

Evaluaciones Sumativa y Formativa en Matemática

Apoyo a las Metas de los Estándares de Núcleo Curricular

Común (Common Core Standards)

Alan H. Schoenfeld

University of California, Berkeley

Alan H. Schoenfeld tiene la cátedra Elizabeth and Edward Conner Professor of Education y es profesor afiliado de Matemática en la Universidad de California, en Berkeley. Su dirección electrónica es [alans@berkeley.edu](mailto:alans@berkeley.edu).

## Resumen

La excelencia en matemáticas implica mucho más que el dominio de hechos y de procedimientos. Implica un buen conocimiento de la disciplina, incluyendo, además, acceso a estrategias de resolución de problemas; disposición a la productividad y creencias específicas del dominio de conocimiento; ser un pensador estratégico; ser apropiadamente metacognitivo. Está amplia gama de metas matemáticas en una parte central de los Estándares de Núcleo Curricular Común (CCS por sus siglas en inglés).

En el contexto de las pruebas estandarizadas (*high-stakes examinations*), los profesores centran su atención en lo que se examina; por tanto, resulta de importancia crítica para las evaluaciones producidas por los dos consorcios nacionales, the Smarter Balanced Assessment Consortium (SBAC) y el Partnership for Assessment of Readiness for College and Careers (PARCC), que reflejen realmente los valores de los Estándares de Núcleo Común y que proporcionen a los profesores asistencia en el aula para lograr tales estándares. Se analizan las especificaciones sobre las pruebas y los reactivos-muestra de evaluación, y también la posibilidad de que las evaluaciones funcionen como palancas positivas del cambio.

Durante más de 20 años, el Proyecto de Evaluación en Matemática (Mathematics Assessment Project) ha centrado su atención en el desarrollo de evaluaciones formativas que ponen énfasis en los procesos y las prácticas matemáticas analizados párrafos arriba. Más recientemente, el proyecto se ha enfocado en la creación de *lecciones de evaluación formativa* (FAL) para ayudar a los profesores a construir la comprensión de los estudiantes mediante el uso de evaluación formativa. En el presente artículo se describe nuestro trabajo reciente. En él se ilustra el andamiaje que proporcionan nuestras lecciones para ayudar al profesor a plantear actividades ricas y a prepararse para responder a la gama de concepciones estudiantiles que es probable que manifiesten sus estudiantes.

## Introducción

Estados Unidos de Norteamérica (EU) se encuentra en el umbral de cambios significativos en la evaluación en matemática, en términos de qué tipo de entendimientos son evaluados y en términos de la creciente homogeneidad de la evaluación en matemática en el nivel nacional. Tales cambios reflejan la continua evolución del “movimiento de estándares” que se puede rastrear hasta la producción del Consejo Nacional de Profesores de Matemática (NCTM) de los *Estándares Curriculares y de Evaluación para la Matemática Escolar* en 1989 y hasta el cambio radical en el contexto de la responsabilidad de alto riesgo que significó el programa gubernamental “No Child Left Behind” (Que Ningún Niño se Quede Atrás) aprobado por el Congreso en 2001; tienen el potencial de tener consecuencias reales para la educación matemática en Estados Unidos. Como lo explico más adelante, en pocos años la gran mayoría de los estudiantes estadounidenses estarán haciendo uno de dos exámenes de alto riesgo (“high stakes examinations” se refiere a exámenes estandarizados cuyos resultados tienen repercusiones en el presupuesto de las escuelas y en las percepciones económicas de los profesores); ambos con la intención de representar los valores matemáticos definidos en los Estándares Nacionales de Núcleo Curricular Común para Matemática o CCSSM (por sus siglas en inglés; Common Core State Standards Initiative, 2010). En la medida en que las evaluaciones representen los valores del CCSSM y en el grado en que la evaluación de alto riesgo guíe la docencia, la enseñanza de la matemática en EU estará mucho más enfocada y será mucho más coherente que en los últimos 25 años.

En esta introducción describo brevemente la doble evolución del contexto de los estándares en matemática y de los exámenes nacionales. Con tal contexto establecido,

examinaremos, entonces, algunos reactivos típicos de examen actuales, y algunos de los reactivos que representan a las evaluaciones que se están produciendo por parte de los dos consorcios nacionales de evaluación, the Smarter Balanced Assessment Consortium (SBAC. Consorcio de Evaluación Equilibrada más Inteligente) y el Partnership for Assessment of Readiness for College and Careers (PARCC, Asociación para la Evaluación de la Preparación para la Universidad y las Carreras [Profesionales]). Como veremos, todavía quedan pendientes algunas cuestiones de alineación con los CCSSM, pero, suponiendo que esto se puede arreglar, las nuevas evaluaciones auguran un cambio significativo. Siendo éste el caso, la cuestión es cómo preparar a los estudiantes y a los profesores para dicho cambio. Una manera de hacerlo es explorada por el Proyecto de Evaluación en Matemática (MAP, por sus siglas en inglés) que se ha dado a la tarea de producir 100 *Lecciones de Evaluación Formativa* dirigidas a apoyar prácticas de clase consistentes con los CCSSM. Se describe una lección muestra y se analizan algunos de retos prácticos que enfrentan los educadores matemáticos.

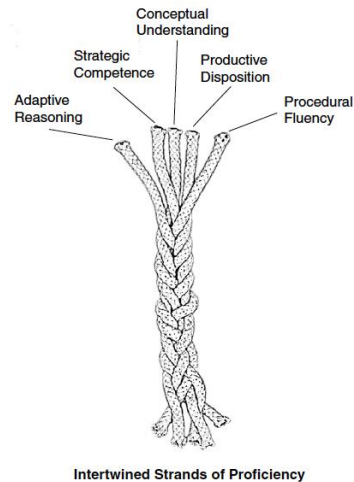
### ***La evolución de los estándares, 1975-2010***

El “movimiento de estándares” inició cuando el NCTM, como reacción al caos curricular que había en todo el país y ante la percepción de que EU se estaba quedando significativamente atrás con respecto a los demás países industrializados en la formación de estudiantes en Ciencia, Tecnología, Ingeniería y Matemática (STEM, por sus siglas en inglés; véase, por ejemplo, el informe de la Comisión Nacional para la Excelencia en Educación de 1983, *Una Nación en Riesgo*), produjo un volumen de estándares matemáticos que representan la visión del NCTM de un currículo y de evaluaciones de alta calidad. Los Estándares de 1989 fueron revolucionarios en el aspecto de que fue el

primer documento curricular nacional que puso una significativa atención a los *procesos* matemáticos (que uno sigue cuando se enrola en tareas matemáticas) así como en los *contenidos* (por ejemplo, álgebra, geometría, probabilidad, estadística y análisis de datos, entre otros). En los *Estándares* se declaraba que en cada intervalo de grados (K-5, 5-8 y 9-12; K se refiere al nivel preescolar) se debía poner especial atención a la matemática como herramienta de resolución de problemas, de comunicación y de razonamiento, y a las conexiones de la matemática. Este enfoque fue confirmado en el volumen de 2000 del NCTM, *Principios y Estándares para la Matemática Escolar*, en donde se reiteró el énfasis en los procesos matemáticos de los *estándares* originales, y en los cuales se agregó un punto de atención en las representaciones matemáticas.

Esta visión “orientada a procesos” de la matemática está fundamentada en décadas de investigación en pensamiento matemático y resolución de problemas, cuyo énfasis está no sólo en el contenido de la matemática, sino en las formas productivas en que la gente aprende y utiliza su conocimiento matemático. Por ejemplo, Schoenfeld (1985, 1992) caracterizó a alguien con competencia matemática (que, presumiblemente, es el objetivo de la educación matemática) como un individuo que, además de tener una cantidad sustantiva de conocimiento matemático, tiene también acceso a estrategias productivas de resolución de problemas, ha aprendido a ser efectivo, en términos metacognitivos, (al usar sus recursos de manera eficiente), y que ha desarrollado creencias productivas y una disposición hacia la matemática y sus propias prácticas como hacedor de matemática. Al enfocarse en los niveles elementales (con implicaciones en todos los niveles), el volumen de 2001 del Consejo Nacional de Investigación, *A sumar*

(*Adding it up*), considera la competencia matemática como formada por cinco ejes que se entrelazan (véase Figura 1).



**Figura 1.** (Consejo Nacional de Investigación, 2001, p. 5) Adaptive Reasoning = adaptable; Strategic Competence = Competencia estratégica; Conceptual Understanding = Entendimiento conceptual; Productive Disposition = Disposición productiva; Procedural Fluency = Fluidez procedimental.

Los *Estándares de Núcleo Curricular Común (Common Core Standards)* representan la evolución natural del trabajo analizado en párrafos anteriores. Dado el contexto de su creación véase más adelante), existen algunas diferencias con éste; pero, sobre todo, uno observa una especificación más detallada en los contenidos y en el continuo énfasis en los procesos, referidos ahora como *prácticas*. Tal vez el cambio más importante sea que en la descripción de contenidos del CCSSM, éstos se especifican para cada grado hasta el octavo, en lugar de hacerlo en bandas amplias de grados. Esto significa que las progresiones de contenidos están mucho más detalladas que antes, por ejemplo, las especificaciones de números racionales y de fracciones se presentan de manera que conducen a discusiones de proporcionalidad y de funciones lineales (los autores de CCSSM afirman que el enfoque y la coherencia son las principales virtudes de los *Estándares*). Esto tendrá consecuencias significativas para el currículo y para la

evaluación, en el sentido de que ambos están mucho más determinados que antes (por ejemplo, “el contenido de séptimo grado” se enseñará y evaluará en el séptimo grado, mientras que podría ser que antes ése contenido se enseñara o se evaluara en sexto u octavo grados, como parte de la “banda de grados”) . Lo que en documentos anteriores era referido como “procesos”, en los CCSSM se refieren a “Estándares para la Práctica Matemática”. Sin embargo, los autores ponen en claro su fuente:

“Los Estándares para la Práctica Matemática... descansan en importantes *procesos* y *competencias* con una larga importancia en educación matemática. Los primeros son los estándares de procedimiento del NCTM sobre resolución de problemas, razonamiento y demostración, comunicación, representación y conexiones. Los segundos son los ejes de competencia matemática especificados en el texto *A sumar*, del Consejo Nacional de Investigación: razonamiento adaptable, competencia estratégica, entendimiento conceptual (comprensión de conceptos, operaciones y relaciones matemáticos), fluidez procedimental (habilidad en desarrollar procedimientos de maneras flexible, precisa, eficiente y apropiada), y disposición productiva (inclinación habitual para ver la matemática como sensible, útil y valiosa, junto con una creencia en la diligencia de la propia eficiencia).”

Las ocho prácticas matemáticas resaltadas en los CCSSM son que los estudiantes:

1. Le encuentren sentido a los problemas y perseveren en resolverlos.
2. Razonen de maneras abstracta y cuantitativa.
3. Construya argumentos viables...
4. Modele con la matemática.
5. Use estratégicamente las herramientas apropiadas.

6. Atienda a la precisión.
7. Busque estructuras y las utilice.
8. Busque y exprese regularidades al usar su razonamiento repetidas veces.

El reto presente y futuro de la evaluación se puede expresar como: ¿Es posible capturar confiablemente la habilidad en tales prácticas?

### ***El contexto curricular y de evaluación, 1975-al presente***

En la década de 1970 y durante la de 1980, un pequeño número de estados tenían estándares matemáticos para todo el estado; un número menor (por ejemplo, California, Nueva York y Texas) tenían evaluaciones alineadas con tales estándares. Esto empezó a cambiar con la emisión de los Estándares del NCTM en 1989; a lo largo de la nación, varios estados adoptaron los estándares de matemática. En términos curriculares, se dio una entropía significativa: por un lado, se ha dicho que hay 15 000 distritos escolares independientes en Estados Unidos –en el sentido que, durante las décadas de 1970 y 1980, los distritos escolares tenían una tremenda laxitud con respecto a la selección de las metas educativas y los medios para lograrlas<sup>1</sup>. Las restricciones verticales, si las hubo, sólo eran en el nivel estatal: formalmente, cualesquiera dos estados adyacentes, en principio, podían seguir sus propias reglas independientes (si tenían reglas). Por el otro, había restricciones reales. Sin importar en donde estuviera localizado el distrito escolar, se tenía un número limitado de opciones de libro de texto, dado que había un número relativamente bajo de editoriales que producían series de libros de texto. Además, el

---

<sup>1</sup> Existieron algunas restricciones. California es un estado “que adopta libros”: los distritos pueden usar cualquier libro que seleccionen, pero el Estado sólo reembolsaba a los distritos los costos de aquellos textos que se hallaban en una lista aprobada por el mismo estado. Y, por supuesto, los exámenes en uso (como el Examen del Regente del Estado de Nueva York) ejercieron un sustantiva presión hacia la homogeneidad.



currículo estaba un poco limitado desde arriba, en el sentido de que los distritos estaban concientes de los requisitos de admisión para la universidad (como requisito mínimo, los estudiantes con intención de entrar a la universidad debían estar listos para llevar Cálculo, lo cual significaba que debieron haber llevado los cursos regulares de Álgebra I-Geometría-Álgebra II-Trigonometría-Precálculo en secuencia. Pero los estándares estaban hechos de retazos no uniformes al igual que las evaluaciones estatales (si las había) para juzgar la competencia de los estudiantes. En la década de 1990, por ejemplo, el principal método de evaluación en Oklahoma era un examen de opción múltiple que se centraba en las habilidades básicas, mientras que en Vermont se usaba una evaluación de los estudiantes mediante portafolios con una colección de problemas resueltos. Grosso modo, debido a la larga tradición estadounidense sobre los derechos de los estados, había poca coherencia en el sistema y poca demanda de ésta.

La situación cambió con la aprobación de la ley “No Child Left Behind (NCLB: Que ningún niño se quede atrás)” de 2001<sup>2</sup>. Para calificar como depositario de fondos federales bajo la ley NCLB, como es sabido, cada estado debía institucionalizar estándares para el desempeño matemático, y evaluar a los estudiantes de manera regular. Estos exámenes eran de “alto riesgo”: la promoción de los estudiantes, el salario de los profesores (y su trabajo), el salario de los administradores (y su salario) y la existencia misma de las escuelas y de los distritos (que podían ser desmantelados si la calificación de los estudiantes no alcanzaba los requisitos cada vez más severos durante un periodo de algunos años) dependían de los resultados de las pruebas. No fue sorprendente que la

---

<sup>2</sup> No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425-2094, 2002.

mayoría de las escuelas se centrara en gran medida en enseñar a resolver las pruebas<sup>3</sup> que eran de una calidad altamente variable. Dado que cada estado tenía sus propios estándares y formas de evaluación, el resultado fue una incoherencia institucionalizada en todo el país.

El contexto continuó evolucionando y se dieron cambios significativos catalizados por el fondo federal de 4 mil millones de dólares Race to the Top (Carrera a la Cima, RTT)<sup>4</sup>, que fue parte de la ley “Recuperación y Reinversión Estadounidense” (American Recovery and Reinvestment act) anunciada por el presidente Obama y por el secretario de educación Duncan el 24 de julio de 2009. La restricción de RTT consistió en que *consorcios* de estados, no estados individuales, solicitaran los fondos. Esta restricción llevó al Concejo de Directores de Escuelas Estatales (Council of Chief State School Officers) y a la Asociación de Gobernadores (Governors Association) a patrocinar la Iniciativa de Estándares de Núcleo Curricular Común (Common Core state Standards Initiative) que produjo los CCSSM. A la fecha, cuarenta y cinco estados y tres territorios han adoptado los CCSSM, estableciendo, en consecuencia, lo que de facto es un conjunto nacional de estándares de matemática.

Además, el Programa de Evaluación Carrera a la Cima<sup>5</sup> “proporcionó financiamiento a [dos] consorcios de Estados para desarrollar evaluaciones que sean una instrucción válida, que apoyen e informen a los educadores, que proporcionen

---

<sup>3</sup> Hace algunos años, Hugh Burkhardt, acuñó la oración “Lo que Pruebas es lo que Obtienes” (WYTIWYG, por sus siglas en inglés) para representar esta realidad. Por razones de espacio no se puede hacer un análisis de WYTIWYG, pero se puede consultar Barnes, Clarke, & Stephens (2000) and Bell & Burkhardt (2001).

<sup>4</sup> See <http://www2.ed.gov/programs/racetothetop/index.html>

<sup>5</sup> See <http://www2.ed.gov/programs/racetothetop-assessment/index.html>

información precisa sobre lo que los estudiantes conocen y pueden hacer, y que midan el logro de los estudiantes con respecto a estándares diseñados para asegurar que todos los alumnos adquieran conocimiento y habilidades necesarios para tener éxito en la universidad y en el ámbito laboral”. Estos consorcios, la Sociedad para la Evaluación de la Preparación para la Universidad y sus Carreras (Partnership for Assessment of Readiness for College and Careers, PARCC<sup>6</sup>) y el Consorcio más Inteligente para una Evaluación Equilibrada (Smarter Balanced Assessment Consortium, SBAC)<sup>7</sup>, han enrolado, cada uno, alrededor de la mitad de los estados que han acordado alinearse con los CSS. Como resultado de esto, ya no habrá una retacería de 50 evaluaciones estatales. Con la excepción de los estudiantes de cinco estados y un territorio que no han signado la iniciativa RTT, los estudiantes del país enfrentarán una de dos evaluaciones, construidas por PARCC o por SBAC y alineadas ostensiblemente con los CCSSM. Dado el WYTIWYG (lo que evalúas es lo que obtienes), y el hecho de que los CCSSM y sus evaluaciones se aplicarán en cada grado, desde preescolar hasta octavo grado (segundo de secundaria en México), habrá un cierto grado de homogeneidad en los currículos y en las evaluaciones, algo sin precedente en la historia de los Estados Unidos.

### **La naturaleza de las evaluaciones en matemática, pasado y posible futuro**

Como vimos, las evaluaciones en matemática en Estados Unidos han variado ampliamente de un estado a otro. Aquí proporcionaré un ejemplo de las Pruebas Estándares de California (CST, por sus siglas en inglés) para ilustrar la realidad de un estado y contrastaré esto con una evaluación más rica de la competencia en esa misma

---

<sup>6</sup> See <http://www.parcconline.org/about-parcc>

<sup>7</sup> See <http://www.smarterbalanced.org/>

área de contenido. Después analizaré las especificaciones de reactivos y reactivos muestra de los dos consorcios nacionales.

En la Figura 2 se tiene un problema representativo de álgebra de octavo grado tomado del CST<sup>8</sup>.

¿Cuál es la intersección con el eje  $y$  de la gráfica de  $4x + 2y = 12$ ?

- (A) -4, (B) -2, (C) 6, (D) 12

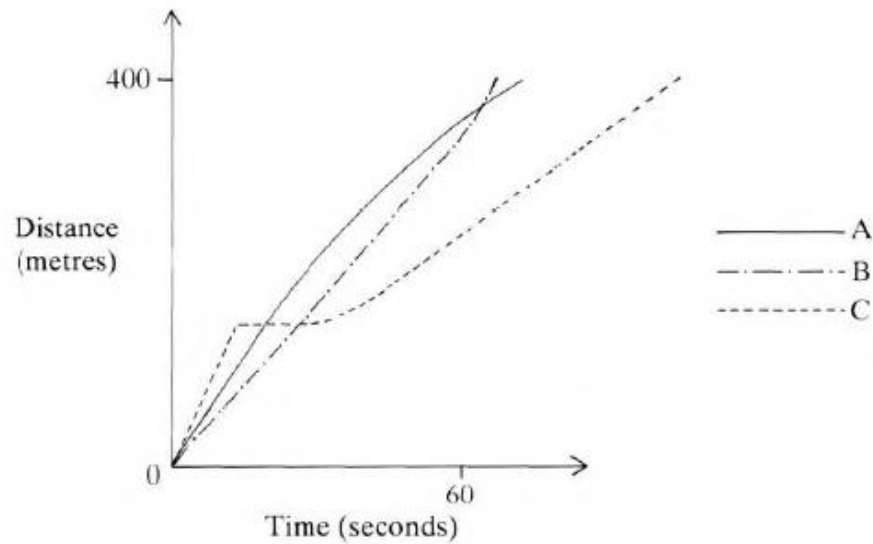
Figura 2. Problema de la CST correspondiente a la prueba de Álgebra I de 8°

Esta tarea, como la mayoría de las de CST, se centra en el conocimiento de contenidos. Hay al menos tres formas directas de obtener la respuesta: sustituyendo  $x = 0$  en la ecuación y resolviendo la ecuación resultante,  $2y = 12$ ; escribiendo la ecuación en forma pendiente-ordenada al origen  $y = -2x + 6$ ; y escribiéndola en su forma simétrica  $x/3 + y/6 = 1$ . En cada caso, el procedimiento es mecánico y la respuesta algo que se obtiene directamente. Aunque se evalúa el conocimiento de contenidos, es difícil argumentar que los estándares para la práctica matemática son evaluados de alguna manera significativa.

En contraste, considérese la “carrera de obstáculos” dada en la Figura 3.

---

<sup>8</sup> Éste es el problema 23 de los problemas de Álgebra I tomado de los exámenes para los Estándares de California, se los puede obtener en [http://www.cde.ca.gov/search/searchresults.asp?cx=001779225245372747843:gpfw5rhxiw&output=xml\\_no\\_dtd&filter=1&num=20&start=0&q=released%20items%20california%20standards%20test](http://www.cde.ca.gov/search/searchresults.asp?cx=001779225245372747843:gpfw5rhxiw&output=xml_no_dtd&filter=1&num=20&start=0&q=released%20items%20california%20standards%20test). Es típico del nivel de dificultad del examen. Se pueden obtener más preguntas muestra en <http://starsamplequestions.org/starRTQ/search.jsp>.



The rough sketch graph shown above describes what happens when 3 athletes A, B and C enter a 400 metres hurdles race.

Imagine that you are the race commentator. Describe what is happening as carefully as you can. You do not need to measure anything accurately.

Figura 3. Carrera de Obstáculos. Swan, M., y el equipo de Shell Centre Team (1985), p. 42.

Reimpreso con licencia (Texto de la Figura: La gráfica describe lo que sucede cuando 3 atletas A, B y C hacen una carrera de obstáculos. Imagina que eres el comentarista de la carrera. Describe qué está sucediendo con todo el cuidado del que seas capaz. No necesitas medir nada de manera precisa.

Esta actividad pide la interpretación de gráficas distancia-tiempo en un contexto real, un componente central de la modelación matemática. Una respuesta completa incluiría lo siguiente:

- Entender que un corredor cuya gráfica aparece “a la izquierda” de la de otro está *adelante* en ese punto, y le tomó menos tiempo recorrer la misma distancia (por tanto B gana la carrera);
- Entender qué significan en este contexto los puntos de intersección (que dos corredores han recorrido la misma distancia al mismo tiempo, de modo que están empatados en ese punto de la carrera);

- Interpretar el segmento de recta horizontal (el corredor no está avanzando, así, en el contexto de la carrera, debió haber tropezado con una valla y caído); y
- Poner todo lo anterior en una narrativa coherente.

Igualmente importante, responder la actividad apropiadamente pide mostrar competencia en (por lo menos) las primeras cuatro de las prácticas matemáticas mencionadas anteriormente. Si tareas con este nivel de complejidad aparecieran en las evaluaciones de los dos consorcios, entonces habría cambios significativos en lo que se evalúa (y en virtud del WYTIWYG, en lo que se enseña) a lo largo del país. Desde mi perspectiva, hay una promesa significativa en el hecho de que los dos consorcios de evaluación puedan mover las cosas en direcciones muy productivas: pero difícilmente se garantiza el progreso. Hay varios puntos en los que las cosas pueden salir mal: en las especificaciones de los exámenes; en las formas en que tales especificaciones se traducen en los exámenes mismos; y en la forma de calificar, por mencionar sólo tres.

### ***Las especificaciones de los exámenes de los consorcios***

En este punto creo que hay bases para tener un gran optimismo<sup>9</sup>. El cambio fundamental en las evaluaciones de SBAC es que informaran sobre cuatro resultados, no sólo uno. Hasta ahora, el resultado de un estudiante en la mayoría de las evaluaciones es un solo número. Esto ha permitido a la gente (incluyendo al mismo estudiante evaluado) decir qué tan bien se desempeñó el estudiante, en una escala absoluta y en una comparativa (la mayoría de los exámenes de alto riesgo proporcionan datos que producen resultados percentiles), pero no dan información sobre lo que el estudiante hizo o no bien

---

<sup>9</sup> Apertura total: Yo fui autor líder de las especificaciones de los contenidos de matemática del SBAC.

(por ejemplo, ¿el estudiante se desempeñó bien álgebra, pero no en geometría, o viceversa? ¿Obtuvo la mayoría de sus puntos en preguntas de procedimiento, en aquellas que piden una cadena extendida de razonamientos o en algunas de ambas?) En contraste, las especificaciones de prueba del SBAC (212, p. 19) pide que se informe sobre cuatro resultados para cada estudiante, correspondientes a las siguientes afirmaciones sobre lo que las evaluaciones deben revelar:

*Afirmación #1, Conceptos y procedimientos:* Los estudiantes pueden explicar y aplicar conceptos matemáticos e interpretar y seguir procedimientos con precisión y fluidez.

*Afirmación # 2, Resolución de problemas:* Los estudiantes pueden resolver una gama de problemas bien planteados en matemática pura y aplicada, haciendo un uso productivo del conocimiento y de las estrategias de resolución de problemas.

*Afirmación # 3, Comunicación de razonamientos:* Los estudiantes pueden clara y precisamente construir argumentos viables para apoyar sus propios razonamientos y para criticar el razonamiento de otros.

*Afirmación # 4, Modelación y análisis de datos.* Los estudiantes pueden analizar escenarios reales complejos y pueden construir y utilizar modelos matemáticos para interpretar y resolver problemas.

En la lectura más optimista, esto augura algo extremadamente positivo en la matemática basada en estándares. Queda bastante claro que con una prueba como la de California, CST, con sólo preguntas de opción múltiple que se centran en conceptos y procedimientos, no se pueden evaluar las Afirmaciones 2, 3 y 4 de una manera significativa. Tareas extendidas de resolución de problemas, de complejidad parecida a la

de la “carrera de obstáculos” mencionada, pueblan las especificaciones de SBAC. Si tales tareas se cuelean en las evaluaciones reales, llevarán la enseñanza en las aulas (mediante el principio WYTIWIG) en la dirección de los CCSSM. Pero hay riesgos; veamos más adelante.

La evaluación de PARCC promete tareas de tres tipos: (1) de evaluación de conceptos, habilidades y procedimientos, (2) de evaluación de la expresión de razonamiento matemático, y (3) de evaluación de modelación y aplicaciones (PARCC, 2013, p. 14). Esto tiene una amplia consistencia con el enfoque adoptado por SBAC y los CCSSM. No queda claro en los documentos disponibles en el sitio electrónico del consorcio (<http://www.parcconline.org/>) cuál será el formato para informar sobre los resultados de los estudiantes, por tanto, no fui capaz de determinar si habrá resultados separados para las tres categorías mencionadas.

### ***Los planes del consorcio para reportar resultados***

Nos encontramos aquí en un terreno un tanto desconocido, y encuentro la perspectiva problemática. El SBAC afirma<sup>10</sup> que “el sistema –que incluye evaluaciones formativas para fines de rendición de cuentas y evaluaciones provisionales opcionales para uso en la instrucción- utilizará tecnologías de prueba adaptables en la mayor medida posible para proporciona una retroalimentación significativa y datos útiles que profesores y otros educadores puedan usar para apoyar el éxito de los estudiantes”. El PARCC “utilizará tecnología en el diseño y en la instrumentación del sistema de evaluación. El diseño completo del sistema de evaluación incluirá una mezcla de reactivos de respuesta construida (opción múltiple), tareas de desempeño y reactivos mejorados y calificados

---

<sup>10</sup> <http://www.smarterbalanced.org/smarter-balanced-assessments/>



mediante computadora. Las evaluaciones de PARCC serán administradas mediante computadora y se empleará una combinación de calificación automatizada y hecha por humanos” (<http://www.parcconline.org/parcc-assessment-design>).

Tengo varias preocupaciones que tienen que ver con la “eficiencia” basada en la computadora. La meta especificada (en persona, si no impresa) de ambos consorcios es desplazarse hacia el punto en el que todas las evaluaciones no sólo sean dadas en computadora, sino completamente calificadas por computadora. Aunque los promotores digan lo contrario, estoy menos que convencido que lo más avanzado de la calificación automática de “preguntas de ensayo (abiertas)” en matemática –en especial aquellas en las que se emplean diagramas y otras representaciones matemáticas- se encuentra cerca del punto en el cual el trabajo del estudiante en preguntas complejas de respuesta abierta pueda ser evaluado precisamente (de hecho, poner la respuesta en la computadora puede ser un problema: en donde uno tenga que dibujar un diagrama y escribir algunas ecuaciones en poco tiempo puede ser un proceso largo, tedioso y distractor).

Igualmente tengo grandes reservas sobre el concepto de calificación adaptable por computadora. Tal calificación no puede ser demasiado peligrosa cuando la meta es simplemente asignar una calificación y el informe sobre el conocimiento y las prácticas del estudiante no es importante (obsérvese que éste no debería ser el caso aquí). Lo que es peor, los estudiantes que tengan un inicio poco tranquilo por dar respuestas incorrectas a los primeros dos problemas de una prueba cuya calificación es por computadora podrían nunca tener la oportunidad de mostrar lo que saben. El principal determinante para obtener la “siguiente” pregunta en una prueba por computadora es la dificultad del reactivo, el objetivo de esto es la rápida convergencia hacia la calificación del estudiante.

Esto puede ser eficiente, pero no apoya la necesidad del estudiante o del profesor de tener información sobre lo que el estudiante sabe y puede hacer.

Para decirlo francamente, la principal preocupación de aquellos que están construyendo los exámenes debe ser la matemática descrita en los CCSSM –no la eficiencia ni las preocupaciones psicométricas como la confiabilidad y la validez. Tales preocupaciones son importantes, desde luego, pero si guían la construcción de las evaluaciones, pueden distorsionarlas bastante.

### ***Reactivo muestra***

Al principio mencioné que el proceso está en sus inicios, y que un pequeño número de reactivos muestra ha sido presentado por PARCC –de modo que es peligroso hacer una extrapolación a partir de lo que se tiene hasta este momento; el lector debería revisar la colección de reactivos de los dos consorcios, PARCC en <http://www.parcconline.org/samples/item-task-prototypes>, y SBAC en <http://sampleitems.smarterbalanced.org/itempreview/sbac/index.htm>, y formar su propia opinión. Mi apreciación general de los reactivos de SBAC es que, mientras que la representación dinámica de algunos de ellos (por ejemplo, reactivo 43025) es superflua, el contenido es razonable y el uso del medio es objetivo y es, en ocasiones, creativo. Por ejemplo el reactivo 42960 (Figura 4) es directo y apunta a un contenido importante. El formato computarizado mejora la forma de hacer el apareamiento que si se tuviera sólo lápiz y papel o actividades computarizadas. Esta muestra de reactivos extendidos pide a los estudiantes que utilicen las prácticas, y el uso de rúbricas para calificar parece enfocado apropiadamente (uno siempre puede tener discusiones sin sentido sobre los detalles, desde luego).

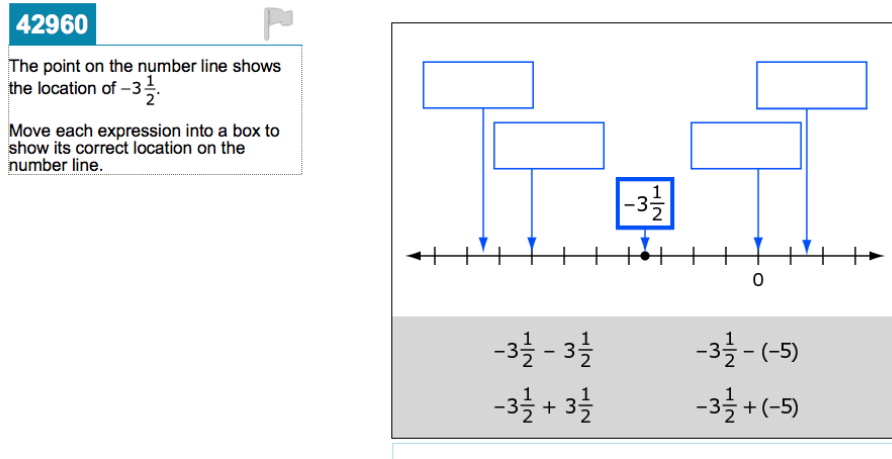


Figura 4. Versión computarizada de un problema de apareamiento. (Traducción de la figura: El punto de la recta numérica muestra la ubicación de  $-3\frac{1}{2}$ . Mueve cada expresión hacia la caja que corresponde a su correcta ubicación en la recta numérica.

PARCC deja a los lectores menos que examinar y algo de lo que ofrece es problemático. Considere, por ejemplo, la tarea de la Figura 5.

Como en la tarea de la Figura 4, en la Figura 5 se pide a los estudiantes que arrastren objetos. Aquí, sin embargo, el uso de la tecnología computarizada parece más una distracción que una ayuda.

En suma, existe causa de preocupación; dado que las pruebas tendrán un impacto poderoso en la

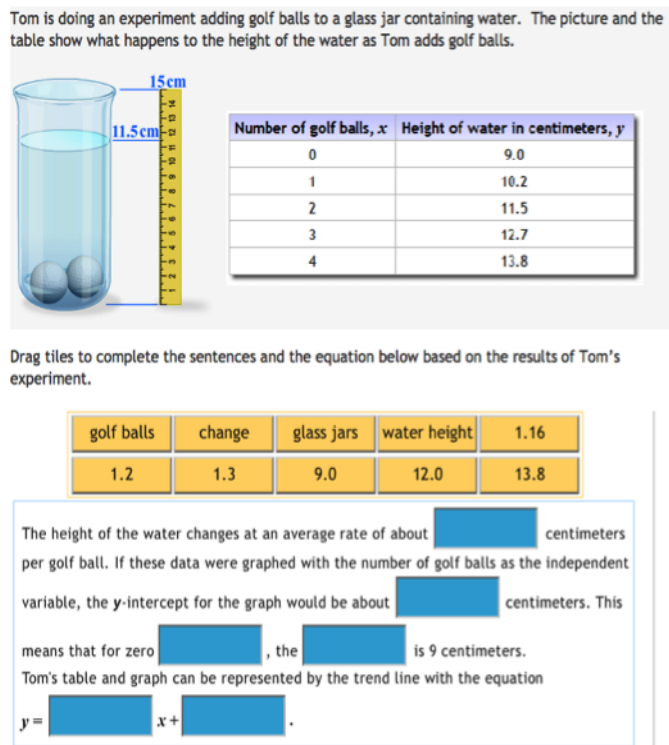


Figura 5. Parte de una tarea de modelación de PARCC.

educación matemática de EU, es importante hacer las cosas bien.

### ***Evaluación Formativa***

Uno de los mayores retos que enfrentarán los profesores, especialmente aquellos cuya formación se ha centrado principalmente en cuestiones procedimentales como el caso del reactivo de la figura 2, es aprender cómo proporcionar a los estudiantes las habilidades y los entendimientos requeridos para llevar a cabo tareas como la de la Figura 3. Parte de este desafío es aprender a tratar productivamente los enfoques de los estudiantes –tanto correctos como incorrectos- a medida que éstos se enfrentan a tareas complejas. Una manera importante de hacerlo se conoce como *evaluación formativa*. El objetivo de la evaluación formativa no se simplemente mostrar lo que los estudiantes “saben y pueden hacer” *después* de un periodo de enseñanza (éste es el tipo de evaluación formativa que analizamos en párrafos anteriores), sino revelar su entendimiento actual de modo que el profesor pueda ayudar al estudiante a mejorar. Existe mucha literatura sobre evaluación formativa, pero voy a saltarme la parte superficial de tal literatura poniendo énfasis en dos puntos. Primero, la evaluación formativa *no es* una evaluación sumaria que se hace con frecuencia. Como se observó, el propósito de la evaluación formativa es proporcionar información sobre el entendimiento del estudiante en un punto en el que profesor y estudiantes puedan actuar productivamente sobre dicho entendimiento. Segundo, la cuestión sobre la evaluación formativa no es asignar calificaciones, sino resaltar las fortalezas conceptuales y los retos que enfrentan los estudiantes. De hecho, calificar el trabajo de los estudiantes (incluso un trabajo sobre el cual se ha comentado mucho) incrementa las posibilidades de que los comentarios del profesor no sean leídos (Black & Wiliam, 1998).

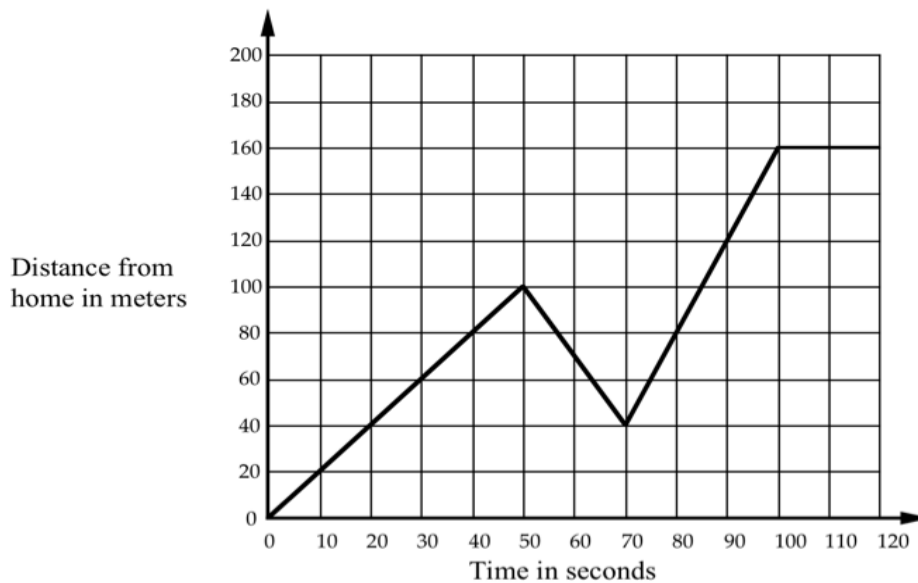
El Proyecto de Evaluación en Matemática (MAP), del cual soy Investigador Principal, ha estado produciendo lecciones de evaluación formativa (FAL) cuyo propósito es apoyar a los profesores en la realización de evaluaciones formativas. Como lo mencioné, 60 FAL (así como un cierto número de otros recursos relacionados con la evaluación) están disponibles en la página web de MAP

(<http://map.mathshell.org/materials/index.php>). Para transmitir el sabor del enfoque adoptado por el proyecto, describiré brevemente la tarea de FAL “interpretación de gráficas distancia-tiempo”

(<http://map.mathshell.org/materials/lessons.php?taskid=208&subpage=concept>).

### Journey to the Bus Stop

Todas las mañanas, Tom camina a lo largo de un camino recto de su casa a la parada de autobús una distancia de 160 metros. La gráfica muestra su recorrido en un cierto día.



1. Describe qué pudo haber pasado.  
Debes incluir detalles como qué tan rápido caminó.
2. ¿Son realistas todas las secciones de la gráfica? Explica completamente tu respuesta.

Figure 6. Diagnóstico para “interpretación de gráficas distancia-tiempo”.

FAL inicia con un problema diagnóstico que los estudiantes trabajan *antes* de la lección, de modo que el profesor obtiene información sobre las posibles fortalezas y debilidades de sus estudiantes. En la Figura 6 se presenta el problema diagnóstico para la lección “interpretación de gráficas distancia-tiempo”.

El plan de lección de FAL sugiere que el profesor responda al trabajo del estudiante, no asignándole una calificación, sino creando una serie de preguntas que se dirijan hacia las cuestiones reveladas por lo que los estudiantes escribieron. Como apoyo al profesor, en el plan se identifican interpretaciones erróneas típicas de los estudiantes y se sugieren preguntas que podrían hacer que los estudiantes piensen más. Algunos de los errores que más se dan son: (a) los estudiantes interpretan la gráfica como dibujo; (b) interpretan la gráfica como rapidez-tiempo; (c) no mencionan distancia o tiempo; (d) no calculan ni representan la rapidez; (e) no interpretan correctamente la escala; y (f) explican poco sobre las razones por la que la gráfica es realista. En la Figura 7 se dan una muestra de las preguntas para la cuestión (a).

Error	Preguntas sugeridas
<p><b>Los estudiantes interpretan la gráfica como dibujo</b></p> <p>Por ejemplo: el estudiante supone que a medida que la gráfica sube y baja, la trayectoria de Tom también sube y baja.</p> <p>O: El estudiante supone que una recta en la gráfica significa que la trayectoria es en línea recta.</p> <p>O: El estudiante piensa que una pendiente negativa significa que Tom tomó un atajo.</p>	<ul style="list-style-type: none"> <li>• <input type="checkbox"/> <input type="checkbox"/> Si una persona camina en círculo alrededor de su casa, ¿cómo se vería la gráfica?</li> <li>• <input type="checkbox"/> Si una persona camina con una rapidez constante sobre una colina, alejándose de su casa, ¿cómo se vería la gráfica?</li> <li>• <input type="checkbox"/> En cada sección de su trayecto, ¿la rapidez de Tom cambia o se mantiene constante? ¿Cómo sabes?</li> <li>• <input type="checkbox"/> ¿Cómo puedes calcular la rapidez de Tom en cada parte de su trayectoria?</li> </ul>

Figure 7. Error que puede cometer el estudiante y preguntas para explorarlo.

El objetivo es que el profesor comente el trabajo del estudiante (de manera individual si el tiempo lo permite o en forma de una lista de preguntas para reflexionar

para toda la clase), de modo que los estudiantes puedan comprometerse más plenamente con el contenido. La clase de 90 minutos empieza con una discusión de todo el grupo del problema de la Figura 8. Se le pide al estudiante que decida cuál de las historias A, B y C corresponde a la gráfica distancia-tiempo de la figura, y hay una discusión global sobre las razones de cada elección. El resultado de esta discusión es una gráfica comentada, que se ve un tanto como en la Figura 9.

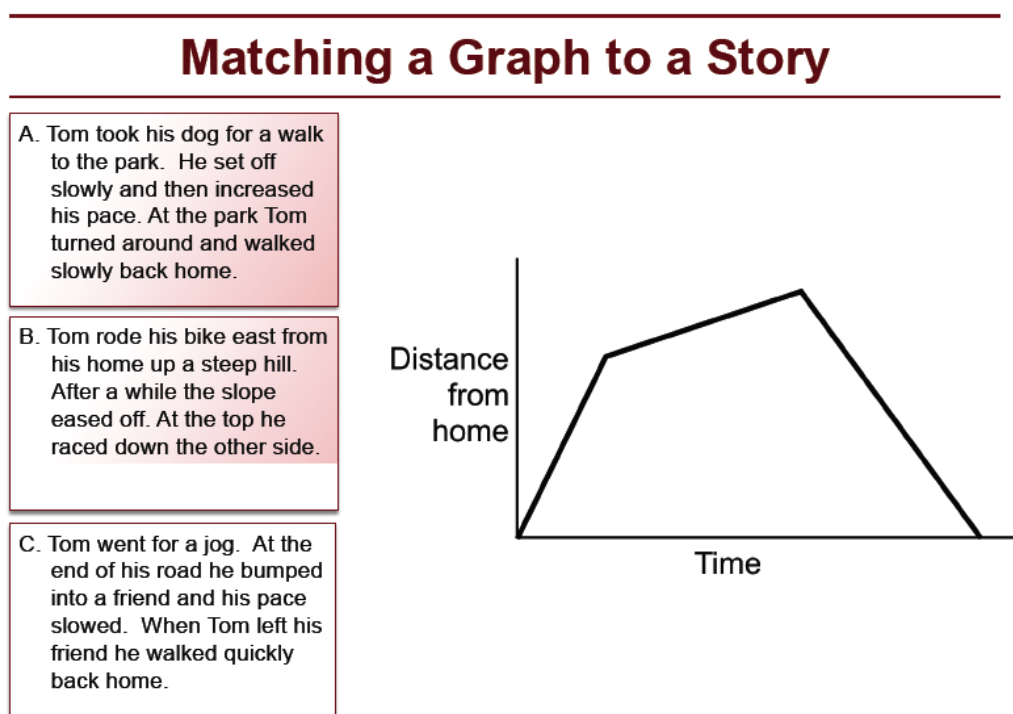


Figure 8. A distance-time graph question to start the lesson. (Traducción: A. Tom sacó a pasear a su perro al parque. Sale lentamente y luego aumenta su paso. En el parque, Tom da vuelta y regresa lentamente a casa. B. Tom viajó en su bicicleta hacia el este desde su casa subiendo una colina muy inclinada. Después de un tiempo la pendiente de la colina disminuye. En la cima, Tom se precipita del otro lado colina abajo. C. Tom salió a trotar. Al final de su recorrido se encontró con un amigo y bajó su rapidez. Cuando Tom se despidió de su amigo, caminó rápidamente a casa.

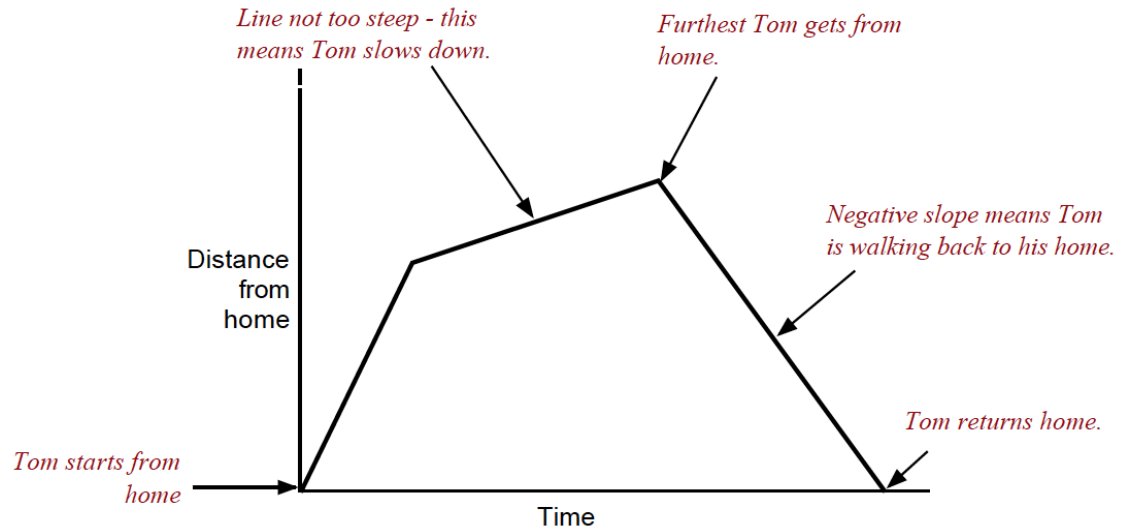


Figura 9. Gráfica comentada. Traducción empezando del origen en sentido horario: Tom empieza en su casa. Distancia a partir de la casa. Recta no muy inclinada: significa que Tom frena su paso. Mayor distancia que se aleja de casa. Pendiente negativa significa que Tom camina de regreso a casa. Tom llega a casa.

Con esto como telón de fondo, inicia la parte principal de la lección, un ejercicio de apareamiento de tarjetas. Se les da a los estudiantes un conjunto de diez gráficas distancia tiempo y diez historias. Se les pide que trabajen en equipos pequeños, emparejando las historias con las gráficas. En la Figura 10 se tiene una muestra de las cuatro primeras gráficas.

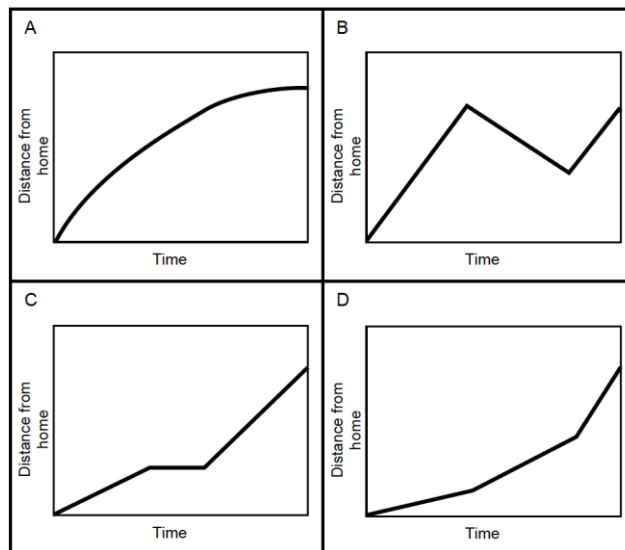


Figura 10. Muestra de gráficas distancia-tiempo para la actividad de las tarjetas.



En la Figura 11 se presentan 4 de las nueve historias. La décima tarjeta dice: haz tu propia historia.

<p><b>1</b> Tom ran from his home to the bus stop and waited. He realized that he had missed the bus so he walked home.</p>	<p><b>2</b> Opposite Tom's home is a hill. Tom climbed slowly up the hill, walked across the top, and then ran quickly down the other side.</p>
<p><b>3</b> Tom skateboarded from his house, gradually building up speed. He slowed down to avoid some rough ground, but then speeded up again.</p>	<p><b>4</b> Tom walked slowly along the road, stopped to look at his watch, realized he was late, and then started running.</p>

Figura 11. Muestra de historias de las tarjetas. Traducción: 1 Tom corrió de su casa a la parada del autobús y esperó. Se dio cuenta de que lo dejó el autobús y se regresó a casa caminando. 2 Enfrente de la casa de Tom hay una colina. Tom la subió lentamente, caminó por la cima y luego corrió rápidamente colina abajo hacia el otro lado. 3 Tom salió en patineta desde su casa subiendo poco a poco su rapidez. Se enfrena un poco para evitar algunos baches, pero después aumenta su rapidez de nuevo. 4 Tom caminó lentamente por el camino, se detuvo a mirar su reloj, se dio cuenta de que se le hacía tarde y entonces empezó a correr.

Conforme los estudiantes trabajan en la tarea de emparejamiento, con frecuencia encuentran situaciones insostenibles, por ejemplo, tienen dos historias desproporcionadas para la misma gráfica, o dos gráficas para la misma historia. Esto da lugar a calurosas conversaciones sobre *por qué* las historias y las gráficas corresponden o no.

En este punto de la clase, el profesor, que ha estado monitoreando las discusiones, inicia una conversación sobre cómo resolver los conflictos. Introduce la idea de construir una tabla a partir de la gráfica. Considere la gráfica de la Figura 8, por ejemplo. Uno puede asignar una escala para el tiempo y para la distancia, y luego usar algunos puntos de la gráfica para generar la tabla (véase la Figura 12, en la cual se han asignado valores al eje horizontal). La tabla completa se puede usar para hacer preguntas como “¿Tom se mueve más rápidamente en el primer tramo o en el segundo de su trayectoria representada en la gráfica?”

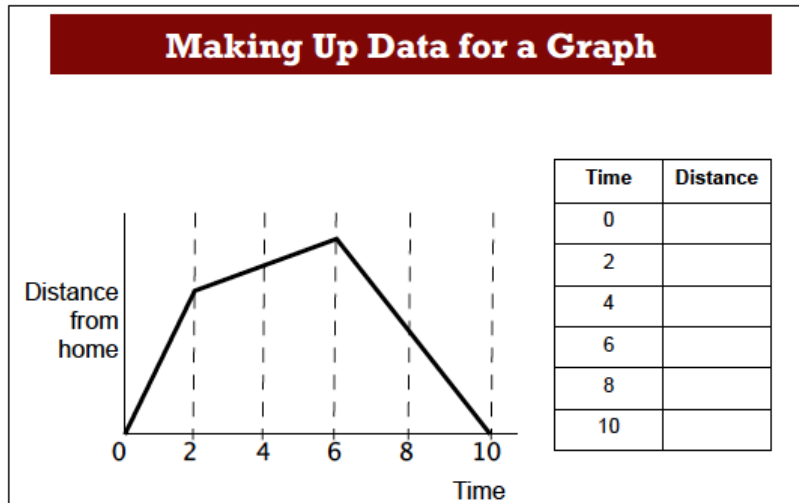


Figura 12. Construcción de una tabla a partir de la gráfica.

Ahora que los estudiantes saben cómo construir tablas a partir de las gráficas y usarlas como instrumentos mediáticos para verificar las historias que creen que están relacionadas con las gráficas, se les da un tercer conjunto de tarjetas que contienen una colección de tablas distancia-tiempo. Su tarea ahora es usar las tablas para reconsiderar su emparejamiento de gráficas con su historia, y para conformar un cartel que muestre las tres tarjetas correspondientes: una historia, una gráfica y una tabla que son mutuamente consistentes. Los estudiantes comparten sus carteles, comparan sus resultados como grupo. La clase termina dando tiempo a los estudiantes para que revisen sus carteles basándose en lo que se ha dicho durante la discusión global.

Cómo se observó, esta lección es una de 60 (de 100 lecciones planeadas) que aparecen en la página web de MAP. La esperanza es que tales materiales sean utilizados por los profesores cuando lleguen a las unidades importantes de su programa.

## Discusión

Estados Unidos se encuentra en una encrucijada con respecto a la educación matemática, en la que la evaluación tiene un papel principal como una palanca potencial para el cambio. Los CCSSM representan la evolución natural de los estándares de matemática, que vienen desde los *Estándares del NCTM de 1989*. El potencial para un cambio significativo viene con (a) la adopción de los CCSSM por parte de 45 de los 50 estados y por tres territorios; (b) el hecho de que los estados que se han alineado con los CCSSM utilizarán una de dos evaluaciones (una producida por el consorcio PARCC y la otra diseñada por SBAC) para evaluar la competencia de los estudiantes en matemática. La condición (a) sugiere que tendremos, por primera vez en EU, un currículo nacional de facto. La condición (b) sugiere que las dos evaluaciones actuales, debido a la implicación de alto riesgo, tendrán un papel fundamental en conformar la manera en que ese currículo cobre vida en las aulas estadounidenses. Si la evaluación se centra en los valores matemáticos propuestos por los CCSSM, hay un gran potencial para un progreso dirigido por la evaluación; pero si las evaluaciones pervierten las intenciones matemáticas de los autores de los CCSSM por razones de costo, facilidad para obtener resultados o consideraciones psicométricas<sup>11</sup>, los resultados pueden ser desastrosos. Los riesgos son, efectivamente, altos para (al menos) la siguiente década de educación matemática en EU.

Las evaluaciones correctas pueden orientar al sistema en la dirección correcta, pero incluso así, hay cuestiones de capacidad del sistema. Es difícil enseñar para el tipo de entendimiento de contenido y para las prácticas matemáticas que plantean los CCSSM. Hablando de manera general, los programas de formación de profesores no han

---

<sup>11</sup> Es absolutamente esencial para la integridad matemática de los estándares llevar a cabo el proceso de construcción de pruebas y después tomar en consideración cuestiones psicométricas, más que hacerlo al revés, como es típico en la construcción de pruebas.

tenido tiempo (ni los recursos) de apoyar a los profesores a ser competentes en evaluación formativa; ni la actual generación de textos proporcionan a los profesores un apoyo adecuado. La evaluación formativa, bien hecha, puede apoyar a los profesores en la construcción de ambientes matemáticos ricos en el aula. Es nuestra esperanza que los tipos de lecciones de evaluación formativa (FAL) descritas en este artículo ayuden a proporcionar dicho apoyo.

### Referencias

- Barnes, M., Clarke, D., and Stephens, M. (2000). Assessment: the engine of systemic curricular reform? *Journal of Curriculum Studies*, 32 (5), 623-650.
- Bell, A., & Burkhardt, H. (2001). Evidence on the influence of assessment on teachers' classroom practices. Downloaded October 31, 2001, from <<http://www.toolkitforchange.org/toolkit/view.php?obj=1034&menu=i>>.
- Black, P., & Wiliam, D. (1998). *Inside the black box: raising standards through classroom assessment*. London: King's College London School of Education.
- California State Department of Education. Standardized Testing and Reporting - STAR Sample Test Questions. Downloadable at <  
<http://starsamplequestions.org/starRTQ/search.jsp>>.
- Common Core State Standards Initiative. (2010). *Common core state standards for mathematics*. Downloaded July 9, 2010, from <http://www.corestandards.org/the-standards>
- National Commission on Excellence in Education. (1983). *A nation at risk*. Washington, DC: U. S. Government Printing Office.
- National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: NCTM.
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- National Research Council. (2001). *Adding it up: Helping children learn mathematics*. Washington DC: National Academy Press.
- Partnership for Assessment of Readiness of College and Careers. (2012a). PARCC

model content frameworks: Mathematics grades 3–11, version 2.0, august 31, 2012  
(revised). Downloaded September 23, 2012 from <<http://www.parcconline.org/parcc-model-content-frameworks>>.

Partnership for Assessment of Readiness of College and Careers. (2012b).

PPT\_Messaging for PARCC Math Sample Items\_FINAL\_revised\_081212\_1.

Downloaded from <<http://www.parcconline.org/samples/item-task-prototypes>> on  
September 23, 2012.

Smarter Balanced Assessment Consortium. (2012). Content Specifications for the

Summative assessment of the *Common Core State Standards for Mathematics*.

Downloaded March 20, 2012, from <<http://www.smarterbalanced.org/smarter-balanced-assessments/>>.

Swan, M., and the Shell Centre Team (1985). *The language of functions and graphs*.

Nottingham, England: The Shell Centre of Mathematics Education.