesta que se califica como buena o mala.

la memorización, sin una conciencia clara con respecto a los aprendizajes y su importancia. Lo hace durante un período corto y determinado, cuando se acerca la fecha de cada prueba. Pero esta realidad le enseña la poca importancia que tienen los aprendizajes, que tiene la conciencia de haber alcanzado.

El maestro como protagonista en la rendición de cuentas

Durante los últimos veinte años se ha producido un movimiento muy importante en Latinoamérica, tendente a establecer sistemas nacionales de evaluación del logro de los estudiantes (Rojas y Esquivel, 1998; Ferrer, 2007). Esta tendencia corresponde a una concepción economicista de la educación, que pretende medir la eficiencia de los sistemas educativos mediante la medición del logro de los aprendizajes de los estudiantes. Algunos pocos sistemas nacionales emplean pruebas estandarizadas, cuyos resultados pueden tener consecuencias graves para los estudiantes, pues sirven para decidir la graduación en un determinado nivel de la educación. Las hay también con resultados cuyas consecuencias son menos graves, pues se emplean como información diagnóstica del sistema educativo.

El hecho de que en estos sistemas se empleen pruebas estandarizadas constituye para el maestro una presión para que imite el tipo y la forma de preguntas que se aplican en estos tests. Significa para el maestro una razón más para elaborar pruebas con preguntas de opción múltiple en la evaluación de las intenciones u objetivos correspondientes a aprendizajes que los estudiantes deben alcanzar. Pero las consecuencias de la administración de estas pruebas son muy preocupantes, sobre todo en aquellos casos en que los resultados significan consecuencias graves, pues incentivan la tendencia a enseñar para la prueba.

Las pruebas estandarizadas ofrecen grandes ventajas para la medición de características psicológicas del ser humano y como determinantes generales del rendimiento académico o del logro de los aprendizajes, y pueden cumplir una función válida de rendición de cuentas en un sistema educativo. Sin embargo, tienen poca utilidad en la evaluación de los aprendizajes en el aula y en la determinación de los aprendizajes complejos de los estudiantes, por las limitaciones impuestas por el tipo de pregunta empleada y por el modelo estadístico que lo fundamenta.

Existen datos que demuestran que el empleo de este tipo de pruebas en el aula tiene un efecto perjudicial en la calidad de la enseñanza y el aprendizaje de los estudiantes (Resnick y Resnick, 1992, citado por Shepard, 2006).

La evaluación según una concepción integradora

Como se señaló en la introducción, la evaluación puede ponerse al servicio de la toma de decisiones de naturaleza pedagógica y didáctica. En otras palabras, se confiere una función pedagógica a la evaluación, que se logra por la aplicación de la evaluación formativa, mediante la cual se evalúa continuamente durante todo el proceso de enseñanza y de aprendizaje. Las tareas y medios de la evaluación son una extensión de este proceso, que enfatiza una caracterización cualitativa y dinámica de los aprendizajes. Esta función pedagógico-didáctica es de carácter integrador, en el sentido de que hace posible la integración de la unidad enseñanza, aprendizaje y evaluación. A esta función pedagógica se contraponen una función social de la evaluación, que separa las actividades de la evaluación de los procesos de enseñanza y de aprendizaje. Esta concepción se caracteriza, entre otras razones, por el empleo de la evaluación sumativa; por el énfasis en los re-

sultados cuantitativos y estadísticos, que redundan en la calificación de los estudiantes al final del proceso de enseñanza y de aprendizaje; por el empleo de pruebas escritas de tipo test, hechas por el profesor, pero que imitan las pruebas estandarizadas. Esta función social hace que la evaluación esté disociada de los procesos de enseñanza y de aprendizaje.

En la realidad del proceso educativo que se da en las aulas de nuestras escuelas existen tensiones entre estas dos funciones, la social y la pedagógico-didáctica, pero predomina la función social. Estas tensiones provienen del fuerte impulso que se ha dado a la teoría constructivista para explicar el aprendizaje y dirigir la enseñanza. Ello es evidente en los documentos curriculares y en el discurso de los curriculistas de los Ministerios de Educación, en contraposición con una evaluación que continúa cumpliendo en la mayoría de los casos una función social por la costumbre, por la falta de formación de los curriculistas y los docentes o por reglamentos o normas de evaluación.

Esta situación debe cambiar. Existe la necesidad de que los teóricos del currículo de los ministerios y los profesores comprendan los elementos básicos de una evaluación con función pedagógico-didáctica, fundamentada en la concepción socio-constructivista del aprendizaje, y la apliquen. De esta manera, la evaluación estará relacionada con un modelo cognitivo que no excluye la ayuda social que el enfoque sociocultural aporta con respecto a cómo aprende el estudiante.

Breve conceptualización del aprendizaje socio-constructivista

La concepción cognitiva del aprendizaje constituye una reacción contra la psicología de las diferencias individuales y el conductismo, cuyo interés fundamental era la adquisición de conocimientos, habilidades y destrezas, pocas veces integradas en competencias complejas, que se evidenciaban en conductas observables. En la concepción cognitiva, los aprendizajes son contruidos por los aprendices al realizar conexiones entre la nueva información y las estructuras previas de conocimiento. Los esquemas mentales desempeñan la función de organizar el conocimiento. De tal forma, estos pueden ser recuperados y utilizados en nuevas situaciones cuando se necesita. Por otra parte, estos esquemas se aprovechan también en los procesos llamados de “metacognición”, que permiten a quienes aprenden monitorear y manejar su propia comprensión y aprendizaje. Otros elementos importantes que enfatizan los cognitivistas son el papel de la comprensión conceptual y la transferencia (el uso del conocimiento en situaciones novedosas), que se hace posible por la adquisición de generalizaciones y el empleo de los esquemas que permiten reconocer similitudes en las distintas tareas que se presentan. Por otra parte, el modelo sociocultural (Vigotsky, 1978) enfatiza la naturaleza social del aprendizaje y “la idea de que la competencia y la identidad de quienes aprenden se desarrollan por medio de la participación mediada socialmente en una actividad práctica llena de significado” (Shepard, 2006, p. 17). Esta autora propone una concepción del aprendizaje socio-constructivista, que reúne las teorías sociocultural y cognitiva. Así pues, la primera establece que los procesos sociales respaldan y definen el aprendizaje, y la segunda permite entender los procesos mentales subsecuentes y recurrentes del individuo.

Caracterización de la evaluación con función pedagógico-didáctica

En párrafos anteriores se estableció una caracterización breve de la evaluación para que corresponda a la función pedagógico-didáctica. Las características necesarias para ello se explican con mayor detalle a continuación:

- Los objetivos o intenciones de aprendizaje deben estar claramente definidos y deben ser entendidos de igual forma por el profesor y sus estudiantes. Estas intenciones deberán ser objeto de discusión en la sala de clase y de una debida apropiación por parte de todos los actores: estudiantes y profesor.
- Además deben estar igualmente claros los criterios con que se juzgarán las tareas, demostraciones, exhibiciones y resolución de problemas con que el estudiante demostrará el dominio de los aprendizajes. La elaboración conjunta o la discusión de una propuesta de las hojas o escalas de calificación o rúbricas entre el profesor y los estudiantes constituye una excelente forma de comprensión de esos criterios.
- El desempeño del estudiante deberá ser juzgado en relación con los criterios definidos y no comparando su desempeño con el de sus compañeros (evaluación referida a criterios).
- Los estudiantes deberán participar en la valoración de sus propios aprendizajes y en la valoración del aprendizaje de sus compañeros. En otras palabras, el profesor incluirá oportunidades de autoevaluación y coevaluación en el planeamiento de su proceso de enseñanza.
- La evaluación deberá ser esencialmente formativa y formadora. Estas características implican que el profesor no solo deberá ofrecer numerosas y continuas oportunidades para que el estudiante se enfrente a situaciones problemáticas que le permitan demostrar sus aprendizajes, sino también establecer el clima apropiado para una clara retroalimentación que sea entendible y que ofrezca al estudiante la oportunidad de llenar las carencias y reflexionar sobre su propio aprendizaje. La evaluación formadora requiere no solo que los estudiantes se apropien de los objetivos o intenciones, como se ha explicitado en una caracterización anterior, sino también que logren supervisar su propio mejoramiento y asumir la plena responsabilidad de su aprendizaje.
- Los estudiantes deben realizar procesos mentales más complejos y retadores, pues las actividades de evaluación los obligan a ejercitarse en estos procesos.
- Es importante la comprensión por parte del profesor de las consecuencias motivacionales de la retroalimentación, pues podrían tener consecuencias negativas el empleo de elogios falsos o exagerados y la retroalimentación negativa directa. Facilitar el aprendizaje es la finalidad fundamental de la retroalimentación; por lo tanto, este propósito debe ser bien comprendido por los profesores y los estudiantes.
- La relación estrecha entre enseñanza, aprendizaje y evaluación se ejemplifica con el diseño y la ejecución de actividades de evaluación que se constituyan de la forma más semejante posible en ejemplos de los objetivos reales de aprendizaje. De tal manera, las actividades de enseñanza se convierten en actividades de evaluación.
- Shepard (2006) propone el concepto de “encarnar” como la caracterización completa de la alineación entre las tareas, los proyectos y las preguntas complejas o problemas a que se enfrentan los estudiantes y todos los elementos constitutivos que se han especificado como comprensiones y capacidades que se quiere que logren.
- La evaluación debe ser continua. Las oportunidades de realizar evaluaciones formativas se planifican de tal forma que se presenten durante todo el proceso de enseñanza y de aprendizaje.

- La evaluación se constituye en un recurso para juzgar la docencia y significa una oportunidad de mejoramiento del proceso de enseñanza.
- La calidad de los instrumentos y medios que se utilizan para evaluar es importante. Esta calidad se podría juzgar, desde la perspectiva de las demostraciones de su validez, por lo siguiente:
 1. El alineamiento que presenten con respecto a los objetivos o intenciones y los criterios con los que se juzgan.
 2. La congruencia con las actividades de enseñanza.
 3. Su representatividad entre las posibles situaciones o problemas que ejemplifican las intenciones u objetivos.
- También, la calidad de los instrumentos y medios se podría juzgar por la confiabilidad de los resultados. Metodológicamente podría calcularse la confiabilidad de la aplicación de los criterios para juzgar un logro determinado:
 1. Por parte de todos los estudiantes y el profesor.
 2. En la autoevaluación por parte del estudiante en diferentes ocasiones.
 3. En la coevaluación por parte de los coevaluadores.
- Ante la realidad de que nuestros sistemas educativos privilegian la evaluación sumativa en su acepción acreditativa, porque lo establecen las normas legales o por la costumbre, se debería procurar lo siguiente:
 1. Que tenga un papel menos protagónico en la educación básica. Que en esta se incremente su empleo conforme aumenta el nivel educativo, empezando después del tercer año escolar.
 2. Que se realice de tal manera que se juzgue el desempeño del individuo respecto a los criterios establecidos y no con respecto al desempeño de los otros estudiantes (evaluación criterial en vez de evaluación normativa).
 3. Que los resultados obtenidos se empleen también para retroalimentar a los estudiantes y para modificar y mejorar la enseñanza.
 4. Que sirvan para verificar el logro de aprendizajes importantes y complejos (competencias) por parte de los estudiantes y que tengan relación con desempeños evaluados mediante la evaluación formativa.

Una de las propuestas sobre evaluación de los aprendizajes que responde –desde la perspectiva de su definición, propósitos, estructuración, instrumentos y medios– a las características antes analizadas es la evaluación del desempeño, conocida también como evaluación de la actuación. Los apartados siguientes de este capítulo se dedicarán a desarrollar sus elementos constitutivos.

EVALUACIÓN DEL DESEMPEÑO

Definición

La definición de la evaluación del desempeño que aquí se ofrece fue elaborada sobre la base de las características establecidas por varios autores:

“Es un enfoque de evaluación en el que los estudiantes deben ejecutar tareas, mostrar ejecuciones o resolver problemas, en vez de simplemente proveer respuestas marcadas, escritas u orales de una prueba. El desempeño se juzga contra criterios preestablecidos y de naturaleza múltiple. Enfatiza la evaluación de conocimientos y habilidades complejas y de alto nivel de pensamiento, en un contexto del mundo real en que se emplean esos conocimientos y habilidades. Emplea diversos instrumentos y medios que requieren un tiempo sustancial de parte de los estudiantes para completarla” (Wheeler y Haertel, 1993, y Swanson, Norman y Linn, 1995).

Es importante destacar varios elementos de esta definición:

- Los estudiantes realizan tareas, muestran sus logros o resuelven problemas. En otras palabras, se hace énfasis en la producción más que en la reproducción. Son importantes los discursos, la elaboración de objetos y el desempeño que se muestre en productos concretos.
- Se establecen previamente criterios con los que se juzga el trabajo del estudiante. Estos criterios dependerán del juicio de los seres humanos involucrados en el proceso de definirlos.
- Se evalúan conocimientos y habilidades complejas. Ello podría interpretarse como la evaluación de competencias. Esta característica puede implicar el empleo de la indagación disciplinada: aplicación de los conocimientos previos, el entendimiento profundo y la integración para producir, ensamblar, interpretar información, formular nuevas ideas y formular críticas.
- Las tareas, los problemas o los proyectos se contextualizan en la realidad del estudiante y su entorno. Esta característica añade a la evaluación del desempeño la denominación de “auténtica”.
- El tiempo que requiere el estudiante para ejecutar las tareas, problemas o proyectos es sustancial. Que puede incluir realizar trabajo fuera de clase y aun por varios días.

La evaluación del desempeño tiene consecuencias para los actores del proceso de enseñanza y del proceso de aprendizaje, por estas razones:

- Es significativa para el estudiante, pues lo convierte en un participante activo de la evaluación. El estudiante debe expresar el conocimiento logrado mediante varias formas de desempeño.
- Al profesor lo obliga a reenfocar su acción didáctica: esta será menos directiva y él se convertirá en un docente más guiador-facilitador del aprendizaje.
- La evaluación se convierte en parte necesaria del proceso de enseñanza y del proceso de aprendizaje, no solo porque la evaluación implica privilegiar la evaluación formativa, sino también porque el estudiante aprende en las actividades de evaluación.
- El currículo se modifica, pues pasa de listados de contenidos y habilidades a descripciones de desempeños, tales como distintas clase de productos, de discursos, de artefactos y de proyectos o tareas que el estudiante debe realizar.
- Los ambientes educativos se transforman, pues se caracterizarán por la colaboración entre estudiantes, por el acceso a recursos de diversa índole, por la flexibilidad en el ritmo del aprendizaje, por la oportunidad de conversar sustantivamente e intercambiar información y por un uso flexible del tiempo.

Medios e instrumentos

Se distinguen varias clases de medios e instrumentos que emplea la evaluación del desempeño:

1. *La observación* del trabajo del estudiante por parte del profesor y por parte del mismo estudiante. La observación será más efectiva si se tiene una guía que dirija la observación y un registro que permita sistematizar las observaciones. También resultan útiles una escala de calificación, una hoja de cotejo o una rúbrica; sin embargo, estos últimos recursos requieren para su aplicación condiciones difíciles de alcanzar en algunos casos.
2. *Muestras de desempeño*. Son documentos o artefactos que indican los logros de los estudiantes. En general, se trata de productos y no de procesos. La valoración de estas muestras requerirá la confección y uso de una rúbrica.
3. *Pruebas*. Estarán constituidas por preguntas de formato abierto, por preguntas analíticas o por problemas. Se trata de pruebas en las que el estudiante podría tener la oportunidad de consultar diversos materiales o que podrían responderse o completarse en su casa.
4. *Tareas o proyectos*. Son trabajos de diversa índole según la disciplina y habilidades de que se trate. Requieren planeamiento, ejecución y la presentación de un informe. Por ejemplo, en Ciencias Naturales podrían estar constituidos por investigaciones sencillas para redescubrir principios y practicar los procesos científicos.
5. *Exposiciones*. Se trata de discursos, obras de arte o artefactos producidos por los estudiantes que se muestran a un público. Pueden ser proyectos tecnológicos o científicos que se exhiben y explican en una feria científica.
6. *Portafolios*. En términos generales, es una carpeta que contiene información acerca de las experiencias y logros del estudiante. Más adelante se realizará una explicación más extensa de este medio.

Las rúbricas

En las explicaciones anteriores se ha establecido que las rúbricas son escalas de calificación que desarrolla el profesor y discute con sus estudiantes. En algunos casos podrían ser elaboradas conjuntamente por el educador y los alumnos. Contienen descripciones detalladas de los niveles de logro que podría alcanzar un estudiante en una tarea, en la elaboración de un producto o de un proyecto. Además presentan una escala, que puede ser cualitativa, cuantitativa o una mezcla de ambas, la cual califica los diferentes niveles de logro. Tal escala está claramente articulada con todos los niveles. Para Mabry (1999), una rúbrica contiene las reglas por las cuales se determina la calidad de las respuestas, los productos o las tareas.

La elaboración de la rúbrica requiere que el profesor tenga claro el marco de referencia conceptual de la tarea que se va a calificar, así como el objetivo o la competencia que se quiere que el estudiante demuestre. Deberá describir los niveles de logro y establecer la escala con que se califica cada uno de esos niveles. Es importante insistir en que la rúbrica debe ser claramente entendida por los estudiantes y que su discusión es necesaria. Más adelante se ofrecerá un ejemplo de rúbrica general para emplearse en Matemáticas.

Ejemplos de evaluaciones de desempeño

A continuación se ofrecerán ejemplos de algunos medios e instrumentos. Estos se tomaron de diversas fuentes, que se señalan en cada uno de ellos.

Matemáticas

TAREA DE TEORÍA DE NÚMERO

Yolanda le contó a su hermano Damián lo que hizo en su clase de Matemáticas.

Yolanda le dijo: “Damián, hoy empleé bloques en mi clase de Matemáticas. Cuando agrupé los bloques en grupos de dos, me sobró un bloque. Cuando agrupé los bloques en grupos de tres, me sobró también un bloque. Y cuando los agrupé en grupos de cuatro, tuve aún un bloque sobrante”. Damián le preguntó: “¿Cuántos bloques tenías en total?”

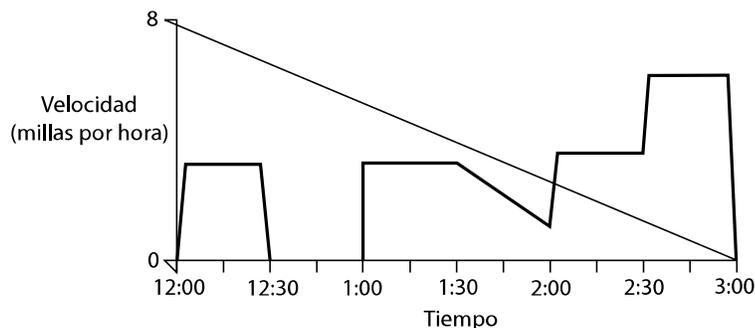
¿Cómo contestó Yolanda la pregunta de su hermano?

Muestre su trabajo

TAREA DE INTERPRETACIÓN DE UN GRÁFICO

Use el gráfico que se muestra más abajo y la siguiente información para escribir una historia o cuento acerca de la caminata de Antonio.

“Al mediodía Antonio empezó a caminar hacia la casa de su abuela, adonde llegó a las 3 pm. El gráfico muestra la velocidad de Antonio, en millas por hora, durante su caminata”.



Escriba una historia acerca de la caminata de Antonio. En su historia, describa lo que pudo haber estado haciendo Antonio en los diferentes momentos de su caminata.

Tomado de Lane, S.; Liu, M.; Ankenmann, R. y Stone, C. (1996)

Según los autores, estas tareas están constituidas por preguntas de final abierto y pretenden evaluar un marco de referencia contextual de cuatro componentes: contenido matemático, proceso cognitivo, modo de representación y contexto de la tarea. Este marco de referencia es la base de la rúbrica que se emplea en el proyecto QUASAR (Lane, 1993).

Ciencias

PULSO CARDÍACO

El sistema circulatorio de una persona se ve afectado por muchas razones. El cambio en el ritmo cardíaco puede depender de factores tales como:

- la cantidad de tiempo en que una persona se ejercite (duración),
- la clase de ejercicio que él o ella hace (tipo),
- cuán fuertemente él o ella se ejercite (intensidad) y
- la condición física general de la persona.

Un valor aproximado de la condición física se puede obtener al computar el tiempo que requiere el ritmo cardíaco para regresar después del ejercicio al nivel de descanso. En su proyecto, usted debe hacer lo siguiente:

1. Escribir un problema y una hipótesis acerca de la relación entre el pulso cardíaco de una persona y la duración, tipo o intensidad del ejercicio que él o ella ejecuten. Usted empleará el pulso cardíaco como medida del ritmo cardíaco.
2. Planear, diseñar y ejecutar un estudio para someter a prueba su hipótesis (podría necesitar recolectar datos fuera de clase; sin embargo, todo el trabajo debe completarse dentro de los períodos designados en clase). Recoja datos entre la gente, porque este estudio no debe ser una simple revisión de bibliografía.
3. Tome nota de al menos cuatro resultados. Emplee tablas, cuadros o gráficos para mostrar sus resultados.
4. Escriba sus conclusiones acerca de la relación, si existe alguna, entre el ritmo cardíaco, la duración, el tipo o la intensidad del ejercicio.
5. Investigue y explique la influencia de la condición física (“fitness”) en la(s) relación(es) que usted encontró entre el ritmo cardíaco y el ejercicio. Fundamente sus explicaciones en los datos recolectados y la información que encontró al investigar sobre el tema en la literatura.
6. Escriba un informe final.

Su profesor calificará su proyecto basándose solamente en su informe. Asegúrese de que este incluya:

1. Un problema claramente definido. Escríbalo en forma de pregunta.
2. Una hipótesis que tenga que ver con la relación entre el ritmo cardíaco y la duración, el tipo o la intensidad del ejercicio.
3. Una descripción de su plan y del diseño de su estudio para investigar la hipótesis establecida. Ello responde a la pregunta: “¿cuáles fueron sus procedimientos?”. Incluya una descripción de sus variables.
4. Una descripción de la información que recolectó para mostrar la relación entre el ritmo cardíaco y la duración, el tipo o la intensidad del ejercicio. Incluya una discusión de las limitaciones del estudio o de los datos recolectados.

5. Un informe de sus resultados que puede ser fácilmente comprendido por cualquier persona que lo lea. Incluya cuadros, gráficos o tablas si es necesario, así como las relaciones que encontró, si las hay, entre el ritmo cardíaco y la duración, el tipo o la intensidad del ejercicio.
6. Las conclusiones o inferencias que usted obtuvo acerca de la relación entre el ritmo cardíaco y el ejercicio, incluyendo la posible influencia de la condición física.
7. Determine y escriba si su hipótesis era la correcta.

Recuerde que quien lea su informe deberá ser capaz de realizar el mismo estudio usando únicamente la información contenida en tal informe.

Tomado de *Kansas Science Assessment*. Grade 8 (1995)

Esta tarea presenta una gran complejidad para jóvenes de octavo año, pues requiere el uso de conocimientos y habilidades relativas al proceso científico. Los estudiantes deben planificar un estudio completo, establecer problemas e hipótesis, buscar y analizar la información bibliográfica referente al problema, definir procedimientos, recolectar información apropiadamente, analizar la información recolectada, interpretar los resultados a la luz de la hipótesis y de lo que las fuentes bibliográficas establecen acerca de la relación estudiada, y comunicar lo realizado mediante un informe completo. En este informe se fundamentará la evaluación, que se realizará empleando una rúbrica previamente definida.

Lenguaje

TAREA DE ESCRIBIR UNA CARTA

La profesora de cuarto grado planea la enseñanza de la elaboración de una carta.

Primer día. Realiza una discusión con los estudiantes, a los que previamente ha agrupado de cinco en cinco. Les reparte ejemplos de cartas y les pide que en cada grupo se identifiquen las partes que en común tienen las cartas. En la discusión con todos los grupos juntos logra que, con la contribución de todos ellos, identifiquen las partes constitutivas de una carta.

Seguidamente, solicita a cada uno de los alumnos ejecutar en sus casas o en su vecindario una tarea muy concreta. Deberán identificar una situación problemática que tengan en su hogar, en el de algún vecino o en el barrio, que amerite escribir una carta: a algún vecino, al representante vecinal ante el gobierno municipal, a algún departamento municipal, al alcalde municipal, a alguna organización comunal, a alguna organización de servicio (Cruz Roja, bomberos, etc.), a alguna empresa de servicio público (empresas de agua, luz o gas), a alguna autoridad de las iglesias, etc. Tomar nota de la situación, de a qué persona deberán dirigir la carta y anotar el nombre de la persona que informó acerca de la situación problemática.

Segundo día. La profesora solicita a cada estudiante que escriba un primer borrador de la carta que dirigirá a la persona sugerida, describiendo la situación y haciendo una solicitud determinada. Cada estudiante escribe el primer borrador. La profesora les sugiere que hagan una revisión de ese borrador y lo corrijan si lo consideran necesario, y les solicita que, en grupos de cinco, cada uno lea su carta a los demás estudiantes del grupo con el propósito de que reciban observaciones de sus compañeros. Después de este ejercicio, les solicita que corrijan nuevamente con base en las observaciones recibidas de sus compañeros. Recoge esas primeras versiones de las cartas y las fotocopias. Devuelve el original de la carta a cada estudiante y les solicita que la lleven a la persona que les indicó la situación problemática, le pidan

que la lea y que les haga alguna observación relativa a la descripción del problema y a la petición concreta. Les solicita que tomen nota en su cuaderno de las observaciones de dicha persona.

Tercer día. La profesora recolecta los originales y les devuelve la fotocopia de la primera versión. Solicita a los estudiantes que, con base en las observaciones de la persona que les sugirió la situación problemática, escriban una versión final de la carta. La profesora recolecta esta segunda versión.

Una semana después. La profesora devuelve las dos versiones de la carta, con observaciones en cada una, acompañadas de una rúbrica en cada caso. Reúne a los estudiantes en grupos de cinco y les solicita que compartan las observaciones recibidas. Se acerca a cada grupo para escuchar y reafirmar en cada grupo aquellos aspectos que deben ser corregidos para alcanzar mejores resultados. Además solicita a todos los estudiantes que escriban unos párrafos indicando sus sentimientos acerca de la experiencia de aprendizaje vivida y cómo miran ellos el aprendizaje alcanzado, enfatizando los aspectos positivos de lo logrado y cómo se podría mejorar la experiencia de aprendizaje y de evaluación.

Tomado de Esquivel, J. M. (2000)

Nótese que, en el ejemplo de Lenguaje, se trata de una tarea de enseñanza que a su vez se convierte en una tarea evaluada. Este es un ejemplo en que la integración entre enseñanza, aprendizaje y evaluación es total, pues los procesos se funden y se confunden entre sí.

Se muestran a continuación los elementos básicos de la rúbrica utilizada para la evaluación de la tarea de Matemáticas. Además se muestra una rúbrica de Ciencias, que se emplea para la evaluación de una tarea para 5.º grado.

Rúbrica de Matemáticas

NIVEL DE PUNTAJE CUATRO

Conocimiento matemático. Muestra comprensión de los conceptos y principios matemáticos, usa apropiadamente la terminología y notaciones matemáticas, y ejecuta los algoritmos correcta y completamente.

Conocimiento estratégico. Podría emplear información de afuera de naturaleza formal o informal; identifica todos los elementos importantes del problema y muestra comprensión de las relaciones entre ellos; refleja una estrategia apropiada y sistemática para resolver el problema; da evidencia clara del proceso de solución y este resulta completo.

Comunicación. Ofrece una respuesta completa y clara, que incluye una explicación o descripción no ambigua; podría ser un diagrama apropiado y completo. Comunica eficazmente a la audiencia identificada; presenta argumentos fuertes y con soporte que son lógicos y completos; podría incluir ejemplos que ilustren la respuesta y ejemplos contrarios.

NIVEL DE PUNTAJE CERO

Conocimiento matemático. No muestra comprensión de los conceptos y principios del problema matemático.

NIVEL DE PUNTAJE CERO

Conocimiento matemático. No muestra comprensión de los conceptos y principios del problema matemático.

Conocimiento estratégico. Podría intentar emplear información de afuera irrelevante; falla en indicar cuáles elementos del problema son apropiados; copia parte del problema, pero no intenta ninguna solución.

Comunicación. Se comunica ineficientemente; las palabras que emplea no reflejan el problema; podría incluir dibujos que no representan la situación que establece el problema.

Existen descripciones para las tres categorías en los niveles de puntaje tres, dos y uno, que no se transcriben por problemas de espacio. Pueden consultarse en Lane, S. (1993) o en Lane, S.; Lui, M.; Ankenmann, R. y Stone, C. (1995).

Rúbrica de Ciencias

COMPONENTES DEL PROCESO

Reconoce y define el problema.

Diseña la estrategia de resolución del problema.

Implementa la estrategia de resolución del problema.

Interpreta y comunica resultados y conclusiones.

Los cuatro componentes del proceso se califican en una escala de cinco valores; por ejemplo, para el componente “Reconoce y define el problema”.

Puntaje uno: La observación y la descripción de objetos y sucesos es limitada. Escribe una pregunta que no es aceptable o no escribe la pregunta.

Puntaje dos: Entre 1 y 3.

Puntaje tres: Observa y describe algunas características singulares de los objetos o los sucesos. Escribe preguntas aceptables, pero no claramente definidas.

Puntaje cuatro: Entre 3 y 5.

Puntaje cinco: Observa y escribe características singulares de los objetos o fenómenos. Escribe una pregunta aceptable para ser objeto de investigación.

Tomado de *Kansas Science Assessment, Fifth Grade, Teacher Manual, 1995*

Portafolios

De las diversas opciones de instrumentos y medios que presenta la evaluación del desempeño, se realiza una explicación detallada del portafolio por ser uno de los medios más empleados. Muchas veces se hace sin respetar los elementos básicos y los propósitos que tiene este medio.

Definición

La siguiente definición de portafolio fue elaborada sobre la base de los elementos ofrecidos por los autores que se señalan más adelante.

“Es una colección comprensiva del trabajo del estudiante que exhibe sus esfuerzos, progresos y rendimientos en una o más áreas. La colección debe incluir la participación del estudiante en seleccionar

los contenidos, los criterios de selección empleados, los criterios para juzgar el mérito de su trabajo y la demostración de autorreflexión sobre el trabajo. Las descripciones de sus logros pueden incluir muestras de grabaciones de audio, vídeo y fotografías que registren desempeños y proyectos, y testimonios de autoridades sobre la calidad del trabajo del estudiante. En resumen, contiene una variedad de información que documenta las experiencias y logros del estudiante” (Paulson, F.; Paulson, P. y Mayer, C., 1991, y Archabald, D. y Newman, F., 1991).

Agregan Paulson, Paulson y Mayer (1991): “Un portafolio es un portafolio cuando provee una visión contextualizada, comprensiva y compleja del desempeño del estudiante. Es un portafolio cuando el estudiante es un participante en vez de ser objeto de evaluación. Pero, sobre todo, un portafolio es un portafolio cuando provee un foro que estimule al estudiante a desarrollar las habilidades necesarias para llegar a ser un aprendiz independiente” (p. 15).

De la definición anterior y de lo expresado por diversos autores se pueden destacar algunos aspectos importantes:

- Al desarrollar un portafolio, se ofrece al estudiante una oportunidad para aprender acerca del aprendizaje.
- El portafolio es un producto hecho *por* el estudiante y no *para* el estudiante.
- El portafolio difiere de una carpeta acumulativa sobre un estudiante.
- Podría tener un carácter formativo durante el año y un carácter sumativo al final del año.
- Podría cumplir propósitos múltiples, pero estos no deben estar en conflicto.
- Debe dar a entender explícita o implícitamente las actividades del estudiante, pero también debe mostrar todos sus logros.
- Debe contener información que ilustre el crecimiento del estudiante.
- Los estudiantes muchas veces requieren modelos y ejemplos de portafolios para producir sus propios portafolios eficientes.

Portafolio sobre producción escrita

Un portafolio que es, a su vez, una selección realizada por el estudiante del contenido del portafolio, que el estudiante ha elaborado durante todo el año, acerca de su capacidad de producción escrita. Eventualmente, el profesor selecciona también alguna producción.

Al estudiante se le pidió que seleccionara cuatro piezas, calificadas así:

- Una pieza importante para él.
- Una pieza que le fue satisfactoria.
- Una pieza no satisfactoria para él.
- Una pieza de selección libre.
- Ocasionalmente, el profesor seleccionaba otra pieza .

Además debe presentar una tabla de contenidos. Un inventario de la producción realizada en el año y una descripción de su experiencia como escritor. Una reflexión final sobre su producción que ofrezca una evaluación personal y la explicación de los cambios que han experimentado durante el año su producción y su capacidad como escritor.

La rúbrica que se empleó tenía tres dimensiones:

- Logros en la producción escrita.
- Uso de procesos y estrategias para escribir.
- Crecimiento, desarrollo y compromiso como escritor.

Cada una de estas dimensiones está acompañada por una descripción de elementos constituyentes que indicaban criterios importantes para calificar la dimensión (no se detallan por falta de espacio).

Cada una de estas dimensiones se calificaba en una escala de siete puntos: desde el puntaje seis, que indicaba excelente desempeño, hasta el puntaje cero, que indicaba desempeño totalmente inadecuado. Cada dimensión se hace mediante una calificación holística.

LeMahieu, P.; Gitomer, D. y Eresh, J. (1995)

CONCLUSIÓN

Los profesores habrán encontrado en este capítulo la justificación sustentada de la necesidad de un paradigma de evaluación que sea capaz de integrar el trinomio enseñanza-aprendizaje y evaluación. También la fundamentación de una evaluación que tenga una función pedagógico-didáctica, que sea consonante con la teoría de aprendizaje socio-constructivista. La evaluación que se ha propuesto para que se emplee en las aulas se caracteriza por: basarse en objetivos y criterios claros, ser participativa, emplear una interpretación criterial, ser formativa y formadora, porque sus medios e instrumentos obligan a los estudiantes a realizar procesos mentales complejos y retadores, retroalimentar e indicar cómo superar las deficiencias, requerir el diseño de actividades de evaluación muy semejantes a las de enseñanza, ser continua y porque sus resultados permiten mejorar la enseñanza.

La visión general sobre la evaluación del desempeño, la explicación de sus medios e instrumentos y los ejemplos que se han ofrecido, se espera que se constituyan en un reto para que los profesores investiguen las fuentes bibliográficas y encuentren ideas que les permitan mejorar la evaluación de los aprendizajes. Finalmente, es importante señalar que la evaluación del desempeño ofrece la fundamentación, los medios e instrumentos para ser una de las mejores opciones para la evaluación de las competencias.