

# CONSIGNAS, DEVOLUCIONES Y CALIFICACIONES: LOS PROBLEMAS DE LA EVALUACIÓN EN LAS AULAS DE EDUCACIÓN PRIMARIA EN AMÉRICA LATINA

*Pedro Ravela\**

**Resumen.** El propósito del artículo es discutir tres aspectos centrales de la práctica de la evaluación en el aula: a) las consignas que los maestros proponen a los alumnos para evaluar su aprendizaje; b) el tipo de devolución que realizan a sus estudiantes; y, c) el uso de las calificaciones. La discusión se apoya, por un lado, en una revisión de literatura sobre el tema y, por otro, en un amplio banco de información recogida en el marco de una investigación realizada en aulas de 6° grado de Primaria en ocho países de América Latina. El artículo ilustra, con evidencia empírica, tres problemas principales de las prácticas de evaluación: el carácter puramente escolar y descontextualizado de las actividades de evaluación, la ausencia de devoluciones de carácter formativa y la arbitrariedad y falta de significado de las calificaciones.

**Palabras clave:** evaluación del aprendizaje, calificaciones, educación primaria, América Latina

---

\* Profesor de Filosofía y Magíster en Ciencias Sociales y Educación. Especializado en temas de evaluación educativa, fue Director de Evaluación en la Administración Nacional de Educación Pública y Coordinador Nacional del Estudio PISA en Uruguay. En 2005 y 2006 integró la Coordinación Técnica del SERCE en UNESCO/OREALC. Actualmente dirige el Instituto de Evaluación Educativa en la Universidad Católica del Uruguay. Es investigador, docente y asesor en temas de evaluación en varios países de la región. Integra el Grupo de Trabajo sobre Estándares y Evaluación de PREAL.

## INTRODUCCIÓN

La evaluación suele ser uno de los aspectos más ingratos del quehacer educativo, tanto para los estudiantes como para los docentes. Para muchos estudiantes, porque las situaciones de evaluación constituyen una instancia de exposición al fracaso. Muchos las viven con estrés, en especial las situaciones de examen. Otros lo hacen con incertidumbre y/o indiferencia, porque se han resignado a no comprender qué es lo que se espera de ellos.

Para los docentes la tarea de evaluar suele ser particularmente desagradable, por varios motivos. Preparar propuestas de evaluación lleva tiempo, al igual que corregir los trabajos de los estudiantes. Revisar trabajos similares de muchos estudiantes es tedioso —en especial en la educación media, porque cada docente atiende varios cientos de estudiantes—. Definir calificaciones de las que dependen las perspectivas personales de los estudiantes, es estresante. El docente está permanentemente enfrentado a una tensión entre su responsabilidad de garantizar públicamente que un estudiante ha logrado ciertos aprendizajes y el riesgo de obstaculizar las posibilidades de desarrollo de un individuo.

Cajiao Restrepo (2008), en el marco de un programa de talleres de discusión sobre evaluación, realizado en 2008 en 25 ciudades de Colombia, con la participación de más de 15.000 docentes, directivos, alumnos y padres de familia, recogió varios miles de frases sobre la problemática de la evaluación en las aulas. Se transcriben a continuación algunas frases que ilustran el carácter conflictivo de la evaluación para docentes, estudiantes y padres, y que tienen vigencia más allá del caso colombiano.

Los docentes señalan que «a pesar de tanta reflexión pedagógica y normatividad vigente, no se han cambiado los paradigmas»; «a los maestros nos falta formación sobre evaluación», «un buen porcentaje de maestros no ha logrado desligarse de la evaluación sumativa, haciendo en la práctica equivalencias entre las evaluaciones cualitativas y cuantitativas. Igual a los padres lo que les importa es si pasó o no el estudiante y qué puesto ocupó dentro del grupo» (29).

Los estudiantes indican que «predomina la memorización de los temas», que las «metodologías y formas de evaluación son anticuadas» y que en la evaluación se pone de manifiesto «la preferencia de los profesores hacia algunos estudiantes». Los padres de familia, por su parte, expresan que «no entendemos los informes que nos entregan en los colegios, por eso al final uno se limita a preguntar si pasó o no pasó»; que la evaluación «limita la creación y capacidad intelectual del niño»; que «no detecta el estado real en que se encuentra un alumno en un área determinada»; y que «en muchos casos genera más conflictos que solución a las dificultades de comprensión del estudiante» (29).

En el contexto de la problemática sucintamente descrita, este artículo tiene como propósito discutir tres aspectos centrales de la práctica de la evaluación en el aula: a) las tareas que los maestros proponen a los alumnos para evaluar su aprendizaje; b) el tipo de devoluciones que los maestros realizan a sus estudiantes

al revisar sus trabajos; y, c) el uso de las calificaciones<sup>1</sup>. El artículo se apoya en una revisión de literatura sobre el tema y en la evidencia empírica recogida en el marco de un trabajo de investigación en aulas de 6° grado de Primaria en ocho países latinoamericanos<sup>2</sup>. El artículo no pretende ser un informe de investigación, sino ilustrar con evidencia empírica la discusión acerca de las propuestas y discursos sobre la evaluación predominantes en las escuelas.

En el primer apartado se plantea la discusión conceptual más general en torno a las denominadas evaluación «formativa» y «sumativa». En el segundo se describe la investigación realizada. Siguen tres apartados temáticos, con la discusión de los tres grandes temas propuestos: tareas de evaluación, devoluciones a los alumnos y sistemas de calificación. Cada uno de estos apartados tiene dos partes: discusión teórica y presentación de evidencia empírica. El artículo se cierra con un apartado de conclusiones y reflexiones finales.

## EVALUACIÓN FORMATIVA Y EVALUACIÓN PARA LA CERTIFICACIÓN

Para el discurso pedagógico<sup>3</sup>, la evaluación tiene principal (sino exclusivamente) una función formativa y es parte de la enseñanza. Desde esta perspectiva, la evaluación es una herramienta que forma parte del conjunto de estrategias didácticas. Su finalidad central es ayudar al estudiante a identificar lo que ha logrado y lo que no, así como permitir al docente reorientar la enseñanza y detectar estudiantes que requieren de explicaciones u otro tipo de apoyos adicionales. La cita que sigue, tomada de una revista de divulgación para docentes, es particularmente interesante como ejemplo de este discurso<sup>4</sup>.

En la evaluación confluyen contradicciones irreconciliables que conviven en el proceso de escolarización. Éstas pueden sintetizarse en las dos formas de entender la evaluación: sumativa y formativa. Por formativa debe

---

1 En nuestra visión, la problemática de la evaluación en el aula incluye un cuarto problema central, el relativo a la explicitación de los objetivos formativos, que no profundizamos en este artículo.

2 El proyecto de investigación «La evaluación de aprendizajes en las aulas de primaria en América Latina. Enfoques y prácticas», fue realizado entre enero de 2008 y abril de 2009 desde el Instituto de Evaluación Educativa de la Universidad Católica del Uruguay. El equipo de investigación estuvo integrado por Graciela Loureiro, Beatriz Picaroni, Ana Atorresi, Liliana Pazos y Pedro Ravela. La investigación contó con el apoyo financiero del Programa de Promoción de la Reforma Educativa en América Latina y su Grupo de Trabajo sobre Estándares y Evaluación (GTEEPREAL). Los productos finales del estudio se encuentran en fase de revisión y edición final en el momento de escribir este artículo. En el momento de la publicación de este artículo estarán disponibles en el sitio web del Instituto de Evaluación Educativa: <http://evaluacioneducativa.ucu.edu.uy>.

3 Con el término «discurso pedagógico» nos referimos a las visiones sobre la evaluación que predominan en los artículos dirigidos a docentes, en los programas de formación docente y entre los mismos docentes cuando se refieren a la evaluación.

4 Si bien la cita de Álvarez Méndez es algo antigua, se pueden ver otros ejemplos más recientes que muestran la vigencia del discurso centrado en el papel formativo de la evaluación en Menéndez (1999), Mancovsky (2007), López y otros (2007).

entenderse aquella que ayuda a crecer y a desarrollarse intelectual, afectiva, moral y socialmente al sujeto. En la práctica se traduce, en el caso de la corrección de un ejercicio o de un examen, en una actividad interactiva que aporta al alumno las informaciones suplementarias donde tiene necesidad de corregir su representación del problema o de aclarar cuestiones confusas. Parte de un análisis, sea de su producción o resultado, sea del proceso seguido, para introducir las modificaciones que debe aportar el alumno a su aprendizaje o los cambios que el profesor debe introducir en sus condiciones de aprendizaje [ . . . ].

La evaluación sumativa está al servicio de intereses que no son propios de la actividad educativa que se lleva a cabo en el aula. Si acaso, arrastra consecuencias colaterales no deseadas —aunque a veces sirven de ocultamiento de responsabilidades— en la misma. De hecho, termina siendo medio de control del alumno en primera instancia; pero es a la vez medio por el cual es controlado el profesor [ . . . ].

La paradoja que podemos encontrar es que el decir sobre la evaluación formativa es una constante omnipresente en el discurso de la evaluación, cuando en la práctica ocupa más bien un lugar marginal. (Álvarez Méndez, 1993: 29)

Las afirmaciones de Álvarez Méndez sobre el carácter formativo de la evaluación no merecen objeción alguna. Simultáneamente, su visión sobre la evaluación «sumativa» es por lo menos discutible. En la cita no es definida, sino que se la descalifica directamente. En otros lugares del artículo se la asocia con notas/calificaciones y con procesos de selección de alumnos. Este texto refleja muy bien la (falsa) oposición, predominante en el discurso pedagógico, entre evaluar procesos (que sería formativo, cualitativo y «bueno») y evaluar resultados (que sería sumativo, cuantitativo y «malo»). Esta trilogía de términos opuestos, procesos-resultados, formativo-sumativo y cualitativo-cuantitativo, resume muy bien el modo en que suele demarcarse conceptual y valorativamente el campo de la evaluación en educación<sup>5</sup>.

Nos interesa especialmente poner en tela de juicio la afirmación «la evaluación sumativa está al servicio de intereses que no son propios de la actividad del aula». Esta visión ignora el carácter social de la educación y la concibe como una actividad privada al interior del aula, en la que participan solamente el educador y sus estudiantes. Pero la educación es una labor social en sentido amplio. El docente desempeña una función que le ha sido encomendada por la sociedad y por las familias de sus alumnos. En los niveles de educación obligatoria, alumnos y docentes están allí como resultado de un mandato social. Este mandato incluye, para el docente, establecer el grado en que cada estudiante está logrando

---

5 Los términos «evaluación formativa» y «evaluación sumativa» fueron acuñados hace ya mucho tiempo en la literatura anglosajona. Airasian y Madaus (1972) proponen la distinción entre «evaluación formativa» (dirigida a monitorear el aprendizaje durante el proceso); «evaluación diagnóstica» (dirigida a identificar dificultades de aprendizaje) y «evaluación sumativa» (dirigida a evaluar los logros al final de la enseñanza).

los aprendizajes definidos como necesarios a través del currículo. A través de la normativa, todos los sistemas educativos atribuyen al docente la función de «certificar»<sup>6</sup> el grado en que cada estudiante ha adquirido los conocimientos y capacidades propuestos.

El término *certificar*, según el diccionario de la Real Academia Española, significa «asegurar, afirmar, dar por cierto algo» y «hacer constar por escrito una realidad de hecho por quien tenga fe pública o atribución para ello». Los docentes tienen la atribución, la fe pública y el mandato normativo de «asegurar» y «hacer constar por escrito» (normalmente a través de las calificaciones) que han logrado aprender sus estudiantes.

Los propios docentes son usuarios principales de esta certificación: parten de la base de que los estudiantes que reciben han incorporado los saberes definidos para el curso previo al propio. Lo mismo ocurre con el resto de las instituciones educativas. Las universidades necesitan una certificación de los estudios realizados y los conocimientos alcanzados por un estudiante que se presenta para cursar una carrera. Las familias de los estudiantes son otro destinatario relevante de la certificación.

Lo expresado no significa defender las calificaciones tal como se utilizan en la mayoría de los sistemas educativos. Un apartado de este artículo está destinado, justamente, a discutir los sistemas de calificaciones y sus usos. Pero el problema no es la certificación, sino el mal uso de malos sistemas de calificaciones para cumplir con una función que, en sí misma, es relevante y necesaria.

La certificación es, además, una herramienta de responsabilización, en la medida en que el aprendizaje no es únicamente una actividad placentera y espontánea, sino que requiere también de esfuerzo y disciplina. Requiere tanto de motivaciones intrínsecas como de exigencias externas. La certificación es una exigencia externa que hace al estudiante responsable de esforzarse por alcanzar determinados logros. El aprendizaje depende tanto de lo bueno que sea el profesor enseñando como del esfuerzo y dedicación del estudiante. Probablemente, esto que llamo «responsabilización», con un sentido positivo, es equivalente a lo que Álvarez Méndez denomina «control», con un sentido negativo.

Como resumen de esta primer entrada conceptual en el tema, es relevante distinguir dos funciones de la evaluación educativa.

- a. La formativa (evaluación para el aprendizaje), cuyo foco principal está puesto en la reflexión sobre lo que se va aprendiendo, en la confrontación entre lo que se intenta y lo que se alcanza, en la búsqueda de nuevos caminos para avanzar hacia los conocimientos y desempeños que se busca lograr.
- b. La certificativa (evaluación del aprendizaje), cuyo foco principal es constatar el aprendizaje y certificarlo públicamente, es decir, dar fe pública de cuáles son los conocimientos y desempeños logrados por cada estudiante.

---

6 Considero más apropiado el uso del término «evaluación para la certificación» en lugar de «evaluación sumativa», porque enfatiza la función social y porque el término «sumativo» induce a una vinculación falaz con lo cuantitativo.

La evaluación certificativa requiere, principalmente, de buenos dispositivos de valoración del trabajo de alumno, que permitan establecer un juicio de valor válido y confiable acerca de sus logros. La evaluación formativa requiere, ante todo, de buenos dispositivos de devolución al estudiante, que le permitan reflexionar sobre lo que está haciendo y buscar los modos para mejorarlo.

Ahora bien, como afirma Álvarez Méndez al final de la cita, una cosa es el discurso sobre la evaluación y otra cosa es la práctica de la misma. Si bien la dicotomía evaluación formativa—evaluación sumativa está fuertemente presente en el discurso pedagógico, no ocurre lo mismo en la práctica. Pero el problema no es que no haya evaluación formativa en las aulas, sino que ambos tipos de evaluación se entremezclan continuamente. Según mostraremos en este artículo, en la práctica se produce una curiosa mezcla de situaciones. Por un lado, existen maestros que califican continuamente todas las tareas y que utilizan la calificación incluso en el marco de evaluaciones formativas. Al mismo tiempo, otros maestros nunca realizan pruebas ni asignan calificaciones, salvo al final del curso, momento en el que califican a partir de su «conocimiento directo» y acumulado del «proceso» de cada alumno.

## DISEÑO DE LA INVESTIGACIÓN

Este artículo se apoya en un estudio empírico de carácter exploratorio, cualitativo y descriptivo, dirigido a indagar las concepciones y prácticas de evaluación en las aulas de Primaria. Inicialmente incluyó escuelas de cinco países latinoamericanos, seleccionados en forma intencional, buscando representar la variedad de resultados en el Estudio SERCE<sup>7</sup> y la diversidad geográfica y cultural. Durante la realización del mismo se incorporaron otros tres países, por acuerdos con instituciones locales<sup>8</sup>. El trabajo de campo en cada país involucró a diez escuelas y a dos maestros de 6° grado por escuela, 160 maestros en total.

RESULTADOS EN SERCE	REGIÓN		
	CONO SUR	ANDINA	NORTE Y CENTROAMÉRICA
Altos	Uruguay		México - Costa Rica
Medios	Argentina	Colombia	El Salvador
Bajos		Perú	Guatemala

Países incluidos en la investigación según región y resultados en SERCE

7 Segundo Estudio Regional Comparativo y Explicativo (SERCE), llevado adelante por la Oficina Regional de Educación (OREALC) de la UNESCO. Evaluó aprendizajes en lectura, escritura, matemática y ciencias en 3° y 6° grados de Primaria, en 16 países de la región.

8 En Colombia se estableció un acuerdo con el Grupo de Investigación sobre Evaluación de la Calidad de la Educación, de la Universidad San Buenaventura de Cali; en México, con el Instituto Nacional para la Evaluación de la Educación (INEE); en Argentina, con el Departamento de Educación de la Universidad Católica Argentina.

Las escuelas fueron seleccionadas de manera intencional, por las unidades de evaluación de cada país, de acuerdo con dos criterios principales:

- a. debían ser escuelas ubicadas en contextos urbanos desfavorecidos, preferentemente públicas;
- b. debían haber tenido resultados educativos en el promedio del país o por encima del mismo, en evaluaciones externas estandarizadas.

El primer criterio obedece a la intención de indagar cómo se evalúa en las escuelas que trabajan con estudiantes de origen social desfavorecido, dado que es en dichos sectores sociales donde se plantean los desafíos educativos más importantes. Existe abundante investigación que avala esta afirmación, incluyendo los resultados del SERCE. Se buscó, a través de este criterio de selección, indagar acerca de cómo son evaluados (y, por tanto, que se prioriza en términos de aprendizaje) los estudiantes provenientes de sectores desfavorecidos.

El segundo criterio obedece a la intención de indagar acerca de las prácticas vigentes en escuelas en las que se producen aprendizajes. La literatura sobre escuelas de contextos desfavorecidos muestra que muchas veces es muy difícil enseñar en estos contextos, que a veces las limitaciones sociales superan las posibilidades de enseñar de los maestros, que muchas veces existe alta rotación del personal y los maestros más experimentados emigran hacia escuelas con mejores condiciones, entre otros fenómenos que afectan negativamente la enseñanza y el aprendizaje. A través de este criterio se buscó trabajar con escuelas en las que hubiese un nivel de logro educativo razonablemente bueno (medido a través de pruebas externas), bajo el supuesto de que en este tipo de escuelas encontraríamos enfoques y prácticas de evaluación interesantes desde el punto de vista didáctico.

Para recoger información se emplearon tres técnicas:

- a. una entrevista en profundidad a cada maestro.
- b. un registro fotográfico de trabajos de evaluación.
- c. un cuestionario autoadministrado al maestro.

Las entrevistas tuvieron alrededor de una hora de duración y se realizaron siguiendo una pauta flexible, orientada a cubrir los siguientes temas: ¿qué sentido tiene para el maestro la evaluación? ¿Cómo la realiza y registra? ¿Qué tipo de consignas e instrumentos emplea? ¿Qué tipo de devolución hace a los estudiantes? ¿Qué aspectos prioriza a la hora de evaluar la lectura y la escritura? ¿Qué aspectos prioriza a la hora de evaluar matemática? ¿Qué criterios tiene en cuenta a la hora de calificar a sus estudiantes? El registro fue realizado mediante grabación digital.

Las entrevistas proporcionaron información sobre las creencias, convicciones y percepciones de los docentes sobre la evaluación. Con el fin de complementar y contrastar el carácter subjetivo y normativo de las percepciones, se realizó un relevamiento fotográfico de trabajos de los estudiantes. Para ello se le solicitó a cada docente, con antelación a la visita, que seleccionase tareas o consignas de evaluación que hubiese propuesto a los estudiantes y que, además, para cada una de ellas

eligiese tres ejemplos de trabajos realizados por buenos alumnos, tres de alumnos medios y tres pertenecientes a alumnos de bajo rendimiento. El procedimiento seguido permite afirmar que se trata de propuestas que tienen un sesgo hacia lo mejor de cada maestro, dado que tuvieron la oportunidad de evitar entregar propuestas que considerasen defectuosas.

El cuestionario autoadministrado estuvo destinado a recoger información sobre el grado de conocimiento y uso de los resultados de evaluaciones estandarizadas.

Como resultado de este relevamiento se conformó un banco de información integrado por 158 archivos de audio digital, 4360 registros fotográficos de propuestas de evaluación y trabajos de estudiantes y 158 cuestionarios autoadministrados.

## LAS TAREAS DE EVALUACIÓN Y SUS CONSIGNAS

### DISCUSIÓN TEÓRICA

Las tareas que los maestros proponen a sus alumnos para evaluar el aprendizaje constituyen uno de los mejores indicadores del currículo implementado: muestran qué es lo que los maestros valoran e indican a los estudiantes qué conocimientos y desempeños es importante alcanzar. Si las tareas requieren principalmente la memorización de conceptos y fechas, hacia allí estará orientado el esfuerzo de los estudiantes. Si en matemática se pide únicamente la realización de operaciones aritméticas, eso es lo que los estudiantes aprenderán.

La evaluación no puede impulsar el aprendizaje si se basa en tareas o preguntas que distraen la atención de los verdaderos objetivos de la enseñanza [. . .] si queremos que los estudiantes sean capaces de razonar y usar el conocimiento científico, entonces debemos darles la oportunidad de explicar cómo funcionan las cosas, realizando investigaciones y elaborando explicaciones con sus propias palabras [. . .]. (Shepard, 2008: 21)

Es importante señalar que las reflexiones contenidas en este apartado, focalizado en el análisis de las tareas de evaluación, aplican tanto a las evaluaciones de carácter formativo como a las que están dirigidas a la certificación.

Shepard menciona una investigación realizada en los años noventa en Estados Unidos que reportó que el 95% de las preguntas de prueba incluidas en libros de texto de matemática evaluaban información elemental, cálculos y uso de algoritmos y fórmulas en problemas rutinarios. Apenas el 5% de las tareas evaluaba habilidades de pensamiento y conocimiento conceptual de alto nivel (34).

El carácter puramente escolar y poco complejo de las tareas a través de las cuales se enseña y evalúa ha dado lugar a enfoques teóricos que enfatizan la necesidad de que las tareas escolares intenten reflejar contextos reales, propios de la vida personal y social.

Debe proponerse actividades que impliquen desafío para los alumnos, contextualizadas, con sentido, de manera que se resuelvan por su valor intrínseco. Es imprescindible atender al factor motivacional que potencia el desempeño en la resolución de la situación planteada. Las situaciones que no impliquen atractivo, que no comprometan social ni cognitivamente, pretendiendo un valor extrínseco inherente a la evaluación, no pondrán en juego la competencia real del alumno y, por lo tanto, los resultados no reflejarán los logros posibles [. . .] Las situaciones de evaluación serán situaciones de producción, de resignificación de conocimientos, no sólo de reproducción [. . .] Tratan de resolver problemas reales y se constituyen, en la medida de lo posible, en procesos de resolución de problemas genuinos y consistentes en el campo de conocimientos de que se trata (Menéndez, 1999: 20).

En la didáctica de la lengua, desde el enfoque comunicacional, se enfatiza la promoción de la lectura de textos reales (etiquetas, carteles, libros) y la escritura de textos en situaciones comunicacionales reales y con destinatarios reales (en lugar de las tradicionales «composiciones» o «redacciones» cuyo destinatario principal es la maestra).

El estudio internacional PISA, uno de los más reconocidos a nivel internacional, utiliza exclusivamente actividades de evaluación que se derivan de una situación o contexto propio de la vida real: un experimento científico, problemas derivados del calentamiento global o la clonación, el uso de conocimientos matemáticos para analizar un artículo periodístico, entre otros. En cada una de las áreas que evalúa (lectura, matemática y ciencias) los contextos en que se utilizan los conocimientos son uno de los componentes centrales del marco conceptual de las pruebas, que están organizadas en conjuntos de entre cuatro y seis preguntas relacionadas con una misma situación.

Uno de los términos acuñados para designar este tipo de tareas o consignas es «auténtico». De acuerdo con Wiggins (1998: 22-30), una actividad auténtica se caracteriza por reproducir los modos en que las personas usan el conocimiento en situaciones reales. Las tareas auténticas tienen las siguientes propiedades:

- a. Propósitos: la tarea tiene una finalidad definida, un producto o una meta a alcanzar.
- b. Destinatarios o audiencias: las actividades están dirigidas a interlocutores reales, que pueden percibir sus resultados y opinar sobre ella.
- c. Incertidumbre: requieren enfrentar y resolver problemas poco estructurados y que pueden tener más de una solución (por oposición al típico «problema» escolar, que incluye todos los datos necesarios y sólo los necesarios, y que tiene una única solución posible).
- d. Restricciones: en la vida real existen siempre limitaciones, lo que hace necesario idear alternativas y tomar decisiones acerca del camino más apropiado o menos malo, en las condiciones dadas.

- e. Repertorio de recursos cognitivos: las situaciones, productos y problemas de la vida real en general no se resuelven a través de *un* conocimiento o procedimiento específico (como la mayoría de las situaciones escolares) sino que requieren de la activación simultánea de varios de ellos, probablemente adquiridos en distintos momentos y disciplinas.

Ejemplos de tareas auténticas son realizar una investigación científica o histórica, escribir un artículo periodístico, diseñar un producto para un destinatario y una finalidad reales, desarrollar una propuesta organizativa para un emprendimiento real, establecer una estrategia de comunicación para promover una agenda social, producir y realizar una obra de teatro, entre otras. Solamente a través de este tipo de desempeños reales, dirigidos a audiencias reales, los conocimientos y habilidades que enseñamos en las escuelas pueden tener sentido para los estudiantes (Wiggins, 1998: 34-42).

Es importante señalar que, en rigor, estamos hablando de situaciones *propias* del mundo real. Por definición, la escuela no es el mundo real (aunque es parte del mismo). Por tanto, la mayor parte del tiempo las situaciones y actividades se dan en un contexto escolar. Lo importante es que, en dicho contexto, se trabaje en el *tipo de situaciones* a las que normalmente se enfrenta un ciudadano, un científico, un trabajador o un artista.

Las actividades auténticas tienen dos virtudes principales: exigen un nivel de pensamiento complejo y resultan motivadoras y desafiantes para los estudiantes. «Las tareas auténticas que requieren un pensamiento de más alto nivel y una activa solución de problemas, también incrementan la motivación del estudiante porque son intrínsecamente más interesantes que la memorización o la aplicación de procedimientos sencillos.» (Shepard, 2008: 34)

Lo antedicho no implica que no sea apropiado, en determinados momentos, trabajar en torno a conocimientos y procedimientos «puros», sin contexto. La escuela, en sí misma, es un tipo de contexto real, con sus propias demandas y exigencias. Del mismo modo, las disciplinas son contextos, comunidades de personas nucleadas en torno a un cuerpo de conocimientos, con sus propias reglas y demandas. Es legítimo, entonces, que en la escuela se trabaje con situaciones puramente «intra-disciplinares», así como en situaciones propias de la lógica escolar (como lo es una prueba). El problema se plantea cuando el conocimiento se escolariza completamente y pierde todo vínculo con las formas de producción y de uso en situaciones propias del mundo real. Cuando esto ocurre, el conocimiento pierde sentido, los estudiantes no entienden «para qué sirve» lo que se les enseña y, como consecuencia, pierden su motivación (o la misma se reduce a la obtención de buenas calificaciones). Rápidamente se acostumbran a un conjunto de «reglas» de los problemas escolares: siempre hay que hacer cuentas, siempre hay que usar todos los datos, siempre hay una única solución. Finalmente, aprenden la regla de juego principal: lo importante es dejar satisfecho al docente.

## EVIDENCIA EMPÍRICA

La información recogida en 160 aulas de 6° grado de Primaria de ocho países latinoamericanos, muestra que la presencia de actividades de evaluación que puedan ser consideradas auténticas es residual. En matemática predominan las actividades de evocación de conceptos y propiedades o de realización de operaciones sin contexto alguno (figuras 1 a 5). En algunos casos, se trata de complicadas operaciones con raíces y potencias (figuras 4 y 5), ante las cuales uno no puede evitar preguntarse si tiene sentido resolverlas manualmente en la era de la informática. Simultáneamente, nunca se proponen situaciones reales en que dichas operaciones sean necesarias.

Un segundo tipo de actividades matemáticas aparentan tener un «contexto», porque no plantean directamente la realización de operaciones, sino que tienen «letra» de problema (figuras 6 y 7). Pero se trata de actividades simples que se

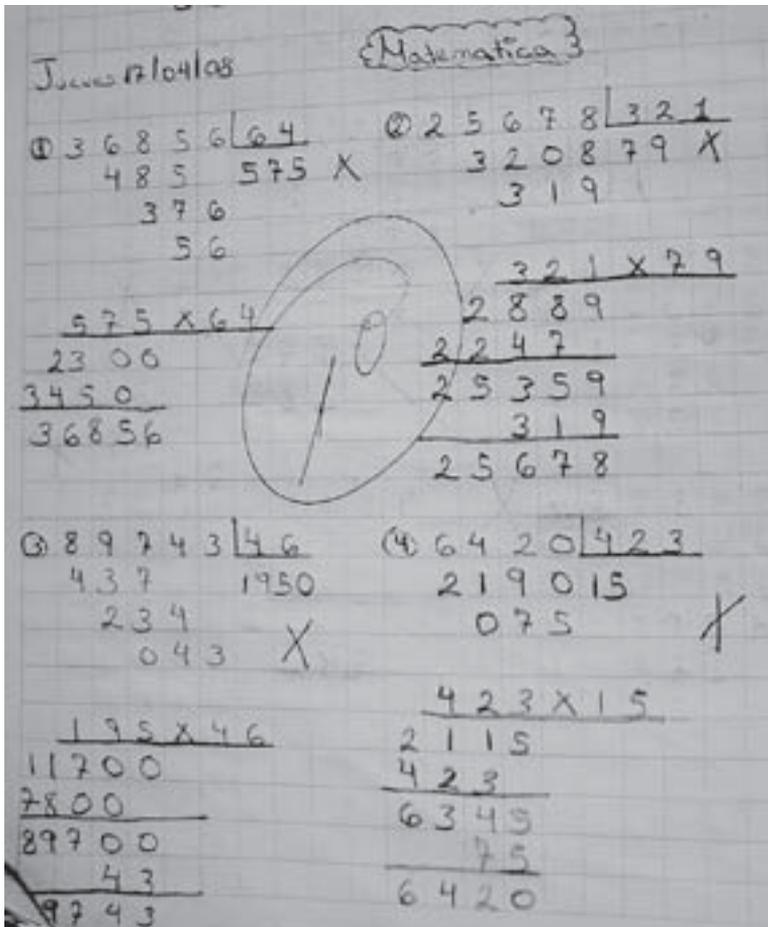


Figura 1

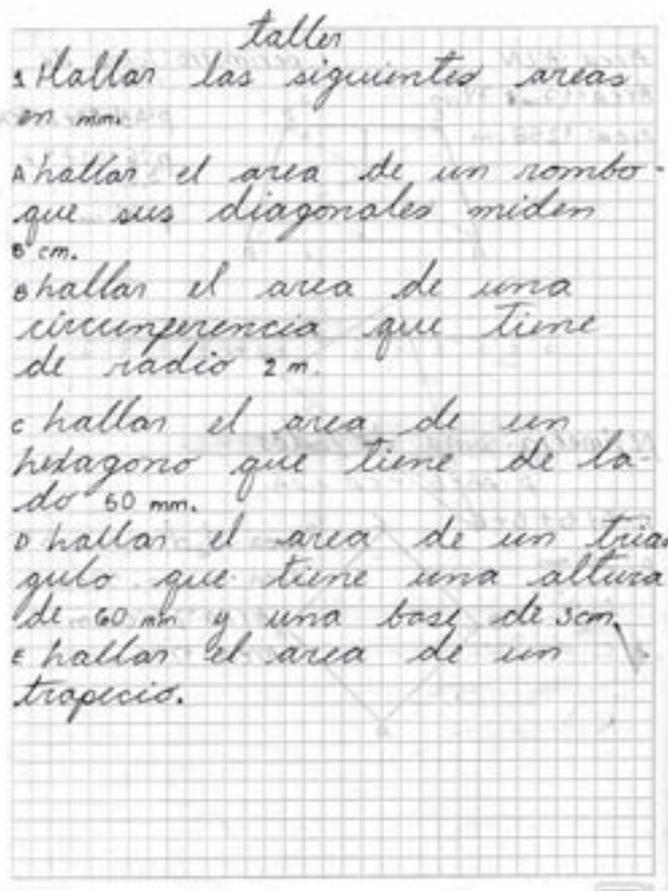


Figura 2

resuelven con una sola operación y con los datos dados, y que tienen poca relación con el uso del conocimiento matemático en la vida real (no es común que alguien ponga atención en que comió cinco novenos de un chocolate). La figura 6 es particularmente interesante. En la imagen se muestran dos consignas, pero la propuesta incluía varias más, todas iguales, modificando la cantidad de metros y la cantidad de moños. Estas actividades muestran que «el hábito no hace al monje»: el hecho de que en la formulación de una tarea se incluyan referencias a elementos de la realidad, no hace de la situación algo auténtico.

En tercer lugar, fue posible identificar intentos fallidos de formular situaciones plausibles desde el punto de vista real. La figura 8, perteneciente a un libro de texto, es un ejemplo interesante. La consigna intenta que el estudiante reflexione sobre diversos aspectos: qué es lo que se le pide, qué datos se le ofrecen, qué operaciones necesita realizar (una sola operación), para que, finalmente, dé una respuesta. La consigna contiene errores. Pide una transformación de decímetros

VALOR 10 Pts.

INSTRUCCIONES: Realiza lo que se te indica a continuación.

1. Encuentra el perímetro de las siguientes figuras.

A)  $123\text{cm}$   $44\text{cm}$   
 $P = 2 \times (123 + 44)$   
 $P = 2 \times 167$   
 $P = 334\text{cm}$

B)  $24\text{cm}$   
 $24\text{cm}$   $24\text{cm}$   $24\text{cm}$   
 $P = 4 \times 24$   
 $P = 96\text{cm}$

C)  $25\text{cm}$   $20\text{cm}$   
 $25\text{cm}$   $20\text{cm}$   
 $P = 2 \times (25 + 20)$   
 $P = 2 \times 45$   
 $P = 90\text{cm}$

D)  $8\text{cm}$   $8\text{cm}$   $8\text{cm}$   $8\text{cm}$   $8\text{cm}$   $8\text{cm}$   
 $P = 6 \times 8$   
 $P = 48\text{cm}$

E)  $20\text{cm}$   $30\text{cm}$   $8\text{cm}$   
 $20\text{cm}$   $30\text{cm}$   
 $P = 20 + 30 + 8 + 8$   
 $P = 66\text{cm}$

2. Encuentra el área de las siguientes figuras geométricas.

A)  $12\text{m}$   $148\text{m}$   
 $A = 12 \times 148$   
 $A = 1776\text{m}^2$

B)  $17\text{m}$   $17\text{m}$   
 $A = 17 \times 17$   
 $A = 289\text{m}^2$

C)  $12\text{cm}$   $8\text{cm}$   
 $A = \frac{12 \times 8}{2}$   
 $A = \frac{96}{2}$   
 $A = 48\text{cm}^2$

D)  $9\text{cm}$   $9\text{cm}$   
 $A = 9 \times 9$   
 $A = 81\text{cm}^2$

3. Realiza la unión de los siguientes conjuntos.

$A = \{ \text{cañes, perralón, corbata, chicho} \}$

$B = \{ \text{blusa, taca, corbata, somer} \}$

Figura 3

**DESAFÍO TU HABILIDAD**

Efectúa:

- $100 : 2 + 2^3 + 4 - \sqrt{16}$
- $\sqrt[3]{64} \times 5 - 18 : 2 + 6^2 : 4$
- $\sqrt{81} : 3 + 10 \times 4^3 - 5^2 \times 20$
- $7^3 - 5 \times 12 - 8^2 \times 3$
- $8^2 + 6^2 : 9 - 7^3 - 6 \times 2$
- $5^2 \times \sqrt[3]{27} - \sqrt{64} : 4 + 16 - 3 \times 2$
- $[2^3 \times 3^2 - (6 + 3 - 2)^2] - [5^2 - 20 : 4]$
- $\{[(16 + 30 - 5) - 2^3 \times 2^3] - [8 - (\sqrt{16} + 2^3)]\}$
- $6^2 : 8 + 2 - [(25 : 5 \times 2^2) + 16]$
- $(4^3 \times 10) - (10 \times 3^2 + 8^2 \times 9^0)$
- $(4^2 \times 2^3 + 3^2 \times 5) - 5^3$
- $\sqrt{5 \times 4^2 + 36 : 6^2} + [10^3 : 50 - (3^2 \times 3^2 - 2^3 \times 9)]$

Figura 4

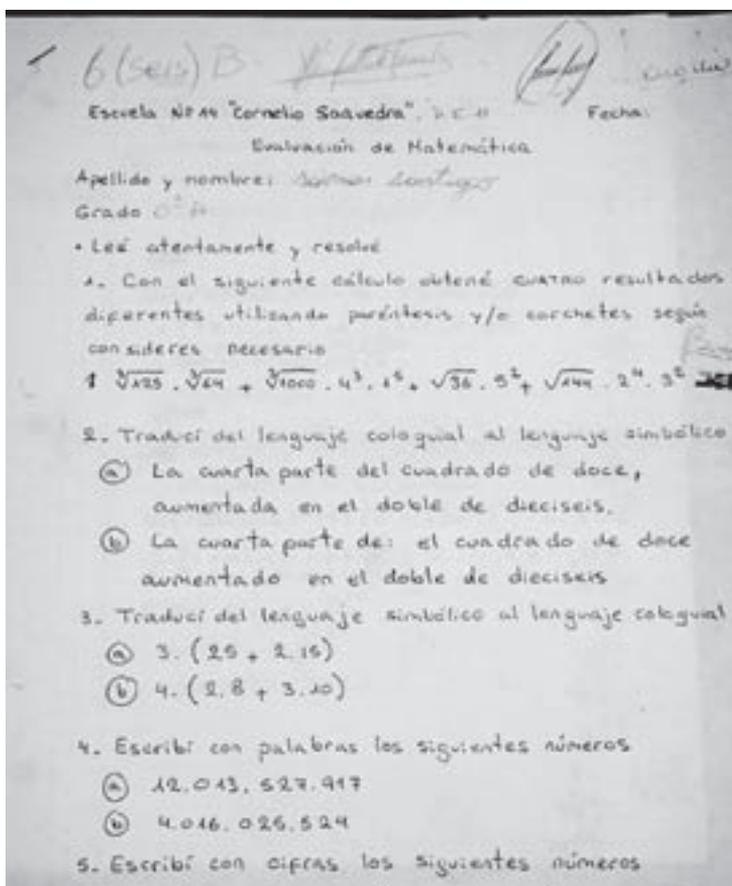


Figura 5

cúbicos a milímetros, a secas. El estudiante transforma la expresión milímetros en mililitros, tal vez a instancias de la maestra, que puede haber detectado el error (note el lector que un milímetro cúbico no es igual a un mililitro, ya que la milésima parte de un litro es un centímetro cúbico). El alumno no realiza ninguna operación, pero su trabajo es calificado con un «muy bien» por su maestra.

La primera duda que surge, desde el punto de vista de la autenticidad de la consigna, es si el nance se envasa en botellas de 8 litros (el nance es una variedad de café de grano amarillo, que se cultiva en zonas cálidas de Costa Rica). No parece razonable, pero podría ocurrir. La segunda duda es por qué utilizar decímetros cúbicos y mililitros como unidades de medida. ¿Alguien las utiliza hoy en la industria, el comercio o la ciencia? En mi vida sólo he visto el uso de hectolitros, decilitros y decímetros en las escuelas. Si el objetivo es trabajar la habilidad para transformar unidades de medición, ¿por qué no trabajar con bytes, kilobytes, megabytes y gigabytes? Estas unidades de medida se utilizan todo el

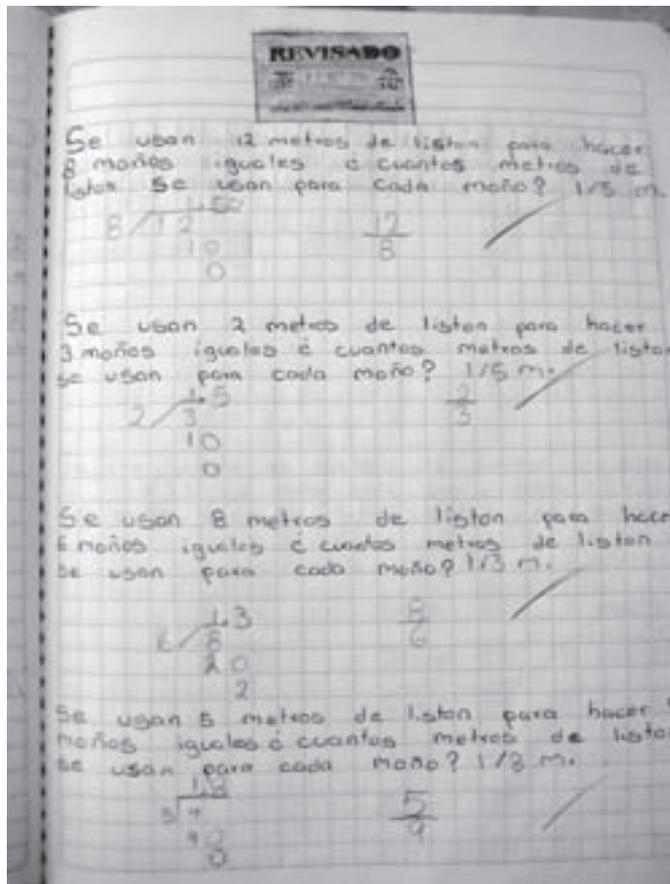


Figura 6

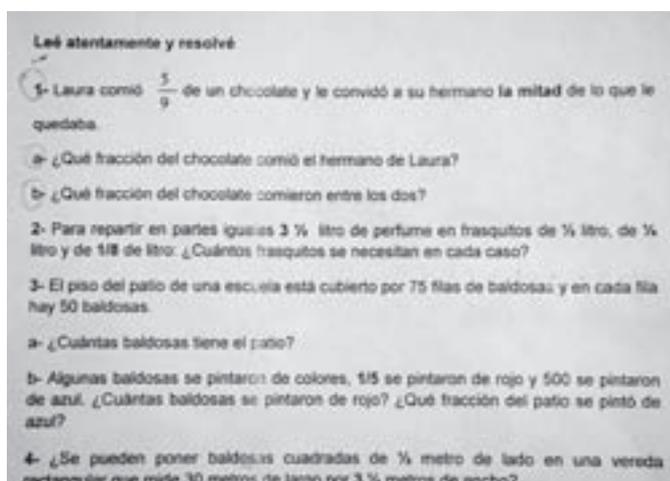


Figura 7

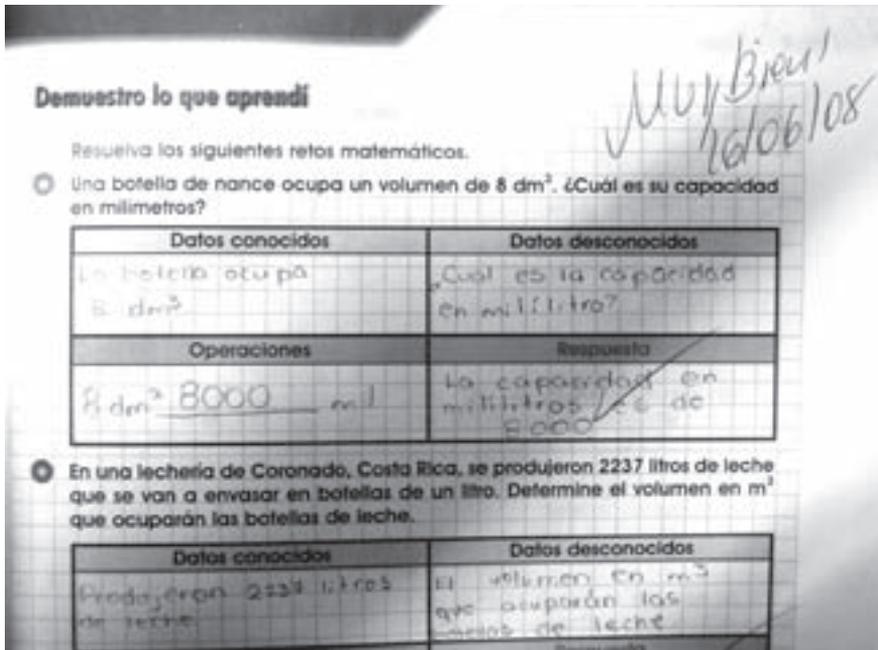


Figura 8

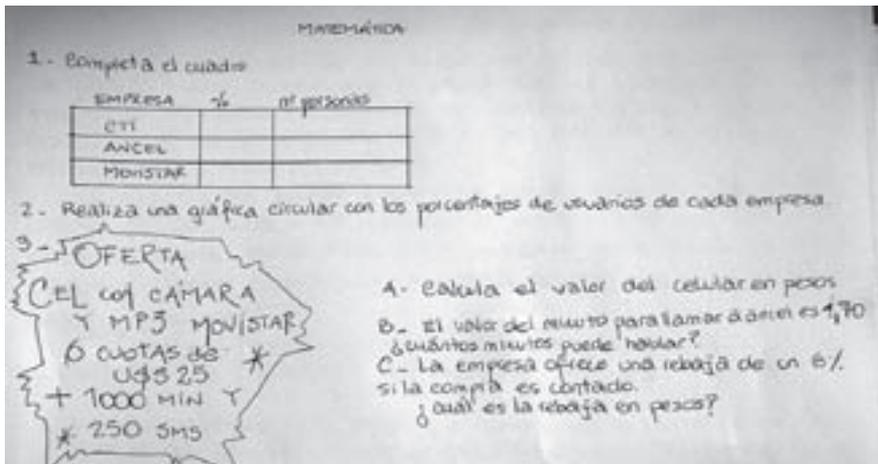


Figura 9

tiempo y uno podría imaginar cientos de problemas concretos y reales vinculados con tamaños de archivos, espacios de almacenamiento en discos y velocidad de transferencia de datos (por ejemplo, la cantidad de canciones que se puede archivar en un disco determinado o el tiempo que requiere «bajarla» de Internet).

Finalmente, tanto esta consigna como la que le sigue, referida a 2237 litros de leche, confunden la capacidad de la botella con el espacio que ocupan. Cuando

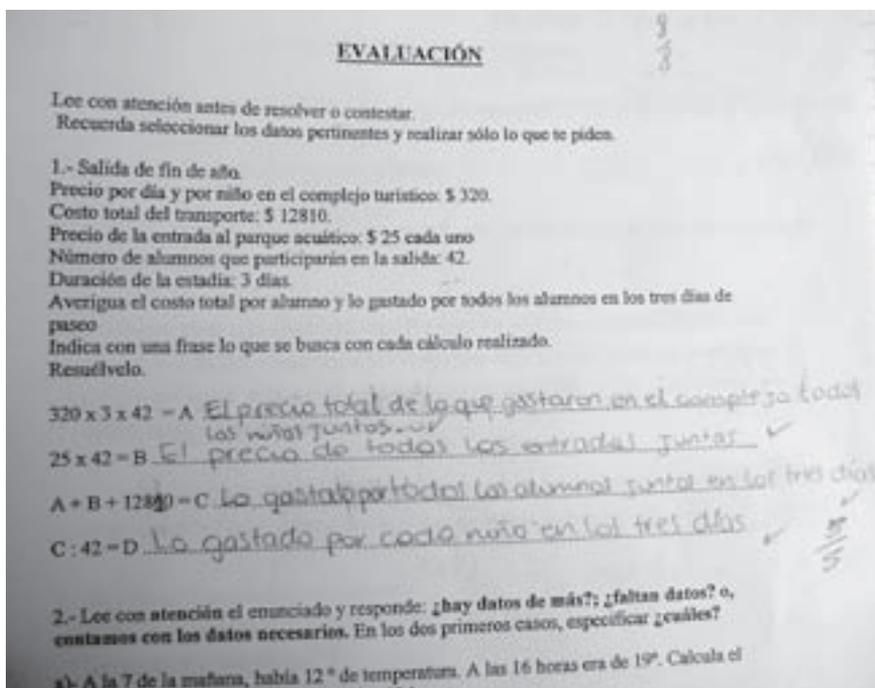


Figura 10

en la segunda consigna se pregunta qué volumen en metros cúbicos *ocuparán* las botellas, se presupone que las botellas son cubos que no ocupan más espacio que el que contienen en su interior. Pero en una situación real, si alguien necesita estimar cuánto espacio será necesario para almacenar las 2237 botellas, se encontrará con que necesita más espacio, simplemente debido a la forma de las botellas.

A lo largo de los registros es posible encontrar muchas intentos similares, de construir situaciones «reales», que no resisten un análisis de sentido común.

Las figuras 9 y 10 muestran lo más aproximado a situaciones auténticas que se encontró en los más de 4000 registros. En un caso, las maestras toman información del mercado de telefonía celular. En el segundo, la situación elegida es un paseo de fin de año. En ambos casos se trata de situaciones plausibles en la vida real, a partir de las cuales se solicita a los estudiantes varias tareas, que requieren del uso de distintos conceptos y habilidades matemáticas, así como la reflexión sobre los procedimientos seguidos.

En lenguaje predomina la evaluación de conocimientos gramaticales, especialmente en Costa Rica (figuras 11 y 12) y el uso de la «composición» o «redacción», un género puramente escolar (figuras 12 a 14). Nótese la consigna «amo la vida» en la Figura 12, IV Parte. No existe ninguna especificación acerca de la extensión del texto, sus propósitos o destinatarios. Las únicas indicaciones acerca de las características que debe tener el texto son normas gramaticales. Ante esta consigna, uno podría escribir «Amo la vida porque es hermosa». Esta frase cumple

Escuela República De Chile  
 Primer Trimestre / Sexto Grado  
 II Prueba de Español / Tiempo Probable: 80 min.

Valor total 50 pts      Pts. Obti. 38      Nota: 76

Nombre: Jatara Salazar Aguilera Sección: 6-1      8-05-08  
 Firma del Padre o encargado: \_\_\_\_\_  
 Adecuación Curricular: ( ) No Significativa ( ) Significativa

**Instrucciones generales:** Escuche con atención las indicaciones que da la maestra, trabaje en orden, no me distraiga, cuido mi prueba para que no se me ensucie, si tengo dudas levanto la mano y espero en mi lugar, El orden y aseo de la prueba se tomará en cuenta. No se acepta el uso del corrector en la prueba.

**I. Parte Selección Única.** Selección única. Marque con equis (X) la opción que antecede a la respuesta correcta., de acuerdo a lo visto en clase. Valor 10 pts. 1 pt cada acierto.

1- Son palabras agudas las que se encuentran en la siguiente opción:

a. brisa - Raúl - delantal  
 X cayó - Raúl - canción  
 c. cayó - Raúl - María  
 d. cayó - Raúl - aéreo

**II PARTE . DESARROLLO ( 9 PTS )**

Lea con atención lo que se le solicita, responda en forma ordenada, letra legible y con limpieza

2. Explique mediante 3 razones la diferencia entre HIATO, DIPTONGO y TRIPTONGO (3pts)

HIATO Es la unión de dos vocales fuertes o una fuerte con una débil y débil  
 DIPTONGO Es la unión de una vocal débil con una fuerte  
 TRIPTONGO Es la unión de una vocal débil fuerte una débil y otra fuerte

3. Explique mediante 3 razones la diferencia entre los tipos de AGUDAS, GRAVES y ESDRÚJULAS ( 3 Pts )

AGUDAS tienen el acento en la última sílaba  
 GRAVES tienen el acento en la penúltima sílaba  
 ESDRÚJULAS tienen el acento en la antepenúltima sílaba

4. Explique con 4 razones cuando debemos de emplear correctamente la letra mayúscula ( 3 Pts )

a. Cuando se inicia una oración  
 b. Después de un punto  
 c. Nombres propios

Figura 11

**III PARTE  
IDENTIFICACIÓN**

En cada oración coloque la tilde en aquellas palabras que así lo requieran de acuerdo a lo estudiado en clase. **Valor 10 puntos. 1 punto cada tilde correcta.**

En la pulpería no hay café ni azúcar. Valor 3 puntos - |

¿te gusta el té con limón ácido? Valor 3 puntos - |

Envío un recado para él. Valor 2 puntos

Tú encanto es único. Valor 2 puntos

**IV PARTE  
DESARROLLO**

**A.** Conteste de acuerdo a los datos del siguiente gráfico de barras la información que se le indica **Valor 6 puntos. 1 punto cada acierto.**

Alumnos aprobados por nivel

**IV PARTE. DESARROLLO. VALOR 18 PUNTOS**

A continuación usted debe escribir un texto, aplicando cada uno de los objetivos temáticos escritos en la tabla. El tema es **"Amo la vida"**.

Tabla para calificar la redacción

Objetivos temáticos	3 puntos No errores	2 puntos 1 a 3 errores	1 punto 4 a 6 errores	Puntos obtenidos
1. Escritura legible en sus trazos. Repetición innecesaria de ideas.				
2. Uso del acento ortográfico: agudas, graves y esdrújulas.				
3. Uso correcto de la tilde diacrítica				
4. Uso del hiato y ley del hiato.				
5. Uso correcto de las consonantes: b, v, c, s, z, m, n, h, rr, r, g, j. Uso gu con o sin diéresis.				
6. Utiliza signos de puntuación. Letra mayúscula.				

Figura 12

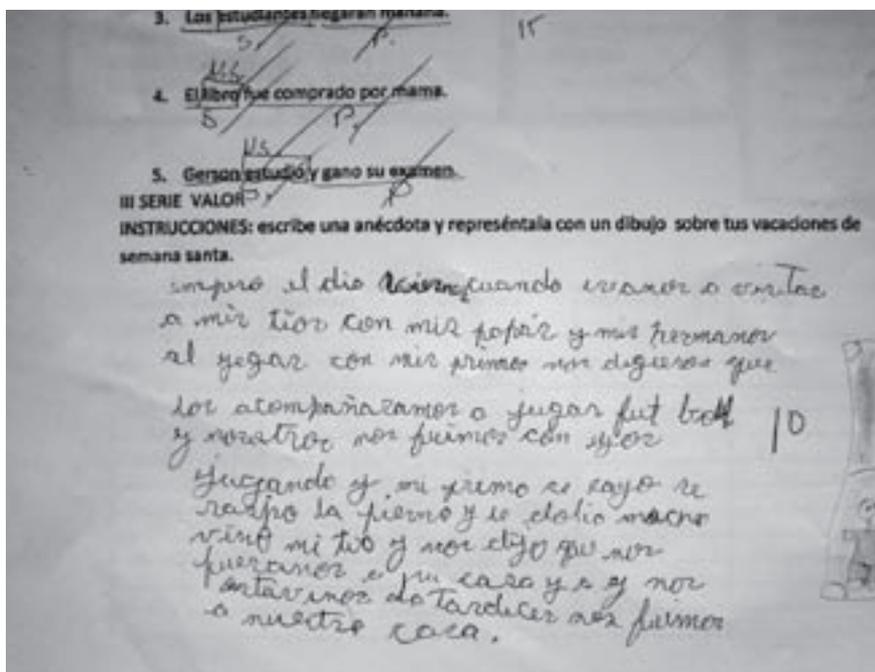


Figura 13

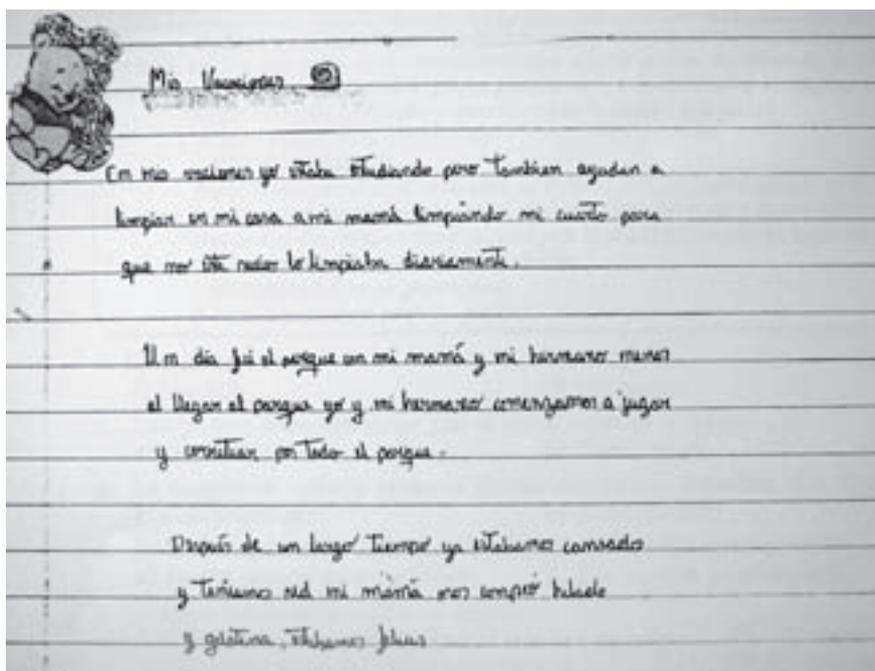


Figura 14

con todos los requisitos establecidos para la tarea: es un texto, no tiene errores, usa mayúscula y puntuación, no repite ideas. El único problema que podría tener es que la escritura no fuese legible. Como contraste, las figuras 15 y 16 muestran dos propuestas de escritura con mayor sentido y complejidad (aunque carecen de contexto, propósito y audiencia explícitos).

Las actividades de lectura suelen estar propuestas a partir de textos poco complejos y de relevancia social al menos discutible, con preguntas poco agudas que no promueven una reflexión en profundidad sobre los textos (figuras 17 a 19). Dan lugar a «explicaciones» tales como «trata de una escuela» (figura 17); a «conclusiones» centradas en cómo pronunciar las letras y mejorar la caligrafía (figura 18); y a «análisis personales» tan poco analíticos como «Esta historia me gustó porque conocí una cultura en una isla llamada Haití» (figura 19). En Argentina y Uruguay fue posible encontrar propuestas de lectura más relevantes y complejas, que incluyen el uso de noticias de prensa (figura 20), textos informativos auténticos (figura 21) y literatura de cierta complejidad (véase el texto de Kafka en la figura 22). Por otra parte, es notoria la mayor agudeza de las preguntas que formulan los maestros (figuras 20 y 23).

En síntesis, a pesar del énfasis que el discurso pedagógico pone en el pensamiento crítico y en la importancia de la reflexión, en las aulas predominan las propuestas de corte netamente escolar, sin contexto real plausible, de discutible relevancia y cuya resolución requiere procesos extremadamente simples.

## LA EVALUACIÓN FORMATIVA Y LAS DEVOLUCIONES A LOS ESTUDIANTES

### DISCUSIÓN TEÓRICA

La evaluación formativa es parte del proceso de enseñanza y su propósito es ayudar al estudiante a avanzar en el aprendizaje. Según afirmamos antes, aunque en el discurso predomina la valoración positiva de la evaluación formativa, ello no se ve reflejado en la práctica. «La mayoría de los maestros en servicio tiene sólo un conocimiento limitado de estrategias de evaluación formativa, y sigue pensando en la evaluación como un proceso que sirve principalmente para calificar.» (Linn, 2000: 17)

Para que exista evaluación formativa son indispensables dos elementos: una buena explicación de qué es lo que se espera que el estudiante logre y una buena devolución<sup>9</sup> a lo largo del proceso de trabajo para alcanzar dichos logros. Según indicamos antes, no nos detendremos en el primero de estos dos elementos. Simplemente consignaremos un par de citas sobre su importancia.

---

9 Utilizaremos el término «devolución» como traducción del término anglosajón «feedback», si bien muchos autores prefieren utilizar el término «retroalimentación».

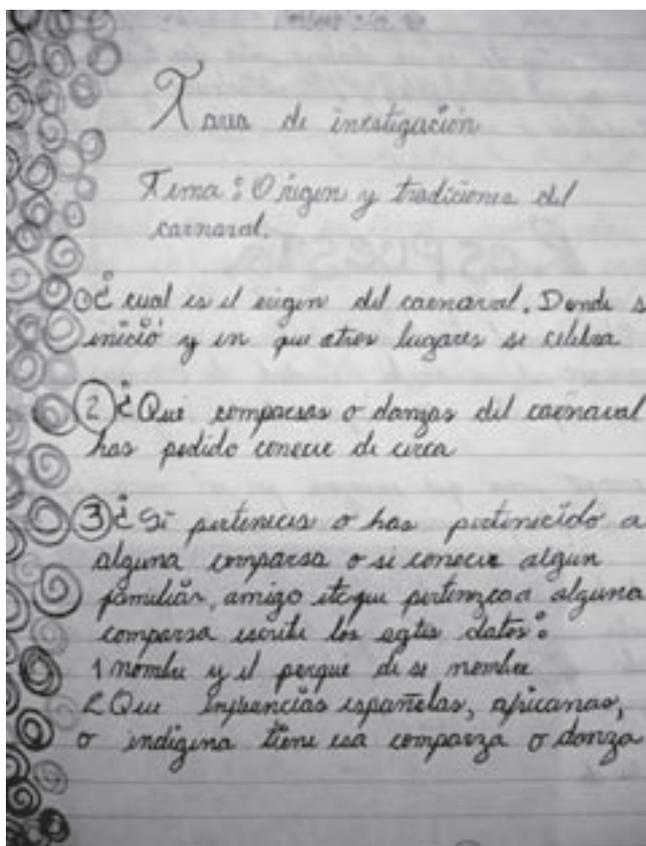


Figura 15

Establecer objetivos claros para aprendizaje por parte del estudiante implica mucho más que anunciar una finalidad de la enseñanza para que los estudiantes la contemplen. También requiere la elaboración de los criterios mediante los cuales será juzgado el trabajo del estudiante. ¿Cómo sabrán el maestro y el estudiante que se ha entendido un concepto? ¿Cómo se evaluará la capacidad del estudiante para defender un argumento? (19)

Sin objetivos, sin metas, sin una finalidad que cumplir, no hay nada que evaluar. Lo que equivale a decir que si los objetivos del proceso de enseñanza no se conocen, o son vagos o nebulosos, la evaluación será sólo una actividad ritual. (Meza, 1991: 14)

En una primera aproximación conceptual al tema que nos ocupa en este apartado, Tunstall y Gipps (1996) distinguen dos tipos: la devolución *valorativa* y la devolución *descriptiva*. La primera asume la forma de juicios de valor, que se expresan de distintos modos y que, normalmente, son de carácter comparativo con el resto de los alumnos. La segunda asume la forma de referencias específicas a los logros o progresos de los alumnos, en relación con los desempeños esperados.

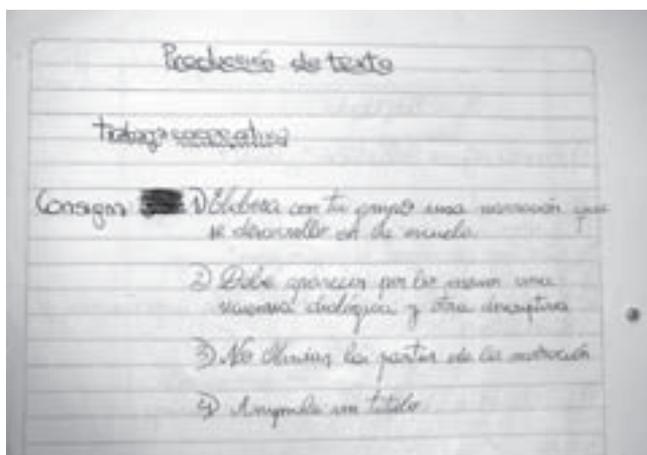


Figura 16

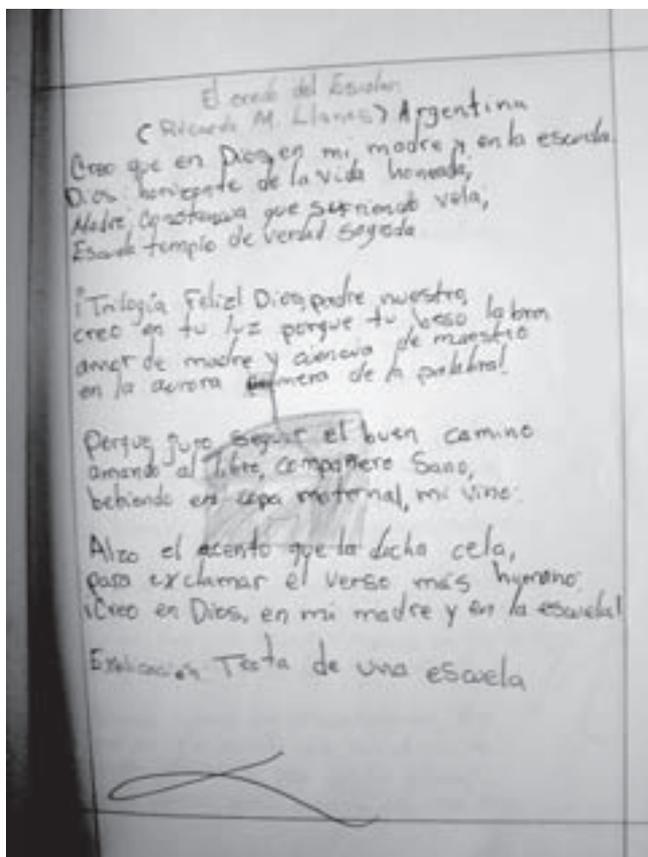


Figura 17

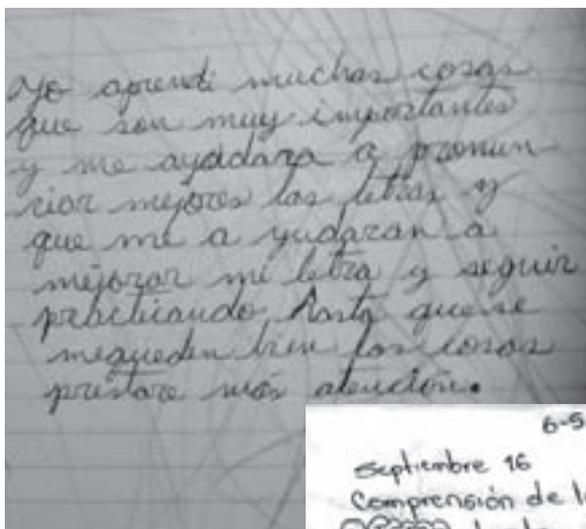


Figura 18

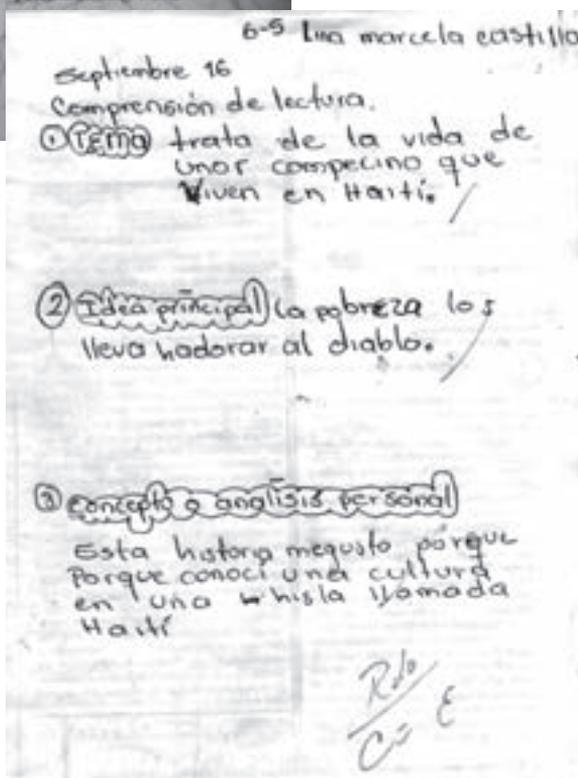


Figura 19

La devolución valorativa se expresa a través de manifestaciones de aprobación o desaprobación. Está orientada a los aspectos afectivos, motivacionales y de esfuerzo en el aprendizaje. La devolución descriptiva se expresa a través de la especificación de los logros o progresos del alumno y ofrece modelos de trabajo o de acción, o procedimientos para la revisión y corrección por parte del alumno (189).

Shepard (2008) sugiere que es más formativa la devolución de tipo descriptivo que la valorativa. «Es más probable estimular el aprendizaje cuando la

retroalimentación se enfoca en ciertos aspectos de la tarea y destaca los objetivos de aprendizaje.» (25) Según este autor, la evidencia derivada de la investigación muestra que es un error común de muchos docentes hacer siempre elogios para estimular a los estudiantes, aún cuando su trabajo no sea bueno. «El modelo de evaluación formativa, consistente con la literatura cognitiva, demuestra que la retroalimentación es especialmente eficaz cuando dirige su atención a cualidades particulares del trabajo del estudiante en relación con criterios establecidos y proporciona una guía sobre qué hacer para mejorar.» (25)

Wiggins (1998, 46-53) va un poco más allá y realiza una interesante distinción entre tres conceptos: *valoración*, *orientación* y *devolución* propiamente dicha. La *valoración* tiene lugar cuando entregamos a los alumnos juicios de valor acerca de su trabajo, sea en la forma de calificaciones, sea a través de expresiones tales como «te felicito», «debes esforzarte más», «insuficiente», etcétera. La *orientación* consiste en consejos o sugerencias acerca de cómo mejorar el trabajo. La *devolución* es información que le permite al alumno comparar lo que intentó lograr con lo que efectivamente hizo. Cuanto más autoevidente sea la devolución, mejor será, porque ayudará al estudiante a darse cuenta por sí mismo de lo que ha logrado y lo que todavía no.

Muchos educadores parecen creer que la devolución significa brindar a los estudiantes mucha aprobación y un poco de desaprobación y consejo [ . . ] Es importante elogiar a los estudiantes porque los satisface y los anima. Los elogios te mantienen en el juego; pero sólo la devolución real te ayuda a mejorar [ . . ] La devolución es información sobre como una persona se desempeñó, a la luz de lo que intentó hacer —intento contra efecto [ . . ]. (46)

Un error todavía más común es la visión de que devolución es lo mismo que orientación. Devolución y orientación son cosas muy diferentes; representan partes complementarias de un sistema de auto-corrección. La devolución te dice lo que resultó de tu acción; la orientación te dice como mejorar la situación [ . . ] En general tendemos a dar demasiada orientación y poca devolución. De hecho, muchas veces saltamos por encima de la devolución y vamos directamente a dar consejos [ . . ]. (50-51)

Para ilustrar su concepción de la devolución, propone el siguiente ejemplo. Un maestro técnico está enseñando en un curso de soldadura. Entrega a sus estudiantes una serie de instrucciones para construir un ángulo recto con metal y los criterios que debe cumplir el producto. Les indica que cuando crean que han finalizado, escriban su nombre en el ángulo que hicieron y lo depositen en una gran mesa de trabajo ubicada al frente del salón. A medida que terminan, los estudiantes escriben sus nombres en sus ángulos y se acercan decididos a la mesa. Pero, al acercarse, ven que sobre ella están dispuestos en hilera una gran cantidad de ángulos hechos por alumnos de cursos anteriores, ordenados desde los más desprolijos hasta los mejor logrados. Antes de depositar el suyo sobre la mesa, los estudiantes lo miran y lo comparan con los que están allí. Muchos vuelven a su mesa de trabajo y comienzan nuevamente su trabajo (50).



Éste es, según Wiggins, un ejemplo del mejor tipo de devolución formativa. No hay juicios de valor, no hay elogios ni críticas. Se trata simplemente de un dispositivo que permite que el propio estudiante compare lo que hizo con lo que se propuso hacer. Obviamente, es relativamente sencillo montar un dispositivo de esta naturaleza en una producción material. Es más difícil llevar esta idea al plano de la producción intelectual. Un camino posible es hacer que los estudiantes analicen tres trabajos escritos, uno muy bueno, uno intermedio y uno insatisfactorio, para luego comparar y analizar el propio trabajo. Otro camino es emplear rúbricas que permitan al estudiante autoevaluar distintas dimensiones de su trabajo.

## EVIDENCIA EMPÍRICA

A lo largo de los varios miles de tareas revisadas, se encontró casi exclusivamente devoluciones de tipo valorativo: marcas de acierto o error, calificaciones, felicitaciones, indicaciones de que algo es incorrecto. Casi no existen devoluciones descriptivas y reflexivas en los trabajos escritos. Si el lector vuelve sobre las figuras presentadas, encontrará esto en todas las que tienen algún tipo de devolución o

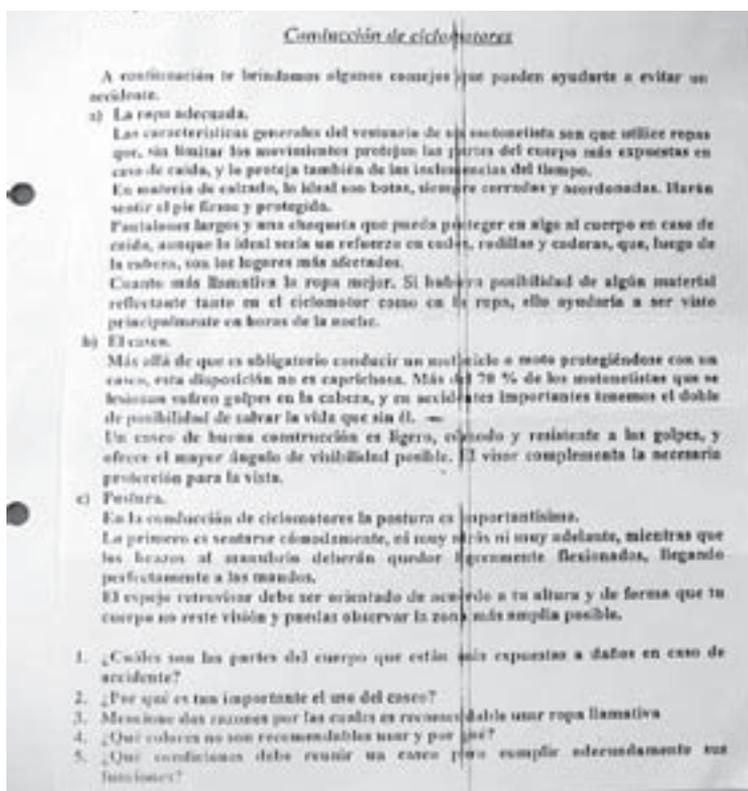


Figura 21

marca escrita del maestro. Hay una sola excepción sumamente interesante. En la figura 24 es posible observar una serie de comentarios y preguntas del maestro al estudiante: «¿por qué aparecen entre comas?», «puede ser», «¿lo llama de algún modo especial?», «¿qué cosa te lo podría asegurar», ¿de quién habla ahí?». Se trata de un maestro excepcional que, de un modo muy sencillo pero sistemático, escribe a todos los estudiantes, incluidos aquéllos que resuelven las tareas de manera correcta, este tipo de comentarios que invitan a la reflexión. Este mismo maestro, en matemática, ante ejercicios bien resueltos, formula por escrito preguntas tales como «¿y qué hubiese pasado si en vez de tal cifra hubiese sido tal otra?». En la entrevista este maestro explicó su enfoque en los siguientes términos:

Les devuelvo los trabajos con preguntas. Generalmente resalto alguna cosa que no se tuvo en cuenta. Por ejemplo, el otro día les había dado una propuesta de evaluación con fracciones a las que les faltaba el numerador o el denominador y les pedía que las completaran para que la fracción fuera mayor a 1. Entonces la devolución sería una pregunta como «¿qué pensás sobre esto?». Es una pregunta para que reflexionen y puedan volver sobre el tema con un objetivo concreto. (E15-URU)

En contraposición con lo que hace este maestro, es llamativa en las figuras 17 a 19, la presencia de respuestas absolutamente simples, que no merecieron ningún comentario del maestro que promoviese una ampliación del horizonte de reflexión del estudiante.

Se podría suponer que la mayoría de los maestros realizan devoluciones descriptivas y formativas de manera oral y no escrita. En las entrevistas en profundidad se preguntó a los maestros de qué modo trabajan con los estudiantes a partir de una evaluación y qué hacen con los resultados de la misma. La evidencia recogida permite identificar dos modalidades principales de devolución.

En primer lugar, muchos maestros explican que la devolución consiste en señalar a los alumnos los errores cometidos en forma escrita u oral. Según se desprende de los pasajes que siguen y de los trabajos incluidos en las figuras presentadas más arriba, se trata de correcciones directas a tareas simples, que no admiten alternativas de resolución.

- Les devuelvo los trabajos y les hago ver las equivocaciones, las palabras mal escritas. (E13-SAL)
- En las redacciones se les devuelve en un círculo lo que está mal para que encuentren el error. Les doy indicaciones: busque en el diccionario. También uno va corrigiendo automáticamente la ortografía. Si hay errores de secuencia se les hace un comentario. (E3-CRI)
- Explico en forma general dónde están las fallas y ellos siguen en sus evaluaciones escritas [ . . . ] (En lenguaje) comunico dando sugerencias en el propio trabajo. Si hay muchos errores, lo mando a hacer de nuevo. Señalo los errores con un círculo. (E16-PER)

En segundo lugar, es interesante la amplia presencia del recurso a estímulos, premios y exhortaciones a un mayor esfuerzo. La idea implícita es que el problema principal es

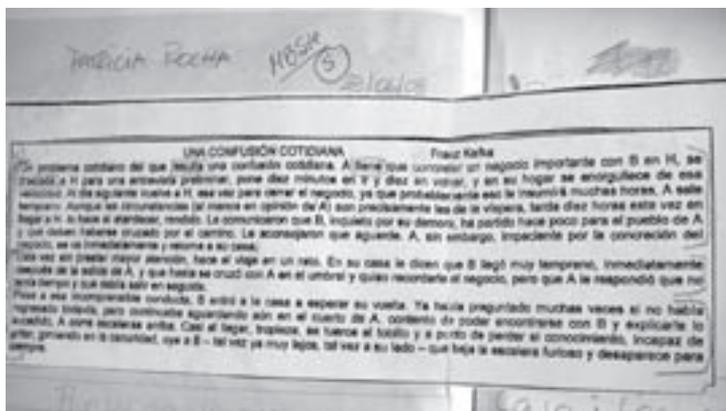


Figura 22

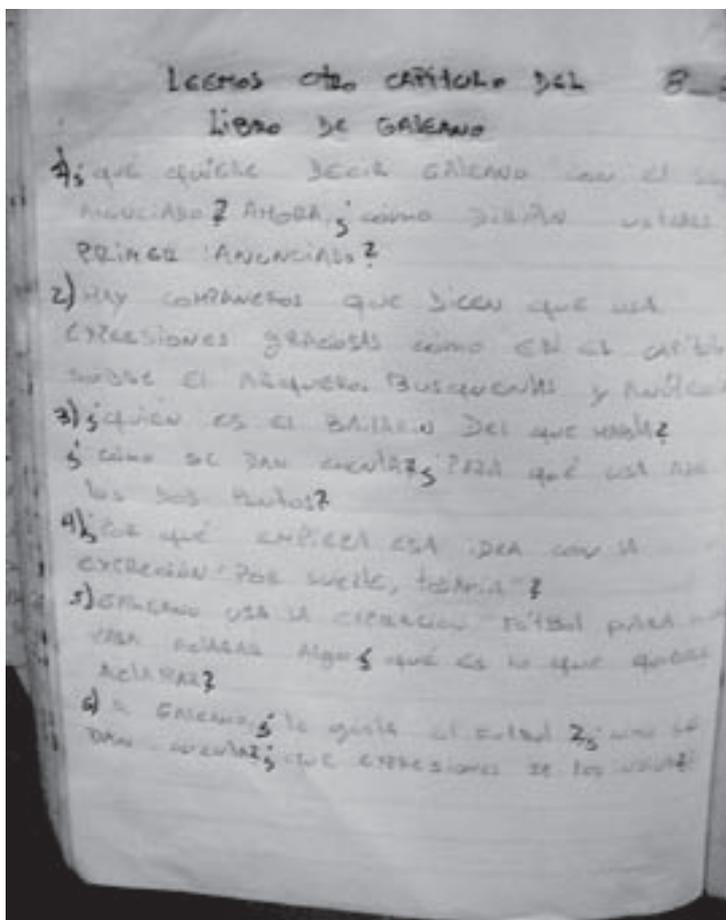


Figura 23

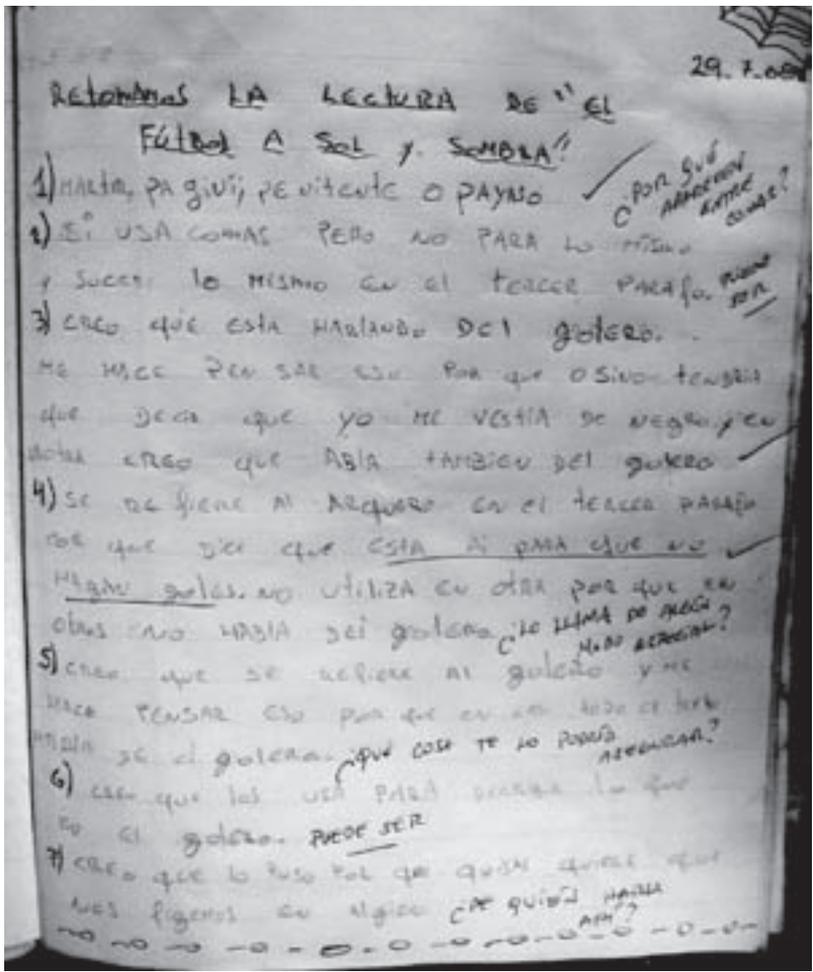


Figura 24

el esfuerzo de los alumnos. Esta postura refleja ausencia de percepción acerca de los procesos cognitivos de los alumnos y de sus dificultades. Como consecuencia, genera ausencia de modos de intervención oportunos para ayudar al alumno a superar sus dificultades. Se confunde evaluación y devolución con motivación y estímulo.

- Doy estímulos positivos diciéndole «tú puedes». (E4-CRI)
- A cada quien se le entrega su papel, su trabajo. Ahí va su nota [ . . . ] Cuando amerita, por ejemplo un alumno que nunca sale bien, le digo «lo felicito, mi amor». Y así se motiva. (E5-SAL)
- Se les entrega la papeleta [ . . . ] Se les pone alguna observación. Por ejemplo al niño que ha trabajado muy bien se le pone una felicitación. Al que ha salido mal le pongo «necesitas esforzarte un poco más», «tu tarea no ha estado completa», «muy bien, te felicito». Siempre nota. (E15-SAL).

- Todo trabajo tiene validez. Si bien es cierto que hay niñas que no trabajan como uno espera, no es correcto decirles que su trabajo está mal. [En esos casos] yo pongo un cartelito «Revisado» y luego cartelitos que dicen, «Corrige ortografía», «Entregue su trabajo ordenado», «Dedíquese», «Debe mejorar», para algunos tengo sellos. (E18-GUA).

- En las escritas le doy los puntos obtenidos y algunos comentarios tales como «esto no es así», «repasa más la resta». [En lenguaje] les hago comentarios escritos en el borrador de redacción y evalúo todo el proceso incluyendo borrador y final. Entrego la hoja final con la nota y algún comentario de ser necesario. (E5-PER)

- Si mejoran al final les ofrezco premios [. . .] Les digo que si mejoran les daré juguetes [. . .] chocolates [. . .] (E9-GUA)

En resumen, en consonancia con lo que se expresa en la literatura sobre el tema, la mayoría de los maestros tiende a dar a sus alumnos mucha devolución de carácter valorativo, a veces a través de calificaciones, a veces de expresiones orientadas a la motivación, y poca descripción de lo producido por el alumno, poca orientación y pocas oportunidades para reflexionar sobre las tareas y sus dificultades. La visión del aprendizaje subyacente es que éste depende sobre todo del esfuerzo (lo cual en parte es cierto), pero desconoce la importancia y peculiaridad de los procesos cognitivos de los estudiantes.

## EL SINSENTIDO DE LAS CALIFICACIONES

### DISCUSIÓN TEÓRICA

Según se explicó al inicio de este texto, calificar el trabajo de los alumnos responde a una función social: certificar que un estudiante ha logrado un conjunto de capacidades y conocimientos, que constituyen el propósito de un período de formación. A través de la certificación, lo que ocurre en el aula sale de ese ámbito aislado y se hace público: se comunica a las familias de los alumnos, al resto de los educadores y a la sociedad.

En el nivel terciario es dónde la necesidad de esta función aparece con mayor claridad: una institución educativa que otorga un título de médico está haciéndose cargo ante la sociedad de que la persona posee los conocimientos y herramientas necesarios para desempeñar esa función. En niveles inferiores, cuando un estudiante aprueba un nivel, el sistema educativo está diciendo al resto del propio sistema que el estudiante ha alcanzado los conocimientos y capacidades necesarias para continuar estudios en el nivel siguiente<sup>10</sup>.

La certificación, en tanto constancia escrita, suele plasmarse a través de puntajes, categorías o calificaciones. En teoría, cada categoría o calificación sintetiza o

10 Lo cual no implica asumir como único modelo educativo válido el que organiza la escolaridad en grados anuales.

representa un determinado nivel de conocimientos y capacidades alcanzado por el estudiante. Como veremos enseguida, esto no es así, ya que los caminos por los cuales los docentes llegan a establecer las calificaciones son variados, arbitrarios y confusos.

Desde el punto de vista conceptual, hay tres enfoques principales para establecer los juicios de valor que dan lugar a las calificaciones: normativo, de progreso y criterial (Ravela 2006, 43-49).

Una calificación tiene carácter *normativo* cuando el juicio de valor acerca del desempeño de un estudiante se construye a partir de la comparación con el resto del grupo. Un alumno con desempeño *mejor* que el resto, recibirá una calificación alta. Este enfoque predominó durante mucho tiempo en los sistemas educativos. Está vinculado a una concepción selectiva de la enseñanza (identificar y reconocer a los mejores estudiantes) y presupone que la motivación principal de los estudiantes radica en el deseo de destacarse y en la competencia con los demás. El problema principal de este enfoque es que promueve la confusión entre *mejor* y *bueno*. El mejor estudiante de un grupo en una determinada escuela puede tener un desempeño pobre en relación con lo que ocurre en otras escuelas o en relación con los conocimientos y capacidades esperados.

El enfoque de *progreso* implica construir el juicio de valor sobre el desempeño del estudiante a partir de la comparación con su propio punto de partida. Un estudiante que mejoró mucho, recibirá una calificación alta. Este enfoque está vinculado a una visión del aprendizaje como un proceso que tiene lugar de distintos modos y con distintos ritmos, según los individuos. Presupone como fuente de motivación principal el deseo de superación personal: el estudiante «compite consigo mismo» para mejorar. Un riesgo de este enfoque es que no necesariamente asegura que el estudiante logre los aprendizajes esperados. Un segundo riesgo, según veremos enseguida, es que puede propiciar la confusión entre progreso y esfuerzo, que son cosas diferentes.

El enfoque es *criterial* cuando el juicio de valor sobre el desempeño del estudiante se construye a partir de la comparación de su desempeño con un conjunto bien definido de conocimientos y capacidades que debe dominar. El énfasis está puesto en los propósitos de la acción educativa en un curso, en aquello que se espera que los estudiantes logren. Requiere describir distintos niveles de desempeño, algunos considerados aceptables y otros no. Presupone la motivación de los estudiantes por alcanzar un buen nivel de dominio de la disciplina o materia.

Desde la perspectiva de la función de certificación de la evaluación, tal como fue presentada en este texto, las calificaciones deberían estar basadas en un enfoque criterial. Cada categoría de calificación debería estar vinculada a una descripción de desempeños. Sin embargo, no es común encontrar buenas descripciones de los desempeños esperados de los estudiantes ni en los programas de estudio, ni en las planificaciones de los maestros.

El principal problema de las calificaciones es, justamente, que los logros a alcanzar no suelen estar definidos con claridad. Según mostraremos enseguida, los documentos normativos de los países establecen la escala de calificaciones y, a

veces, los procedimientos e instrumentos que deberá emplear el maestro, así como su peso relativo en la calificación. Pero no especifican qué es lo que se espera de un alumno para que sea considerado aceptable, destacado o insuficiente y reprobado.

En estas condiciones y teniendo en cuenta su importancia social, la calificación muchas veces deja de ser un instrumento de certificación y se transforma en un instrumento de disciplinamiento por parte del docente. Esto puede tergiversar el sentido de la labor educativa, colocando a las calificaciones como el objetivo central y desplazando al aprendizaje (o limitándolo a aquello que será objeto de calificación). Este problema no se puede resolver eliminando las calificaciones sino empleándolas de manera apropiada.

Shepard (2008) plantea esta problemática en los siguientes términos:

La evaluación sumativa y la calificación constituyen una seria amenaza para los objetivos de aprendizaje declarados por la evaluación formativa [ . . . ] las calificaciones pueden minar el proceso de aprendizaje de varias maneras [ . . . ] los estudiantes concentran su atención y esfuerzo sólo en la porción calificada del currículo [ . . . ] el uso de calificaciones como premio o como castigo puede socavar la motivación intrínseca de aprender [ . . . ] aquellos estudiantes para quienes los criterios de las calificaciones les parecen fuera de su alcance pueden reducir su esfuerzo y su ulterior aprendizaje. (30)

Sin embargo, hay cierta evidencia de beneficios cognitivos positivos de las evaluaciones sumativas [ . . . ] lo más importante, los estudiantes parecen estudiar y aprender más si esperan que les hagan una prueba [ . . . ] la experiencia misma de presentar pruebas pone a los estudiantes a procesar mentalmente el contenido, *si bien esto depende mucho de la calidad de las preguntas en la prueba* (35, énfasis nuestro).

## LA EVIDENCIA EMPÍRICA

En el marco del estudio se realizó un relevamiento de las normas de evaluación vigentes en cada país y de los currículos de Primaria. En todos los casos existe una normativa que define el tipo de calificaciones que deben ser utilizadas (puntajes, categorías, números, letras). Algunas normativas establecen también los procedimientos de evaluación que deben ser implementados y los pesos que cada uno debe tener en la calificación final de un estudiante. En ninguno de los países se encontraron definiciones de las calificaciones en términos de los desempeños o conocimientos específicos que corresponden a cada una de ellas. Veamos algunos ejemplos.

En Costa Rica el docente debe calificar a cada estudiante a través de un puntaje en una escala de 0 a 100, que se conforma de la siguiente manera:

- 50% con resultados de pruebas objetivas.
- 30% con la observación del trabajo cotidiano.
- 15% con la valoración del trabajo extra clase.
- 5% con el «concepto» del docente sobre el estudiante.

En El Salvador se debe dar a cada estudiante una calificación trimestral, que debe resultar del promedio ponderado de tres tipos de actividades:

- Una actividad integradora de resolución de problemas o de aplicación de contenidos conceptuales, procedimentales y actitudinales (35%).
- La revisión de cuadernos, trabajos grupales y tareas (35%).
- Una o más pruebas (30%).

En Guatemala rige una escala numérica de 1 a 100 puntos, con cuatro categorías:

- Excelente: cuando el estudiante alcanza entre 90 y 100 puntos en las actividades de evaluación registradas.
- Muy bueno: entre 80 y 89.
- Satisfactorio: entre 70 y 79.
- Debe mejorar: cuando el alumno logra menos de 70 puntos.

En México se utiliza una escala de 5 a 10 puntos, siendo 6 la nota de aprobación. La calificación final de cada asignatura es el promedio de las calificaciones parciales que el docente debe asignar en cinco momentos del año. En Uruguay se emplea una escala de 13 categorías, que van desde el «deficiente» hasta el «sobresaliente». La calificación para aprobar es «bueno».

Como se puede apreciar, los formatos y escalas de calificación difieren entre los distintos países. Pero todos tienen una propiedad en común: no incluyen descripciones de qué debe saber y ser capaz de hacer un estudiante para obtener cada calificación. Como consecuencia, las mismas dependen principalmente del juicio subjetivo del maestro. Nótese que, aún cuando se especifiquen instrumentos y ponderaciones, el grado de complejidad de las pruebas y tareas dependen del docente.

[. . .] los criterios son diversos, a un mismo niño con el mismo conocimiento, uno le puede poner ocho y el otro maestro le puede poner diez o algún otro puede decir que le pone siete, porque influye mucho el criterio de los maestros [. . .] (E5-MEX)

Preguntados acerca de los criterios que emplean a la hora de calificar a sus estudiantes, los maestros indican que proceden de varias maneras. Algunos elaboran una suerte de algoritmo matemático para llegar a la calificación, lo cual le da una suerte de «objetividad» al proceso, aunque esconde el problema de los desempeños.

[. . .] que los alumnos resuelvan en parejas durante dos horas vale 50% de la nota del mes. (E5-SAL)

[. . .] las calificaciones del examen las registro aquí en mi lista, y luego ya eso sería una parcial y para la total tomo en cuenta, mi registro con sus tareas, con sus trabajos; [para que al] final de cuentas ya tenga tres calificaciones, las sumo y ya la sumatoria es lo total. (E8-MEX)

En cualquier materia, por ejemplo, yo tomo un valor en el examen de siete puntos y tres puntos de escala, es una escala que llevamos aquí que considera los trabajos, y tomando muchos otros puntos, por ejemplo la asistencia, la disciplina, la puntualidad, higiene y varias actividades, les damos un porcentaje para completar los tres puntos que son, más los siete puntos del examen. (E19-MEX)

[. . .] con una escala numérica. No es la misma para todos [los trabajos]: 30 puntos, 10 puntos, 8 puntos. Esto con la traducción a una nota, porque ellos cuando ven la nota entienden más. Digo de tal puntaje a tal tiene tal nota, de tal otro a otro tiene esta otra [. . .] (E3-URU)

Una parte de los maestros no utiliza instancias formales de evaluación tipo prueba o examen para establecer las calificaciones, sino que éstas se derivan de un juicio global que, supuestamente, es el resultado de todo lo que hace el alumno. Dentro de este grupo de maestros que no utilizan pruebas formales, hay dos categorías. Unos califican continuamente todas las tareas y actividades, a partir de lo cual llegan a una calificación al final del proceso.

Sumo todas las notas que tienen durante el trimestre. No hago pruebas específicas para evaluar ni en matemática ni en lenguaje. (E5-PER)

[. . .] todas [las evaluaciones] son para calificar [. . .] todas se califican y luego se promedian con la responsabilidad. (E15-PER).

[. . .] Todo lleva al resultado final [calificación], hasta la asistencia [. . .] El trabajo cotidiano que uno hace con los chiquitos, tareas cortas, tareas extra clase [. . .] Todas las evaluaciones sirven para ver cómo seguir trabajando y para calificar. (E2-CRI)

Otros no califican nunca los trabajos de sus alumnos. La calificación final es el resultado de «todo lo que yo sé del alumno».

No pongo evaluaciones para calificarlos, los pone nerviosos, no sirven. Uso más las evaluaciones permanentes. (E9-URU)

Para calificar uso las evaluaciones de todos los días. La nota que va al boletín es el resumen de todo lo que yo sé del alumno. No hago nada en forma especial. (E18-URU)

Otro aspecto importante es que, a la hora de calificar, los maestros no tienen en cuenta únicamente el desempeño, sino que consideran también aspectos actitudinales (véanse los pasajes recién transcritos E19-MEX, E15-PER y E2-CRI). Los testimonios que siguen lo hacen más explícito aún:

Evalúo constantemente: conducta, higiene, trabajo, participación, interés. (E6-GUA)

Califico lo afectivo, la colaboración, la participación, el aseo, cómo se comportan, lo cognoscitivo. (E14-GUA)

Un tipo de evaluación que hago es evaluar conducta. Los niños son inquietos en conducta; hablan, no ponen atención. Me gusta fomentar valores, la autoestima. (E5-CRI)

A partir de una revisión de trabajos de investigación sobre el tema de las calificaciones, Shepard (2008) señala que la consideración de factores ajenos al desempeño a la hora de calificar, crean un «sesgo de buenos resultados» .

Los expertos dan argumentos en contra de considerar el esfuerzo, la capacidad y la actitud cuando se califica, porque minan la validez de las calificaciones como indicadores de desempeño. Además, estos factores crean desigualdades, invitan a los estudiantes a fingir y confunden a la mayoría o a todos los públicos acerca del significado de las calificaciones [ . . . ] También debemos tomar en consideración si los maestros utilizan calificaciones de esfuerzo para *controlar* la conducta de los estudiantes, que no es lo mismo que crear un ambiente de aprendizaje que los *motive*. (33, todos los énfasis en el original)

La práctica de promediar diversos factores y aspectos en las calificaciones, sin distinguirlos, da lugar a una concepción de la «escuela como trabajo y esfuerzo». Los maestros generan así «un elaborado sistema para estar informados del trabajo del estudiante, pero sin evaluar la calidad o el contenido de ese trabajo» (33).

## CONCLUSIONES

A lo largo de este artículo, hemos intentado construir una visión sobre el modo en que se desarrolla la evaluación en las aulas de Primaria, apoyados tanto en literatura sobre el tema como en evidencia empírica recogida en ocho países latinoamericanos. Es importante recordar que la evidencia empírica recogida tiene un sesgo hacia lo positivo, dado que corresponde a maestros de escuelas que se encuentran en el promedio del país o por encima del mismo, quienes, además, tuvieron la posibilidad de seleccionar los materiales que mostraron a los investigadores.

Las consideraciones teóricas y la evidencia empírica presentadas permiten ilustrar y explicar por qué la evaluación es una actividad tan controvertida en el ámbito educativo, según daban cuenta las expresiones recogidas por Caijao Restrepo entre docentes, familias y estudiantes colombianos: «predomina la memorización de los temas», «las formas de evaluación son anticuadas», los maestros no han «cambiado los paradigmas», la evaluación es arbitraria y pone de manifiesto «la preferencia de los profesores hacia algunos estudiantes», los padres no entienden los reportes de evaluación y creen que la misma «limita la creación y capacidad intelectual del niño» y que «genera más conflictos que solución a las dificultades».

En primer lugar, hemos discutido e ilustrado el problema de la excesiva simplicidad de las actividades escolares (tanto las de evaluación como las de enseñanza). Asumiendo que la evaluación que propone el maestro es un excelente indicador de lo que enseña, es posible afirmar que la labor escolar, para la mayoría de los niños de la región, está fuertemente centrada en tareas simples y

descontextualizadas, que no apelan al uso del conocimiento en situación, sino al aprendizaje de rutinas. Hemos tenido también la oportunidad de ver propuestas diferentes, más exigentes y reflexivas, cuya presencia en las aulas es minoritaria, pero que indican un camino por el que es posible y necesario transitar.

En este punto es importante señalar que el problema no radica en los currículos, ya que éstos han sido modificados y actualizados (en una perspectiva «constructivista») en prácticamente todos los países durante la última década. El problema es que entre la «reforma curricular» en el papel y el cambio en las prácticas en el aula hay un gran trecho. Modificar las prácticas de enseñanza requiere mucho más que cambios en los programas de estudio. Requiere, principalmente, de tiempos y espacios de formación en servicio en los cuáles los docentes tengan la oportunidad de revisar sus prácticas con otros y experimentar alternativas, según se indica en el cierre de este artículo. Uno de los problemas de la región es que el diseño curricular es entendido como escritura de planes y programas, pero no incluye otros aspectos centrales de una política curricular como guías de apoyo a los docentes, instancias de formación en servicio, oportunidades para la experimentación de los nuevos enfoques, entre otras.

En ausencia de actividades y metas de aprendizaje complejas, la evaluación formativa inevitablemente «hace agua». Es difícil ofrecer a los estudiantes devoluciones que inviten a la reflexión y oportunidades de autorregulación de procesos de aprendizaje, cuando éstos son de carácter mecánico y memorístico. Predomina, en cambio, el recurso a frases motivadoras, la apelación al esfuerzo, la indicación del carácter correcto o incorrecto de la respuesta dada, el señalamiento de errores y el uso permanente de la calificación de pequeñas tareas, como forma de obtener «información» para la calificación final.

Según fue posible observar, en la práctica la evaluación formativa y la evaluación para la certificación se confunden permanentemente. Todas las actividades y tareas reciben calificación. Las calificaciones finales no son el resultado de un análisis sistemático de los desempeños de los estudiantes, sino de una suerte de «hechicería» matemática que suma y promedia puntos de pruebas y actividades diversas, a lo cual se agregan, en muchos casos, valoraciones de las actitudes de los estudiantes o aspectos como la higiene personal. En forma paralela, una parte menor de los maestros se saltea estos procedimientos y deriva las calificaciones directamente de la impresión subjetiva que ha ido construyendo a lo largo del año, acerca de cada uno de sus estudiantes. Más allá de que algunos maestros lo hagan bien y otros mal, el procedimiento no deja de ser arbitrario y poco transparente para los destinatarios de la información.

En este marco y en ausencia de descripciones claras de los desempeños esperados, es casi imposible que las calificaciones tengan un genuino sentido de certificación. Como consecuencia, se transforman en algo ambiguo y poco predecible, pero que por su peso social terminan constituyéndose en el eje de las motivaciones de los estudiantes y de sus familias. Toda la actividad educativa empieza entonces a girar en torno a este eje carente de sentido.

Una agenda para enfrentar estos problemas, debería incluir la formación en servicio y la orientación continua de los maestros, en torno a cuatro ejes principales:

1. aprender a pensar y planificar los cursos y actividades en términos de los desempeños que deben lograr los estudiantes, en lugar de en términos de «temas a dar», y a comunicar de manera efectiva a los estudiantes cuáles son dichos desempeños.
2. aprender a diseñar tareas auténticas y complejas, que exijan la utilización del conocimiento y los procedimientos para la resolución de situaciones planteadas en contextos plausibles en la vida real.
3. aprender a desarrollar dispositivos de devolución de carácter descriptivo, reflexivo y orientador, que ayuden a los estudiantes a autoevaluarse y a reflexionar acerca de las respuestas que dan a las tareas y acerca de la calidad de los que producen.
4. aprender a desarrollar dispositivos para definir calificaciones que estén basados en buenas descripciones de los desempeños esperados y en evidencia empírica apropiada, que sean comprensibles para estudiantes y familias. Ello incluye distinguir claramente la función de certificación de aprendizajes de la formulación de juicios de valor sobre las actitudes y conducta de los estudiantes.

El último de estos desafíos debería, en realidad, ser asumido previamente en el nivel de conducción del sistema educativo, a través de un alineamiento explícito entre los sistemas de calificaciones y descripciones de los niveles de desempeño y comprensión que los estudiantes deben alcanzar. Sin este alineamiento, decir que una determinada calificación es la mínima para aprobar o que otra refleja un desempeño destacado, carece de todo sentido. Este paso debería ser anterior a cualquier normativa sobre los procedimientos para establecer calificaciones.

Para terminar, es importante destacar que un proceso de cambio en estos ejes sólo puede tener algún efecto en el marco de un espacio de formación y aprendizaje para los docentes, en el que se les brinden oportunidades de reflexión teórica y experimentación práctica en torno al tema.

La investigación sobre el aprendizaje y desarrollo profesional de los maestros nos proporcionan varios principios generales para respaldar el cambio [ . . . ] necesitamos tratar a los maestros como personas que aprenden, de la misma manera como les pedimos que traten a sus estudiantes [ . . . ] debemos estar conscientes de que los maestros harán contribuciones y también modificarán las herramientas y prácticas de la comunidad en la que participan. Para que los docentes hagan cambios significativos en las creencias pedagógicas y en las prácticas concomitantes, ellos mismos necesitarán experimentar y reflexionar sobre nuevos procedimientos, en el contexto de sus propias aulas. (Shepard, 2008: 42)

No existen atajos para el cambio en las prácticas educativas. Dicho cambio no se producirá a través de reglamentaciones ni de lecturas. Tampoco se producirá en lapsos cortos. Solamente se puede producir si existe un espacio institucional, un esfuerzo sostenido y un acompañamiento apropiado que ayuden a «re-visar» (volver a mirar con cierta distancia) las prácticas vigentes y a experimentar, en forma progresiva y acumulativa, nuevas propuestas de evaluación en las aulas.

## BIBLIOGRAFÍA

- Álvarez Méndez, Juan Manuel. «El Alumnado. La evaluación como actividad crítica de aprendizaje» en *Cuadernos de Pedagogía*, n.219, 1993, pp. 28-32.
- Airasian, Peter y Madaus, George. «Functional types of student evaluation» en *Measurement and Evaluation in Guidance*, n.4, 1972, pp. 221-33.
- Caijao Restrepo, Francisco. «Un recorrido por la geografía de la evaluación» en *Revista Internacional Magisterio*, n.35, 2008, pp. 28-31.
- Chaiklin, Seth y Lave, Jane (compiladores) *Estudiar las prácticas. Perspectivas sobre actividad y contexto*. Buenos Aires: Amorrortu, 2001. Primera edición en inglés: Cambridge: Cambridge UP, 1996.
- Gipps, Caroline, McCallum, Bet y Margaret Brown. «Models of teachers assessment among primary school teachers in England» en *The Curriculum Journal*, vol. 7, n.2, 1996, pp. 167-83.
- Linn, Robert y Gronlund, Norman. *Measurement and Assessment in Teaching*. Nueva Jersey: Prentice-Hall, 2000.
- López Pastor, Víctor et al. «¿Debe el alumnado participar en la evaluación? Propuestas y experiencias de autoevaluación, coevaluación y evaluación compartida» en *Novedades Educativas* Año 18, n.195, 2007, pp. 24-29.
- Mancovsky, Viviana, «De quién es la evaluación? Hacia una concepción ética de las prácticas evaluativas» en *Novedades Educativas* Año 18, n.195, 2007, pp. 4-7.
- Menéndez, Irma. «Evaluación» en *Quehacer Educativo* n.38, 1999, pp. 19-20.
- Meza, Iván et al. *Evaluación Educativa. Manual para Educadores*. Santiago de Chile: Universidad Católica de Chile, Facultad de Educación, 1991.
- Ravela, Pedro. *Fichas didácticas. Para comprender las evaluaciones educativas*. Santiago de Chile: PREAL, 2006.
- Shepard, Lorrie. *La evaluación en el aula*. México: Instituto Nacional para la Evaluación de la Educación, 2008. Traducción del original en inglés, «Classroom Assessment» en Brennan, Robert (ed.) *Educational Measurement*, 4ta. Edición. ACE/Praeger Westport, 2006, pp. 623-46.
- Tunstall, Pat y Gipps, Caroline. «How does your teacher help you to make your work better? Children's understanding of formative assessment» en *The Curriculum Journal* vol. 7, n.2, 1996, pp. 185-203.
- . «Teacher Feedback to Young Children in Formative Assessment: a typology», en *British Educational Research Journal*, vol. 22, n.4, 1996, pp. 389-404.
- UNESCO/OREALC. Los aprendizajes de los estudiantes de América Latina y el Caribe. Primer reporte de los resultados del Segundo Estudio Regional Comparativo y Explicativo. Santiago de Chile, 2008.
- Wiggins, Grant. *Educative Assessment. Designing Assessments to Inform and Improve Student Performance*. San Francisco: Jossey-Bass, 1998.