

Por qué no existe la evaluación formativa

Dylan Wiliam, Emeritus Professor of Educational Assessment - UCL Institute of Education, UK

¿Qué es exactamente la evaluación formativa? Para algunos es una evaluación de poco riesgo, que se da al alumnado como una forma para preparar una tarea más importante. Por ejemplo, se puede animar al alumnado a enviar una redacción antes de la fecha límite para que tenga la posibilidad de recibir retroalimentación que pueda utilizar para mejorar la redacción antes de que finalice. Este uso del término formativo describe el momento del proceso de enseñanza y aprendizaje en el que se ocurre la evaluación. Así, una evaluación formativa es cualquier evaluación que se lleve a cabo antes de «la grande».

Para otros las evaluaciones formativas son útiles para comprobar el progreso del alumnado. La idea es recopilar pruebas de los logros del alumnado antes de que finalice el bloque didáctico, para que se pueda comprobar si el alumnado está progresando lo suficiente como para alcanzar el nivel requerido o esperado. Es, por supuesto, es algo bueno. Es desastroso descubrimos que el alumnado no ha alcanzado los niveles requeridos por primera vez al final del bloque, así que debemos comprobar que el alumnado está aprendiendo y recordando lo que se le ha enseñado. Podríamos llamar a este tipo de evaluación «evaluación sumativa de alerta temprana».

Sin embargo, llevarlo a cabo no suele ser fácil. Si

evaluamos el progreso del alumnado a mitad de una unidad, ¿evaluamos la unidad al completo o solo lo que se les ha enseñado hasta ahora? Si evaluamos la unidad al completo, es obvio que el alumnado no lo hará muy bien, ya que se le estaría evaluando sobre cosas que aún no se le han enseñado. Pero, en el otro lado de la moneda, gran parte del alumnado vería que su evaluación final sería más alta que la evaluación intermedia. Si, por el contrario, solo se evalúa lo que se ha tratado hasta ese momento, las puntuaciones del alumnado reflejarían lo bien que ha aprendido lo que se le ha enseñado, pero si la segunda parte de la unidad es más exigente, esto puede dar una impresión engañosa sobre su nivel final de rendimiento. No hay respuestas fáciles en este caso; simplemente hay que compensar. Lo importante es que, si tenemos claras las compensaciones, podemos tomar decisiones que probablemente representen los mejores intereses del alumnado.

Además de identificar al alumnado que podría necesitar más apoyo para seguir el ritmo del resto de la clase, si se utilizan de forma inteligente, estas evaluaciones también pueden identificar desajustes entre lo que el alumnado necesita aprender y lo que se está enseñando. Por ejemplo, si tenemos una hoja de cálculo con los nombres del alumnado en un lateral y los números de las preguntas a lo largo de la parte superior, y cada celda contiene la puntuación que ha obtenido el alumnado en esa pregunta (cero o uno para las preguntas correctas/incorrectas, o la puntuación real para las preguntas con créditos parciales), si sumamos las puntuaciones de forma horizontal tendremos la puntuación de cada estudiante en el examen. Sin embargo, si sumamos las puntuaciones en vertical,

conoceremos cómo de bien se ha respondido a cada pregunta. Las puntuaciones bajas podrían indicar que se trata de un tema que todavía no se ha enseñado, pero también podrían indicar que hay un elemento que el alumnado debería haber sido capaz de responder, pero no lo ha sido así. Podría significar que hay algún error en la pregunta o que el enfoque que se ha adoptado para enseñar



lo que esa pregunta evalúa ha sido inadecuado. Lo importante es que la recopilación de datos sobre lo que aprende el alumnado ayuda al profesorado a que su enseñanza responda más a las necesidades del alumnado.

Sin embargo, si pensamos en la evaluación formativa de esta manera (descubrir lo que va mal después de que ya haya ido mal) se pierden aspectos que podrían tener un impacto mucho mayor en el aprendizaje del alumnado.

Mientras que descubrir si el alumnado ha aprendido algo es importante, la evaluación puede desempeñar un papel importante a la hora de ayudar al alumnado a ser más activo en la gestión de su propio aprendizaje, lo que Richard Stiggins denomina «evaluación en el aula con la implicación del alumnado». Si el alumnado entiende lo que debe aprender y es capaz de supervisar su propio progreso hacia esos objetivos, es probable que sea capaz de mejorar su propio aprendizaje.

La evaluación también puede ayudar al profesorado a tomar mejores decisiones en su enseñanza. El objetivo de toda enseñanza es aumentar la capacidad a largo plazo. Y es que, si el alumnado no puede recordar algo una semana después de que se le haya enseñado algo, es que no lo ha aprendido realmente. Pero el profesorado tiene que tomar decisiones de forma constante mientras enseña (no puede esperar dos semanas para ver si el aprendizaje se ha «fijado») y, si tiene mejores pruebas de lo que pasa por la cabeza del alumnado, es probable que tome mejores decisiones sobre lo que debe hacer después. El hecho de que el alumnado pueda hacer algo al final de la lección de hoy no garantiza que sea capaz de hacerlo dentro de dos semanas, pero si no puede hacer algo al final de la lección de hoy, es bastante improbable que pueda hacerlo dentro de dos semanas. Mejores pruebas llevan a mejores decisiones, lo que a su vez conduce a un mejor aprendizaje.

Si pensamos en la calidad de las pruebas que tiene el profesorado para tomar sus decisiones, hay dos aspectos especialmente importantes: la amplitud y la profundidad. Por ejemplo, un docente pregunta a una clase de niñas y niños pequeños cuáles de las siguientes opciones son seres vivos:

- A. Pájaro
- B. Árbol
- C. Gato
- D. Mesa

El docente pide al alumnado que levante la mano si sabe la respuesta. Entonces si un estudiante

elige la letra A y otro la letra C, es tentador concluir que la clase ha entendido lo que es un ser vivo. Sin embargo, si el docente obtiene respuestas únicamente de dos estudiantes, tiene una idea muy limitada de lo que está pasando por la cabeza del resto de la clase. Por eso es una buena idea obtener respuestas de cada uno de los estudiantes de forma periódica (es decir, cada 20 o 30 minutos de enseñanza en grupo), utilizando un sistema de respuesta para todo el alumnado, como minimizarras de borrado en seco, tarjetas con letras o incluso votaciones con los dedos (1 para A, 2 para B, etc.). Algunas empresas tecnológicas recomiendan utilizar sistemas de votación electrónica, que afirman que son particularmente útiles porque permiten registrar la respuesta de cada uno de los estudiantes. Sin embargo, no estoy seguro de que sea una buena idea. Si queremos crear aulas en las que el alumnado se sienta bien al cometer errores, lo último que deberíamos hacer es registrar cada uno de ellos. Por eso la tecnología a veces puede ser un obstáculo y la versión de baja tecnología suele ser mejor. Y esa es la razón por la que a veces bromeo diciendo que las minimizarras de borrado en seco son el avance más importante en tecnología educativa desde la pizarra.

Además de obtener pruebas del alumnado al completo, también es importante prestar atención a la calidad de las pruebas que se obtienen, y por eso es muy importante diseñar con detenimiento las preguntas que se formulan. Si todas y todos los estudiantes de una clase responden A y C a la pregunta anterior, podríamos concluir que saben lo que es un ser vivo. Por desgracia, es poco probable que esta sea una conclusión válida, ya que sabemos que muchos niños pequeños tienen una idea errónea sobre los seres vivos: que se mueven. Un niño con esta idea errónea respondería correctamente a la pregunta anterior. En cambio, si se utiliza la siguiente pregunta, saldrá a la luz el concepto erróneo:

¿Cuáles de los siguientes son seres vivos?

- A. Hierba
- B. Autobús
- C. Árbol
- D. Ordenador



La calidad de las decisiones que tomamos depende en gran medida de la calidad de las pruebas que recogemos, y esto depende, a su vez, de la calidad de las preguntas que formulamos. Al fin y al cabo, si el alumnado con una idea errónea y el alumnado con una idea correcta dan la misma respuesta a una pregunta, no es una pregunta muy útil. Por eso es importante que el profesorado piense las

preguntas que va a hacer en clase (a ser posible con un colega) como parte de la planificación de sus clases. Como me dijo una vez un profesor: «No puedes pensar en buenas preguntas tú solo. Siempre serás víctima de tu propia forma de pensar».

De lo anterior puede deducirse que la evaluación formativa puede llevarse a cabo en distintas escalas de tiempo, desde una o dos veces al semestre, en un extremo, hasta minuto a minuto y día a día, en el otro. Pero también es importante destacar que lo que hace que una evaluación sea formativa no es el momento en que se lleva a cabo, ni siquiera el tipo de evaluación que utilizamos. Es lo que hacemos con la información.

Por ejemplo, supongamos que examino a un alumno sobre las operaciones de multiplicación desde 1×1 hasta 10×10 seleccionando al azar veinte de las cien operaciones numéricas posibles.



Si el alumno responde correctamente a diez de las veinte, puedo estar relativamente seguro de que conoce aproximadamente el 50 % de las operaciones numéricas. Esa sería una conclusión sumativa, ya que estoy utilizando las pruebas para hacer una afirmación sobre el estado actual del conocimiento de ese alumno. Sin embargo, si veo que parece tener una dificultad concreta con la tabla del siete, eso me da, como profesor, algo con lo que trabajar. Sé qué hacer a continuación. Es una conclusión formativa. Lo importante aquí es que la misma evaluación, e incluso la misma información de evaluación, puede utilizarse de forma sumativa o formativa. Esto significa que «formativa» y «sumativa» no pueden ser propiedades de una evaluación, sino de los usos que hacemos de la información, o de la función que cumple la evaluación. Si pensamos en las evaluaciones como

procedimientos para sacar conclusiones (damos al alumnado cosas que hacer, observamos lo que hace y sacamos conclusiones), entonces las palabras «formativa» y «sumativa» se consideran mejor como descripciones de las conclusiones que sacamos. A veces sacamos conclusiones sumativas (este alumno sabe el 50 % de las operaciones de multiplicación) y otras veces sacamos conclusiones formativas (tengo que ayudar a este alumno con las tablas de multiplicar). Ahora bien, si aceptamos que las palabras «formativa» y «sumativa» describen las conclusiones que sacamos, está claro que no puede existir una «evaluación formativa», sino una evaluación que arroje pruebas que se utilicen de manera formativa.

Para concretarlo, podemos tener en cuenta la prueba intermedia de la que hablamos al comienzo de este artículo. Si hacemos un examen a al alumnado a mitad del bloque, el hecho de que sea formativo o no depende de lo que hagamos después con las pruebas de la evaluación. Si puntuamos la evaluación y utilizamos esa puntuación para contribuir a la nota final del semestre, estará funcionando de forma sumativa, pero si también ofrecemos al alumnado retroalimentación sobre lo que debe mejorar, entonces también estará funcionando de manera formativa. El problema, por supuesto, es que la presencia de la puntuación puede impedir a menudo que el alumnado se fije en la retroalimentación sobre el modo de mejorar. Mira su propia puntuación... y luego mira la puntuación del vecino: lo sumativo desplaza a lo formativo. Todas las

evaluaciones pueden utilizarse tanto de manera formativa como sumativa, pero normalmente una función interfiere con la otra, por lo que generalmente es mejor decidir desde el principio el propósito de la evaluación: ¿es para ayudar al alumnado a mejorar o para decirle lo bueno que es? Es muy complicado hacer las dos cosas a la vez.

