**Que evaluar en su lugar**

**The Boston Globe**

**Una nueva ola de diseñadores de la prueba cree que pueden medir creatividad, resolución de problemas y la colaboración - y que un examen más inteligente podría cambiar la educación.**

Cuando la Universidad de Harvard anunció el mes pasado que estaba investigando a 125 estudiantes por hacer trampa en un examen para llevar a casa, la mayor parte del alboroto público que siguió estaba centrado en los alumnos: si eran estudiantes que metieron la pata antes las exigencias de una impredecible, como lo alegaban, o astuto triunfadores motivados a “cortar esquinas” por alguna combinación de ambición y pereza.

Pero más allá de la cuestión moral de los estudiantes de Harvard, había otro jugador en el drama: la prueba en sí.

El examen final, de acuerdo con las instrucciones, había sido "libro completamente abierto, nota abierta, acceso a Internet, etc." Lo único que prohibió a los estudiantes era trabajar juntos, requisito que algunos comentaristas, como el columnista de Slate Farhad Manjoo, argumentó como absurdo. Si el propósito de la educación es ayudar a los jóvenes a desarrollar habilidades que necesitarán en el futuro, Manjoo escribió después de que estallara el escándalo, no tiene sentido impedir arbitrariamente que se demuestren precisamente esas habilidades cuando se está tomando una prueba.

Ese punto no exime al acusado: Si había una regla y se rompió, hicieron trampa. Pero sí plantea una pregunta más profunda: ¿Qué esta la prueba tratando de lograr? ¿Qué es exactamente lo que queremos que nuestras pruebas prueben?

Tener éxito en el mundo de hoy, como todos reconocen ahora, requiere algo más que la capacidad de pensar con rapidez y recordar hechos en el comando. Y nuestro sistema educativo, sin embargo a ratos, se movió para hacer frente a esos valores. El problema es que nuestras pruebas aún están por detrás. Un examen final como la de Harvard uno, por ejemplo, los intentos para poner a prueba el ingenio de los estudiantes ya que les permite consultar libros y sitios web-pero entonces, porque todavía necesitan profesores para calificar el desempeño individual, dibuja una línea dura en el trabajo en grupos, que a su el mundo real es una parte importante del ingenio también. Que los estudiantes fueron tentados para usar estas habilidades ilícitas de puntos de colaboración en el problema: Hay un amplio espectro de habilidades cruciales pensamiento-creativo, resolución de problemas, la comunicación con los demás, que los educadores todavía están luchando para poner a prueba en un objetivo justo, , de manera eficiente.

Pruebas de reingeniería se ha convertido en una especie de convocatoria de un grupo de educadores e investigadores de todo el país. Con millones de dólares de fondos del gobierno federal, la Fundación Bill y Melinda Gates, así como de empresas como Cisco Systems, Intel y Microsoft, que han comenzado a repensar lo que un examen puede hacer, lo que puede parecer, y ¿Qué cualidades se puede evaluar. Con nuevas ideas pruebas como las simulaciones por ordenador, juegos y monitoreo sigilo, que están tratando de tomar lo que ellos creen que es

un enorme salto y necesario cambio de la prueba tal como la conocemos de una medida fija de lo que un estudiante puede recordar en un día en particular, a algo mucho más dinámico e informativo. Si la investigación da sus frutos, los estudiantes universitarios del futuro pueden encontrarse tomar pruebas de que aquellos de nosotros que recordamos libros azules y scantrons no se reconocen como pruebas en absoluto.

Los investigadores en la vanguardia del diseño de pruebas también tienen un sueño más grande, enraizada en la idea de que las pruebas no son sólo una parte estática de la educación, sino que activamente puede dar forma a lo que los profesores enseñan y lo que los estudiantes aprenden. Si realmente se puede construir más inteligente, pruebas más sofisticadas, dicen, puede cambiar la propia educación.

"[Pruebas son] la cola que mueve al perro", dijo David Williamson Shaffer, profesor de la Universidad de WisconsinMadison que estudia la psicología de la educación. "Y el problema es que tenemos la cola equivocada en estos momentos. Contamos con una cola que se desarrolló literalmente hace 100 años. "

Cuando la mayoría de nosotros imaginamos tomar una prueba, nos imaginamos a un estresante, de altas apuestas juicio en el que un profesor o un empleador nos pone en el lugar para averiguar lo buenos que son y lo mucho que conocemos. Pruebas de determinar los grados, rangos y títulos, son uno de los principales usos de las herramientas de la sociedad para determinar quién va dónde y quién puede hacer qué.

Pero aquellos cuyo trabajo es diseñar pruebas de entenderlos como algo más: una conjetura sobre el futuro. Cuando probamos, realmente estamos investigando las cualidades-la mezcla particular de ciertos conocimientos y habilidades, que nos dice un estudiante está listo para seguir adelante, o un empleado va a ser un activo para la empresa.

Tales predicciones requieren un sentido claro de las cualidades que una persona necesita para prosperar. Durante las últimas décadas, los educadores han cambiado de opinión acerca de lo que esas cualidades son, como habilidades como la colaboración se han vuelto más importante en la fuerza de trabajo. "La naturaleza de lo que el mundo está pidiendo está cambiando", dijo Randy Bennett, un científico de investigación en el ETS compañía de pruebas. Añadido a su colega, Robert Mislevy: "Hay trabajos sólo un número mucho menor en el que no estamos haciendo la búsqueda de información, interpretación, resolución de problemas y la comunicación que en el pasado".

Cuando se trata de medir las cosas, es evidente para los investigadores que las pruebas escritas tradicionales-incluso aquellas que son más de forma libre

y no a libro abierto, son a la altura. "¿Crees que de verdadero / falso, selección múltiple, tal vez a juego y tal vez algo de llenar los espacios en blanco y tal vez una pregunta de desarrollo", dijo Valerie Shute, profesor de psicología de la educación y de aprendizaje en la Florida State University. "Ese tipo de cosas son extraordinariamente estrecho en cuanto al alcance de lo que se puede conseguir."

Para ampliar el alcance, investigadores como Shute está tratando de ingeniero de pruebas-que prefieren el término "evaluaciones"-que requieren las personas para ejercer un conjunto de habilidades complejas, al mismo tiempo, no sólo regurgitar información, sino utilizarla para resolver problemas. Kathleen Scalise, profesor asociado en la Universidad de Oregon que estudia cómo las computadoras pueden utilizarse en el aprendizaje, ha trazado una taxonomía de probar innovaciones que incluye una amplia gama de tipos no tradicionales de preguntas. Una pregunta a los estudiantes para mover un par de luces de la calle alrededor de modo que la mujer que aparece en la pantalla proyecta dos sombras, otra muestra 15 burbujas que contienen palabras como "congresistas", "Presidente", "Corte Suprema" y "jueces", y pide estudiantes a los conectan entre sí mediante flechas y organizarlos en grupos. En la progresiva Misión Nueva York la escuela para aprender, los estudiantes se colocan en pequeños grupos y encargó la construcción de máquinas de Rube Goldberg para demostrar una comprensión de la física básica y la capacidad de trabajar en equipo.

Pero los investigadores en el campo de diseño de la prueba están invirtiendo más energía en la elaboración de los programas de ordenador que se aprovechan de la evaluación de sigilo llamada, un método de juzgar a los examinados sin decirles exactamente lo que está siendo juzgado. Jody ClarkeMidura y Chris Dede en Harvard, por ejemplo, desarrolló un videojuego 3D para poner a prueba las habilidades científicas: En una evaluación, los agricultores de una aldea encontrar una rana muerta con seis patas, y los estudiantes deben ayudarles a descubrir la causa de la aberración por a pie de finca en finca la recopilación de datos. Para probar la comprensión de los estudiantes de la física, por su parte, Valerie Shute y su equipo diseñaron un juego de ordenador llamado Patio de Newton. El programa ofrece a los estudiantes un problema y les pide que utilice su ratón para dibujar herramientas, tales como péndulos, rampas y palancas, con el fin de resolverlo. Obtención de los asuntos respuesta correcta, pero al ver cómo los niños van a tratar de resolver los problemas, la prueba también puede evaluar si han aprendido conceptos importantes como la inercia y la fuerza, así como su capacidad de ser persistente y pensar creativamente. La prueba, en otras palabras, se abre una ventana a la forma en la mente del estudiante realmente funciona.

El avance clave conceptual incrustado en las pruebas de sigilo, como juegos de Newton es que los estudiantes ya no se está evaluando sólo en sus respuestas finales, pero el proceso que pasar para atacar un problema. Ese es también el principio que impulsa el trabajo de David Shaffer, quien ha construido un juego multijugador sobre el diseño de una ciudad que los estudiantes calificaciones en sus habilidades de colaboración mediante el control y registro de sus interacciones. "Realiza un seguimiento de todo lo que hacen", dijo Shaffer. "Cada correo electrónico que envían, cada vez que se comunica con alguien a través del sistema de chat, todos los informes que escriben, cada página que miran .... Eso es una gran cantidad de información que podemos utilizar para averiguar qué tan bien están haciendo lo que están haciendo. "

Pero si Shaffer y otros de nueva generación diseñadores de la prueba comparten un sueño de sustitución de lápiz y papel con exámenes orientado a los procesos de resolución de problemas ejercicios, también comparten un complicado desafío: Las habilidades que están tratando de medir son mucho más difíciles de detectar cuantificar y que, por ejemplo, si alguien conoce la fórmula cuadrática. "No es justo que [las] ​​habilidades complejas son más difíciles de aislar-es que no existen de manera aislada", dijo Shaffer. El desglose de estas habilidades multifacéticas en cualidades comprobables es difícil, y es algo que los educadores han estado tratando de hacer y de no más de medio siglo.

El primer presidente de ETS, que durante mucho tiempo administra el SAT, establecido en 1948 para desarrollar una prueba que pueda evaluar a un estudiante de la resistencia intelectual habilidad de llevarse bien con los demás, y así sucesivamente, pero la compañía finalmente llegó a la conclusión que era demasiado difícil medir de una manera fiable. Más recientemente, a finales de 1980 y 90, la Universidad de Harvard Howard Gardner psicólogo del desarrollo participaron en un esfuerzo por diseñar nuevos tipos de pruebas en las humanidades que podrían ser calificadas objetivamente. En última instancia, se encontró con que el matiz necesario para medir las habilidades blandas chocaron con las exigencias de la normalización. Cuando una prueba tiene que comparar con fiabilidad los estudiantes de todas las escuelas y distritos, "hay una presión para simplificar, tener acorazado rúbricas, esencialmente avanzar hacia opciones múltiples", escribió en un correo electrónico Gardner.

La tecnología ha abierto la puerta a aflojar esas demandas, como las computadoras son ahora capaces de rastrear todo lo que una persona hace al tratar de resolver un problema y lo traduce en números. Pero eso todavía deja abierta la pregunta más difícil de lo que la mezcla de los correos electrónicos y las pulsaciones de teclado y los clics nos dice que alguien es, digamos, un buen colaborador, y lo que constituye una señal de alerta. Diseñadores de prueba tiene algunas maneras de tratar de averiguarlo.

Uno de ellos implica tener grupos de principiantes y expertos completar una simulación, y la construcción de un sistema de clasificación basado en la forma en que sus enfoques son diferentes; otro implica actuaciones coincidentes en una prueba experimental con un largo tiempo de uso intensivo de evaluación de personal. Si eso suena difícil y caro bien, hay una razón de opción múltiple ha durado tanto tiempo.

"[Históricamente], la industria de las pruebas, porque era pragmático, sólo prueba lo que era fácil de probar", dijo James Paul Gee, profesor de la Universidad Estatal de Arizona que diseña las pruebas que toman la forma de juegos. "Pero como padre, no quiero que pruebes sólo lo que es fácil de probar, quiero probar lo que es importante para poner a prueba".

A medida que crecemos, las pruebas que damos forma a nuestra comprensión de lo que importa en el mundo. Usted sabe que el control de su punto ciego cuando se cambia de carril es importante en parte porque si no lo hacemos, el tipo DMV no le pase en el examen de conducir. O pensemos en el MCAS, que los estudiantes de secundaria en Massachusetts debe pasar para graduarse: Cubre sólo matemáticas Inglés, y la ciencia, lo que refleja-y en última instancia, hacer cumplir-la creencia de vuelta a lo básico que sean competentes en esas áreas es la base de una educación, mientras que una comprensión de la historia, educación cívica, o las artes no lo es. De un modo muy directo, entonces, una prueba expresa un conjunto de valores, una visión de cómo queremos que sea la gente. "Una prueba se convierte en un cartel", dijo Randy Bennett. "Se convierte en un ejemplo de lo que luchar."

Esto es lo que motiva a muchos diseñadores de prueba, cuyo objetivo final es ampliar lo que se enseña en las escuelas mediante la ampliación de lo que se puede probar. Si tuviéramos mejores pruebas, se imaginan-los que realmente podrían evaluar habilidades como la creatividad y la colaboración-de repente la puerta se abría a una visión de la educación en la que aparentemente habilidades "blandas" se puede enseñar con el mismo rigor que lo que consideramos lo básico.

Así las cosas como son, nos erizan ante la idea de los maestros sentirse presionados a "enseñar para la prueba", por temor a sus alumnos de lo contrario se obtienen malas calificaciones en escala nacional obligatorios los exámenes estandarizados. Pero ¿y si la enseñanza para la prueba no tiene por qué significar mecánicamente hijos de capacitación para poner en marcha las respuestas a las preguntas inventadas? Esa es la promesa de una mejor prueba: Al dibujar un mapa que refleja con mayor precisión nuestro mundo, podemos descubrir los caminos más prometedores para llegar a donde queremos ir.