

jos resultados de lectura se vería sometida a un período de prueba y se enfrentaría a la amenaza de cierre, con el consiguiente despido o redesignación del personal. El CPS también suprimió lo que se conoce como promoción social. En el pasado, sólo un estudiante completamente inepto o difícil repetía curso. Ahora, para pasar de curso, todo estudiante en tercero, sexto y octavo debe lograr una puntuación mínima en el examen estandarizado de tipo test conocido como Examen de Aptitudes Básicas Iowa.

Los que defienden las pruebas de alto índice argumentan que éstas elevan los parámetros de aprendizaje y proporcionan a los alumnos un mayor aliciente para estudiar. Asimismo, si el examen evita que los malos estudiantes avancen sin hacer mérito para ello, éstos no entorpecerán cursos superiores y lentificarán a los buenos alumnos. Entretanto, a quienes se oponen a ello les preocupa que determinados estudiantes se vean injustamente penalizados si resulta que no obtienen buenos resultados en el examen, y que los profesores se concentren en los temas para el examen en detrimento de lecciones más importantes.

Los estudiantes, por supuesto, han dispuesto de alicientes para hacer trampas desde que existen los exámenes. Pero las pruebas de alto índice han cambiado los incentivos para los profesores de un modo tan radical que ahora son ellos quienes tienen un motivo añadido para engañar. Con las pruebas de alto índice, un profesor cuyos alumnos obtienen malos resultados puede ver cómo se lo censura o se le niega un aumento de sueldo o un ascenso. Si la escuela en su totalidad obtiene malos resultados, corre el riesgo de que se le retiren fondos federales; si la escuela entra en período de prueba, el profesor se enfrenta al despido. Las pruebas de alto índice también plantean a los profesores algunos incentivos positivos. Si sus alumnos lo hacen lo bastante bien, recibirán elogios, ascensos y recompensas económicas: en un momento dado el estado de California introdujo bonificaciones de 25.000 dólares para los profesores que produjeran grandes aumentos en las calificaciones de las pruebas de alto índice.

Y si un profesor resuelve examinar este nuevo panorama incentivado y considerar de algún modo inflar los resultados de sus alumnos, contaría, para convencerse de ello, con un incentivo final: el engaño llevado a cabo por profesores rara vez se persigue, difícilmente se detecta y casi nunca se castiga.

¿Qué podría hacer un profesor para engañar? Existen infinidad de posibilidades, desde el descarado hasta la sutileza. Recientemente, una alumna de quinto curso de Oakland regresó a casa y muy suelta de lengua le dijo a su madre que su superamable profesor había escrito las respuestas del examen estatal directamente en la pizarra. Por supuesto, tales ejemplos son raros, porque poner su destino en manos de treinta testigos prepubescentes parece un riesgo que ni el peor profesor correría. (Como era de esperar, el profesor de Oakland fue despedido.) Existen formas más sutiles de inflar los resultados de los estudiantes. Un profesor puede sencillamente dar más tiempo a los alumnos para realizar el examen. Si precisamente obtiene una copia del examen —es decir, por medios legítimos—, tiene la posibilidad de prepararlos para preguntas concretas. En términos más generales, puede «enseñar el examen» basando sus lecciones en preguntas de años anteriores, lo cual no se considera engañar, pero puede violar perfectamente el espíritu de la prueba. Puesto que estos exámenes son de opciones múltiples, sin penalización alguna por las respuestas incorrectas, un profesor podría instruir a sus alumnos para rellenar al azar los blancos cuando el tiempo apremie, quizás introduciendo una larga serie de B o un patrón que alterne B y C. Incluso podría rellenar los blancos por ellos cuando abandonasen la clase.

Pero si un profesor realmente deseara hacer trampas —y mereciese la pena hacerlo— podría recoger los exámenes de sus alumnos y, en la hora aproximada antes de entregarlos para ser corregidos mediante un escáner electrónico, borrar las respuestas erróneas e introducir las correctas. (Pensar que usted siempre creyó que el lápiz del número 2 era para que «los niños» modificasen sus respuestas.) Si este tipo de trampas por parte de los profesores realmente está produciéndose, ¿cómo podría detectarse?

Para atrapar a un tramposo resulta muy útil pensar como él. Si usted pretendiese cambiar las respuestas erróneas de sus alumnos e introducir las correctas, lo más probable es que no las cambiase todas. Ése sería un indicio claro. Tampoco querría modificar respuestas en los exámenes de todos los alumnos (otro indicio). Lo más seguro es que no tuviese tiempo suficiente, porque las hojas de examen deben entregarse poco después de que éste haya terminado. De modo que lo que podría hacer es seleccionar una serie de

ocho o diez preguntas consecutivas e introducir las respuestas correctas en las hojas de, digamos, la mitad o dos tercios de sus alumnos. Podría memorizar con facilidad un breve patrón de respuestas correctas, y resultaría mucho más rápido borrar y cambiar ese patrón que revisar cada hoja de respuestas por separado. Incluso podría pensar en enfocar su actividad hacia las últimas preguntas, que suelen ser más difíciles que las primeras. De esa forma, tendría más probabilidades de sustituir respuestas erróneas por correctas.

Si la economía es una ciencia que básicamente se ocupa de los incentivos, también es —afortunadamente— una ciencia con herramientas estadísticas para calcular cómo responden las personas a esos incentivos. Lo único que se necesita es algunos datos.

En este caso, se ofreció el sistema de Escuelas Públicas de Chicago, que dispuso una base de datos de las respuestas de los exámenes de cada alumno del CPS desde tercero hasta séptimo y desde 1993 hasta 2000. Esto significa aproximadamente treinta mil estudiantes por curso y año, más de setecientas mil series de respuestas y casi cien millones de respuestas individuales. Los datos, organizados por clase, incluían la serie de respuestas, pregunta por pregunta, de cada estudiante en los exámenes de lectura y matemáticas. (Las hojas de respuestas de papel reales no se incluían; generalmente se destruían poco después del examen.) Los datos también aportaban alguna información acerca de cada profesor e información demográfica de cada estudiante, además de sus resultados pasados y futuros, un elemento clave para detectar el engaño del profesor.

Entonces había que idear un algoritmo capaz de extraer algunas conclusiones de esta masa de datos. ¿Cuál sería el aspecto de la clase de un profesor tramposo?

En primer lugar habría que buscar patrones de respuestas inusuales en una clase determinada: bloques de respuestas idénticas, por ejemplo, especialmente entre las preguntas más difíciles. Si diez alumnos muy brillantes (como indican sus resultados pasados y futuros) proporcionaron las respuestas correctas a las primeras cinco preguntas del examen (generalmente las más fáciles), un bloque idéntico no debería considerarse sospechoso. Pero si diez estudiantes malos respondieron correctamente a las cinco últimas preguntas del examen (las más difíciles), merece la pena examinarlo. Otra clara señal sería un patrón extraño dentro del examen de un estu-

diente —como acertar las respuestas difíciles y fallar las fáciles—, especialmente si se compara con los miles de estudiantes de otras clases con similares puntuaciones en el mismo examen. Además, el algoritmo buscaría una clase llena de alumnos cuyos resultados fueran mucho mejores de lo que podía suponerse por sus puntuaciones anteriores y que al año siguiente empeoraran de manera considerable. Un pico pronunciado en un año en los resultados de los exámenes podía atribuirse inicialmente a un buen profesor; pero si iba seguido de una caída brusca, era muy probable que el pico procediera de medios artificiales.

Consideremos ahora las series de respuestas de los alumnos de dos clases de sexto curso de Chicago que realizaron exactamente el mismo examen de matemáticas. Cada fila horizontal representa las respuestas de un alumno. Las letras *a*, *b*, *c* o *d* indican una respuesta correcta; un número indica una respuesta errónea, con el 1 correspondiente a *a*, el 2 a *b*, etcétera. Un 0 representa una respuesta en blanco. Una de estas clases, casi sin duda, tenía un profesor tramposo, mientras que la otra no. Intente distinguirlas, aunque le advierto de que no resulta fácil a simple vista.

Clase A

```
112a4a342cb214d0001acd24a3a12dadbc4a0000000
d4a2341cacbddad3142a2344a2ac23421c00adb4b3cb
1b2a34d4ac42d23b141acd24a3a12dadbc4a2134141
dbaab3dcacb1dadbc42ac2cc31012dadbc4adb40000
d12443d43232d32323c213c22d2c23234c332db4b300
db2abad1acbdda212b1acd24a3a12dadbc400000000
d4aab2124cbddadbcb1a42cca3412dadbc423134bc1
1b33b4d4a2b1dadbc3ca22c00000000000000000000
d43a3a24acb1d32b412acd24a3a12dadbc422143bc0
313a3ad1ac3d2a23431223c000012dadbc400000000
db2a33dcacbd32d313c21142323cc300000000000000
d43ab4d1ac3dd43421240d24a3a12dadbc400000000
db223a24acb11a3b24cacd12a241cdadbcb4adb4b300
db4abadcacb1dad3141ac212a3a1c3a144ba2db41b43
1142340c2cbddadb4b1acd24a3a12dadbc43d133bc4
```

214ab4dc4cbdd31b1b2213c4ad412dadbbcb4adb00000
 1423b4d4a23d24131413234123a243a2413a21441343
 3b3ab4d14c3d2ad4cbcac1c003a12dadbbcb4adb40000
 dba2ba21ac3d2ad3c4c4cd40a3a12dadbbcb400000000
 d122ba2cacbd1a13211a2d02a2412d0dbcb4adb4b3c0
 144a3adc4cbddadbbcb2c2cc43a12dadbbcb4211ab343
 d43aba3cacbddadbbcbca42c2a3212dadbbcb42344b3cb

Clase B

db3a431422bd131b4413cd422a1acda332342d3ab4c4
 d1aa1a11acb2d3dbc1ca22c23242c3a142b3adb243c1
 d42a12d2a4b1d32b21ca2312a3411d00000000000000
 3b2a34344c32d21b1123cdc00000000000000000000
 34aabad12cbdd3d4c1ca112cad2ccd000000000000000
 d33a3431a2b2d2d44b2acd2cad2c2223b40000000000
 23aa32d2a1bd2431141342c13d212d233c34a3b3b000
 d32234d4a1bdd23b242a22c2a1a1cda2b1baa33a0000
 d3aab23c4cbddadb23c322c2a22223232b443b24bc3
 d13a14313c31d42b14c421c42332cd2242b3433a3343
 d13a3ad122b1da2b11242dc1a3a12100000000000000
 d12a3ad1a13d23d3cb2a21ccada24d2131b440000000
 314a133c4cbd142141ca424cad34c122413223ba4b40
 d42a3adcacbddadbbcb42ac2c2ada2cda341baa3b24321
 db1134dc2cb2dadb24c412c1ada2c3a341ba20000000
 d1341431acbddad3c4c213412da22d3d1132a1344b1b
 1ba41a21a1b2dadb24ca22c1ada2cd32413200000000
 dbaa33d2a2bddadbbcbca11c2a2accda1b2ba20000000

Si ha adivinado que la clase A es la que hizo trampas, enhora-
 buena. A continuación aparecen de nuevo las series de respuestas
 de la clase A reordenadas ahora por un ordenador que ha aplica-
 do el algoritmo falseado y señalado patrones sospechosos.

Clase A

(Con la aplicación del algoritmo)

1. 112a4a342cb214d0001acd24a3a12dadbbcb4a0000000
2. 1b2a34d4ac42d23b141acd24a3a12dadbbcb4a2134141
3. db2abad1acbdda212b1acd24a3a12dadbbcb400000000
4. d43a3a24acb1d32b412acd24a3a12dadbbcb422143bc0
5. 1142340c2cbddadb4b1acd24a3a12dadbbcb43d133bc4
6. d43ab4d1ac3dd43421240d24a3a12dadbbcb400000000
7. dba2ba21ac3d2ad3c4c4cd40a3a12dadbbcb400000000
8. 144a3adc4cbddadbbcb2c2cc43a12dadbbcb4211ab343
9. 3b3ab4d14c3d2ad4cbcac1c003a12dadbbcb4adb40000
10. d43aba3cacbddadbbcbca42c2a3212dadbbcb42344b3cb
11. 214ab4dc4cbdd31b1b2213c4ad412dadbbcb4adb00000
12. 313a3ad1ac3d2a23431223c000012dadbbcb400000000
13. d4aab2124cbddadbbcb1a42cca3412dadbbcb423134bc1
14. dbaab3dcacbd1dadbbcb42ac2cc31012dadbbcb4adb40000
15. db223a24acb11a3b24cacd12a241cdadbbcb4adb4b300
16. d122ba2cacbd1a13211a2d02a2412d0dbcb4adb4b3c0
17. 1423b4d4a23d24131413234123a243a2413a21441343
18. db4abadcacbd1dad3141ac212a3a1c3a144ba2db41b43
19. db2a33dcacbd32d313c21142323cc3000000000000000
20. 1b33b4d4a2b1dadbbcb3ca22c00000000000000000000
21. d12443d43232d32323c213c22d2c23234c332db4b300
22. d4a2341cacbddad3142a2344a2ac23421c00adb4b3cb

Observe las respuestas en negrita. ¿Consiguieron quince de veintidós alumnos recitar de un tirón las mismas seis respuestas consecutivas correctas (la serie d-a-d-b-c-b) solos?

Existen al menos cuatro razones para que esto resulte improbable. Una: esas preguntas, cerca del final del examen, eran más difíciles que las primeras. Dos: se trataba principalmente de alumnos con un nivel inferior, de los cuales pocos consiguieron seis respuestas correctas consecutivas en otra parte del examen, lo que hace aún más improbable que acertaran las mismas seis preguntas difíciles. Tres: hasta esta parte del examen las respuestas virtualmente eran no correlativas. Cuatro: tres de los alumnos (números 1, 9 y 12) dejaron más de una respuesta en blanco antes de la serie sospechosa y terminaron el examen con otra serie de respuestas en blanco. Esto sugiere que se rompió una larga serie ininterrumpida de respuestas en blanco, y no por parte del alumno, sino del profesor.

Existe otro aspecto extraño en torno a la serie de respuestas sospechosa. En nueve de los quince exámenes, las seis respuestas correctas van precedidas por otra serie idéntica, 3-a-1-2, que incluye tres respuestas incorrectas de cuatro. Y en los quince exámenes, las seis respuestas correctas van seguidas de la misma respuesta incorrecta, un 4. ¿Por qué demonios se preocuparía un profesor tramposo de borrar la hoja de respuestas de un alumno e introducir la respuesta errónea?

Quizá se trata de simple estrategia. En el caso de que se lo descubriera y se lo condujera al despacho del director, podría señalar las respuestas incorrectas como prueba de que no falsificó los exámenes. O quizás —y ésta es una respuesta menos benévola, pero igual de probable— él mismo desconoce la respuesta correcta. (Con los exámenes estandarizados, lo habitual es que el profesor no tenga una clave para la respuesta.) Si éste es el caso, entonces disponemos de una pista bastante sólida de por qué sus alumnos necesitan notas infladas para empezar: tienen un mal profesor.

Otro indicador de las trampas en la clase A es el resultado general de la clase. Cuando el octavo mes del curso académico los alumnos de octavo realizaron el examen, necesitaban alcanzar una puntuación media de 6,8 para considerarse al nivel de los estándares nacionales. (Los alumnos de quinto que realizan el examen el

octavo mes necesitan una nota de 5,8, los de séptimo 7,8, etcétera.) La nota media de los alumnos de la clase A en los exámenes de sexto fue de 5,8, lo que significa que todo el curso se hallaba por debajo del nivel exigible. Así que, obviamente, se trata de malos estudiantes. Un año antes, no obstante, estos mismos estudiantes lo hicieron aún peor, obteniendo una media de sólo 4,1 en los exámenes de quinto. En lugar de mejorar por un punto entero entre quinto y sexto, como cabría esperar, mejoraron en un 1,7, el valor de casi dos cursos. Pero esta milagrosa mejora resultó pasajera. Cuando llegaron a séptimo, tuvieron de media un 5,5, más de dos niveles por debajo del estándar y peor incluso que en sexto. Obsérvense las calificaciones irregulares año tras año de tres alumnos en particular de la clase A:

	<i>Nota de quinto curso</i>	<i>Nota de sexto curso</i>	<i>Nota de séptimo curso</i>
Alumno 3	3,0	6,5	5,1
Alumno 5	3,6	6,3	4,9
Alumno 14	3,8	7,1	5,6

Las puntuaciones de tres años de la clase B, por otro lado, también son malas, pero al menos indican un esfuerzo sincero: 4,2, 5,1 y 6. De modo que una clase entera de niños, la clase A, de repente se volvió muy inteligente un año y muy tonta al siguiente, o más probablemente, su profesor de sexto consiguió realizar algún truco de magia con su lápiz.

Cabe destacar dos puntos acerca de los alumnos de la clase A, al margen del engaño en sí. El primero es que evidentemente se hallan en una forma académica baja, lo que los convierte en los estudiantes a quienes se supone que más ayudan las pruebas de alto índice. El segundo punto es que esos alumnos (y sus padres) sufrirían un *shock* terrible al llegar a séptimo. Todo lo que sabían era que habían pasado de curso con éxito debido a los resultados de sus exámenes. (Ningún niño queda atrás, en efecto.) No fueron ellos quienes aumentaron artificialmente sus calificaciones; probablemente esperaban hacerlo muy bien en séptimo, y fracasaron penosamente. Éste es quizás el aspecto más cruel de las pruebas de alto