

Dificultades para implementar la evaluación formativa

Revisión de literatura

FELIPE MARTÍNEZ RIZO*

El enfoque de evaluación del aprendizaje denominado *evaluación formativa* llama la atención por los beneficios que se pueden esperar de su aplicación en las escuelas, pero no siempre hay conciencia de las dificultades que implica. Esto puede conducir a esfuerzos de intervención que carezcan de la solidez necesaria para dar frutos. Suponiendo conocida la oposición entre evaluación formativa y sumativa, y con base en una revisión de la literatura al respecto, el artículo presenta un panorama de la variedad de factores que inciden en la práctica docente y de las dificultades que eso trae consigo para la transformación de dicha práctica, en particular en cuanto a la adopción del enfoque formativo de la evaluación.

The learning assessment approach called formative assessment seems useful because of the benefits that its implementation in schools can provide, but the people who show interest in it do not always know the difficulties that this system entails. This can lead to intervention efforts that lack in the necessary soundness to bear fruit. Supposing that the opposition between formative and accumulative assessment is already known, and based on a revision of the literature on this matter, this paper offers an overview of the variety of factors that have an incidence on teaching practice and of the difficulties that this entails for the transformation of this practice, and more particularly for what regards to the adoption of the formative assessment approach.

Palabras clave

Evaluación formativa
Rendimiento académico
Educación básica
Actualización de maestros
Programas de intervención

Keywords

Formative assessment
Academic performance
Basic education
Teachers' update
Intervention programs

Recepción: 2 de noviembre de 2011 | Aceptación: 27 de febrero de 2012

* Licenciado en Ciencias Sociales por la Universidad de Lovaina. Profesor-investigador del Departamento de Educación de la Universidad Autónoma de Aguascalientes. Miembro del Sistema Nacional de Investigadores y la Academia Mexicana de Ciencias. Director general fundador del Instituto Nacional para la Evaluación de la Educación (2002-2008). Su trabajo de investigación cubre todos los niveles educativos, en temas de calidad, planeación, evaluación y desigualdad. Actualmente desarrolla una línea de investigación sobre evaluación en aula en educación primaria. Publicaciones recientes: (2012), *La evaluación en el aula*, vol. 1: *Promesas y desafíos de la evaluación formativa*, Aguascalientes, Universidad Autónoma de Aguascalientes; (2012), *La escuela y el futuro. Alegato por la esperanza*, Aguascalientes, Universidad Autónoma de Aguascalientes. CE: fmartin@correo.uaa.mx

INTRODUCCIÓN¹

Un elemento central de la práctica docente es la evaluación que hace el maestro del avance de sus alumnos para alcanzar buenos niveles de aprendizaje. Según el uso que se haga de los resultados, como toda evaluación, la que hace un docente en el aula puede tener el enfoque llamado sumativo (si se limita a llegar a una conclusión sobre el sujeto evaluado, como la de promoverlo al grado superior o reprobado, asignarle una calificación u otra), o bien el denominado formativo, si el resultado se utiliza para orientar al alumno para que su aprendizaje mejore, y al mismo maestro para que reoriente su docencia.

Hoy se reconoce también que, si el docente adopta el enfoque *formativo*, el impacto de sus evaluaciones puede ser más positivo que con uno sumativo. La investigación muestra que esas ideas tienen sustento, pero que no hay que ver de manera simplista la evaluación formativa (EF), es decir, como una panacea que resolverá todos los problemas. La investigación muestra también que no es fácil cambiar prácticas profundamente arraigadas entre los maestros. El propósito de este trabajo es sintetizar los hallazgos de la investigación sobre este tema.

La persistencia de prácticas tradicionales de evaluación en el aula no sorprende si se considera que los esfuerzos por transformarlas no han sido importantes ni sistemáticos. La atención se ha concentrado en las evaluaciones en gran escala, a las que se dedican muchos recursos, mientras que las evaluaciones que más influyen en el desempeño de los alumnos se descuidan casi por completo. La formación inicial de los maestros mexicanos ha sido particularmente débil en lo que se refiere a evaluación, y las actividades de actualización profesional que se ofrecen a los docentes en servicio han descuidado también este

aspecto, que sólo recientemente comienza a ser atendido, aunque en forma limitada.

Esto ocurre no sólo en México, sino también en otros sistemas educativos en los que se podría esperar más avance. En relación con Estados Unidos, Stiggins dice:

El principal desafío que enfrentamos actualmente en las escuelas en lo relativo a evaluación es asegurar que las prácticas sanas lleguen a cada aula, que las evaluaciones se usen para beneficiar a los alumnos... Este reto ha permanecido sin ser atendido por décadas y parece que ha llegado el tiempo de conquistar esa última frontera de la evaluación: el uso efectivo de evaluación formativa para apoyar el aprendizaje (2007: 10).

Para entender por qué no bastan unas horas de actualización para transformar las prácticas docentes hay que tener en cuenta el arraigo de conductas que se han llevado a cabo una y otra vez durante mucho tiempo. Dylan Wiliam hace esta reflexión: “Un maestro o maestra con 20 años de experiencia debe haber formulado alrededor de medio millón de preguntas a lo largo de su carrera. Cuando uno ha hecho algo de cierta manera medio millón de veces es difícil hacerlo de otra forma” (2007: 196).

Modificar prácticas de enseñanza y evaluación en aula no se refiere a conductas que se llevan a cabo de vez en cuando, con tiempo para pensar con calma la manera de actuar en cada caso; se trata de transformar conductas que se repiten muchas veces al día, debiendo reaccionar en segundos a lo que se observa en el grupo y en cuanto a cierto alumno. Esto es mucho más complicado que cambiar solamente las ideas que se tienen sobre el tema.

Las aulas son lugares ruidosos. Los maestros deben tomar decisiones rápidamente,

¹ Se agradece el apoyo del Consejo Nacional de Ciencia y Tecnología para la realización del proyecto del que se deriva este artículo. Una versión más amplia se publicará en la obra *La evaluación en el aula*, vol. 1: *Promesas y desafíos de la evaluación formativa*, Universidad Autónoma de Aguascalientes.

con poco tiempo para reflexionar, y su atención es atraída en múltiples direcciones, por todos y cada uno de los alumnos, por el director y los demás maestros, por las políticas y mandatos locales y estatales (Wilson, 2004: 265).

Una obra de la Organisation for Economic Cooperation and Development (OECD) desarrolla el tema de las *barreras* que obstaculizan el cambio de las prácticas de evaluación de los maestros. Además de problemas logísticos (v. gr. grupos numerosos), los docentes encuentran difícil modificar sus prácticas porque implica cambiar muchas cosas, como la forma de interactuar con los alumnos, de planear sus clases, de atender las diferencias de los alumnos y hasta la manera de concebir el éxito de éstos (OECD-CERI, 2005: 71).

El Center for Educational Research and Innovation, de la OECD (CERI), destaca otro obstáculo que dificulta la extensión de la evaluación formativa:

La tensión entre la evaluación del aprendizaje de enfoque formativo que se hace en el aula y la visibilidad de las pruebas sumativas, así como las evaluaciones nacionales o regionales en gran escala que buscan responsabilizar a las escuelas por alcanzar ciertos estándares, pueden tener consecuencias para las de bajo rendimiento. Con demasiada frecuencia estas pruebas de gran visibilidad determinan lo que sucede en las aulas... Muchos maestros, si es que no todos, perciben esas evaluaciones externas como algo que está en conflicto o incluso se opone a las prácticas de evaluación formativa. Pruebas externas de baja calidad, ordenamientos de escuelas (*league tables*) difundidos en los medios, que usan un reducido conjunto de datos para comparar el desempeño de las escuelas, y la desconexión entre pruebas y currículo también pueden inhibir la innovación (OECD-CERI, 2005: 24).

Sutton señala que cambiar las rutinas básicas del aula mostró ser mucho más difícil de lo que los maestros esperaban, al grado de que se preguntaban si la idea podría funcionar. Los directores comprendieron que se trataba de un juego de largo plazo, que requiere objetivos claros, pasos pequeños, retroalimentación, práctica y perseverancia... Algunos maestros simplemente se sentían espantados por las implicaciones que tendría para su carga de trabajo dar retroalimentación de buena calidad a todos sus alumnos; otros veían que sería inevitable verse abrumados por las demandas de otros actores, como la exigencia de calificaciones por los padres de familia y la competencia entre los alumnos por las calificaciones para acceder a la universidad. Todo eso lleva a los maestros a decir: “Tenemos las manos atadas. Adoptar prácticas de evaluación formativa, por muy deseable que sea, es imposible en términos pragmáticos” (Sutton, 2010: 356).

El Assessment Reform Group sintetiza los obstáculos a enfrentar como sigue:

Las prácticas prevalecientes que prestan más atención a la calificación y la asignación de los alumnos a niveles de desempeño que a retroalimentarlos sobre la forma de mejorar; la falta de conciencia que prevalece entre los maestros sobre las necesidades de aprendizaje de los alumnos; y el alto impacto de los resultados de las pruebas nacionales y estatales, que animan a los maestros a centrar la atención en el contenido de las pruebas y en preparar a los alumnos para ellas (Harlen, 2007: 131).

Harlen retoma de Black la explicación de que modificar las prácticas puede tener como consecuencia que un maestro que tiene confianza en sí mismo se sienta incompetente, lo cual es amenazante; añade que, después de trabajar con maestros que cambiaron tanto sus creencias como sus prácticas, se observó que tanto maestros como alumnos sentían alegría y confianza a medida que los primeros

transferían a los segundos el control sobre su aprendizaje (Harlen, 2007: 132).

Tierney resume la problemática del cambio en cuestión diciendo que:

Cambiar las prácticas de evaluación no es simplemente cuestión de aumentar los conocimientos de los docentes sobre el tema por medio de talleres de actualización, sino un proceso mucho más amplio, que exige una transformación conceptual por parte de todos los involucrados (2006: 259).

Unos autores señalan que, como el nivel de conocimientos que suelen tener los maestros sobre evaluación es limitado, para que sus prácticas cambien hay que capacitarlos. Green, Smith y Brown señalan que este aspecto ha sido descuidado:

Los maestros dedican probablemente entre una tercera parte y la mitad de su vida profesional a actividades asociadas con la evaluación, pero tienen habilidades inadecuadas al respecto. Sólo la mitad de estados americanos exigen formación en evaluación como requisito para la certificación de los docentes, y pocos programas de formación de maestros exigen que los alumnos lleven un curso sobre evaluación, lo que trae como consecuencia que los maestros se sientan mal preparados para responder a las exigencias de la evaluación en el aula (2007: 39).

Otros indican que para que las prácticas cambien no basta con dar capacitación; hace falta el apoyo de otros cambios. Retomando el modelo utilizado en la implementación del enfoque de “evaluación para el aprendizaje” en Hong-Kong, Carless propone un conjunto de factores, agrupados en tres niveles, para lograr cambios en las prácticas de evaluación:

Nivel de los conocimientos y creencias de los maestros. Conocimiento de los principios de

la evaluación formativa. Congruencia entre valores y creencias de los profesores y los principios de la evaluación formativa.

Nivel de la escuela (micro). Cambios en los usos de la evaluación y en la cultura evaluativa; apoyos a profesores, adecuación de aulas, etc. Cambios en la percepción de las funciones y usos de la evaluación por parte de los padres de familia. Apoyo a los maestros por parte de las instancias responsables del currículo.

Nivel del sistema educativo (macro). Existencia de grupos en que los profesores discutan sus prácticas. Clima de reforma curricular. Iniciativas de cambio por parte de las instancias pertinentes. Políticas educativas congruentes con evaluación formativa (Carless, 2005).

La coincidencia de las opiniones en el sentido de que las prácticas de evaluación formativa están poco presentes en las aulas, pese a que en general se considera que su uso trae consigo grandes beneficios, hace obligado preguntarse por las razones de esta aparente paradoja. Otro investigador plantea lo siguiente:

Las experiencias en evaluación a que fueron sometidos los maestros básicamente fueron varias formas de enfoque sumativo. Como alumnos en primaria y secundaria aprendieron que la prioridad era sacar altas puntuaciones en pruebas externas... Es difícil abandonar esos valores si el maestro se ve bombardeado constantemente en los medios y por los padres de familia sobre la necesidad de que sus alumnos consigan otro tanto. Los directores presionan a los maestros para que se concentren en alcanzar altos resultados académicos en sus grupos... En algunos países... los que no lo consiguen enfrentan consecuencias serias, incluyendo el retiro de fondos clave para la escuela o, en algunos casos, el cierre de ésta.

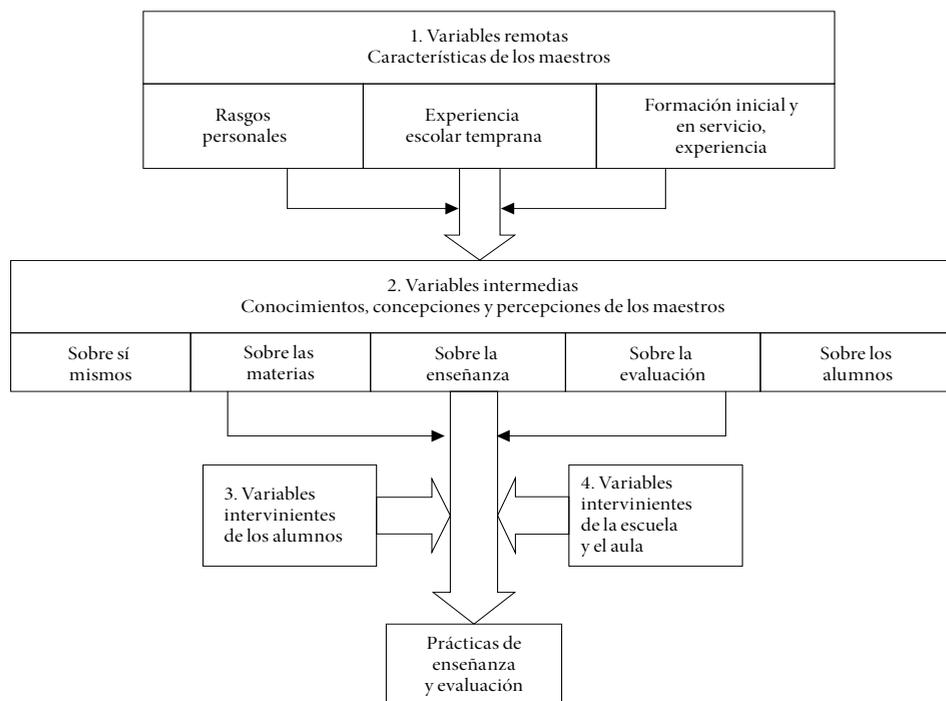
Las pruebas son hoy sumamente destacadas en muchos países... penetran todos

los aspectos de la vida social, empresarial y política... Aunque la intención de tales legislaciones... pueda ser admirable, muchos educadores consideran que su impacto sobre alumnos y maestros es desastroso... Muchos sistemas dan premios a alumnos de alto desempeño, que casi sin excepción se conceden con base en resultados en evaluaciones sumativas. Muchas culturas elogian a los alumnos que alcanzan un resultado o nivel, pero dan poco reconocimiento a los procesos que implica conseguirlo... Lo mismo puede decirse de los documentos curriculares que utilizan los maestros... aunque en ellos se presta algo de atención a los procesos de aprendizaje, el foco principal se pone sobre los conocimientos y

habilidades medidos por exámenes sumativos (Marsh, 2007: 27-28).

FACTORES QUE INFLUYEN EN LA IMPLEMENTACIÓN DE LA EVALUACIÓN FORMATIVA

La dificultad para poner en práctica en las aulas los principios de la evaluación formativa parece deberse, pues, a la influencia sobre las prácticas de factores de diferente peso y antigüedad, que individualmente podrían considerarse menores, pero en conjunto conforman un entramado que se ha mostrado muy difícil de desmontar, y que se trata de describir mediante la figura siguiente:



La figura muestra cuatro bloques de variables que influyen de manera directa o indirecta sobre las prácticas de enseñanza y de evaluación de los maestros. La influencia más directa sobre las prácticas de enseñanza-evaluación proviene del segundo bloque (conocimientos, concepciones y percepciones de los docentes),

pero se ve afectada por las variables intervinientes de los alumnos, así como por las de la escuela y el aula. Las variables del segundo bloque están determinadas por factores que se agrupan en el primero: rasgos personales, experiencia escolar temprana y formación inicial y en servicio de los maestros.

Características de los maestros

- *Rasgos personales*: edad, sexo, estado civil, extracción social y habilidades cognitivas. No parece haber relación sistemática entre variables como edad, sexo o estado civil de los maestros y el desempeño de sus alumnos; hay indicios de que algunos rasgos se asocian con cierto tipo de trabajos, como el sexo femenino con la enseñanza en preescolar.

Darling-Hammond señala que “desde la década de 1940 algunos estudios encontraron correlaciones positivas entre el desempeño de los maestros y medidas de su inteligencia o su habilidad académica, pero en general pequeñas y no significativas estadísticamente”. La autora señala que se encontraron relaciones más fuertes con la habilidad verbal, y que se conjetura que eso puede deberse a que esta última “puede ser una medida más sensible de la capacidad de los maestros para transmitir ideas en forma clara y convincente” (1999: 6).

En igual sentido apuntan otros hallazgos: variables sobre el nivel de habilidad intelectual de los maestros (por ejemplo de competencia lectora o matemática o su desempeño en pruebas de acceso a la universidad) se asocian positivamente con los resultados de los alumnos, pero los resultados tampoco son consistentes (Wayne y Youngs, 2003: 97-100).

- *Experiencia escolar temprana*: la que cada maestro vivió a su paso por la escuela, el tipo de enseñanza que practicaron los maestros que tuvo, las evaluaciones a que fue sometido. Wiliam, después de comentar lo difícil que es hacer de manera distinta algo que se ha hecho de cierta forma medio millón de veces, se refiere al peso de la experiencia

escolar temprana de los maestros sobre su futura práctica profesional en los siguientes términos:

Pero hay una razón más profunda por la que el cambio resulta difícil, aun para maestros sin experiencia. Los maestros aprenden la mayor parte de lo que saben sobre la enseñanza antes de llegar a los 18 años de edad. Así como la mayoría de nosotros aprendemos lo que sabemos sobre cómo ser padres de familia por la experiencia con los nuestros, así también los maestros han interiorizado los “guiones” de lo que se hace en la escuela a partir de su experiencia como alumnos. El mejor programa de licenciatura en educación básica encontrará difícil sobreponerse a los modelos de práctica que sus alumnos, los futuros docentes, aprendieron en los 13 o 14 años que pasaron en la escuela (Wiliam, 2007: 196).

No se encontraron resultados precisos de estudios sobre la posible influencia de esas experiencias tempranas en el desempeño de los maestros, pero no debería sorprender que las relaciones que aparecieran fueran también débiles y poco consistentes ya que, al igual que las características personales de los docentes, se trata también de variables muy distantes de las prácticas. Por ello, su influencia sobre otras variables que ocupan posiciones más cercanas a las prácticas puede ser importante y no se debe de perder de vista, recordando también que se trata de elementos particularmente difíciles de modificar.

- *Formación inicial y en servicio, experiencia docente*: duración de la formación, tipo de institución y programa en que se realizó; número de actividades de actualización profesional que ha llevado

y duración de las mismas; años de servicio. La práctica de los maestros también se puede ver influenciada por la formación recibida cuando estaban en formación, así como por el mismo ejercicio de la profesión. Se consideran sólo aspectos formales como duración de la formación inicial, de las actividades de actualización y antigüedad en el ejercicio profesional, dejando fuera aspectos como el enfoque del programa cuando el maestro se preparaba para su trabajo, lo establecido en el currículo y lo realmente cubierto con más o menos profundidad, la proporción de formación teórica y práctica, la de conocimientos a enseñar y la de elementos pedagógicos, etc. Desde luego el enfoque y la calidad de esos aspectos es importante, pero los estudios no suelen captarlos. Por ello se tratará después de los conocimientos, concepciones y percepciones de los maestros, cuya influencia sobre las prácticas es muy directa, con independencia de si se deben a la formación inicial recibida, a la actualización o al esfuerzo autodidacta de los docentes, a su interacción con otros colegas, o a sus experiencias tempranas.

Según Wayne y Youngs, los primeros estudios sobre la influencia de la formación de los docentes en su desempeño posterior no dieron resultados concluyentes, pero tampoco lo consiguen otros más recientes en relación con áreas curriculares como lectura, escritura o ciencias naturales. Los pocos resultados consistentes se refieren a matemáticas y a maestros de enseñanza media; en este caso el hecho de haber llevado más cursos especializados de matemáticas durante la formación sí parece asociarse consistentemente con mejores resultados de los alumnos (Wayne y Youngs, 2003: 101-103). Sin embargo, estos autores reportan que

algunos estudios sí encuentran una relación positiva entre la calidad del programa en el que se formó un maestro y los resultados de sus estudiantes (2003: 95-96). Según Darling-Hammond:

De las variables sobre la calificación de los maestros, la relación más fuerte se encontró con los exámenes para dar la licencia para enseñar, que miden tanto habilidades básicas como conocimientos para la docencia... Si se controlan las características de los alumnos, la relación entre la calificación de los maestros y el desempeño de los alumnos es aún más fuerte. Un estudio de escuelas de alto y bajo rendimiento de Nueva York, con alumnos similares demográficamente, encontró que la diferencia en la calificación de los maestros (grados en educación, estatus de certificación y experiencia) explicaban aproximadamente 90 por ciento de la varianza total en el desempeño de los alumnos, en el nivel de la escuela, en lectura y matemáticas, en todos los grados estudiados (1999: 12-13).

Por lo que se refiere a la experiencia docente, la misma investigadora señala:

Otros estudios... encuentran relación entre la eficacia del profesor y sus años de experiencia, pero no siempre significativa ni lineal. Si bien muchos estudios han establecido que los maestros sin experiencia (menos de tres años) son típicamente menos efectivos que otros con más antigüedad, el beneficio de la experiencia parece alcanzar un máximo después de unos cinco años... (1999: 9-10).

Por lo que se refiere a los cursos de actualización, William afirma lo siguiente:

Si bien la inversión en actividades de actualización profesional para maestros ha sido durante muchos años un rasgo del panorama educativo, la evidencia de que producen alguna diferencia en el desempeño de los alumnos es deprimentemente escasa. “Nada ha prometido tanto y ha sido un desperdicio tan frustrante como los miles de talleres y conferencias que no trajeron cambio significativo en la práctica una vez que los maestros regresaron a sus aulas” (Fullan) (Wiliam, 2007: 187).

En México no hay información para analizar el desempeño de los egresados de diferentes instituciones que preparan maestros de educación básica. Sabemos que los de primarias públicas casi todos cuentan ya con licenciatura, pero también que hay sectores del sistema en que hay muchos docentes con niveles de preparación inferiores, como en las primarias comunitarias del CONAFE y algunas telesecundarias. Es relevante señalar que el nivel de los alumnos de programas de formación de maestros de educación básica suele ser inferior al de los estudiantes de otros programas de nivel universitario. Se sabe que año con año se ofrece a los maestros en servicio una gran cantidad de cursos de actualización, incluyendo los que obligatoriamente deben llevar todos los maestros al inicio del ciclo escolar, y también que sólo una fracción de los maestros acreditan los cursos no obligatorios; y hay fundamento para afirmar que, en muchos casos, la calidad de dicha oferta es muy baja.

Conocimientos, concepciones y percepciones de los maestros

Por su ubicación más lejana respecto de la práctica, las variables de los tres grupos del

primer bloque influyen más directamente sobre los elementos del segundo; la influencia de estos últimos sobre las prácticas de los maestros, en cambio, es más directa.

- *Sobre sí mismos*: autopercepción y autoestima del maestro en general; manera en que concibe su función docente y, muy especialmente, su función evaluadora. Aunque no se cuente con estudios que analicen estos aspectos, parece razonable esperar un mejor desempeño de un maestro que se vea a sí mismo como un profesional y tenga confianza en sí mismo; que considere que su profesión es digna y atractiva y no la única opción que tuvo a su alcance para buscar cierta movilidad social; y que conciba su función como docente y como evaluador como una tarea cuyo fin último es el bien de sus alumnos.
- *Sobre las materias*: concepción del lenguaje, las matemáticas, las ciencias de la naturaleza y demás áreas curriculares que tenga el maestro, así como la valoración que tenga de ellas, su interés, desinterés o eventual rechazo de algunas, etc. Será difícil que un docente consiga que sus alumnos desarrollen el hábito de la lectura si él o ella no lo tienen; o que consiga buenos resultados en matemáticas o ciencias si él mismo es alérgico a los números o tiene una concepción pre-científica del mundo. Sin embargo, los estudios no muestran una relación clara entre conocimientos del maestro y su práctica.

El conocimiento de la materia es otra variable que se podría pensar se relaciona con la eficacia del docente. Aunque esta suposición encuentra algún respaldo, los hallazgos no son tan fuertes y consistentes como uno podría suponer... La mayoría de los

estudios muestra relaciones pequeñas, no significativas estadísticamente, y tanto positivas como negativas... Los estudios han mostrado una influencia un poco más fuerte y consistente de los cursos de contenido pedagógico sobre la efectividad de los docentes (Darling-Hammond, 1999: 6-8).

Parece haber una relación más fuerte de los conocimientos del maestro con el rendimiento de los alumnos en el caso de matemáticas, y trabajos recientes muestran también influencia significativa del conocimiento que tiene el maestro de matemáticas, pero no de sus conocimientos “puros” sino, muy concretamente, de su dominio de lo que Hill, Rowan y Ball (2005) llaman “conocimiento de las matemáticas para la enseñanza” (*Mathematics knowledge for teaching, MKT*) (cfr. Wiliam, 2007: 186).

- *Sobre enseñanza*: incluye tanto las ideas teóricas sobre los diferentes enfoques de la enseñanza, como las concepciones prácticas que el maestro tenga respecto de lo que es viable o no en las condiciones concretas en que trabaja, lo que funciona, etc.
- *Sobre evaluación*: ideas del maestro sobre el papel de la evaluación en el proceso de enseñanza-aprendizaje, formas de realizarla, uso de resultados tanto para retroalimentar su propia práctica como para dar información a los alumnos y a los padres de familia, sobre el impacto afectivo de la evaluación, etc.
- *Sobre los alumnos*: forma como el maestro los percibe en general (*v. gr.* como sujetos activos de su aprendizaje o como receptores pasivos de conocimientos) y en particular (los de este grupo y este curso, o incluso este alumno o alumna), como capaz o incapaz, listo o no, aplicado o negligente, dócil o rebelde, etc.

Muchos factores influyen en las prácticas de evaluación, que dependen de los contenidos de que se trate, los recursos de que se disponga, el apoyo de los padres y las características del entorno, etc. Especial importancia tiene la influencia de las concepciones de los mismos docentes sobre sus propias decisiones relativas a qué actividades de evaluación asignar a los alumnos, así como sobre la lectura que hacen luego de los resultados:

Conocer las concepciones de los profesores a la hora de definir las actividades de evaluación de los aprendizajes de los alumnos y sus estrategias de enseñanza nos permitirá comprender sus criterios de evaluación y la jerarquización que realizan de los contenidos... El conocimiento profesional (asociado al conocimiento pedagógico general, al disciplinario y al didáctico del contenido) y la experiencia profesional son elementos importantes en estas construcciones mentales que influyen en las prácticas. Estas concepciones serán las que definan las decisiones que se tomarán en relación con las evaluaciones... (Katzkovicz, 2010: 114-117).

Retomando los resultados de trabajos de otros investigadores sobre las creencias y concepciones de los maestros en relación con las prácticas de evaluación, la autora mencionada señala que se ha encontrado que muchos docentes tienen bajas expectativas en cuanto al rendimiento de los chicos de medio desfavorable por lo que saben al respecto, pero que eso a su vez incide negativamente en el rendimiento de esos mismos alumnos.

Otras investigaciones muestran que la forma en que los docentes conciben la ciencia los lleva a privilegiar el uso de acercamientos que supuestamente miden con mucha precisión lo que domina cada alumno —aunque se trate de aspectos superficiales, como ocurre muchas veces con las pruebas conformadas por preguntas estructuradas— y a descuidar estrategias de evaluación que pueden dar lugar

a respuestas imprecisas sobre aspectos fundamentales del trabajo científico, como las que se refieren a las relaciones ciencia-técnica-sociedad-ambiente, entre otras, que contribuyen a la transmisión de una imagen deformada de la ciencia (Katzkovicz, 2010: 115). La misma autora cita a Moreno y Azcárate: “Las concepciones docentes se constituyen como organizadores implícitos referidos a creencias, significados, conceptos, proposiciones, imágenes mentales y preferencias que influyen tanto la manera de percibir la realidad como las prácticas que implementan” (Katzkovicz, 2010: 116).

Variables de los alumnos

Como muestra el esquema, este bloque de variables y el siguiente ocupan una posición lateral respecto de la relación principal, que va de las variables remotas (características de los maestros), e intermedias (conocimientos, concepciones y percepciones), a las prácticas.

Así pues, las variables de los alumnos, así como las de la escuela y el aula, tienen un estatus de variables intervinientes que modifican o modulan la influencia de las variables remotas (las características de los maestros) y las próximas (sus conocimientos, concepciones y percepciones) sobre las prácticas docentes y, en particular, las de evaluación.

El bloque relativo a las variables de los alumnos comprende las características reales de los estudiantes —vs. la percepción de ellos que tenga el maestro— individualmente o en grupo: niños de familias urbanas o rurales, acomodadas o pobres, indígenas, de rendimiento normal o no, que asisten con regularidad o faltan con frecuencia, homogéneos o no, etc.

La naturaleza interviniente de estas variables respecto de la relación entre las próximas y remotas y las prácticas de los maestros, significa que la forma en que los docentes se conducen en el aula (sus prácticas docentes y de evaluación) depende directamente de sus características personales y de sus conocimientos y concepciones, pero que se ve afectada de

alguna manera por el tipo de alumnos que tienen. El maestro debe ajustar sus prácticas teniendo en cuenta si sus alumnos aprenden con más o menos rapidez, si son especialmente inquietos o tranquilos, entre otros muchos aspectos, como el tipo de motivaciones a que son sensibles. Al respecto Brookhart cita estudios que han encontrado que los alumnos tienen cuatro tipos de patrones motivacionales, según que estén orientados al aprendizaje o al éxito, o bien que sean no comprometidos o elusivos (*avoidant*).

La existencia de diferentes patrones del enfoque de motivación que tienen los alumnos respecto al aprendizaje tiene consecuencias para la evaluación formativa en aula. La retroalimentación que se da a los alumnos que no tienen éxito llega en un momento en que se están sintiendo mal por su fracaso. La retroalimentación debe hacer algo para tener en cuenta esos sentimientos negativos si quiere romper el círculo vicioso del fracaso, de lo contrario no servirá para hacer que la atención de los alumnos se concentre en el trabajo (Kluger y DeNisi)... Los alumnos necesitan tener tanto habilidad para tener éxito en el aula, como voluntad para ello (*skill & will*) (Brookhart, 2007: 52-53).

Variables de la escuela y el aula

El cuarto bloque comprende otras variables intervinientes: si la escuela es grande o chica; de organización completa o incompleta; pública o privada; bien o mal dotada de mobiliario, infraestructura, equipamiento y material didáctico; en medio favorable u hostil; en que los padres de familia participan activamente o no; con trabajo o no; con o sin apoyo del director e instancias de supervisión, etc. Igualmente, si los alumnos son muchos o pocos; de uno o varios grados; con necesidades especiales, de bajo rendimiento, repetidores, etc.

Es claro el carácter modulador o mediador de este tipo de variables en la relación que

va de las características, conocimientos y concepciones del profesor a su práctica: aunque el maestro tenga ideas muy buenas sobre cómo enseñar o cómo evaluar, no lo podrá hacer de la forma que quisiera con un grupo muy grande, o sin un mínimo de materiales y apoyos.

Otras variables de este grupo se refieren a las orientaciones que debe seguir el maestro por la normatividad o las directrices marcadas por el director y la supervisión. En primer lugar el currículo, muchas veces demasiado amplio y con exceso de contenidos, lo que propicia que los maestros enseñen superficialmente en vez de tomarse tiempo para desarrollar habilidades cognitivas de mayor complejidad. Las orientaciones sobre evaluación, que muchas veces privilegian usos sumativos e instrumentos estandarizados, tanto los que se aplican en escala nacional o estatal por parte de las autoridades, como los que se manejan en escala menor y de manera menos formal, pero no por ello menos influyente.

Por lo que se refiere al currículo, Rindone y McQuarrie afirman que

...sigue habiendo serios obstáculos a la implementación de prácticas de evaluación formativa en la enseñanza cotidiana. Hasta cierto punto esto se debe a la confusión sobre lo que es la evaluación formativa, pero el mayor obstáculo es el antiquísimo problema de la amplitud *versus* la profundidad. Los maestros de educación básica o superior deben cubrir demasiados contenidos con poca o ninguna profundidad, sin tiempo para la práctica y sin evaluación formativa (2010: 325).

En cuanto a la evaluación, Black señala el peso que tienen en muchos sistemas educativos las pruebas sumativas externas, que tienen consecuencias de gran envergadura, y añade:

La necesidad de enseñar tácticas superficiales que mejoren el desempeño en pruebas igualmente superficiales limita la libertad de maniobra del maestro, así como la

confianza en él para que desempeñe un papel más responsable en la evaluación de sus alumnos (Black, 2010: 363).

La participación de los padres, en principio positiva, puede tener consecuencias negativas, pues los maestros se pueden ver presionados a utilizar formas de evaluación más distantes del enfoque formativo pero que los padres consideran más objetivas, con lo que las decisiones basadas en ellas son más defendibles y se evitan conflictos (McMillan, 2003: 38).

LECCIONES PARA LAS INTERVENCIONES EN LA PRÁCTICA DOCENTE

Así pues, hay dificultades considerables para poner en práctica la evaluación formativa, al menos en el contexto anglosajón en que se han conducido las investigaciones revisadas. Es fácil perderse en el paso de los principios a las políticas y de éstas a la práctica; no bastan estrategias como utilizar terminología nueva (*assessment for learning* en lugar de *formative assessment*) cuando se constata que los esfuerzos previos no dan los resultados esperados, “con la improbable esperanza” de que el nuevo término no se asocie con los anteriores fracasos. Utilizando como ejemplo un trabajo de Smith y Gorard, Stobart señala que evitar calificar el trabajo de los alumnos, haciendo sólo comentarios elogiosos o para animarlos, refleja una pobre concepción de la noción de retroalimentación lo que, además de no dar resultados, condujo a la “suprema ironía del alumno que dijo: maestra, yo quiero conocer mis calificaciones porque los comentarios no nos dicen mucho” (Stobart, 2005: 3-4).

En seguida se revisan experiencias que podrán orientar los esfuerzos de intervención que busquen modificar las prácticas de los maestros en lo relativo a evaluación, para no caer en errores detectados en esfuerzos anteriores y aprovechar experiencias positivas de formación que incluyen lo relativo a evaluación, pero que pueden tener propósitos más amplios.

LIMITACIONES DE ESFUERZOS POR INTRODUCIR NUEVAS PRÁCTICAS

La experiencia de dos sistemas educativos muy distintos, como son el del Reino Unido y el de Francia, coinciden en que en ambos se ha encontrado considerable dificultad para que el enfoque de la evaluación formativa se generalice. Muchos maestros ingleses creyeron que podrían hacer que sus evaluaciones fueran formativas, pero “en los primeros tiempos de las reformas, las dimensiones y la complejidad de la tarea de llevar registros y recoger las evidencias amenazaban con abrumar a los maestros”, de manera que se reportaba que los cambios en las evaluaciones seguían siendo intuitivos, más que basados en evidencias. En Francia, las leyes de 1989 y 2005 intentaron provocar un cambio cultural para pasar de una concepción de la evaluación sumativa a una formativa, pero “hasta ahora estas iniciativas de arriba hacia abajo parecen haber tenido escaso impacto sobre las prácticas de aula, según describen la investigación académica y los informes de los supervisores” (Osborn y Raveaud, 2010: 425).

Al abordar el tema de las variables relacionadas con la formación de maestros se han citado opiniones sobre el “frustrante resultado” de muchas actividades de actualización profesional. Wiliam se pregunta cómo podremos hacer que el desempeño de los alumnos mejore, y responde:

Muchas veces confiamos en soluciones rápidas, que rara vez tienen éxito. Para elevar exitosamente el desempeño de los alumnos debemos mejorar la calidad de los maestros... en particular los que ya están en la escuela. Tenemos que analizar con cuidado los costos y los beneficios de las posibles reformas (Wiliam, 2007: 184).

Wiliam considera dos estrategias: una *gradual*, elevando los requisitos que se exigen para aceptar a los nuevos docentes en la

profesión, lo que llevaría años para tener efectos; y otra *rápida*, permitiendo el acceso a la profesión de maestro a personas que no hayan recibido formación para ello, pero señala que la investigación ha mostrado que este tipo de docentes no resulta mejor que los formados de la manera tradicional (2007: 186-187). Una estrategia más tiene como elemento clave “ayudar a los maestros a desarrollar prácticas de evaluación formativa” lo que, según Wiliam, “tiene una relación costo-beneficio mejor, pero cambiar lo que los profesores hacen día a día no se puede lograr efectivamente con métodos tradicionales, como talleres de verano” y añade:

Ahora es claro que la principal razón por la que la actualización profesional no ha conseguido en general tener impacto en el rendimiento de los alumnos es porque no ha implementado lo que la investigación muestra que hace la diferencia en cuanto al aprendizaje de los alumnos (2007: 187).

Después de estas consideraciones, nuestro autor concluye: “Después de muchos falsos comienzos y callejones sin salida he llegado al convencimiento de que la mejor forma de apoyar a los maestros a adoptar evaluaciones formativas... es mediante comunidades de aprendizaje de maestros basadas en la escuela” (Wiliam, 2007: 196).

A una conclusión idéntica, en contexto latinoamericano, llega Rebeca Katzkovicz, quien, tras analizar la importancia que tienen las concepciones de los maestros sobre sus prácticas, subraya “la necesidad de contar con ámbitos de formación y reflexión profesional para los docentes” para mejorar sus estrategias en el aula y, a través de ellas, los aprendizajes de todos sus alumnos y sus posibilidades de éxito escolar. En concreto recomienda:

Favorecer los espacios de reflexión profesionales en las instituciones para trabajar la cultura de la evaluación en ellas, así como

para develar las creencias, las racionalidades y las concepciones de los profesores en relación con sus prácticas de enseñanza y evaluación (2010: 123-124).

NECESIDAD DE PROCESOS MÁS ADECUADOS DE CAPACITACIÓN

Black sintetiza algunos problemas para implementar evaluaciones formativas: falta de una definición clara del concepto; conflicto entre la evaluación formativa y las presiones de las pruebas sumativas dirigidas a responsabilizar a las escuelas; y no aceptar que la evaluación formativa exige un cambio difícil en cuanto al papel del maestro (2010: 359).

Black da ejemplos de interpretaciones superficiales de la evaluación formativa: un docente puede formular preguntas abiertas, pero luego califica las respuestas en términos de correcta o incorrecta, en vez de usarlas para reorientar el trabajo de los alumnos; puede pedir a los alumnos que se autoevalúen, pero no usa la información resultante para dar retroalimentación. Finalmente, señala que el obstáculo más grande es que implementar prácticas formativas nuevas es una tarea muy demandante para los maestros que, en muchos casos, implica un cambio profundo de sus creencias sobre su papel en el aula, y que esos cambios no se pueden conseguir con una breve sesión de entrenamiento, sino que necesitan el apoyo sostenido de la colaboración colegiada. Un obstáculo más es que los maestros necesitan una formación básica en temas de evaluación, que incluya una buena comprensión de los criterios de validez y confiabilidad (Black, 2010: 363).

Conviene reflexionar sobre esto, porque una inadecuada comprensión puede llevar a una orientación equivocada de los esfuerzos de formación. Hay que precisar la idea de que implementar evaluaciones formativas supone un cierto dominio de temas de evaluación en general por parte de los maestros. Popham se

pregunta si la *alfabetización en evaluación* que se espera tengan los maestros es una moda o realmente algo importante (2009).

Popham distingue los conocimientos sobre evaluación que se refieren al aula y los que se refieren a pruebas en gran escala, que son aquellas que se usan para la rendición de cuentas por escuelas y maestros. En su opinión, en ambos casos un conocimiento inadecuado puede traer consigo efectos negativos para la calidad, por lo que un conocimiento básico (*literacy*) al respecto parece una condición *sine qua non* para los educadores. Este autor propone 13 temas a incluir en las actividades de actualización, y señala que este tipo de formación en servicio será necesario hasta el momento en que los programas de formación inicial lo hagan superfluo porque todos los futuros docentes salgan con una buena preparación sobre estos aspectos. Añade que el uso de “comunidades profesionales de aprendizaje” puede complementar las actividades tradicionales de actualización, o sustituirlas (Popham, 2009: 8-10).

El error en cuanto al enfoque de este tipo de formación es orientarlo como si se dirigiera a investigadores especialistas en evaluación y no a maestros, cuyas necesidades son distintas. Hace ya una década Stiggins había hecho una severa crítica que los responsables de cursos de actualización en evaluación para maestros deben tener presente: después de repasar el “preocupante panorama” que presentaban las evaluaciones de aula, en contraste con los sofisticados avances de las pruebas en gran escala, Stiggins cuestiona que el especialista en pruebas eluda su parte de responsabilidad al respecto, atribuyendo todo el fenómeno al desinterés de otros actores, en especial, maestros y autoridades educativas. En su opinión, es cierto que los especialistas han planteado el tema, pero se pregunta y responde:

¿Por qué no hemos sido escuchados en las salas de las escuelas? En mi opinión porque no nos hemos educado nosotros mismos

en las realidades de la vida de las aulas, de manera que seamos capaces de traducir nuestras ideas a aplicaciones que puedan ser puestas en práctica por los maestros, en beneficio de sus alumnos (2001: 6-7).

Con base en los resultados de su propio trabajo etnográfico (Stigglins y Conklin, 1992) sobre lo que pasa en las aulas, y en relación con las actividades de desarrollo profesional sobre evaluación dirigidas a maestros, Stigglins afirma que:

Los resultados dejaron claro el reto a enfrentar. Comprendimos lo que los maestros necesitan saber sobre evaluación y no era lo que tratábamos de enseñarles, sin éxito, según reflejan los textos más comunes de introducción a la medición... Durante décadas hemos escrito textos para maestros que incluyen la expectativa de que ellos van a apreciar la confiabilidad de una prueba, a calcular índices de validez de la misma y a hacer análisis de ítems para determinar su dificultad y poder de discriminación. Pienso que los autores que tienen tales expectativas revelan su propia ingenuidad sobre la vida en las escuelas y las aulas (2001: 9-14).

Según Stiggins, sería bueno que los maestros hicieran eso, pero que no ocurrirá, y añade:

¿Quiere esto decir que los maestros no pueden aprender lo relativo a las fuentes que afectan la confiabilidad o la validez y a evitarlas? Por supuesto que no. Los maestros pueden y deben aprender esas cosas, pero tenemos que ayudarles a hacerlo en términos que puedan entender y aplicar en su contexto (2001: 14).

Un rasgo que aparece en todas las recomendaciones sobre las características de un buen proceso de actualización para maestros en servicio es el que consiste en la conformación de una *comunidad profesional*

de aprendizaje que, según la propuesta de Popham, debe ser un complemento fundamental de las actividades tradicionales, o incluso puede sustituirlas.

El punto de partida para modificar prácticas de evaluación debe ser un proceso participativo de reflexión que involucre a maestros y otros actores en comunidades de aprendizaje en donde sea posible cuestionar los conocimientos previos y entender los nuevos enfoques y su sustento, teniendo en cuenta que el cambio no solamente debe referirse a los conocimientos de los profesores, sino también a sus actitudes, de forma que estén dispuestos a experimentar nuevas formas de trabajo, con el riesgo que ello supone (Aschbacher, 1993).

Otro punto fundamental es que la actualización para la evaluación formativa no se puede quedar en aspectos teóricos. Ya se han visto ejemplos de las consecuencias de que sólo se cuente con comprensiones superficiales de conceptos clave que ocultan la ausencia de cambios reales bajo la apariencia de una nueva terminología. Allal y Mottier López señalan que:

La búsqueda de referentes teóricos puede llevar a una visión cada vez más abstracta de la evaluación formativa, alejada de las realidades de la práctica en el aula. Por ello es esencial articular el trabajo teórico con el estudio de cómo se pone en práctica realmente la evaluación en el aula (2005: 251).

La obra de la OCDE en la que aparece el texto anterior precisa:

Los maestros necesitan traducir ideas abstractas —como las del aprendizaje centrado en el niño— en prácticas concretas. Los programas vagos puramente conceptuales tienen pocas probabilidades de llegar lejos o durar mucho, especialmente porque los maestros son personas ocupadas que enfrentan una demanda creciente sobre su tiempo (OECD-CERI, 2005: 89).

EL PUNTO CLAVE: LA BUENA RETROALIMENTACIÓN Y SU DIFICULTAD

El uso de alguna retroalimentación por parte de los maestros forma parte de las tradiciones más antiguas del oficio, en el marco de tradiciones diferentes de las corrientes pedagógicas contemporáneas, y todavía presentes en muchas aulas escolares. Según Resnick *et al.*:

Buena parte de la instrucción de enfoque asociacionista, al menos en las sociedades occidentales, se origina en formas tempranas de educación religiosa en la Iglesia. Primero los protestantes, y luego los católicos, desarrollaron una instrucción elemental que se desarrolla oralmente en la forma de preguntas y respuestas, y se conoce como catecismo... La forma de catecismo todavía está presente en lo que ahora llamamos "enseñanza estándar basada en recitación"... (2010: 400).

Luego contrasta ese tipo de interacción con la que tiene lugar en las aulas actuales, las cuales se organizan según concepciones interpretativas. Según Resnick *et al.* (2010), las nuevas formas de *conversación* constituyen un importante cambio en comparación con el catecismo y su recitación, aunque reconocen que el papel del maestro en la conducción del proceso sigue siendo importante y señalan que, por ello, hay oportunidad para que la evaluación de conocimientos de los alumnos se integre como parte del proceso. Sin embargo:

Lo que cuenta como conocimiento se amplía; los bloques a aprender y de los que se habla son mucho mayores; hay más oportunidades de que el maestro se vea sorprendido por las comprensiones (e incomprensiones) de los alumnos, y es claro que se trata de mucho más que de encontrar la respuesta única que el maestro espera... Las oportunidades de evaluación están integradas en las actividades de aprendizaje y es difícil pensar que se puedan separar como ítems

individuales de una prueba (Resnick *et al.*, 2010: 401).

La retroalimentación tipo catecismo pasó a las aulas laicas y se desarrolló en los modelos de enseñanza inspirados en el asociacionismo y el conductismo, con refuerzos positivos o negativos. Con Bloom y el *mastery learning* el neo-conductismo llegó a una concepción de la retroalimentación que la identificaba con la diversificación de la instrucción, de forma que se adaptara al ritmo y el estilo de aprendizaje de cada alumno. Bloom señalaba que era frecuente que la instrucción diversificada se redujera a que el maestro repitiera más despacio y en voz más alta lo que había explicado una primera vez (cfr. Guskey, 2007: 67).

Por su parte, Sadler elabora lo relativo a la retroalimentación como sigue:

...en su primera fase, la teoría de estímulo-respuesta identificaba la noción de retroalimentación con el conocimiento de los resultados, que podía llevar a corrección (cuando la respuesta era incorrecta) o refuerzo (cuando era correcta), con lo que la motivación aumentaría y se conseguiría un mejor rendimiento...

La segunda fase tomó un camino distinto: retroalimentar tenía que ver con elogiar el esfuerzo, lo que llevaría a mayor auto-estima, más esfuerzo y, finalmente, mejor rendimiento... en casos más complejos se requieren formas más sofisticadas de retroalimentación... [incluyendo] descripciones de rasgos del trabajo del alumno, comentarios evaluativos, sugerencia de caminos o arreglos alternativos...

Una opción alternativa es disociar la retroalimentación relacionada con el ego y la relativa a la tarea, promoviendo sobre todo que se alcancen logros verdaderos, lo que a su tiempo llevará a mejorar la autoestima... es la calidad y no la cantidad de retroalimentación lo que merece nuestra mayor atención (1998: 78-84).

En forma similar, Anijovich distingue retroalimentación afectiva y cognitiva, individual y grupal, cantidad y calidad. La autora señala que los maestros usan distintos códigos para retroalimentar, que van desde sólo marcar los errores, o bien, además de ello, anotar la respuesta correcta; escribir comentarios al respecto y sugerir cómo mejorar o incluso dialogar sobre la retroalimentación con los alumnos, y añade que los estudios indican que:

La retroalimentación que se brinda asume en pocas ocasiones la función de orientar. No es frecuente el planteo de preguntas o la invitación a compartir reflexiones con los estudiantes. Un factor crítico es el tiempo disponible de los docentes para ofrecer retroalimentaciones sustantivas a todas las producciones de los alumnos. Para evitar el “todo o nada” es posible, por ejemplo, definir criterios y determinar en forma explícita qué trabajos, rotativamente, van a recibir retroalimentación escrita, con qué frecuencia, etc. (Anijovich, 2010: 17).

Según esta autora, al hablar de la retroalimentación habría que preguntarse sobre qué es lo que impacta “los conocimientos de los alumnos, sus procesos de aprendizaje, su autoestima, su motivación o sus acciones futuras”. Anijovich señala que algunos estudios distinguen la retroalimentación que se centra en la autoestima y la que tiene como foco la tarea, y apunta:

...se ha observado que este tipo de retroalimentaciones puede provocar mejoras en el aprendizaje... a partir de un incremento en la autoconfianza y en la capacidad de logro, pero también puede inducir un efecto contrario, es decir que, apoyado en comentarios positivos, el alumno no reconozca que necesita mejorar algunos aspectos de su trabajo. Para superar este problema, algunos autores señalan que la retroalimentación es más productiva si se centra en la tarea, en

cómo el alumno la resuelve, y cómo autorregula su aprendizaje (2010: 16).

Por lo que se refiere a la cantidad y calidad de retroalimentación, la autora señala:

...los buenos alumnos, con buenas producciones, recibían un comentario genérico como “muy buen trabajo”, “buen nivel de análisis”. Estas frases implican muy poca retroalimentación, vinculada a aspectos específicos de la producción evaluada. Los alumnos con dificultades importantes también reciben poca información específica. Suelen obtener señalamientos generales como “rehacer el trabajo” o comentarios que sirven más para justificar la calificación que para retroalimentar la producción o el desempeño del alumno. Es notable que sean los alumnos de nivel intermedio los que reciben más cantidad de información en concepto de retroalimentación. Esta es, además, más específica, y contiene indicaciones como “justificar la pregunta 2 apoyándose en la bibliografía”, “rehacer el diagrama revisando los conceptos incluidos”, etc. (Anijovich, 2010: 16-17).

Anijovich se pregunta si la retroalimentación debería ser individual o grupal, y comenta que, en caso de retroalimentación grupal, si el compromiso del grupo es bajo, *posiblemente nadie se sienta dueño de la información que aporta el docente y ésta se perciba como ajena, pasando inadvertida para los alumnos*. Pero si el compromiso y el conocimiento que tienen los alumnos de sus propias fortalezas y debilidades son mayores, entonces *las devoluciones grupales enriquecerán el intercambio de ideas, estrategias y recursos*. En cuanto a los modos en que se ofrece retroalimentación, la autora señala que dependen de factores

...como las experiencias anteriores en relación con el tipo de respuesta obtenida de los alumnos, el tipo de creencia sobre la

capacidad de los estudiantes, la construcción de un clima de confianza en el aula, y el tipo de comunicación y de vínculo establecido entre los educandos (Anijovich, 2010: 17).

Llegamos así a la manera en que se concibe la retroalimentación en evaluación formativa, basada en la *revolución cognitiva*, con énfasis en la autorregulación del aprendizaje por los alumnos. Conviene reiterar que una buena evaluación formativa no se puede reducir a no manejar evaluaciones sumativas, lo que no mejora por sí mismo la retroalimentación que necesitan los alumnos para orientar sus esfuerzos. Recordemos *la suprema ironía* que se aprecia en el trabajo de Smith y Gorard, cuya estrategia de no dar calificaciones provocó reacciones adversas de los alumnos, porque sin calificaciones no tenían información suficiente sobre su desempeño, no había buena retroalimentación “Quiero conocer mis calificaciones porque los comentarios no nos dicen mucho” (Stobart, 2005: 4).

Otros autores reflexionan sobre la dificultad de la evaluación formativa y, en especial, de formular preguntas que permitan dar luego retroalimentación eficaz:

Los alegatos en pro de la evaluación formativa... parecen dar por supuesto que el proceso de evaluación es enteramente transparente, con profesores capaces de hacer que los alumnos produzcan respuestas claras en forma inequívoca. Un buen número de estudios sugieren que no es así, sobre todo entre alumnos pequeños... esto no quiere decir necesariamente que los alumnos no sepan la respuesta o no entiendan la pregunta, sino que piensan que se les pregunta algo diferente, o que tienen que responder más bien una pregunta implícita inferida a partir del contexto, y no la pregunta explícitamente formulada (Torrance y Pryor, 1995: 309).

Según estos autores, los alumnos intentan siempre interpretar las preguntas de los

maestros y “dar sentido” a lo que se les pregunta, más que entender lo que se podría considerar el significado “obvio” de una pregunta en particular. Además:

...la adquisición de nuevos conceptos parece depender principalmente de las experiencias previas de aprendizaje y de lo que el aprendiz puede estar listo para aprender con la ayuda de un adulto o un compañero. Por ello, para que el aprendizaje sea más efectivo, debe tener lugar en la “zona de desarrollo próximo” [con una visión constructivista de la educación] ...la evaluación debería implicar ofrecer al alumno experiencias que pueda captar con las ideas y habilidades que ya tiene, pero que también exigen una extensión o modificación de esas ideas y habilidades (Torrance y Pryor, 1995: 316).

Un requisito para que los alumnos puedan mejorar es que sean capaces de “apreciar qué es un trabajo de buena calidad y que tengan las habilidades evaluativas necesarias para comparar con objetividad la calidad de sus propias producciones con el estándar más elevado” (Torrance y Pryor, 1995: 307).

Para que el alumno desarrolle la capacidad de identificar productos de buena calidad con los cuales contrastar su trabajo son útiles los “mapas de progreso” (*learning progressions*).

Una *learning progression* es una secuencia de habilidades particulares (*subskills*) y conocimientos habilitadores que el alumno debe manejar en el proceso de llegar a dominar un objetivo curricular más amplio. En un sentido casi literal, las *learning progressions* son mapas que orientan sobre cómo hacer evaluación formativa... ¿Qué deben evaluar los maestros? Esas habilidades particulares y cuerpos de conocimiento habilitador que el mapa de progreso identifica como claves para que los alumnos lleguen a dominar el objetivo curricular final. ¿Cuándo deben

evaluar esas habilidades y conocimientos habilitadores? Antes de pasar al siguiente bloque constructivo en el trayecto, cuyo dominio se supone que depende del dominio de su predecesor (Popham, 2008: 24-30).

Los alumnos son actores clave de la evaluación formativa. Por su acceso inmediato y constante a sus propios pensamientos, acciones y trabajos, son la fuente fundamental de una buena retroalimentación; esta postura no es novedosa para los estudiosos que trabajan con el enfoque del aprendizaje autorregulado, puesto que:

...se sabe desde hace mucho que el aprendizaje eficaz tiende a monitorear y regular su propio aprendizaje y que por ello aprende más y tiene más éxito académico... Además de tener mucho en común entre sí, las teorías de la autorregulación y la autoevaluación comparten también muchos puntos con la reciente investigación sobre evaluación formativa (Andrade, 2010: 90-95).

Andrade retoma un ejemplo de Wiliam que combina la autorregulación con la autoevaluación:

En sus clases de matemáticas y ciencias los alumnos utilizaban objetos de color verde, naranja y rojo que ponían sobre el escritorio para indicar su propia apreciación de qué tan bien habían comprendido el contenido que se estaba estudiando durante una clase o exposición. El maestro podía así reaccionar de inmediato a las diferencias que manifestaban los alumnos en cuanto a la confianza en su propia comprensión, por ejemplo poniendo a trabajar juntos a alumnos que hubieran puesto el símbolo verde con los de símbolo naranja, para que aclararan los puntos confusos comentando entre ellos, mientras el maestro ayudaba a los alumnos de símbolo rojo (Andrade, 2010: 99).

Topping comenta que “en la mayoría de las aulas hay más alumnos que maestros, por lo que la retroalimentación de los compañeros puede ser más inmediata, oportuna e individualizada que la del docente”. Señala que será necesario entrenar a los alumnos para desarrollar habilidades apropiadas, pero sostiene que

...una mejor comprensión de cómo se producen los efectos positivos de la coevaluación permitirá... diseñar formas más efectivas, pero por muchos años la evaluación por pares... se ha apoyado sólo en antiguos dichos como el que dice que “enseñar es aprender dos veces” (2010: 62-63).

Stiggins recuerda que “si un alumno decide que cierto aprendizaje está fuera de su alcance, o que el riesgo de fracaso público es demasiado grande o amenazador, entonces, hagamos lo que hagamos los adultos, el aprendizaje termina” (2008: 8).

El primer reto para que el docente haga evaluación formativa es *redefinir la dinámica motivacional de la evaluación*. De aquí se derivan algunas consecuencias importantes: si alumnos y maestros son quienes toman las decisiones más importantes en el aula, entonces ambos deberán manejar los tres pasos de Sadler: tener claro en qué punto está cada aprendiz, a qué distancia está de la meta establecida y cómo es posible reducir esa distancia. Puede parecer obvio, pero tomarlo en serio no es fácil: supone que cada maestro domine las competencias que se espera alcancen los estudiantes y tenga seguridad; si los maestros no cumplen con ello, las consecuencias pueden ser expectativas inadecuadas o confusas sobre el rendimiento de los estudiantes, y entonces será imposible una evaluación confiable. Stiggins añade:

Evaluar para el aprendizaje requiere que se redacten mapas curriculares basados en estándares, en versiones inteligibles para los alumnos y sus padres, de manera que a

todos quede claro el trayecto de aprendizaje a recorrer... Otra estrategia es ofrecer a los alumnos de manera regular retroalimentación descriptiva, en contraposición a la evaluativa, es decir, información que les ayude a entender cómo mejorar la calidad de su propio trabajo (Stiggins, 2007: 15-18).

Para Davies el fruto de la investigación sobre ventajas de la evaluación formativa, formas de retroalimentación, motivación y evaluación sumativa, son “cuatro piedras angulares” que sustentan la idea de involucrar al alumno en la evaluación en aula. Distingue dos tipos de retroalimentación, cada una con diferente impacto sobre el aprendizaje:

Retroalimentación específica, descriptiva. Puede ocurrir durante el aprendizaje o después de él. Es formativa. El trabajo del alumno se compara con criterios, rúbricas, modelos, ejemplos, muestras o descripciones de trabajos excelentes. Los alumnos aprenden qué puntos cumplen las expectativas de calidad y dónde deben aprender más y mejorar; comprenden más fácilmente esta retroalimentación por su relación con el aprendizaje. Una retroalimentación menos eficaz simplemente juzga el resultado (“buen trabajo” o “necesita mejorar”), mientras la retroalimentación específica y descriptiva, relacionada con criterios claros, informa a los aprendices sobre lo que han hecho bien y lo que necesitan hacer de manera distinta.

Retroalimentación evaluativa. Este segundo tipo tiene lugar al final del aprendizaje. Es sumativo. Dice al aprendiz cómo se ha desempeñado en relación con otros (referencia a norma estadística) o con lo que debía aprender (referencia a un criterio). Se comunica mediante letras, números u otros símbolos que forman un código. Al recibir retroalimentación evaluativa el alumno comprende si necesita mejorar o no; sin

embargo, a no ser que pueda decodificar la retroalimentación, podrá no tener suficiente información para saber cómo mejorar. Los investigadores reportan que las letras, grados y otros símbolos que comunican retroalimentación evaluativa pueden tener un efecto negativo sobre el aprendizaje de todos los alumnos, y que esos efectos negativos son más marcados en el caso de alumnos de bajo rendimiento (Davies, 2007: 32-33).

Davies describe cuatro estrategias para involucrar a los alumnos en evaluación: 1) definir los puntos de llegada del aprendizaje de forma que el alumno entienda bien los objetivos a alcanzar; 2) involucrar al alumno como socio en la construcción conjunta de criterios para evaluar los resultados, 3) multiplicar la cantidad de retroalimentación que recibe el alumno para prever el desarrollo subsecuente del aprendizaje, 4) involucrar al alumno en la recolección, selección, discusión y presentación de evidencias de su aprendizaje (2007: 36).

CONCLUSIÓN

Dada la multiplicidad de variables que inciden en las prácticas docentes y de evaluación, las profundas raíces de algunas de ellas, y el hecho de que modificar otras no esté al alcance de los maestros mismos, sino que involucra a otras instancias del sistema educativo y a los padres de familia, las actividades de actualización que tengan como propósito transformar en profundidad las prácticas de evaluación de los maestros no podrán consistir únicamente en talleres breves, esporádicos y superficiales. La conclusión de Schneider y Randel, es que:

...los esfuerzos de actualización sobre evaluación formativa en aula deben tener una duración sostenida con suficientes horas presenciales para presentar los conceptos y ofrecer apoyo sustancial que permita poner en práctica la aplicación de nuevas habilidades. La duración óptima de los programas

de actualización está por determinarse. Supovitz y Turner han encontrado que los maestros necesitan entre 40 y 80 horas de actualización para que sus prácticas cambien respecto a las del promedio... los maestros que trabajan en escuelas de bajos resultados pueden modificar su base de conocimientos sobre evaluación formativa en aula en 30 a 40 horas; sin embargo, el número de horas necesario para cambiar los conocimientos de los maestros no es necesariamente el mismo que bastará para que el rendimiento de los alumnos aumente (Schneider y Randel, 2010: 272-273).

La última idea es fundamental: es necesario modificar los conocimientos del maestro sobre evaluación para que sus prácticas cambien, pero esto no basta; tampoco será suficiente llevar a cabo talleres de 40-80 horas de duración. Es indispensable un esfuerzo continuado durante uno o dos ciclos escolares, basado en el trabajo conjunto de una comunidad de aprendizaje formada por los maestros de una o varias escuelas. En este sentido, Black y Wiliam dicen:

Los maestros necesitan una variedad de ejemplos vivos para la implementación de tales prácticas por parte de colegas con los que puedan identificarse y de los que puedan derivar la convicción y la confianza de que ellos también pueden hacerlo mejor, y constatar qué quiere decir en la práctica “hacerlo mejor” (2004: 21).

James recuerda que en general son fallas de implementación lo que impide el éxito de iniciativas para implantar prácticas efectivas de evaluación, y que por ello hay que estudiar las condiciones necesarias para hacerlo bien.

Para generalizar innovaciones probadas en experimentos en pequeña escala hay que prever que se contará con un apoyo mucho menos intensivo. Las condiciones adecuadas incluyen “el desarrollo profesional de

los maestros, así como estructuras organizacionales y procesos culturales adecuados” (James, 2010: 169).

Entre las advertencias de este autor en relación con la difusión de prácticas de evaluación formativa se pueden mencionar las siguientes:

Si bien los maestros aprecian las recomendaciones prácticas, las acciones de evaluación formativa se pueden volver mecánicas y rituales si no se promueve la reflexión sobre los principios que las sustentan.

Los valores, las creencias y las prácticas no son uniformes entre los maestros; aunque la mayoría tiene valores educativos claros y positivos, la mayoría encontró dificultad para alinear las nuevas prácticas y sus valores.

La influencia más importante resultó ser la indagación colaborativa entre los maestros sobre sus prácticas... que se puede extender más allá del aula por medio de redes dentro de la escuela y entre escuelas.

Lo anterior depende en buena medida de las estructuras organizacionales, la cultura y el liderazgo. El reto clave para los directivos es, pues, crear el espacio y el clima para que el personal de las escuelas pueda reflexionar sobre su práctica y compartir esa reflexión (James, 2010: 169-170).

Para que puedan extenderse las nuevas prácticas será necesario, pues, que las instancias de dirección y supervisión ofrezcan un apoyo consistente, además de que estén presentes condiciones de trabajo y recursos de la escuela y el aula mínimamente suficientes, y que se remuevan los obstáculos que presentan una normatividad inadecuada y evaluaciones externas de enfoque incompatible con el de evaluación formativa bien entendido. Así mismo, habrá que contar con el apoyo de unos padres de familia enterados de la importancia y el sentido de los cambios. Conviene subrayar tres puntos:

- Que la puesta en práctica dé buenas evaluaciones de enfoque formativo es incompatible con un currículo demasiado extenso, sobrecargado de contenidos, que obliga a los maestros a dedicar poco tiempo a cada tema y les impide actuar sobre los complejos procesos que implica modificar las concepciones de los alumnos, desarrollar la confianza en sus propias posibilidades de aprendizaje y las habilidades de autoerregulación que son fundamentales para que jueguen un papel activo y productivo como corresponsables de la evaluación.
- Que una buena evaluación formativa no se distingue de una buena enseñanza, una que sea verdaderamente congruente con los principios de la pedagogía derivada de la *revolución cognitiva*, el constructivismo.
- Y que la buena evaluación formativa y/o la buena enseñanza constructivista suponen otra condición indispensable: que los maestros mismos tengan la

preparación necesaria para identificar la etapa de desarrollo cognitivo en que se encuentran sus alumnos, y para sustituir las estructuras y los preconceptos inadecuados por otros mejores.

En muchos casos no se puede dar por hecho que estos tres puntos se cumplen, y conseguirlo constituye un reto mayúsculo. No es otra cosa lo que implica hacer buenas evaluaciones formativas.

¿Y cómo se podrá saber si los esfuerzos están dando resultados? Brookhart nos habla de lo que se puede considerar un indicio inequívoco al respecto: “Con buenas evaluaciones formativas basadas en el desempeño la pregunta clave que plantean los alumnos deja de ser si cierto contenido vendrá en la prueba, para dar lugar a otra: ¿ahora qué más puedo aprender?” (2007: 56).

No habrá que olvidar, en todo caso, que: “...la mejor evaluación anual del mundo no puede superar los problemas producidos por evaluaciones cotidianas de baja calidad” (Stiggins, 2001: 13).

REFERENCIAS

- ALLAL, Linda y Lucie Mottier Lopez (2005), “Formative Assessment of Learning: A review of publications in French”, en OECD-CERI, *Formative Assessment. Improving learning in secondary classrooms*, París, OECD-Centre for Educational Research and Innovation (CERI), pp. 241-264.
- ANDRADE, Heidi (2010), “Students as the Definitive Source of Formative Assessment: Academic self-assessment-self regulation of learning”, en Heidi Andrade y Gregory Cizek (eds.), *Handbook of Formative Assessment*, Nueva York, Routledge, pp. 90-105.
- ANIJOVICH, Rebeca (2010), “La retroalimentación en la evaluación”, en Rebeca Anijovich (ed.), *Evaluación significativa*, Buenos Aires, Paidós, pp. 16-17.
- ASCHBACHER, Pamela R. (1993), “Issues in Innovative Assessment for Classroom Practice: Barriers and facilitators”, Report 359, Los Angeles, University of California (UCLA)-National Center for Research on Evaluation, Standards, & Student Testing (CRESST).
- BLACK, Paul y Dylan Wiliam (2004), “The Formative Purpose: Assessment must first promote learning”, en Mark Wilson (ed.), *Towards Coherence between Classroom Assessment and Accountability*, 103rd Yearbook, Part 2, Chicago, National Society for the Study of Education, University of Chicago Press, pp. 20-50.
- BLACK, Paul (2010), “Formative Assessment”, en Penelope Peterson, Eva Baker y Barry McGaw (eds.), *International Encyclopedia of Education*, Amsterdam, Elsevier-Academic Press, vol. 3, pp. 359-364.
- BROOKHART, Susan (2007), “Expanding Views about Formative Classroom Assessment: A review of the literature”, en James H. McMillan (ed.), *Formative Classroom Assessment: Theory into practice*, Nueva York, Teachers College Press, pp. 43-62.
- CARLESS, David (2005), “Prospects for the Implementation of Assessment for Learning”, *Assessment in Education: Principles, Policy & Practice*, vol. 12, núm. 1, pp. 39-54.

- CIZEK, Gregory (2010), "An Introduction to Formative Assessment. History characteristics and challenges", en Heidi Andrade y Gregory Cizek (eds.), *Handbook of Formative Assessment*, Nueva York, Routledge, pp. 3-17.
- DARLING-Hammond, Linda (1999), *Teacher Quality and Student Achievement: A review of state policy evidence*, Document R-99-1, Seattle, University of Washington-Center for the Study of Teaching and Policy.
- DAVIES, Anne (2007), "Involving Students in the Classroom Assessment Process", en Douglas Reeves (ed.), *Ahead of the Curve. The power of assessment to transform teaching and learning*, Bloomington, Solution Tree Press, pp. 31-57.
- GREEN, Susan K., Julian Smith y E. Kenyon Brown (2007), "Using Quick Writes as a Classroom Assessment Tool: Prospects and problems", *Journal of Educational Research & Policy Studies*, vol. 7, núm. 2, pp. 38-52.
- GUSKEY, Thomas (2007), "Formative Classroom Assessment & Benjamin Bloom: Theory, research and practice", en James H. McMillan (ed.), *Formative Classroom Assessment: Theory into practice*, Nueva York, Teachers College Press, pp. 63-78.
- HARLEN, Wynne (2007), "Formative Classroom Assessment in Science and Mathematics", en James H. McMillan (ed.), *Formative Classroom Assessment: Theory into practice*, Nueva York, Teachers College Press, pp. 99-115.
- HILL, Heather C., Brian Rowan y Deborah Loewenberg Ball (2005), "Effects of Teachers' Mathematical Knowledge for Teaching on Student Achievement", *American Educational Research Journal*, vol. 42, núm. 2, pp. 371-406.
- JAMES, M. (2010), "An Overview of Educational Assessment", en Penelope Peterson, Eva Baker y Barry McGaw (eds.), *International Encyclopedia of Education*, Amsterdam, Elsevier-Academic Press, vol. 3, pp. 161-171.
- KATZKOWICZ, Rebeca (2010), "Diversidad y evaluación", en Rebeca Anijovich (ed.), *Evaluación significativa*, Buenos Aires, Paidós, pp. 114-122.
- MARSH, Colin J. (2007), "A Critical Analysis of the use of Formative Assessment in Schools", *Educational Research Policy & Practice*, vol. 6, pp. 25-29.
- MCMILLAN, James H. (2003), "Understanding and Improving Teachers' Classroom Assessment Decision Making: Implications for theory and practice", *Educational Measurement: Issues and Practice*, vol. 22, núm. 4, pp. 34-43.
- Organisation for Economic Cooperation and Development (OECD)-Centre for Educational Research and Innovation (CERI) (2005), *Formative Assessment. Improving learning in secondary classrooms*, París, OECD.
- OSBORN, Marlyn y Maroussia Raveaud (2010), "Classroom Assessment in Policy Context (England & France)", en Penelope Peterson, Eva Baker y Barry McGaw (eds.), *International Encyclopedia of Education*, Amsterdam, Elsevier-Academic Press, vol. 3, pp. 425-431.
- POPHAM, W. James (2008), *Transformative Assessment*, Alexandria, Association for Supervision and Curriculum Development.
- POPHAM, W. James (2009), "Assessment Literacy for Teachers: Faddish or fundamental?", *Theory into Practice*, vol. 48, pp. 4-11.
- RESNICK, Lauren B., Dylan Wiliam, Rosita Apodaca y Elizabeth Rangel (2010), "The Relationship between Assessment and the Organization and Practice of Teaching", en Penelope Peterson, Eva Baker y Barry McGaw (eds.), *International Encyclopedia of Education*, Amsterdam, Elsevier-Academic Press, vol. 3, pp. 397-402.
- RINDONE, Douglas A. y Duncan McQuarrie (2010), "Strategies and Policies for Incorporating Formative Assessment into Comprehensive and Balanced State Assessment Systems", en Heidi Andrade y Gregory Cizek (eds.), *Handbook of Formative Assessment*, Nueva York, Routledge, pp. 316-327.
- SADLER, D. Royce (1998), "Formative Assessment: Revisiting the territory", *Assessment in Education: Principles, Policy & Practice*, vol. 5, núm. 1, pp. 77-84.
- SCHNEIDER, M. Christina y Bruce Randel (2010), "Research on Characteristics of Effective Professional Development Programs for Enhancing Educators' Skills in Formative Assessment", en Heidi Andrade y Gregory Cizek (eds.), *Handbook of Formative Assessment*, Nueva York, Routledge, pp. 251-276.
- STIGGINS, Richard J. (2001), "Unfulfilled Promise of Classroom Assessment", *Educational Measurement: Issues & Practice*, vol. 20, núm. 3, pp. 5-15.
- STIGGINS, Richard J. (2007), "Conquering the Formative Assessment Frontier", en James H. McMillan (ed.), *Formative Classroom Assessment: Theory into practice*, Nueva York, Teachers College Press, pp. 8-27.
- STIGGINS, Richard J. (2008), *Assessment Manifesto: A call for the development of balanced assessment systems*, Portland, ETS Assessment Training Institute.
- STIGGINS, Richard J. y Nancy Faires Conklin (1992), *In Teachers' hands: Investigating the practice of classroom assessment*, Albany, State University of New York Press, Sunny Series in Educational Leadership.
- STOBART, Gordon (2005), "Lost in Translation: Moving from principles to policy to practice in formative assessment", *Assessment in Education: Principles, Policy & Practice*, vol. 12, núm. 1, pp. 3-5.

- SUTTON, Ruth (2010), "Challenges of Developing and Implementing Formative Assessment Practices in Schools", en Penelope Peterson, Eva Baker y Barry McGaw (eds.), *International Encyclopedia of Education*, Amsterdam, Elsevier-Academic Press, vol. 3, pp. 353-358.
- TIERNEY, Robin D. (2006), Changing Practices: Influences on classroom assessment, *Assessment in Education*, vol. 13, núm. 3, pp. 239-264.
- TOPPING, Keith J. (2010), "Peers as a Source of Formative Assessment", en Heidi Andrade y G. Cizek (eds.), *Handbook of Formative Assessment*, Nueva York, Routledge, pp. 61-74.
- TORRANCE, Harry y John Pryor (1995), "Investigating Teaching Assessment in Infant Classrooms: Methodological problems and emerging issues", *Assessment in Education*, vol. 2, núm. 3, pp. 305-320.
- WAYNE, Andrew y Peter Youngs (2003), "Teacher Characteristics and Student Achievement Gains: A Review", *Review of Educational Research*, vol. 73, núm. 1, pp. 89-122.
- WILLIAM, Dylan (2007), "Content Then Process: Teacher learning communities in the service of formative assessment", en Douglas Reeves (ed.), *Ahead of the Curve. The Power of Assessment to Transform Teaching and Learning*, Bloomington, Solution Tree Press, pp. 183-204.
- WILSON, Suzanne (2004), "Student Assessment as an Opportunity to Learn in and From One's teaching Practice", en Mark Wilson (ed.), *Towards Coherence between Classroom Assessment and Accountability*, 103rd Yearbook, Part 2, Chicago, National Society for the Study of Education, University of Chicago Press, pp. 264-271.