

La evaluación formativa en 2050: oportunidades y desafíos

Dylan Wiliam

Instituto de Educación de la UCL

Cuando defiendo que se incrementa el uso de la evaluación formativa en las aulas, me suelen hacer tres preguntas. La primera es que **cómo sé que la evaluación formativa no es otra moda pasajera en la educación**; una que dentro de uno o dos años se sustituirá por otra cosa que llame la atención de la gente.

Mi respuesta es que, como indicó David Ausubel hace más de 50 años, la buena enseñanza parte del punto en el que se encuentra el alumnado, más que del punto en el que nos gustaría que estuviera. Y dado que lo que este aprende de sus experiencias en el aula es muchas veces impredecible, tenemos que descubrir en qué punto se encuentra su aprendizaje antes de decidir qué hacer a continuación. Es decir, la evaluación es el puente entre la enseñanza y el aprendizaje. Solo a través de la evaluación podemos conocer la repercusión que han tenido nuestras actividades didácticas en el alumnado. Mientras que el profesorado analice de forma sistemática la relación entre lo que hace y lo que aprende el alumnado, siempre podrá mejorar su práctica. En estos momentos, los estudios disponibles indican que este es el mejor enfoque para el desarrollo profesional del profesorado.

La segunda pregunta que me hacen es si, tras más de treinta años trabajando en esta área, ha cambiado **mi opinión sobre la evaluación formativa**. La verdad es que **no ha cambiado**. Hace treinta años, ahora y dentro de treinta años, mi idea es la de unas aulas en las que el profesorado adapte continuamente su enseñanza en función de los resultados del alumnado, que son a la vez profundos y generales. Cuando hablo de profundidad, me refiero a hacer preguntas que lleguen al centro de lo que piensa el alumnado. En las clases de ciencias de Educación Secundaria, por ejemplo, en vez de preguntar si el peso de una persona sería el mismo en la luna, a lo que la mayoría del alumnado responde correctamente, es

mucho más útil preguntar si la masa de una persona sería la misma en la luna, ya que muchos se equivocan.

Con estudiantes más jóvenes, al presentar los decimales, el profesorado suele hacer preguntas que estos pueden responder correctamente aunque no lo hayan entendido bien o de forma completa, como por ejemplo: «¿Qué número es más grande: 0,2 o 0,3?» Sería mejor preguntar qué número es mayor entre 0,3 y 0,25, ya que muchos escolares no tienen en cuenta la coma decimal y creen que 0,25 es mayor que 0,3 porque 25 es mayor que 3.

Además de formular preguntas útiles, el profesorado debe obtener respuestas de todos los estudiantes del grupo, y no únicamente de aquellos que levantan siempre la mano. Al fin y al cabo, es poco probable que conocer lo que piensan los estudiantes elocuentes y



seguros de sí mismos sea una buena indicación de lo que piensa el resto. Mejores pruebas llevan a mejores decisiones, que a su vez llevan a un mejor aprendizaje.

La tercera pregunta que me suelen hacer es sobre **el papel de la tecnología** y, en este caso, antes de nada, debo decir que soy escéptico. Muchos docentes utilizan sistemas electrónicos de votación o «pulsadores» para registrar las respuestas de cada alumno, pero

considero que estos sistemas no suelen ofrecer mucho más que los métodos más tradicionales y que, además, presentan una serie de desventajas.

Si un docente quiere tomar una decisión sobre qué hacer a continuación en una clase, me parece que una pregunta de opción múltiple bien diseñada a la que el alumno responda votando con los dedos (1 por A, 2 por B, etc.) es igual de eficaz que utilizar pulsadores y, como la mayoría del alumnado tiene dos manos, el docente puede incluir preguntas con dos respuestas correctas, algo que no sería posible con los pulsadores. Si hay varias respuestas correctas, se reduce al mínimo la probabilidad de que el alumnado adivine la respuesta correcta, y si una de las respuestas correctas es menos obvia que la otra, se consigue estimular y desafiar al alumnado de mayor rendimiento. He utilizado la votación con los dedos de forma eficaz con grupos de más de mil estudiantes, y me da una rápida idea de lo que piensa el alumnado.

Por supuesto, hay muchas otras formas en las que la tecnología podría ayudar al profesorado, pero creo que el potencial de estas herramientas —al menos a corto y medio plazo— es limitado por múltiples razones.

En primer lugar, tenemos que recordar que nuestro objetivo como docentes es lograr cambios a largo plazo en lo que el alumnado puede hacer: si pueden hacer lo que les hemos enseñado al final de la clase, pero no la semana que viene, entonces nuestra enseñanza no habrá sido especialmente exitosa. Es aquí donde entra en juego la distinción que hacen los psicólogos entre rendimiento y aprendizaje. El rendimiento se refiere a la forma en que los alumnos completan una tarea de aprendizaje y el aprendizaje es el cambio resultante a largo plazo. Esto es importante porque es posible que el alumnado complete una tarea instructiva sin ayuda y, sin embargo, que no aprenda lo que la tarea pretendía enseñarle. Podemos tener rendimiento sin aprendizaje. Y, lo que quizá resulte más extraño, también podemos tener aprendizaje sin rendimiento: el alumnado puede estar aprendiendo algo y, sin embargo, no dar pruebas de que algo haya cambiado. Por si esto no fuera suficientemente difícil, los trabajos de Robert Bjork y otros investigadores demuestran que el alumno suele aprender más cuando encuentra «dificultades deseables» para completar una tarea instructiva. Como indica Daniel Willingham, «la memoria es el residuo del pensamiento». La tecnología que se centra en el rendimiento del alumnado en una tarea concreta probablemente sea una mala referencia de lo que será eficaz a largo plazo.

En segundo lugar, en la actualidad contamos con numerosas pruebas de que la enseñanza que parece tener éxito a corto plazo puede ser menos eficaz, o incluso contraproducente, a largo plazo. El aprendizaje automático suele requerir grandes cantidades de datos y breves bucles de retroalimentación, que se obtienen fácilmente en juegos como el ajedrez o el go. Sin embargo, en la enseñanza, los bucles de retroalimentación pueden durar años, por lo que es poco probable que el aprendizaje automático sea de gran ayuda.

En tercer lugar, los tipos de modelos que necesitamos para que nos ayuden a comprender el pensamiento y el aprendizaje del alumnado son mucho más complejos y detallados de lo que podemos crear en la actualidad. Por ejemplo, para construir un modelo de aprendizaje del alumnado en el ámbito de las ciencias en secundaria que fuera útil para el profesorado, probablemente se necesitaría información sobre más de 400 capacidades específicas diferentes y se necesitarían numerosos datos para saber cuánto ha progresado un estudiante en cada una de esas capacidades. La cantidad de datos que se necesitarían para crear modelos útiles

Para las preguntas de respuesta elaborada, las minipizarras son muy eficaces; de hecho, a veces bromeo diciendo que las minipizarras son el avance más importante en tecnología educativa desde la pizarra: los mejores docentes ya las utilizaban hace cientos de años.

Lo que no me ofrecen las pruebas efímeras obtenidas mediante votación con los dedos o minipizarras es un registro de las respuestas del alumnado, pero eso me parece una característica, no un fallo. Al fin y al cabo, si queremos crear aulas en las que el alumno se sienta cómodo cometiendo errores, lo último que debemos hacer es registrar cada uno de ellos. Y como yo utilizo las respuestas del alumno para decidir en tiempo real qué hacer, el hecho de no tener un registro permanente no es un problema. De hecho, cualquier dato que recopilara probablemente quedaría desfasado en pocos minutos, sobre todo si estoy enseñando bien.



del pensamiento del alumnado superan con creces la cantidad de datos que probablemente tengamos a nuestra disposición.

En cuarto lugar, existe un problema ético en la recogida de datos sobre el aprendizaje del alumnado. Como ya se ha mencionado, los modelos útiles necesitarían que se recopilaran continuamente datos sobre el alumnado y, aunque algunos investigadores han estudiado este tipo de evaluación «encubierta», creo que es muy importante que el alumnado sepa cuándo se le está evaluando. Esto quedó perfectamente reflejado en una conversación que mantuve con un niño de 12 años llamado Lester al que di clases hace algunos años. En una clase sobre probabilidad, dijo que la probabilidad de que al lanzar una moneda saliera cara era del 50 %. Como era el capitán del equipo de fútbol de la escuela, le pregunté qué elegía al comienzo del partido cuando se lanzaba la moneda. Me dijo que «cruz». Le pregunté la razón y respondió que cruz salía más veces. Una evaluación encubierta de la conversación de Lester concluiría que no entendía la probabilidad, pero entiende perfectamente las matemáticas; simplemente no cree que se apliquen en la vida real.

Entonces, ¿qué hay que hacer?

Quizá lo mejor que pueden hacer los responsables educativos es mostrarse escépticos sobre el uso de la tecnología para mejorar el aprendizaje. La tecnología sin duda jugará un papel importante en nuestros esfuerzos por mejorar la educación, pero lo cierto es que, en palabras de Larry Cuban, la tecnología lleva más de 50 años «sobrevalorada e infrutilizada». Dada la larga trayectoria de grandes pretensiones y resultados decepcionantes, es necesario demostrar los efectos sostenidos en las aulas ordinarias antes de plantearse siquiera soluciones tecnológicas para mejorar la educación.

También es importante conocer qué problemas trata de solventar la tecnología educativa. Demasiadas veces, se recogen, analizan y presentan los datos del alumnado al profesorado con la esperanza de que haga algo útil con la información, sin tener una idea clara de cómo esa información puede mejorar la enseñanza. En vez de tomar decisiones en función de los datos, tenemos que pasar a recopilar datos en función de las decisiones, empezando por las decisiones que ya toma el profesorado y descubriendo qué datos pueden ayudarle a tomar mejor esas decisiones.

Por último, los responsables educativos deben ser conscientes de que la propia naturaleza de la enseñanza crea grandes dificultades para desarrollar soluciones tecnológicas eficaces, porque lo que aprendemos en un contexto determinado puede ser una mala referencia sobre lo que hay que hacer en otro contexto similar. Puede que en el futuro sea posible comprender y automatizar las decisiones de los expertos, pero en el futuro próximo es probable que el juicio profesional del profesorado experto sea superior a cualquier decisión que pueda tomar incluso la máquina más inteligente. Por este motivo, la adopción de la tecnología debe quedar relegada a un segundo plano frente al desarrollo de los conocimientos del profesorado, a través de un desarrollo profesional sostenido y centrado en la práctica.