

**La reforma de la evaluación de la alfabetización:  
cambios de opinión, enunciación de principios  
y nuevas prácticas emergentes  
(Segunda parte)**

**Robert J. Tierney\***

*En el número de marzo incluimos la primera parte de este artículo donde aparecen enunciados los seis primeros principios formulados por el autor en relación con la evaluación de la lecto-escritura.*

**PRINCIPIO 7: Los procedimientos de evaluación no deben ser estandarizados para poder adaptarse así a cada persona.** Cuando los docentes tratan de brindar a los estudiantes todas las oportunidades a su alcance, están haciendo ajustes constantes en su "lectura" del alumno: su disposición, habilidad verbal, familiaridad con ciertos temas, necesidades, etc. Buscamos formas de maximizar el aprendizaje de los distintos alumnos tomando en cuenta que cada uno de ellos requiere diferentes tipos de apoyo y debe ser alentado de distintas maneras. Si estandarizamos nuestra enseñanza ya sabemos cuál será el probable resultado final: algunos alumnos con un potencial maravilloso sólo mostrarán una parte de sí mismos y no alcanzarán a desarrollar ese potencial o, incluso, no revelarán quiénes son y qué podrían aportar y aprender.

Permitir que un individuo, o inclusive un grupo, pueda expresar las diferencias da por tierra con el deseo de estandarizar la evaluación. La estandarización enfoca a cada individuo y a cada grupo de la misma forma –es decir, los alumnos realizan las mismas tareas al mismo tiempo, y luego se evalúan sus respuestas usando los mismos criterios–. Pero si los repertorios del aprendizaje de los distintos alumnos son diferentes y estos alumnos utilizan distintas estrategias y tienen distintos valores, entonces la forma de evaluar adecuada para uno puede ser inapropiada para otro.

Los estudios interculturales, de aula en aula y dentro de una misma aula, sugieren que los distintos alumnos responden de modos diferentes a las diferentes formas de evaluación, según sus propias historias culturales, escolares o personales (...) Green y Dixon (1994) han subrayado que los alumnos construyen "repertorios localizados asociados con modelos particulares... no modelos genéricos" (p. 237). Contamos con demostraciones más que suficientes de cómo el grado de respuesta de distintos grupos e individuos en situaciones de evaluación depende de su visión de la dinámica social de la situación (Basso, 1970; Ogbu, 1988; Philips, 1983).

De hecho, siempre existe una tensión entre la necesidad de uniformar las personas y los grupos y el uso de procedimientos que sean sensibles a la diferente evolución de la alfabetización del alumno, como así también a la propia predisposición a responder de manera diferente a distintas personas y

---

\* El autor es profesor y director de la School of Teaching and Learning, Ohio State University, Ohio, Estados Unidos. Ha realizado una extensa labor de capacitación con docentes y alumnos sobre cuestiones relacionadas con la evaluación.

de diferentes modos en diferentes ocasiones. En numerosas oportunidades, mi evaluación de algunos estudiantes ha sido revisada buscando un modo de respuesta, buscando establecer diferentes tipos de asociaciones con ellos o de observarlos interactuar en el tiempo ante variadas situaciones con personas o grupos diferentes. A su vez, lo que puede servir como instrumento para descubrir el nivel de alfabetización de un alumno puede ser un método inadecuado para descubrir el de otro, o el del mismo alumno en otro momento. Los docentes deben estar dispuestos a utilizar diferentes medios con los diferentes alumnos, ya sea que estén evaluando o enseñando.

El proceso de toma de decisiones también puede verse complicado por ciertas predilecciones. Junto con mi trabajo sobre carpetas (*portfolios*), siempre me sorprenden los análisis que hacen los alumnos de su propio progreso y los tipos de objetivos que se proponen. Siempre asignan un significado a los elementos en sus portafolios que yo puedo haber pasado por alto o no detectado. Y, a menudo, su decisión acerca de cómo proceder es diferente de lo que yo hubiera sugerido.

**PRINCIPIO 8: Los resúmenes simplistas, las comparaciones y los puntajes deben ser desplazados por enfoques que reconozcan la naturaleza compleja e idiosincrática de la evolución de la alfabetización. Las comparaciones directas entre individuos son normalmente arbitrarias, sesgadas y poco amplias.** Si se asume un enfoque de la evaluación con una nueva apertura hacia la complejidad, respeto por la diversidad, e interés por formarse una imagen más cercana y amplia de cada alumno, ¿cómo se toman las decisiones sobre los estudiantes? Aquellas que exigen reflexionar sobre el progreso individual y las perspectivas de cada uno seguramente abundarán en datos provechosos. Los docentes que buscan un punto de vista abierto y multiforme del alumno tropezarán con escasas dificultades para negociar nuevas áreas de búsqueda con cada uno de ellos y para cada uno de ellos. Las decisiones que exigen comparar una persona con otra presentarán un problema, pero estas dificultades no son insuperables. Demandan un deseo de enfrentarse con incertidumbres, de considerar las posibilidades y negociar las decisiones, inclusive aceptar la eventualidad de que no siempre habrá consenso. El problema con las comparaciones es que no aclaran las cosas cuando se asume que una cierta continuidad o un puntaje en particular logran definir al alumno adecuadamente.

Las comparaciones basadas en puntajes también resultan problemáticas por una multiplicidad de razones: a) cada alumno se desarrolla de un modo que le es único y particular; b) la alfabetización de un alumno difiere de la de otro, y aun una misma alfabetización tiene diferentes matices; y c) algunas de estas facetas serán exclusivas y típicas de una única situación. La evolución de la alfabetización es tan diferente de un alumno a otro que los tipos de comparaciones que se establezcan serán complejos y multifacéticos. La expresión **habilidades de lecto-escritura** (*literacy abilities*) antes que **habilidad de lecto-escritura** (*literacy ability*) parece imponerse. Si tratamos de ilustrar el carácter de estas evoluciones, podríamos encontrarnos ante la duda de describir o no a cada individuo en función de sus propias características individuales. Lamentablemente, los términos de comparación que ofrecen las evaluaciones estandarizadas y aquéllas implícitas en muchos

de los intentos de asignar puntajes a los portafolios y otros datos basados en el trabajo del aula son, a menudo, insensibles a tal complejidad. El analizar a individuos diferentes con un único puntaje enmascara la variación y la individualidad. Nuevamente, quienes preparan las evaluaciones cometen un error al simplificar demasiado: algo parecido a una "conspiración de conveniencia" masiva (Spiro y otros, 1987, p.180).

La tendencia a la uniformidad está bastante generalizada. Nuestros programas de instrucción y evaluación con frecuencia incluyen largas listas de habilidades que son el resultado de evaluar, enseñar y dominar. Se asume que las habilidades constituyen un prolijo paquete discreto y que cada quien hace una contribución uniforme a la evolución de la alfabetización. Se asume que los alumnos adquieren estas habilidades hasta dominarlas y que su capacidad de utilizarlas es uniforme en toda la gama de situaciones que hacen a la lectura y escritura. En situaciones de lectura y escritura auténticas dentro de las cuales se persiguen objetivos genuinos, esto es poco probable. En toda situación literaria, ciertos atributos tienen más probabilidades de entrar en la lista agrupados que uno a uno o separadamente.

Con demasiada frecuencia los alfabetizadores han ignorado la complejidad de estas cuestiones y, en vez de explorar las distintas posibilidades, han caído en la conveniencia. Tomemos, si les parece, los intentos por armonizar los datos surgidos de la evaluación del rendimiento (p.e., los portafolios) a través de consignas. Los datos generados a partir de un portafolios pueden involucrar una riquísima gama de muestras u observaciones sobre el trabajo de los alumnos en distintas ocasiones y situaciones. Con seguridad, estas muestras representan el esfuerzo del alumno por cumplir los distintos objetivos que se ha impuesto, el uso de diferentes recursos, bajo condiciones cambiantes. En cierto modo, las muestras del trabajo del estudiante pueden variar tanto como las obras de arte de un mismo artista. Cada ejemplo puede representar logros y procesos muy dispares. Cuando uno las tiene ante sí, las examina y discute su significado, uno está en contacto con el verdadero instrumento y no con algún derivado distante.

Es en este punto donde algunos dirán que podemos utilizar una consigna para asignar un puntaje o puntajes al trabajo del alumno. Pero es necesario examinar una cuestión que es la inversa de lo que casi siempre se pregunta. En lugar de preguntar cómo calificar el portafolios, debiéramos preguntar si la consigna está al nivel del portafolios o de la evaluación del rendimiento complejo. Es más, en el aula, ¿necesitamos de una medida lejana cuando tenemos las fuentes primarias a nuestro alcance, las verdaderas muestras que podemos examinar y volver a examinar usando una multiplicidad de lentes o perspectivas? Si bien yo hablo a favor de la naturaleza específica del contexto en toda evaluación, quienes defienden la consigna parecen querer dejar de lado idiosincrasias y variedad, es decir, no están dispuestos a considerar a los individuos según sus propios méritos o sus propias modalidades. A no ser que las consignas se utilicen para analizar cómo abordar una tarea o para iniciar una conversación sobre los portafolios del alumno, yo no veo grandes ventajas en su uso.

A veces la evaluación de la lecto-escritura se torna rebuscada si se incorporan subpuntajes. Se supone que dichos puntajes brindan un retrato completo y detallado del grado de alfabetización de una persona y, a su vez, pueden ser empleados como base de comparación en la toma de decisiones. Sin embargo, es de lamentar que estos puntajes no sean exhaustivos, que no permitan una graduación precisa, y cuando es así, no conviene incorporarlos. Para poder hacerlo, deberíamos actuar del siguiente modo:

- 1) Incluir todos los atributos o asegurarnos de que la lista parcial que hemos desarrollado es representativa;
- 2) determinar cómo se configuran estos atributos en las distintas situaciones;
- 3) asumir que se ofrecerá amplia evidencia para evaluar estos atributos;
- 4) desarrollar escalas para evaluar los atributos; y
- 5) generar un algoritmo que funcione con todas las personas y que nos permita combinar elementos y sus dimensiones.

Personalmente, me atrevería a postular que no contamos con tales muestras, procedimientos de muestreo, modos de procurar evidencias, escalas adecuadas o un algoritmo. Y es problemático asumir que un algoritmo que simplemente representa sumas pudiera resultar suficiente alguna vez. La complejidad de la alfabetización es tal que no podemos asumir una base para generar o combinar puntajes.

Las evaluaciones de la alfabetización no pueden ni deben ser tan rígidas. Tal vez haya algunos patrones de comparación que puedan considerarse apropiados para todos los alumnos. Tal vez haya patrones de comparación apropiados para ciertos lectores y no para otros. Pero tales patrones de comparación probablemente representen una imagen parcial de la alfabetización de un alumno dado. El uso de puntajes y de una serie de continuos como modo de simplificar y comparar tiende a enmascarar la subjetividad de la evaluación y a crear entre quienes preparan evaluaciones una ilusión de objetividad. El uso de puntajes y de tales continuos no implica mayor objetividad, sino arbitrariedad. Guba y Lincoln (1989) sugirieron que cambiar para aceptar la inevitabilidad del relativismo y la complejidad de entornos diferentes puede exigir la búsqueda ecuménica, constante y reiterativa de posibilidades compartidas, en vez de un conjunto único de verdades absolutas.

**PRINCIPIO 9: Ciertas cosas que los evaluadores pueden calificar de modo fidedigno no vale la pena que sean evaluadas; ciertas cosas que vale la pena evaluar, pueden ser difíciles de evaluar de modo fidedigno excepto si lo hace un mismo evaluador.** Con frecuencia, quien prepara las evaluaciones y el investigador se preguntan si pueden o no medir ciertas capacidades con coherencia. Hablan de la confiabilidad como criterio principal para decidir si una prueba es válida o no. El resultado final es que ciertas cosas que vale la pena medir son descartadas y otras que no vale la pena medir o valorar alcanzan gran importancia. Es típico que el aprendizaje complejo e individual tienda a ser la moneda de cambio mientras que con otras monedas del aprendizaje más fáciles de definir se produce un fenómeno de inflación. P.e., en la evaluación escrita, ciertas construcciones tales como el

estilo o la voz pueden sufrir un cambio poco favorable, mientras que la ortografía y la puntuación son "infladas". En la lectura, cuestiones como las autopreguntas, el compromiso y la interpretación se dejan de lado mientras que la velocidad, la evocación fáctica y el vocabulario pueden verse jerarquizados.

Lamentablemente, la confiabilidad se equipara a que dos personas distintas califiquen y evalúen la misma cosa del mismo modo. Salvo que los evaluadores demuestren un alto grado de consenso, los redactores de las pruebas considerarán que la medición no es confiable y, por lo tanto, cuestionarán su valor. Al hacerlo, pueden estar cometiendo el error de asumir que la confiabilidad equivale a consenso cuando en realidad la verificabilidad puede ser un abordaje mejor.

Deberíamos estar dispuestos a aceptar diferencias de opinión en cómo se califican o comprenden ciertas capacidades. Algunas capacidades y estrategias son difíciles de identificar con definiciones operativas claras. Los distintos evaluadores, o aun los mismos evaluadores en distintos momentos, seguramente desarrollarán interpretaciones diferentes de los mismos fenómenos. A veces estos cambios surgen como resultado de una distinta predisposición del evaluador. A veces, surgen a medida que los distintos evaluadores (o un mismo evaluador) van tomando en cuenta diferentes facetas de un mismo fenómeno. A veces surgen porque los alumnos logran ciertas habilidades de forma diferente, lo cual no debe sorprender, porque esto coincide con dos postulados básicos de las dos corrientes actuales en aprendizaje: el concepto de una naturaleza interpretativa constante del conocimiento; y la naturaleza situacional específica del aprendizaje. Es lógico que existan diferencias y que se crucen dentro de la alfabetización individual (p.e., leer el diario para encontrar un aviso versus leer una novela por placer), variando también de un individuo a otro. En otras palabras, algunos rasgos pueden o no aplicarse a la alfabetización de ciertos alumnos y algunas facetas pueden aplicarse únicamente a nivel individual.

Uno no debería dejarse seducir por la idea de que sólo las variables fáciles de definir deben ser analizadas excluyendo aquellas que son difíciles de evaluar. Sería una tontería excluir ciertas facetas porque son difíciles de evaluar o porque difieren de un estudiante a otro o de una situación a otra o bien porque los evaluadores sienten que son distintas. Del mismo modo, uno no debería dejarse seducir por la idea de que todo acto de lectura y escritura es igual y que implica las mismas variables. Si las únicas facetas de la alfabetización que reciben puntajes son las comunes a todos los alumnos, y si se consideran sólo las que pueden recibir puntaje con alta confiabilidad en las respuestas de los distintos alumnos, entonces, indefectiblemente, se asignará a ciertas facetas más peso del que merecen y otros aspectos importantes podrán resultar excluidos.

**PRINCIPIO 10: La evaluación debe ser evolutiva y sostenida antes que fraccionada y de poco alcance.** Para evaluar si un alumno va bien, nuestra perspectiva o punto de vista debe cambiar. Si los objetivos de la evaluación están ligados a la evolución, entonces necesitamos analizar los patrones y objetivos a largo plazo. Lo que vemos o buscamos en una única selección o en

un solo caso puede no ser útil para analizar patrones de caso en caso, o en distintas circunstancias o con distintas selecciones. P.e., cuando un lector o escritor lee y escribe una única selección, nosotros podríamos buscar compromiso y participación activa. En otras situaciones, podríamos considerar en qué medida se mantienen el interés y el compromiso cuando se enfrenta a una amplia gama de materiales con diferentes propósitos. También podría interesarnos saber en qué medida el alumno le ha asignado un valor a la lectura o escritura, lo cual se refleja en cómo utiliza la lectura y la escritura dentro y fuera del aula, en qué lecturas elige o en si escribe de *motu proprio* y con diferentes propósitos.

Dentro de áreas tales como la capacidad de leer con comprensión, nuestro objetivo para una única elección puede ser verificar hasta dónde el lector comprende la idea principal o el tema o en qué medida puede sacar conclusiones usando detalles seleccionados, etc. Con las distintas selecciones o en el largo plazo, podríamos interesarnos en cómo se van usando diferentes libros para contribuir a la comprensión total ligada con unidades temáticas o proyectos o con el propio desarrollo de una comprensión del mundo. O bien, podríamos interesarnos en la autoevaluación. Con una única selección, podríamos centrarnos en la capacidad del lector o escritor de supervisar la lectura y la escritura, de fijar objetivos para una selección específica y para resolver problemas y esforzarse en entender significados. Con distintas selecciones, podríamos interesarnos en el lector o escritor y su capacidad de establecer objetivos y evaluar el progreso tomando varias selecciones. Al analizar varias selecciones, uno no debería esperar que los alumnos siempre muestren el mismo nivel de eficacia con ciertas habilidades o estrategias o que, necesariamente, usen las mismas habilidades y estrategias. Mayores detalles en cuanto a contrastes a corto y largo plazo pueden verse en esta tabla.

## CONTRASTES A CORTO Y LARGO PLAZO EN LA EVALUACIÓN

Corto plazo/única vez	Largo plazo/múltiples veces
<i>Afectividad</i>	
Comprometido	Valor
Activo	Autobúsqueda permanente
Pensativo	Hábito
<i>Estrategias</i>	
Planificación	Flexible, reflexivo, coordinado, selectivo,
Resuelve problemas, arregla situaciones	a la medida de la situación
Establece relaciones	
Mira hacia atrás, adelante y más allá	
Colaborador	Fortalece la comunidad
<i>Resultados</i>	
Idea principal	Comprensión integral, relaciones intertextuales
Detalles	Proyectos
Conclusiones	Aplicaciones
Implicancias	Rango de problemas y actividades
	Comprensión integral y temas
<i>Autoevaluación</i>	
Autosupervisión	Autoescrutador, fija objetivos, autodeterminación
Resolución de problemas al momento	Objetivos integrales, progreso, patrones

Un cambio hacia una evaluación que examine a los alumnos en el tiempo coloca a la evaluación en un pie de igualdad con las prácticas que buscan el compromiso sostenido en el aula y que apuntan a inducir a la comprensión de estructuras. Transforma nuestra enseñanza y aprendizaje en posibilidades a largo plazo en lugar de orientarnos hacia los objetivos específicos y a corto plazo de una lección.

**PRINCIPIO 11: La interpretación de resultados por lo general no resulta clara. Debería considerarse que la evaluación es constante, que sugiere, que no es fija ni definitiva.** En más de un sentido, la enseñanza implica desarrollo constante, experimentación continua, ajustarse a planes, instrucciones y futuros objetivos. Para apreciar las complejidades y la sofisticación de la enseñanza, pensemos en un deportista. En ciertos deportes (p.e., el baseball, el tenis) que implican coordinación de la mano y la vista con la raqueta o el bate, los jugadores comienzan el giro y constantemente van haciendo ajustes muy sutiles a medida que juzgan las diferentes velocidades, rotaciones y ángulos con que se arroja la pelota. Pero el deporte empalidece ante la dinámica de la interacción maestro-alumno: los ajustes, la toma de decisiones en el momento preciso, la marea de actividades que se van suscitando. Los docentes se encuentran ante alumnos a quienes pueden estar tratando de dar respuesta, motivar, movilizar, desarrollar e instruir mientras

intentan entender sus necesidades, opiniones, estrategias y modos posibles de responder, a medida que interactúan unos con otros y siguen con el resto de sus vidas. No sorprende entonces que los docentes sean una mezcla de ecologista, constructor, abogado, entrenador, jugador, actor-director, director de escena, alcalde y, a veces, consejero. Los maestros siempre están planificando y reconociendo la necesidad de hacer ajustes constantes a lo que están haciendo y a lo que pueden hacer a continuación.

A tales efectos, los datos típicos de evaluación (p.e., los informes sobre los alumnos provistos a través de los registros escolares, la medición previa de capacidades, las evaluaciones estandarizadas o aun informales) pueden brindar una guía limitada a los docentes en cuanto a la toma de decisiones momento a momento y también sobre la planificación para el día siguiente, o la semana o el mes siguiente. Demasiado a menudo los registros típicos del alumno parecen ser tan limitados como una mala foto carnet: uno puede identificarlo (dependiendo de la capacidad para ver parecidos) pero no siempre con claridad. Ciertamente, esta foto no nos permitirá apreciar el carácter del alumno ni nos ayudará a comprender el rango de cosas que puede hacer, ni sustentará nuestra capacidad negociadora, ya sea en los objetivos de aprendizaje a largo plazo como en los de corto plazo.

La mayoría de las evaluaciones sustentadas en el aula son más prometedoras pero siguen siendo limitadas. Los procedimientos de evaluación en el aula pueden dar al maestro una noción más clara de cómo procederán los alumnos en circunstancias similares y pueden también brindar una imagen más completa del alumno a lo largo del tiempo. Los portafolios, p.e., son equivalentes a los álbumes de recuerdos que incluyen una variedad de instantáneas del alumno en una cantidad de contextos. Tales evaluaciones pueden darnos una idea más cabal del alumno y de su patrón de desarrollo, pero los juicios –en especial las recetas– nunca son tan simples como aparentan. La posibilidad de obtener un panorama completo del alumno se ve complicada por nuestra incapacidad de monitorearlo constantemente, hurgar en sus pensamientos más hondos e interpretarlos, y lograr más de una perspectiva sobre él. Ésta también está vinculada a la naturaleza siempre cambiante del aprendizaje. Aparte del hecho de que nuestras instantáneas del aprendizaje en el aula tienden a ser imágenes muertas del alumno, tales imágenes están ligadas a un lugar y momento que se ha vuelto más histórico que actual. Estas limitaciones podrían entenderse como un problema si insistiéramos en querer tener una receta de qué hacer a continuación con un alumno y estar seguros de seguir el rumbo prefijado. En lugar de esto, las imágenes previas deben ser vistas como base tentativa de dónde y cuándo comenzar. Si bien podemos desarrollar planes y objetivos a corto y largo plazo, no debiéramos enfocar nuestra enseñanza como si nuestras recetas no pudieran ser modificadas, como si la evaluación debiera ser fija y/o como si las orientaciones fueran algo más que meras sugerencias.

De igual modo, no deberíamos enfocar la evaluación como si nuestros resultados fueran definitivos, ni decidir nuestra posterior acción como si las decisiones fueran algo más que un presentimiento. Deberíamos evitar asumir que nuestras evaluaciones hacen algo más que brindarnos una información que debemos considerar. Ninguna evaluación debería usarse de modo



restrictivo o rígido como las decisiones del juez, aunque me temo que muchas lo son. En cambio, deberíamos fortalecer lo que debe suceder en el aula: el ajuste constante, los cambios, la toma constante de decisiones de los maestros que siempre están observando, aprendiendo, instruyendo y respondiendo a los alumnos, pares y otros.

**PRINCIPIO 12: Las posibilidades de aprendizaje deben ser negociadas con los alumnos y otros interesados antes que impuestas mediante estándares y evaluaciones preestablecidas, prescriptas u obligatorias.**

El Estado en el que vivo (además de muchos otros estados) ha sido inducido a pensar que establecer estándares puede ser la respuesta a la expectativa de una mejor educación porque asegura que los maestros enseñen y los alumnos aprendan ciertas habilidades básicas. Me desalienta el que nuestras asociaciones profesionales hayan aceptado tal forma de pensar. Históricamente, establecer estándares (y las pruebas de idoneidad que generan) tendió a restringir el acceso y la experimentación al mismo tiempo que a avalar programas ligados a la exclusión.

El establecimiento de estándares y la industria de las pruebas puede perpetuar el punto de vista de que es posible formular objetivos fáciles de alcanzar. Lamentablemente, asumir que el desarrollo simplemente establece un rumbo para el alumno que va de A a B puede resultar problemático, especialmente si A no es tomado en cuenta y B está ligado a puntos de vista que buscan en los resultados más bien la experiencia que la evaluación individual de la evolución. Sin considerar ampliamente dónde se encuentran los alumnos, cómo y por qué evolucionan y cuáles son sus aspiraciones, es probable que ubiquemos nuestros objetivos en el lugar equivocado y que nuestros caminos hacia el aprendizaje estén mal orientados. Estando yo presente en una de las muchas sesiones sobre estándares patrocinada por la **Asociación Internacional de Lectura (IRA)** y el **National Council of Teachers of English** (*Consejo Nacional de Maestros de Inglés*), un orador comenzó a hablar sobre estándares usando una analogía con un jugador de basquet del calibre de Michael Jordan, tomándolo como el "estándar". A medida que el orador discurría sobre la importancia de fijar estándares en función de nuestras aspiraciones, yo me preguntaba sobre mi estatura y mi habilidad y sobre qué podría hacer para mejorar en ese deporte. Luego recordé por qué juego al basquet y en qué nivel puedo jugar dados mis antecedentes basquetbolísticos: juego al basquet por placer, para compartir con mis hijos y por hacer ejercicio. ¿Hace falta decir que debemos tener presente quién es el que decide? ¿Los estándares de quién estamos representando? En cierto modo, la búsqueda de una mejor educación a través de estándares y pruebas que demuestran nuestra idoneidad premian la uniformidad más bien que la diversidad y favorecen un aprendizaje prearmado en lugar de favorecer el surgimiento de nuevas posibilidades.

Del mismo modo, quienes defienden los estándares subrayan la importancia de abrir juicios por comparación con el patinaje olímpico y otras actividades donde el éxito se mide por los trofeos que se logran o por los puntajes que se obtienen. Yo creo que es necesario objetar esta metáfora y cuestionar el énfasis que se pone en la necesidad de abrir juicios en lugar de ofrecer apoyo. Prefiero pensar que un maestro es entrenador y no juez,

alguien que apoya y aconseja y no juez o árbitro o calificador. Me gustaría que los maestros consideraran que su rol es el de guiar, llevar de la mano, y hacer comentarios y no decir A, B o C o asignar un puntaje. En mi visión de un mundo más ideal, veo maestros, alumnos y padres funcionando en una suerte de esfera pública donde son parte del equipo negociador de un mejor "yo". En este sentido, me fascinan ciertos proyectos del aula: el tipo de autorreflexión y análisis que se da en una comunidad preescolar de Reggio Emilia (Forman, 1993, 1994) donde los maestros, alumnos y la comunidad toda trabajan juntos para desarrollar e instrumentar planes curriculares, donde se formulan las preguntas adecuadas a fin de instar a los alumnos a la reflexión, se desarrolla la intuición y se aprende; me fascina también el trabajo de Short, Harste y Burke (1996) sobre el desarrollo de inquietudes en las escuelas de Indianápolis (comprometen a alumnos y maestros en la consideración de anomalías, patrones y formas de verse); fascinado estoy también con el trabajo del grupo de Discurso en el Aula de Santa Bárbara (1992a, 1992b), una comunidad de docentes, investigadores y alumnos interesados en comprender cómo se interpreta la vida en el aula y cómo las expectativas y prácticas influyen en las oportunidades de acceso, logro y aprendizaje escolar; y también lo estoy con el trabajo de Fenner (1995), quien utiliza una forma genérica del análisis de Toulmin (1958) –un análisis argumentativo para examinar las conversaciones en el aula y las autoevaluaciones del alumno con sus portafolios– que contempla las formas de ayudar al estudiante a entenderse a sí mismo según sus propias pruebas, suposiciones, pretensiones y metas. El abordaje que Fenner hace de la autoevaluación nos aleja de la típica lista que pide al alumno que detalle de modo bastante vago e insustancial sus puntos fuertes y sus objetivos en una especie casi de adivinanza.

Lamentablemente, en lugar de que el lenguaje sugiera una visión del aula evolutiva y creadora, lo que hacen las metáforas adoptadas por quienes evalúan la idoneidad y establecen los estándares parece más apropiado para desarrollar productos de consumo conectados con pautas prescriptas y procedimientos uniformes de inspección. Es decir, parece que encajaran con nuestra perspectiva de la industria y no con nutrir la potencialidad humana (Wile y Tierney, 1996). Con esto *in mente*, sugiero que deberíamos evaluar la evaluación en función de que esté o no en sintonía con una sociedad de derechos humanos y con nuestros puntos de vista sobre los derechos individuales, oportunidades y libertades.

Me temo que los estándares sólo perpetuarán la situación develada cuando Ellwein, Glass y Smith (1988) analizaron los antecedentes del efecto suscitado por distintas pruebas que medían idoneidad de conocimientos a nivel estadual, que de hecho no hacían sino observar en lugar de mejorar las oportunidades. En realidad, en Ohio sospechamos que los hallazgos de Ellwein y otros (1988) están siendo replicados en otros estados. Con la introducción de la evaluación de idoneidad de conocimientos, más alumnos abandonan la escuela. Es irónico, se suponía que las pruebas iban a mejorar la instrucción, pero cada vez menos alumnos rinden estos exámenes, lo que a su vez sugeriría que más alumnos las pasan. Haciendo que estos abandonos permanezcan en la oscuridad, quienes defienden este tipo de pruebas, y los legisladores mismos, sostienen que la reforma está teniendo efectos positivos

–es decir, que mientras que más alumnos abandonan la escuela con odio o descorazonados por la situación, los legisladores y defensores (incluidos los medios) erróneamente sugieren o publicitan falazmente que cada vez más alumnos logran pasar esas pruebas–.

## **Comentarios finales**

Mis principios para la evaluación emanan de una mezcla de puntos de vista sobre la enseñanza centrados en el niño, puntos de vista evolutivos del niño, perspectivas constructivistas del conocimiento, perspectivas teóricas críticas de la transferencia de poder y concepciones pluralistas de la sociedad. Yo los considero sugerencias y pautas para pensar el porqué, el cómo, el dónde y el cuándo, el quién y el qué de la evaluación.

### **¿Por qué?**

Para desarrollar evaluaciones sensibles a la cultura y no aculturadas.

Para ligar la evaluación a la enseñanza y al aprendizaje.

Para ligar la evaluación a la fijación permanente de objetivos, la toma de decisiones y la evolución del alumno.

Para estar mejor informado y tomar mejores decisiones.

Para desarrollar una evaluación que esté actualizada con la enseñanza y el aprendizaje.

### **¿Cómo?**

Colaborativa, participativa, centrada en el cliente.

Instructiva, contenedora y permanente antes que sentenciosa, discriminatoria y definitiva.

Suplementaria y complementaria antes que calificatoria y sumatoria.

Individualmente, diversamente y no predeterminada.

Juiciosamente.

Evolutivamente.

Razonada.

### **¿Dónde y cuándo?**

Integrada dentro de la vida de los alumnos.

En cada una de las actividades y programas cotidianos.

Dentro y fuera de la escuela.

Oportunamente, periódicamente, continuamente.

### **¿Quién?**

Alumnos, maestros e interesados.

### **¿Qué?**

Aprendizaje permanente: desarrollo, recursos y necesidades.

Complejidades.

Personas y grupos.

Evidencias de progreso y toma de decisiones.

Programas, grupos, personas.

Al describir la esencia de mi propuesta, me gustaría volver al punto en que comencé. Creo que un principio primordial, que tal vez sea mi número trece o quizás el que impide percibir dónde empiezan y terminan los demás es que **la evaluación debería ser evaluada en función de su relación con la**

**enseñanza y el aprendizaje, incluyendo las oportunidades que se ofrecen a quienes aprenden y los derechos y el respeto que se les acuerda.**

Los cambios en mi propia forma de pensar la evaluación comenzaron a producirse cuando me pregunté: Si yo tuviera que evaluar la evaluación, ¿qué criterios emplearía? Mi respuesta a esta pregunta fue que las prácticas evaluadoras deberían otorgar poder a maestros, alumnos y padres por igual. En otras palabras, las prácticas de evaluación deberían enriquecer el proceso de enseñanza como así también el de aprendizaje. Al explorar de qué forma podrían utilizarse las pruebas existentes como herramientas de esta transferencia de poder a maestros y alumnos, comencé a interesarme en si este tipo de evaluación realmente ayudaba a docentes y alumnos (y también a los padres, profesores auxiliares, directores y otros) a lograr una visión más amplia del aprendizaje del estudiante. También quise entender si las pruebas contribuían a desarrollar objetivos y a formular planes de acción, lo cual sugeriría que las prácticas de evaluación realmente otorgaban poder. Mi visión de este proceso de poder incluye:

- ▲ Maestros con una más acabada comprensión (ampliada, refinada, diferente) de las capacidades, necesidades y posibilidades de instrucción de los alumnos.
- ▲ Alumnos con una más acabada comprensión de sus propias capacidades, necesidades y posibilidades de instrucción.
- ▲ Maestros que integren la evaluación con la enseñanza y el aprendizaje (esto implicaría el uso dinámico y permanente de prácticas de evaluación como así también una evaluación a medida de la vida del aula); acomodamiento, adaptación, ajustes, adecuación (cambio de las prácticas de evaluación para armonizarlas con los alumnos y su aprendizaje y ajustando la enseñanza de acuerdo con la información resultante de las evaluaciones).
- ▲ Alumnos comprometidos con su propia autoevaluación a la vez que fijan, buscan y supervisan sus propios objetivos de aprendizaje en colaboración con otros, incluyendo a sus pares, a los maestros y a los padres.
- ▲ Comunidades de docentes, alumnos y padres organizadas alrededor de este proceso de evaluación, mutuamente formadoras y sustentadoras.

El uso de pruebas estandarizadas, de las pruebas que acompañan los programas de lectura impresos, y aun las pruebas preparadas por los maestros mismos, no amplía la perspectiva que el maestro se va formando sobre sus alumnos con el correr del tiempo, ni sugiere al maestro de qué forma puede ayudarlos. Tampoco tales pruebas se hallan integradas en la vida dentro del aula. Más bien tienden a desplazar que a realzar las actividades de enseñanza y aprendizaje.

Del mismo modo, los alumnos rara vez parecen estar comprometidos con aprender a evaluarse a sí mismos. Cuando mis colegas y yo entrevistamos a docentes con quienes comenzamos a trabajar en el tema de la evaluación hace ya diez años, la mayoría de los maestros no conceptualizaba el objetivo de una evaluación como algo que ayudaba a los alumnos a reflexionar y obtener *feedback* sobre su propio progreso, ni tampoco visualizaba las

pruebas como algo que ayudaba a los alumnos a establecer, refinar o alcanzar objetivos de aprendizaje. Cuando entrevistábamos alumnos, encontrábamos que los que estaban en estas clases tendían a presentar una visión limitada y más bien negativa de sí mismos y que se habían impuesto pocos objetivos de aprendizaje. Los intentos por examinar el impacto de evaluaciones basadas más en el alumno produjeron resultados bastante contradictorios. En las aulas en las cuales los portafolios se estaban transformando en una parte integral de la vida en el aula, tanto docentes como alumnos habían desarrollado una comprensión más exhaustiva de sus propias capacidades (Carter, 1992; Carter y Tierney, 1988; Fenner, 1995; Tierney y otros, 1991).

Un estudio efectuado por Shavelson, Baxter y Pine (1992) nos ofrece otra confirmación de lo valioso que es alinear la evaluación con la enseñanza y el aprendizaje en el aula. En sus intentos por examinar las variaciones en los programas de instrucción, llegaron a la conclusión de que las observaciones directas y los nuevos procedimientos emergentes captaban los cambios en el aprendizaje mientras que los métodos más tradicionales (el *multiple choice*, las respuestas breves) no lo hacían. Tales hallazgos no deberían sorprender a aquellos de nosotros que hemos investigado los efectos de la enseñanza sobre el aprendizaje; es decir muy pocos investigadores en alfabetización confiarían en una prueba estandarizada para medir la efectividad de una estrategia didáctica en particular con distintos alumnos. En cambio, sí es probable que buscáramos un rango de medidas, y algunos de nosotros jamás utilizaría medidas *a priori*. De hecho, se ha demostrado más de una vez el poder de los nuevos enfoques en evaluación para evaluar y guiar el desarrollo de un programa y el cambio efectivo del docente (Tierney y otros, 1993).

El diseño de estos nuevos abordajes de la evaluación tiene que ver con una nueva forma de enseñar, evaluar y entender que aquélla está en línea con un conjunto de valores diferentes de los que se propugnaron y aún se propugnan entre la mayoría de los educadores reformistas. Lamentablemente, el poder de algunos psicometristas y de sus arraigados valores en relación con la evaluación hacen que el surgimiento de procedimientos alternativos de evaluación se vea dificultado. De hecho, yo veo el cambio como un cambio que implica una transformación cultural, un vuelco para abandonar una práctica que considero algo totalitaria y ligada a la "vieja ciencia" y a las metáforas que equiparan el aprendizaje del alumno con el control de calidad.

Mike Rose (1995) sugiere en **Vidas posibles** que las aulas son espacios creados, y que las que son exitosas crean espacios donde los alumnos se sienten seguros y a salvo; son las aulas en las cuales los alumnos están dispuestos a ampliar, arriesgar y buscar su autoridad interpretativa por sí mismos y con los otros. En una línea similar, Kris Gutiérrez y sus colegas (Gutiérrez y otros, 1995), al tratar el discurso maestro-alumno, sostienen la necesidad de espacios donde los alumnos y docentes puedan relacionarse o interactuar unos con otros, y no simplemente pasar frente a los demás. La clave es encontrar modos de lograr la participación y la interrelación, evitando el distanciamiento y las respuestas monolíticas.

La evaluación debe encarar la posibilidad de hacer posible un futuro y debe ser asequible, no imposible o improbable. Debemos crear espacios donde

alumnos, docentes y otros puedan lograr su futuro y los espacios en los que las dinámicas y prácticas sean tales que impliquen un desafío, pero que no minen la ecología de la persona que son y en que podrán transformarse.

## Referencias bibliográficas

- Basso, K. (1970) "'To Give Up on Words': Silence in Western Apache Culture." En **Southwest Journal of Anthropology**, **26**, 213-230.
- Bober, S. (1995, julio) **Portfolio Conferences**. Presentación en el Lesley College Literacy Institute. Cambridge, MA, Estados Unidos.
- Carter, M. (1992) "Self-assessment Using Writing Portfolios." Disertación doctoral no publicada. The Ohio State University, Columbus.
- Carter, M. y R.J. Tierney (1988, diciembre) "Writing Growth: Using Portfolios in Assessment." Ponencia presentada en la **National Reading Conference**. Tucson, Arizona, Estados Unidos.
- Ellwein, M.C.; G.V. Glass y M.L. Smith (1988) "Standards of Competence: Propositions on the Nature of Testing Reforms." En **Educational Researcher**, **17** (8), 7-9.
- Fenner, L. (1995) **Student Portfolios: A View from Inside the Classroom**. Disertación doctoral no publicada. The Ohio State University, Columbus, Estados Unidos.
- Forman, G. (1993) "Multiple Symbolizations in the Long Jump Project." En C. Edward, L. Gandini y G. Forman (eds.) **The Hundred Languages of Children** (171-188). Norwood, NJ, Ablex.
- Forman, G. (1994) "Different Media, Different Languages." En L. Katz y B. Cesarone (eds.) **Reflections on the Reggio Emilia Approach** (41-54). Urbana, IL, ERIC/EECE.
- Green, J. y C. Dixon (1994) "Talking Knowledge into Being: Discursive and Social Practices in Classroom." En **Linguistics and Education**, **5**, 231-239.
- Guba, E.G. e Y.S. Lincoln (1989) **Fourth Generation Evaluation**. Newbury Park; Ca., Sage.
- Gutiérrez, K.; B. Rymes y J. Larson (1995) "Script, Counterscript, and Underlife in the Classroom: James Brown versus Brown v. Board of Education." **Harvard Educational Review**, **65**, 445-471.
- Ogbu, J. (1988) "Literacy and Schooling in Subordinate Cultures: The Case of Black Americans." En E. Kintgen, B. Kroll y M. Rose (eds.), **Perspectives on literacy**, Carbondale, IL, Southern Illinois University Press, 227-242.
- Ogbu, J. (1991) "Cultural Perspective and School Experience." En C. Walsh (ed.), **Literacy as Praxis: Culture, Language and Pedagogy**, 25-50. Norwood, NJ, Ablex.
- Phillips, S. (1983) **The Invisible Culture: Communication and Community on the Warm Springs Indian Reservation**. New York, Longman.
- Purves, A. (1982) "Conclusion to an International Perspective to the Evaluation of Written Composition." En B.H. Choppin y T.N. Postlethwaite (eds.) **Evaluation in education: An international review series**. (vol. 5, 343-345.) Oxford, Inglaterra, Pergamon Press.
- Rose, M. (1995) **Possible Lives**. Boston, Houghton Mifflin.
- Santa Barbara Classroom Discourse Group (1992a) "Constructing Literacy in Classrooms; Literate Action as Social Accomplishment." En H. Marshall (ed.) **Redefining Student Learning: Roots of Educational Change**. Norwood, NJ, Ablex, 119-150.
- Santa Barbara Classroom Discourse Group. (1992b) "The Referential and Intertextual Nature of Classroom Life." En **Journal of Classroom Interaction**, **27** (2), 29-36.

- Shavelson, R.; G.P. Baxter y J. Pine (1992) "Performance Assessment: Political Rhetoric and Measurement Reality." En **Educational Researcher**, **21** (4), 22-27.
- Short, K.G.; J.C. Harste y C. Burke (1996) **Creating Classrooms for Authors and Inquirers**. Portsmouth, NH, Heinemann.
- Spiro, R.J.; W.L. Vispoel, J. Schmitz, A. Samarapungavan y A. Boerger (1987) "Knowledge Acquisition for Application: Cognitive Flexibility and Transfer in Complex Content Domains." En B.C. Britton y S. Glynn (eds.) **Executive Control Processes** (177-200). Hillsdale, NJ, Erlbaum.
- Tierney, R.J.; M. Carter y L. Desai (1991) **Portfolio Assessment in the Reading-Writing Classroom**. Norwood, MA, Christopher Gordon.
- Tierney, R.J., J. Wile, A.G. Moss, E.W. Reed, J.P. Ribar y A. Zilversmit (1993) **Portfolio Evaluation as History: Evaluation of the History Academy for Ohio Teachers** (occasional paper). **National Council of History Education**.
- Toulmin, S. (1958) **The Uses of Argument**. Cambridge, Inglaterra, Cambridge University Press.
- Wile, J. y R.J. Tierney (1996) "Tension in Assessment: the Battle over Portfolios, Curriculum and Control." En R. Calfee y P. Perfumo (eds.) **Writing Portfolios in the Classrooms: Policy and Practice, Process and Peri**. Hillsdale, NJ, Erlbaum, 203-218.

*Este artículo fue publicado originalmente en **The Reading Teacher**, Vol. 51, 5, febrero de 1998, **Asociación Internacional de Lectura** (p. 374-390), y traducido por Graciela Mestroni para **LECTURA Y VIDA**.*