

**2014**

**La evaluación en las aulas de secundaria básica  
en cuatro países de América Latina,**

**por Pedro Ravela, Julia Leymonié, Jennifer Viñas y Carmen Haretche,  
Propuesta Educativa Número 41 – Año 23 – Jun. 2014 – Vol 1 – Págs 20 a 45**

---

# La evaluación en las aulas de secundaria básica en cuatro países de América Latina<sup>1</sup>

PEDRO RAVELA\*,  
JULIA LEYMONIÉ\*\*,  
JENNIFER VIÑAS\*\*\*  
Y CARMEN HARETCHE\*\*\*\*

## Introducción

Las maneras, criterios y tareas utilizadas para evaluar a los estudiantes en las aulas han sido, al igual que los modos de enseñar y las actividades de aprendizaje, aspectos poco observados en América Latina. Lo que ocurre cotidianamente dentro de las aulas sigue siendo algo poco estudiado y analizado, tanto por parte de los investigadores, como por parte de los propios actores educativos (supervisores, directivos y docentes). Tampoco suele estar presente en las discusiones públicas sobre educación ni en las agendas de política educativa.

Las prácticas de enseñanza y las experiencias concretas de aprendizaje de los estudiantes suelen ser objeto de estudio en muy pequeña escala, enfocados en unos pocos casos. Es poco, muy poco frecuente, que los docentes se visiten mutuamente en las aulas. Las visitas de directivos y supervisores suelen estar rodeadas de cierta artificialidad, sobre todo cuando de ellas se deriva una calificación para el docente.

Obviamente quienes trabajan en las instituciones educativas tienen cierto grado de conocimiento de lo que acontece dentro de las aulas y de los diferentes modos de trabajar de los docentes. Pero no existe la práctica de analizar y convertir la enseñanza, de manera sistemática, en objeto de estudio y experimentación. El trabajo dentro del aula suele realizarse en condiciones de aislamiento y corresponde al ámbito de lo "privado" de cada docente. El hecho de ser observado durante la conducción de una clase no es algo natural y cotidiano en las instituciones educativas, más bien genera resistencia.

La falta de apertura de las aulas a las miradas externas y al escrutinio de los colegas explica, en buena medida, por qué la mayoría de los esfuerzos por innovar, cambiar y/o mejorar la educación tienen un impacto limitado: no logran modificar los modos básicos e instalados de hacer las cosas dentro de las aulas.

Desde sus comienzos en 2007 el Instituto de Evaluación Educativa (IEE-UCU) asumió como

\*Prof. de Filosofía y Mg. en Ciencias Sociales y Educación, FLACSO Argentina. Desde hace más de veinte años ha trabajado en el campo de la evaluación educativa como investigador, docente, asesor y director de proyectos. Actualmente es el Director Ejecutivo del Instituto Nacional de Evaluación Educativa (INEEd) de Uruguay. E-mail: pravela@ineed.edu.uy

\*\*Lic. en Ciencias Biológicas, Universidad de la República, Uruguay; Mg. en Educación, Universidad Católica del Uruguay. Especialista en Didáctica de las Ciencias, Universidad de Lovaina; y en Enseñanza para la Comprensión, Harvard Graduate School of Education (Project Zero). Prof. Agregado (G4) en la Unidad de Enseñanza de la Facultad de Agronomía de la Universidad de la República, Uruguay. E-mail: jleymonie@fagro.edu.uy

\*\*\*Mg. en Ciencias de Desarrollo y Culturas, Katholieke Universiteit Leuven (Bélgica); Lic. en Comunicación Social, Universidad Católica del Uruguay. Inv. del Instituto Nacional de Evaluación Educativa (INEEd) Uruguay. E-mail: jennifer.vinas@gmail.com

\*\*\*\*Lic. en Sociología, Mg. en Ciencia Política y candidata a Dra. en Sociología, Universidad de la República, Uruguay. Directora de la Unidad de Evaluación de Aprendizajes y Evaluación de Programas del Instituto Nacional de Evaluación Educativa (INEEd). Prof. de Métodos Cuantitativos aplicados a Ciencia Política en la Facultad de Ciencias Sociales de la Universidad de la República, Uruguay. E-mail: charetche@ineed.edu.uy



línea de trabajo el registro y estudio de lo que ocurre al interior de las aulas, bajo el supuesto de que se trata de un aspecto clave para la mejora de la educación y para la formación de los docentes. Entre los proyectos más relevantes cabe destacar uno dirigido a crear un banco de registros digitales de clases en cinco países de la región<sup>2</sup> y otro relacionado con el uso de las XO en el aula, en Uruguay<sup>3</sup>.

Como parte de esta línea de investigación, en el año 2008 el IEE realizó, en conjunto con el GTEE-PREAL, un estudio sobre las propuestas de evaluación que utilizan en sus aulas los maestros de 6° grado de primaria en ocho países de la región<sup>4</sup>. En el marco de este estudio fue posible constatar tres problemas principales en las prácticas de evaluación de maestros de primaria:

- i) el tipo de tareas de evaluación que los docentes proponen y su bajo nivel de “demanda cognitiva”;
- ii) la ausencia de ciclos de devolución o retroalimentación como parte del apoyo al aprendizaje de los estudiantes; y,
- iii) la falta de significado de los procedimientos y criterios que se utilizan para calificar a los estudiantes (Ravela, 2010a; Ravela, 2010b; Picaroni, 2009; Loureiro, 2009)<sup>5</sup>.

El presente estudio se propuso realizar una réplica mejorada del recién mencionado. Se optó por trabajar en educación media básica, un nivel mucho menos estudiado que primaria; y en ciencias naturales, por su relevancia para la formación de los estudiantes y para el desarrollo de nuestras sociedades, así como por la posibilidad de contar con la aproximación de PISA a la cultura científica como referente. Este artículo está enfocado en el análisis de las prácticas de evaluación de los docentes.

## Breve descripción del estudio<sup>6</sup>

El universo de estudio fue delimitado en los siguientes términos:

- a. profesores de ciencias naturales: física, química, biología o similares;
- b. educación media básica: principalmente 9° y 8° grados;
- c. cuatro países elegidos por la regularidad de su participación en PISA y por cierta diversidad regional: Chile, Colombia, Perú y Uruguay<sup>7</sup>.

En cada país fueron seleccionados en forma intencional diez establecimientos de secundaria general, ubicados en distintos contextos socioculturales, evitando tanto el extremo inferior (porque suelen verificarse problemas de ausencia de enseñanza) como el superior de la distribución social (para evitar situaciones atípicas). Se incluyó una mayoría de centros públicos y algunos privados. Los establecimientos fueron contactados a través de una institución colaboradora en cada país<sup>8</sup> y su participación fue de carácter voluntario, al igual que la de los profesores.

El relevamiento de información incluyó:

- a) La revisión de los programas de estudio y normas de evaluación para los grados y disciplinas escogidos.
- b) La realización de una entrevista de alrededor de 45 minutos de duración (un período de clase) con cada profesor.
- c) La toma de registros fotográficos digitales de pruebas y propuestas de evaluación seleccionadas por cada profesor y ejemplos de trabajos de los estudiantes.
- d) La aplicación de un cuestionario autoadministrado a los estudiantes, enfocado en sus experiencias y percepciones acerca del modo en que son evaluados.

En la Tabla 1 se consignan los datos sobre el “corpus” de evidencia empírica recogido: 76 entrevistas en profundidad; una colección de 353 propuestas de evaluación, que incluyen un total de 4359 tareas; y 1824 cuestionarios respondidos por estudiantes. A lo largo de todo el estudio utilizamos el término ‘propuesta’ para designar una prueba o pauta de trabajo, y el término ‘tarea’ para designar cada una de las preguntas, ejercicios o actividades incluidas en la propuesta. En la Tabla 2 se puede apreciar la distribución de las propuestas recogidas por asignatura y por grado

en cada país. La denominación de las asignaturas varía en cada país.

Tabla 1 - Cuantificación de la evidencia empírica recogida (números absolutos)

	Cali	Lima	Montevideo	Santiago de Chile	Total
Profesores	24	18	22	12	76
Propuestas	96	91	105	61	353
Tareas	842	985	1120	1412	4359
Alumnos	704	444	267	409	1824

Tabla 2 - Distribución de las propuestas de evaluación recogidas según asignatura y grado, en cada país (números absolutos)

		PAIS				
		CHI	COL	PER	URY	Total
ASIGNATURA	Biología	5	26	9	45	85
	Ciencias Físicas	0	0	0	15	15
	Ciencias Naturales	48	39	0	0	87
	CTA	0	0	74	0	74
	Física	0	14	0	21	35
	Química	8	17	8	24	57
	Total	61	96	91	105	353
GRADO	6°	0	2	0	0	2
	7°	0	3	4	0	7
	8°	31	19	50	30	130
	9°	30	67	37	75	209
	10°	0	5	0	0	5
	Total	61	96	91	105	353

El estudio tuvo un carácter descriptivo, exploratorio y comparado. Se propuso reunir un "corpus" suficientemente amplio de propuestas y tareas empleadas para la evaluación de los estudiantes, así como de concepciones y percepciones de los docentes, con el fin de describir y analizar las habilidades y conocimientos de los profesores con relación a la evaluación de los estudiantes y a la enseñanza de las ciencias.

Es exploratorio, dada la escasez de estudios sobre el tema y la limitación del tamaño de las muestras. Más allá de esta limitación, sostenemos que el "corpus" de evidencia obtenido permite una aproximación suficientemente robusta a los enfoques y propuestas de evaluación en ciencias predominantes en cada país. Esta suposición se basa en el marco conceptual desarrollado por Stigler y Hiebert (1999) acerca del carácter cultural de las prácticas de enseñanza. Los autores muestran que, más allá de las diferencias individuales entre los docentes, en cada país existe una matriz cultural y un conjunto de "guiones" (*scripts*), que determinan el tipo de prácticas posibles y predominantes<sup>9</sup>. Desde esta perspectiva, si bien nuestro estudio no pretende tener precisión estadística, permite una primera caracterización de las prácticas predominantes en cada país y una comparación de las principales diferencias entre países<sup>10</sup>.

Se pidió a los profesores muestras de las propuestas de evaluación que emplean en el aula, tanto con fines formativos como para calificar a sus estudiantes. Este pedido fue realizado con antelación a la visita y los profesores pudieron elegir qué muestras presentar. Por tanto, en caso de existir un sesgo en la selección de propuestas, el mismo es hacia lo que cada profesor considera lo mejor de su trabajo.

Las propuestas recogidas tienen diversidad de denominaciones según los países: pruebas de control, escritos mensuales, pruebas parciales, pruebas finales, exámenes, recuperatorios, prácticas calificadas, talleres, actividades de diagnóstico, actividades formativas.

## Marco conceptual

Como se ha señalado, este artículo analiza las prácticas de los docentes para evaluar a sus alumnos, enfocándose en los tres aspectos centrales identificados como problemáticos en el estudio realizado en escuelas primarias en 2008:

- a) las tareas de evaluación;
- b) las prácticas de evaluación formativa; y,
- c) los criterios y procedimientos para calificar a los estudiantes.

Estos tres aspectos son parte esencial de la labor docente y, por tanto, de las habilidades y conocimientos necesarios para llevar adelante la enseñanza en forma efectiva.

Uno de los principales problemas en la evaluación es la confusión que se produce, en la práctica, entre evaluación formativa y evaluación para la certificación (comúnmente denominada evaluación "sumativa"). Si bien en el discurso pedagógico se insiste en la importancia de la primera, dentro de las aulas predomina la segunda (Ravela, 2009; Picaroni, 2009). Ambas son importantes, pero tienen finalidades y lógicas diferentes.

La evaluación formativa es parte del proceso de aprendizaje. Los alumnos aprenden en la medida en que reciben retroalimentación durante su trabajo. Sin embargo, es necesario precisar qué se entiende por "retroalimentación". Wiggins (1998) distingue tres conceptos diferentes: devolución, orientación y valoración. Normalmente los docentes "valoran" el trabajo de sus alumnos: los felicitan, les dicen si está bien o mal, les asignan una calificación (Wiggins, 1998; Ravela, 2009; Picaroni, 2009). En palabras del propio Wiggins (1998),

*"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" (pp. 46-47).*

La devolución bajo la forma de calificaciones o elogios no es propiamente formativa. La mayor parte del tiempo, los alumnos con bajos resultados no comprenden cabalmente qué es lo que se espera de ellos y por qué obtuvieron una nota baja. Esto es así porque los desempeños a alcanzar y los criterios de valoración no suelen estar explícitos. Pero, además, porque los docentes no suelen establecer verdaderos procesos de devolución dirigidos a que los estudiantes reflexionen sobre su propia tarea, comparen lo que hicieron con lo que se les solicitó y analicen sus logros y carencias, por sí mismos o en conjunto con otros estudiantes.

Para Wiggins (1998)

*"la devolución te dice lo que hiciste y lo que no, y te permite realizar auto-ajustes. Por tanto,*



Foto de Olga Díez



*cuanto más auto-evidente sea la devolución, mayor será la autonomía que desarrollará el estudiante... La devolución es información sobre cómo una persona se desempeñó, a la luz de lo que intentó hacer –intento contra efecto, desempeño real contra desempeño ideal–” (p. 48).*

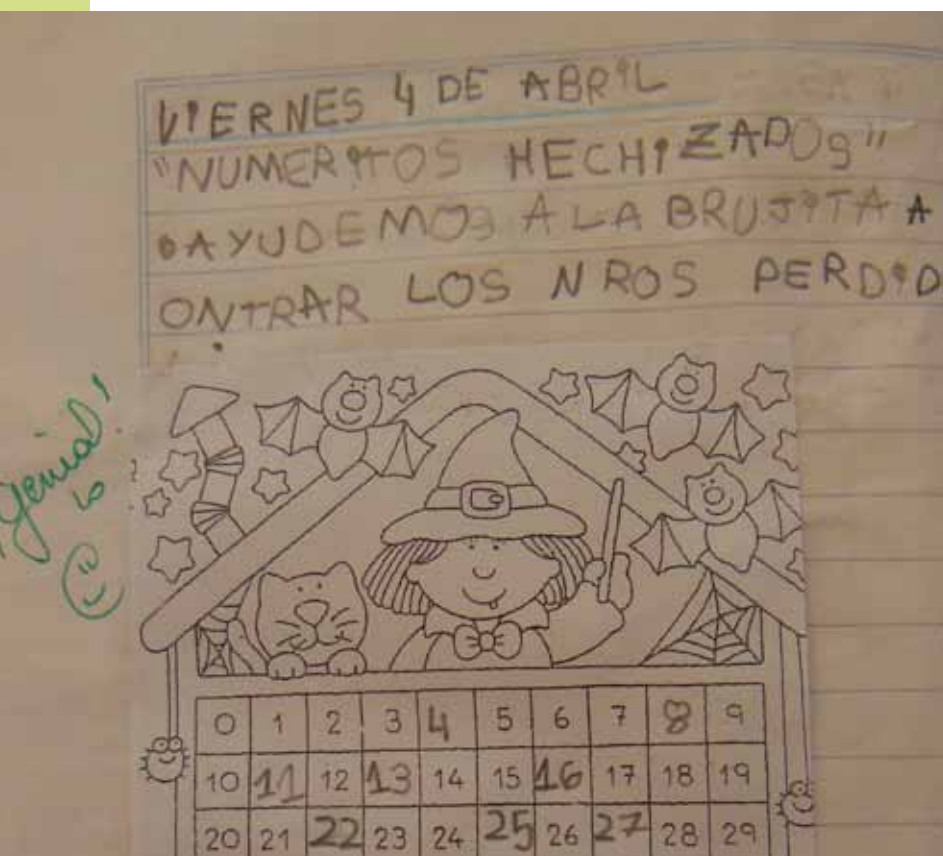
Destaca que devolución y orientación son cosas diferentes y complementarias. La orientación está constituida por recomendaciones y consejos acerca de cómo avanzar. El problema es que muchas veces los docentes ofrecen orientación antes de que los estudiantes hayan comprendido qué se espera de ellos y qué es lo que no han logrado.

*“En general tendemos a dar demasiada orientación y poca devolución. Muchas veces saltamos por encima de la devolución y vamos directamente a dar consejos. Esta orientación a menudo no tiene mucho significado para el estudiante, si no es una respuesta lógica derivada de una devolución” (Wiggins, 1998: 50-51).*

En el estudio realizado en 2008 en primaria (Ravela, 2009; Ravela, 2010b), se constató la ausencia de devoluciones, en el sentido antes definido. Se pudo observar, en cambio, la omnipresencia de la práctica de asignar calificaciones o notas a todas las actividades de los estudiantes, así

como a otros aspectos tales como el aseo, la conducta, la puntualidad y la asistencia. Todos estos ‘números’ finalmente son resumidos en una única nota o promedio final.

La evaluación para la certificación tiene una finalidad y una lógica propias y diferentes a las de la evaluación formativa. Cuando el propósito de la evaluación es calificar a los alumnos, lo importante es desarrollar procedimientos e instrumentos que permitan al docente dar cuenta, en forma pública, del grado en que sus estudiantes están logrando lo que se espera de ellos en el curso. El docente tiene un mandato social para hacerlo, en la medida en que la sociedad, las familias y el resto de las instituciones educativas requieren esta información y confían en el juicio del docente (Ravela, 2009). Para poder cumplir con este mandato en forma



apropiada, el docente necesita:

- a) ser capaz de explicitar expectativas de aprendizaje en forma clara y comprensible para sus alumnos;
- b) definir niveles de desempeño;
- c) construir actividades e instrumentos que aporten evidencia empírica adecuada y suficiente para ubicar a cada estudiante en alguno de dichos niveles.

Foto de Mariana Nóbile

Lo anterior está en la base de un sistema apropiado de calificaciones. Sin embargo, en la práctica cotidiana la evaluación de carácter formativo y la orientada a la certificación, se entremezclan

continuamente. Las expectativas acerca del desempeño de los estudiantes no son explícitas y las calificaciones que se consignan en los boletines son el resultado de un conjunto de procedimientos aritméticos (suma de puntos por ejercicios y promedios de notas por variados conceptos) que dan una apariencia de “objetividad” al proceso de calificar. Pero las calificaciones resultantes son completamente ambiguas y carecen de significado sustantivo (Picaroni, 2009; Ravela, 2009).

### Las tareas que proponen los profesores para evaluar ciencias naturales en educación media

A los efectos del análisis, las **propuestas** fueron clasificadas en dos grandes categorías, según su finalidad (indicada por cada docente): formativas y certificativas. Las primeras son parte del trabajo de aprendizaje e incluyen actividades de diagnóstico, recuperatorias, talleres y prácticas. Las segundas incluyen aquellas cuya finalidad es calificar al estudiante. Incluyen exámenes finales y pruebas mensuales, parciales o semestrales. Las Tablas 3 y 4 permiten apreciar, respectivamente, la distribución de las **propuestas** y de las **tareas** incluidas en ellas, según estas dos grandes categorías. Esta información es simplemente ilustrativa del tipo de materiales en que se basa el análisis. No pretende reflejar la proporción de **propuestas** y **tareas** que utilizan los docentes en cada país, dado que cada docente escogió qué mostrarnos.

Tabla 3 - Distribución de las **PROPUESTAS** según finalidad de la evaluación (en porcentajes)

		PAIS				Total
		CHI	COL	PER	URY	
ASIGNATURA	Biología	5	26	9	45	85
	Ciencias Físicas	0	0	0	15	15
	Ciencias Naturales	48	39	0	0	87
	CTA	0	0	74	0	74
	Física	0	14	0	21	35
	Química	8	17	8	24	57
	Total	61	96	91	105	353
GRADO	6°	0	2	0	0	2
	7°	0	3	4	0	7
	8°	31	19	50	30	130
	9°	30	67	37	75	209
	10°	0	5	0	0	5
	Total	61	96	91	105	353

Tabla 4 - Distribución de las **TAREAS** según finalidad de la evaluación (en porcentajes)

	CHI	COL	PER	URY	Total
Certificativa	66	50	55	72	61
Formativa	34	50	45	28	39
Total	100	100	100	100	100

El hecho de que en todos los países los porcentajes de **tareas** de tipo certificativo sea superior al de **propuestas** devela simplemente que las **propuestas** de evaluación certificativa contienen mayor cantidad de **tareas** que las formativas.

Las **propuestas** de evaluación recogidas incluyen **tareas** de todo tipo. Una primera categorización consistió en distinguir entre:

- i. pruebas compuestas exclusivamente por preguntas de opción múltiple;
- ii. pruebas integradas por preguntas de respuesta “cerrada” (normalmente combinan opción múltiple con preguntas del tipo “verdadero-falso”, apareamiento de términos o búsqueda de palabras en una “sopa de letras”);

- iii. pruebas que solamente incluyen preguntas de respuesta abierta (tanto breve como extensa);
- iv. pruebas que combinan varias de las anteriores.

La Tabla 5 muestra la distribución de las *propuestas* según el tipo de *tareas* que incluyen. En Chile y Colombia se encontró la mayor proporción de propuestas que incluían exclusivamente preguntas de respuesta cerrada: el 43% en Chile y el 49% en Colombia. La mayor proporción de propuestas integradas exclusivamente por preguntas de respuesta abierta se verificó en Uruguay (68%). En este país los docentes entregaron muy pocas *propuestas* que solo incluyeran preguntas cerradas. En Perú la mayor proporción de *propuestas* combinan preguntas de respuesta abierta y cerrada (52%).

Tabla 5 - Distribución de las *PROPUESTAS* según tipo de *tareas* incluidas, por país y finalidad de la evaluación (en porcentajes)

	CHI	COL	PER	URY	Total
<b>Certificativa</b>	82	58	60	82	73
<b>Formativa</b>	18	42	40	18	27
<b>Total</b>	100	100	100	100	100

En todos los países la proporción de *propuestas* con preguntas de respuesta cerrada es mayor en las evaluaciones certificativas que en las formativas. Esto, por un lado, tiene cierta lógica, dado que en las instancias formativas es menor la exigencia de "objetividad" en la corrección de los trabajos y hay más flexibilidad en el uso del tiempo. Sin embargo, cabe notar que esta lógica está vinculada a una concepción de la certificación como asignación de puntajes (ampliaremos este aspecto más adelante).

Simultáneamente, encierra una contradicción: en instancias formativas el alumno realiza un tipo de *tarea* que luego no aparece en la instancia de calificación. Es probable que el mensaje para los estudiantes sea: *'lo importante es aprender a responder las preguntas cerradas, que son las que cuentan'*. Esta 'contradicción' es particularmente notable en Chile, donde las propuestas abiertas y combinadas pasan de 81% en las pruebas formativas a 46% en las certificativas, y en Colombia, donde pasan de 69% a 33%.

Un panorama más preciso se obtiene observando directamente los tipos de respuesta que requerían las *tareas*. Las 4.359 *tareas* fueron codificadas en siete categorías: opción múltiple, completamiento, verdadero/falso, apareamiento de término o frases, respuesta abierta breve, respuesta abierta extensa y otros (ver Tabla 6).

Tabla 6 - Distribución de las *TAREAS* según tipo de respuesta requerida, por país y finalidad de la evaluación (en porcentajes)

TIPO TAREA	CHI			COL			PER			URY			Total		
	Certificativa	Formativa	Total	Certificativa	Formativa	Total	Certificativa	Formativa	Total	Certificativa	Formativa	Total	Certificativa	Formativa	Total
Opción múltiple	79	37	72	70	24	51	22	34	27	1	6	2	44	27	40
Completamiento	4	4	4	5	6	5	19	14	17	4	1	4	7	7	7
Verdadero/Falso	2	6	3	8	1	5	14	6	11	13	5	11	8	5	7
Apareamiento	2	0	2	2	0	1	5	2	4	1	1	1	3	1	2
Abierta breve	9	33	13	9	34	20	34	41	37	44	39	43	24	37	27
Abierta extensa	3	16	5	4	29	15	3	2	3	26	31	27	10	18	12
Otros	0	2	1	1	6	3	2	1	2	11	18	12	4	5	4
<b>Total</b>	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

En Chile el 79% de las *tareas* incluidas en pruebas para certificación son de opción múltiple, a las que se debe agregar un 8% de preguntas cerradas de otros tipos. Las preguntas abiertas breves y extensas constituyen apenas el 12% en las pruebas certificativas, en tanto son el 49% en las pruebas formativas. En Colombia el panorama es similar: 70% de las preguntas para certificación



son de opción múltiple y el total de cerradas alcanza el 85%. En cambio, en las *propuestas* formativas el 63% de las preguntas son abiertas breves o extensas.

En Perú se destaca el hecho de que hay mayor proporción de preguntas de opción múltiple en las pruebas formativas (34%) que en las de certificación (22%). Sin embargo, crecen los otros formatos de preguntas de respuesta cerrada: de 22% en las formativas a 38% en las certificativas. Por otra parte, las preguntas abiertas en Perú son principalmente de respuesta breve y casi no se registraron de respuesta extensa (3%).

En Uruguay las *tareas* de respuesta cerrada pasan de 13% en las *propuestas* formativas a 19% en las certificativas. En este país se registró la mayor proporción de preguntas de respuesta abierta extensa: 31% en pruebas formativas y 26% en pruebas certificativas.

Un segundo aspecto relevante observado en las *tareas* de evaluación, es el tipo y cantidad de información que el alumno debe procesar para responderlas. Según se podrá observar, la mayoría absoluta de las preguntas de evaluación no tienen más información que el texto y los datos de la propia pregunta. La proporción de tareas que requieren que el alumno procese información textual, datos o información gráfica, es mínima.

Para analizar este aspecto las *tareas* fueron codificadas de la siguiente manera:

- i) Se consideró si la consigna incluía algún tipo de información adicional a la pregunta, que el alumno debía leer y comprender. En el caso de las preguntas de opción múltiple no se consideró como información adicional las alternativas de respuesta.
- ii) Un primer tipo de información adicional que las tareas pueden incluir es texto. En estos casos se codificó cada tarea teniendo en cuenta la extensión del texto: hasta 70 palabras, entre 71 y 200, más de 200.
- iii) Un segundo tipo de información codificada fue la presencia de datos numéricos. En este caso se consideró, por un lado, la existencia de datos en la propia consigna o tarea y, por otro lado, la presentación de datos en tablas o gráficos.
- iv) En tercer lugar, se codificó la presencia de infografías: esquemas visuales, ilustraciones, mapas y todo tipo de elemento gráfico.

Los datos se consignan en la Tabla 7. Uruguay y Colombia fueron los países en que se registró la mayor proporción de *tareas* que incluían información adicional (57% y 49% respectivamente). En general se trata de textos breves o de información visual. Este aspecto es importante teniendo en cuenta los requerimientos de pruebas como PISA, que requieren del manejo de textos extensos y abundante información visual y numérica. Cabe aclarar que este comentario no está dirigido a promover la preparación de los estudiantes para dichas pruebas, sino a poner de relieve lo limitado de las demandas y expectativas hacia los estudiantes en nuestros países con relación a la OCDE.

**Tabla 7 - Porcentajes de TAREAS por país que incluyen distintos tipos de información adicional a la pregunta o consigna**

	CHI	COL	PER	URY	Total
La tareas Incluye algún tipo de información además de la pregunta	24	49	25	57	37
La tarea incluye un texto breve (hasta 70 palabras)	4	31	8	29	16
La tarea incluye un texto de extensión media (entre 71 y 200 palabras)	2	1	3	3	2
La tarea incluye un texto extenso (hasta 70 palabras)	2	2	0	2	2
La tareas Incluye datos como parte de la consigna	4	14	16	18	12
La tarea incluye datos bajo la forma de tablas o gráficos	4	5	3	5	4
La tareas Incluye información visual (esquemas, ilustraciones, mapas, etc.)	15	18	11	29	18

Un tercer aspecto codificado fue la existencia de *tareas* que plantearan a los estudiantes algún tipo de ‘*situación*’, entendiendo por tal algo más que información. Una ‘*situación*’ incluye un conjunto de elementos que constituyen un problema a resolver, que puede ser propio de la

vida social, de la vida cotidiana o de la actividad en la disciplina. Es algo más que datos aislados o que un texto disciplinar del cual extraer o inferir información. Se codificaron siete tipos de situaciones:

- i) Sociales: problemas ambientales, de energía, economía, pobreza u otros relacionados con la vida en sociedad;
- ii) Cotidianas: problemas de la vida diaria de los estudiantes, deportes, música, vida en el centro educativo, compras, amigos, familia;
- iii) Escolares: situaciones que no son plausibles en la vida real, si bien incluyen elementos de ella; son intentos típicamente escolares de acercarse a la realidad;
- iv) Disciplinarios-escolares: problemas o ejercicios propios de la disciplina, pero que tampoco son plausibles en la vida real como, por ejemplo, el típico problema de física en que el auto que va siempre a la misma velocidad;
- v) Disciplinarios-reales: problemas o ejercicios propios de la disciplina, plausibles en situaciones de investigación en biología, física o química;
- vi) Experimentales simuladas: problemas que ponen al estudiante ante un experimento narrado y que requieren razonamiento o diseño de tipo experimental;
- vii) Experimental real: propuestas que incluyen la observación y/o registro de datos de la realidad, y/o la manipulación de materiales reales como, por ejemplo, la disección de un encéfalo de res.

No se consideró como situación preguntas tales como "La afinidad electrónica disminuye al bajar en los grupos y aumenta en los períodos al ir de izquierda a derecha" (formato verdadero/falso). Tampoco preguntas como la que se muestra en la Figura 1. La Figura 2 muestra una tarea que incluye una situación disciplinar-escolar. Ambas corresponden a Colombia.

Las Figuras 3 y 4 ilustran las típicas preguntas de opción múltiple, sin información adicional ni situación. Ambas pertenecen a Chile. La Figura 5, también de Chile, ilustra una situación científica simulada en la que el alumno debe emplear un texto relativamente extenso e información numérica. La Figura 6, correspondiente a Uruguay, contiene una situación de carácter social, con un texto breve.

Figura 1

3. La configuración electrónica de un elemento es  $1s^2 2s^2 2p^6 3s^2 3p^4$ , esto nos permite conocer que:

- a. Su número atómico es 16.
- b. Presenta 4 electrones en el último nivel de energía.
- c. Perteneció al período 6.
- d. Se encuentra en el grupo 6.

Figura 2

10. Observa en la siguiente tabla las características de cada uno de los elementos imaginarios y ubícalos dentro de la tabla periódica.

	Elemento A	Elemento B	Elemento C
Estado de oxidación	+7	+3	+4
A	20	50	25
Distribución electrónica	$3s^2 3p^5$	$5s^2 5p^1$	$4s^2 4p^2$

Según lo anterior, el elemento que se encuentra en la tabla periódica dentro del grupo 7 y el período 3 es:

- a. El elemento A.
- b. El elemento B.
- c. El elemento C.
- d. Ninguno de los anteriores.

Figura 3

13. ¿Cuál de las siguientes relaciones es correcta?

- A) Lisosomas - Respiración celular.
- B) Ribosomas - Síntesis de glúcidos.
- C) Membrana plasmática - Permeabilidad selectiva.
- D) Vacuola - Síntesis de proteínas.
- E) Mitocondrias - Transporte de sustancias

Figura 4

14. Lugar dentro de una célula eucarionte animal, donde se sintetiza una mayor cantidad de ATP

A) citoplasma.	D) retículo endoplasmático liso.
B) cloroplasto.	E) aparato de Golgi.
C) mitocondria.	

Figura 5

**III. Lectura científica (9 puntos)**  
 Lee atentamente el siguiente artículo y responde las preguntas.

**¿Cómo se forma un organismo pluricelular?**

Unos investigadores analizaron por microscopio huevos fecundados artificialmente de una especie de erizo. Los huevos fueron puestos en medios que contenían nutrientes y compuestos requeridos para su desarrollo. Al cabo de unas horas observaron la formación de dos células y posteriormente, nuevo aumento en el número de células: 4, 8, 16 32, y así progresivamente. Luego de unos días, se observaron organismos formados por muchísimas células que tenían la apariencia de una larva con capacidad de nadar, alimentarse y responder a diferentes estímulos. Después de 20 días, las larvas se transformaron en pequeños erizos y luego entre los días 25 y 30 se convirtieron en organismos adultos. Los investigadores tomaron muestras de tejidos de los erizos adultos, para observarlos con detención y describir las estructuras. Una de las observaciones fue que había una gran diversidad de células, las que diferían en forma y tamaño. Finalmente, de estos organismos adultos, se extrajeron gametos y se realizaron nuevas fecundaciones. Con los nuevos huevos fecundados se reinició el experimento, obteniendo los mismos resultados que en la generación previa.

De acuerdo a estos antecedentes y basándote en lo que conoces de la teoría celular, responde las siguientes preguntas:

1. Durante las primeras horas se observaron embriones de 2, 4, 8, 16 y 32 células, ¿cómo podrías explicar este aumento en el número de células?(3)  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_
2. Respecto al cambio de estado de huevo a adulto en los erizos, ¿es correcto afirmar que se debe solo por el aumento del número de células? Explica.(3)  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_
3. ¿Qué importancia tiene el hecho de que en el organismo adulto hay una gran diversidad de células? (3)  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_

Figura 6

2. Analiza el siguiente texto:  
*“En un área verde, con gran variedad de especímenes de flora y fauna autóctona, se ha previsto la construcción de un Shopping. El proyecto implica por un lado, la destrucción de las tres cuartas partes de la flora del lugar y del hábitat de las especies animales, pero, por otro lado, es una fuente importante de mano de obra porque se contrataría gran cantidad de obreros del ámbito local, que, dado el reciente cierre de fábricas han quedado sin trabajo”.*

- a. ¿Cuál es el problema que se plantea en el texto?
- b. Desde la perspectiva del desarrollo sustentable, ¿qué posibles soluciones darías al problema? Explica

En la Tabla 8 se muestra la frecuencia con que las *tareas* incluyen algún tipo de situación. Ello ocurre en el 35% de los casos en Uruguay, 27% en Colombia, 11% en Perú y 9% en Chile. Ma-

yoritariamente se trata de situaciones de carácter disciplinar escolar. Es prácticamente nula la presencia de situaciones sociales o reales, lo cual devela el carácter abstracto y poco aplicado de la enseñanza que se brinda a los jóvenes. Esto, presumiblemente, tiene implicancias negativas en cuanto a la motivación de los estudiantes hacia el aprendizaje.

**Tabla 8 - Distribución de las tareas por país según tipo de situación que proponen al estudiante (en porcentajes)**

	CHI	COL	PER	URY	Total
No incluye	91	73	89	65	81
Social	1	0	1	4	1
Cotidiana	2	2	2	9	4
Escolar	0	0	0	1	0
Disciplinar-escolar	2	21	4	14	9
Disciplinar-real	1	2	0	2	1
Experimental simulada	2	1	0	0	1
Experimental real	0	1	3	1	1
Otras	0	0	0	4	1
Total	100	100	100	100	100

En relación a los procesos cognitivos involucrados, en este artículo solo realizaremos un análisis general sobre los niveles de "demanda cognitiva" de las *tareas* de evaluación propuestas a los estudiantes. Para ello construimos un índice muy simple, que consiste en sumar la cantidad de procesos requeridos para resolver una *tarea*, sin contar el más elemental, que era "recordar o evocar un conocimiento, definición, nombre o propiedad". Las *tareas* más demandantes desde el punto de vista cognitivo son aquellas que requieren de la puesta en juego de varios procesos.

Según se puede apreciar en la Tabla 9, la mayor proporción de *tareas* que únicamente requerían recordar o evocar un concepto se registró en Chile (65%), lo que está relacionado con la prevalencia de preguntas de opción múltiple. El país con mayor proporción de *tareas* que requerían de 3 o más procesos para su resolución fue Uruguay (35%), seguido por Chile (22%). Cabe señalar que muchas *tareas* que fueron codificadas como de recordar, en realidad se resuelven por sentido común. La Figura 7, correspondiente a Perú, ilustra esta situación. Resulta obvio que, por ejemplo, "los científicos - 1er paso del método científico" no tiene sentido alguno. Sin embargo, como se puede apreciar, la *tarea* ha sido resuelta, corregida y calificada. Los ejemplos de este tipo son abundantes.

**Tabla 9 - Distribución de las tareas por país según cantidad de procesos cognitivos requeridos para su resolución (en porcentajes)**

	CHI	COL	PER	URY	Total
Sólo recordar	65	42	56	24	48
Dos	13	49	30	41	31
Tres	12	7	11	22	14
Cuatro y más	10	2	3	13	8
Total	100	100	100	100	100

**Figura 7**

I. RELACIONAR LAS COLUMNAS	
a) Ciencia	( C ) ✓
b) Cuerpo	( E ) ✓
c) Observación	( D ) ✓
d) Los científicos	( b ) ✓
e) Elasticidad	( a ) ✓

1º paso del método científico  
Los cuerpos se estiran y vuelven a su estado normal  
Hacen sus investigaciones en los laboratorios  
Es una porción de materia  
Conjunto de conocimientos

El panorama trazado muestra un claro predominio de *tareas* simples, con alta prevalencia de preguntas de respuesta cerrada. Uruguay se diferencia de los demás países, con mayor presencia de *tareas* de respuesta abierta, que incluyen información para analizar y que son cognitivamente más demandantes.

### Las concepciones de los docentes sobre la evaluación

Las ideas y concepciones de los profesores sobre la evaluación fueron exploradas a través de entrevistas en profundidad. En el marco de las mismas se les preguntó cuál era la primera palabra que les venía a la mente al escuchar el término evaluación. La Figura 8 sintetiza las respuestas, mostrando con mayor tamaño las palabras mencionadas con más frecuencia en cada país.

En Chile las respuestas fueron variadas. Hay una fuerte asociación entre evaluación y prueba, calificación y nota-. También se menciona la evaluación como la oportunidad de valorar a las personas, el aprender y el saber. La evaluación es conceptualizada como un proceso que consiste en contrastar lo aprendido con los objetivos fijados, que permite realizar una retroalimentación al docente y al alumno respecto a sus desempeños.

En Colombia la mayoría de los docentes concibe la evaluación como un proceso que permite saber, para cada tema, el conocimiento asimilado por el estudiante, con énfasis en su desempeño y participación. Si bien muchos admiten que lo primero que piensan al hablar de evaluación es en la prueba o examen escrito, hay consenso entre los docentes en cuanto a que “*la evaluación no es solamente un examen*” (COL-7). En el aula el “*trabajo continuo [...] se encamina a formar competencias*”.

Varios contrastan la concepción tradicional de la evaluación –“*medir qué recordaba de lo que me habían enseñado*” (COL-13)– con una concepción actual que “*va más allá*” de ello: “*no solo mirar la parte cognitiva si no también el desempeño que está teniendo*” (COL-14). El foco debería estar puesto en la asimilación del conocimiento: “*estamos viendo que los estudiantes repiten y repiten como loros [...] el estudiante no necesita aprenderse de memoria los temas, sino hay que saber aplicarlos*” (COL-13). Por otro lado, el término desempeño en Colombia incluye, además de lo cognitivo, aspectos actitudinales, personales o sociales.

Figura 8





En Perú los docentes asocian especialmente el término evaluación con la palabra aprendizaje. Es una medición que toma en cuenta los conocimientos, las capacidades y el trabajo de los alumnos, evidenciados tanto en el salón de clase como en el laboratorio. La evaluación sirve para analizar cómo va el proceso de aprendizaje de los alumnos. *"Conocimiento, su aprendizaje, lo que han aprendido, qué conocimientos..."* (PER-18). *"Todo lo que yo les he dado a ellos quiero ver en qué sentido lo han captado, como analizan los conceptos que yo les he dado, en base a eso siento si han captado o no, si me he sentido satisfecha o no, si tengo que esmerarme, cambiar mis prácticas de enseñanza para que ellos aprendan"* (PER-6).

Otro concepto que se repite en el discurso de los docentes peruanos es la evaluación permanente. *"Tenemos una consigna importante en ciencias: la evaluación es permanente"* (PER-15). *"La evaluación es constante, no es un día examen y ahí voy a evaluar. Ahí puedo cuantificar, pero constantemente evalúo las respuestas de los chicos"* (PER-9); *"ahora los exámenes los tomamos en cada clase"* (PER-17).

En Uruguay predomina la identificación de la evaluación con una carga de trabajo, vinculada a la corrección de tareas, lo cual puede estar relacionado con el uso más frecuente de pruebas de respuesta abierta. Por otro lado, aparece claramente la tensión entre la concepción de la evaluación como calificación o medición del alumno y aquella que la considera un proceso. Se asocia la evaluación con una instancia de trabajo escrito u oral que tiene lugar en el contexto del salón de clase y permite conocer y calificar el aprendizaje. *"Escrito es lo que se me viene primero a la mente"* admite un docente (URY-4). Algunos docentes encontraron algún modo elegante de disfrazar lo primero que vino a su mente, pero pronto se sinceraron. Tal es el caso de una docente que, luego de pensar por un instante dijo *"aprendizaje"* pero enseguida admitió que *"lo primero fue medición. Se mide, aunque no es ese el objetivo; pero que aparece, aparece. Me siento psicoanalizada acá"*, URY-8).

## ¿Qué entienden los profesores por evaluación formativa?

Uno de los aspectos relevantes de la evaluación para el aprendizaje es la práctica de la devolución. Wiggins (1998) la define como un dispositivo o procedimiento dirigido a que el estudiante se dé cuenta, por sí mismo, de la distancia entre lo que se esperaba que lograra y lo efectivamente realizado.

El análisis de este aspecto tiene dos limitaciones. Una es propia del estudio: en la medida en que no se observaron clases, no es posible acceder a las prácticas de devolución reales. Otra es propia de las prácticas de evaluación vigentes: en la medida en que la enorme mayoría de las tareas son preguntas simples, de un solo paso y con una sola respuesta correcta, las posibilidades de realizar una devolución reflexiva quedan absolutamente limitadas.

En la codificación de las muestras de trabajos se registró qué tipo de intervención de los docentes había en los trabajos, cuando estos estaban respondidos por los alumnos. La evidencia es contundente en el sentido de que la devolución por escrito se limita a marcas de correcto/incorrecto o a la asignación de puntos y calificaciones. Apenas el 4% del total de las *propuestas* (es decir, considerando tanto las respondidas por alumnos como las no respondidas) de Colombia y el 3% en Uruguay, tenían algún tipo de comentario escrito del profesor. Estos comentarios, por otra parte, son bastante vagos: *¿Por qué no buscó otro video?, revisalo mejor, falta el dibujo, no clasificó cada enlace, falta unir los puntos* (en una gráfica), *completa lo que falta*. Algunos (pocos) comentarios más específicos fueron:

*Bien, pero no aclaras cómo cuidar la piel, lee la consigna.*

*Sí pero ¿cómo? Qué hay en la piel que nos permite percibir los estímulos como el cambio de temperatura o la vibración, etc... no es esta la razón, tiene que ver con la parte a), la piel puede ser gruesa pero muy sensible.*

*Para la próxima: explica de manera completa y detallada la información sobre cada aspecto. Tú no tienes dificultades pero debes hacer rendir más el tiempo de clase.*



## La devolución como análisis de errores

En el transcurso de las entrevistas en profundidad se consultó específicamente a los docentes acerca del tipo de devolución que dan a sus alumnos con el fin de ayudarles a aprender de lo que hicieron y no se queden solamente con la nota o calificación. Muy pocos docentes expresaron una orientación clara hacia la revisión reflexiva de los trabajos por parte de los estudiantes. Las devoluciones suelen enfocarse en la revisión de errores.

En Chile las devoluciones consisten en la revisión de las pruebas de múltiple opción: *“Revisando la prueba completa, pregunta por pregunta. ¿Por qué esa es la alternativa correcta y por qué las otras estaban incorrectas? Vemos en qué se equivocaron”* (CHI-4). *“Todas las evaluaciones que aplico tienen que tener retroalimentación, análisis de qué preguntas estuvieron débiles y en cuáles se concentró mayor cantidad de errores. Tengo que tomar el contenido de esas preguntas y en clase trabajarlo [...] no sirve de nada aplicar prueba si solo retroalimentas con nota e individual. Aplicar y ver por qué fue el problema y por qué cayeron en un distractor”* (CHI-11).

En Colombia la mayoría de los docentes trabaja con los errores más frecuentes del grupo. *“Si yo veo que en la prueba les fue mal yo digo bueno, qué les pasó, en qué punto pecaron”* (COL-12). *“Se socializa para ver el error, para que esa falla se vuelva un aprendizaje”* (COL-16). *“Yo inicio, leemos la pregunta, les digo la respuesta y miramos en que se equivocó”* (COL-13).

En Perú algunos docentes expresan no hacer ninguna devolución específica, más allá de la entrega de los trabajos corregidos. *“Yo solo entrego, nada más... Hago cada mes un examen escrito, lo necesito como registro, como nota, aunque todo es nota; las exposiciones, los mapas, todo es nota”* (PER-14). *“Se les entrega a los alumnos y ellos ven sus errores, dónde se han confundido, en qué están mal”* (PER-18); *“Les voy revisando y sellando”* (PER-12). Otros docentes relatan resoluciones de las tareas en forma colectiva, con registro en el pizarrón, para ser copiado en el cuaderno. En la mayoría de los casos, éstas se realizan únicamente cuando la mitad de la clase no logró resultados aceptables. *“Corrijo, saco el promedio y empezamos a ver sobre todo las preguntas que no han sido resueltas por la mayoría”* (PER-1). *“Se entrega la prueba, la analizan. Yo les pido que en primer lugar cuenten su puntaje, que observen la prueba por todos lados, las respuestas, etc. Luego cuando yo veo más del 50% desaprobado hago un solucionario de la prueba, en su cuaderno de trabajo ponen solución de la prueba y empezamos a desarrollar todos juntos. Se dan cuenta en qué han fallado y cómo fallaron”* (PER-4).

En Uruguay, como son más frecuentes las preguntas de respuesta abierta, la revisión de los trabajos ofrece más posibilidades. *“La clase en que se hace la corrección del escrito se comenta cómo debería haberse resuelto y se recurre a las mismas respuestas que hicieron los compañeros, a veces elijo a quienes contestaron mejor para que cuenten a los demás compañeros cómo podrían haber contestado. Quizás escuchando un compañero que hace las cosas lo entienden mejor que si yo lo digo directamente, les llega de otra manera”* (URY-16). *“Sugiero que vean el escrito o, más bien, el informe de alguien a quien le fue bien...”* (URY-11). *“Busco un trabajo término medio, una respuesta*



Foto de Mariana Nóbile

que no es la mejor ni es la peor; busco la contestación media, 'lee tú lo que pusiste'. Más que nada hago lectura, ya no más pizarrón" (URY-8).

Las percepciones de los estudiantes confirman lo anterior. En el cuestionario fueron incluidas varias afirmaciones relativas a la devolución por parte de sus profesores. Ante cada afirmación los estudiantes debían responder con qué frecuencia ello ocurría en sus clases de ciencias. Los resultados, consignados en la Tabla 10, muestran que marcar errores y 'cosas que están mal' es la experiencia más común de los estudiantes en todos los países. En cambio, la menos común es recibir comentarios escritos. En Chile y Uruguay, los estudiantes tienen menos oportunidades para rehacer sus trabajos.

**Tabla 10 - Estudiantes que informan que frecuentemente o casi siempre sus profesores hacen lo indicado en la frase cuando les devuelven una evaluación escrita (porcentajes)**

	CHI	COL	PER	URY
1. Solamente te pone la nota en el trabajo	59,7	62,6	65,3	50,2
2. Te marca los errores y cosas que están mal	86,1	86,2	84,7	85,6
3. Te escribe un comentario explicando la diferencia entre lo que hiciste y lo que se esperaba que hicieras.	39,4	35,3	46,1	36,7
4. Te dice lo que tienes que mejorar	55,2	59,2	71,6	49,8
5. Te dice que debes esforzarte y estudiar más	64,0	66,7	72,5	61,7
6. Te da la oportunidad de rehacer o mejorar tu trabajo	28,6	49,9	52,4	28,8

El conjunto de evidencia presentado confirma la existencia de una relación entre el tipo de tareas que se propone a los alumnos y las posibilidades de poner en marcha procesos de evaluación formativa realmente vinculados con el aprendizaje. En otras palabras, la posibilidad de llevar adelante evaluaciones formativas depende fuertemente del tipo de tareas que se proponga a los alumnos. En otros términos, la 'evaluación para el aprendizaje' depende de qué se defina como aprendizaje. En tanto las tareas de evaluación sean predominantemente breves, de respuesta cerrada, carentes de situaciones relevantes y de baja demanda cognitiva, inevitablemente las devoluciones estarán centradas en la revisión de errores concretos, dado que no hay en juego aprendizajes más complejos sobre los cuales reflexionar.

### La devolución como estímulo positivo para los alumnos

Una segunda manera de concebir la devolución aparece fuertemente en Perú y Colombia, donde la mayoría de los docentes enfatiza la importancia de dar a los alumnos valoraciones positivas y evitar las negativas.

En Perú los docentes expresan que la devolución siempre debe ser positiva. Más aún, no se debe entregar devoluciones negativas a los alumnos cuyo desempeño resultó insuficiente, dado que no les ayuda a mejorar. "Felicitó a los alumnos que han sacado nota sobresaliente, por ejemplo 16. A los otros no les critico pero solo felicito a los que lo han hecho bien" (PER-13). "A los que están bien sí los felicito, que aplaudan y todo. A los que no, otra vez será" (PER-17). "Se aprende con el tiempo a tener mucho cuidado con lo que le decimos a los estudiantes. Si queremos reforzar una buena actitud tenemos que felicitar... Si hizo la tarea mal y yo le digo 'pensé que la ibas a hacer mejor', lo estoy desanimando. Mejor decirle, 'tu tarea está muy bien, estoy seguro que la puedes hacer mejor la próxima clase'... Recurrimos al reforzamiento emocional" (PER-5) (el subrayado es nuestro).

En Colombia se nota esta misma preocupación por no desmotivar a los estudiantes. "Cuando terminan de exponer se les da sus aplausos y entonces ya yo les doy las sugerencias. Les digo: 'bueno en los carteles el título podría ser de esta forma' o 'para la próxima puedes mejorar esto, mira el margen' o que no le coloque tanta letra y 'colóquele más imágenes para que no canse'" (COL-14).

La preocupación por dar estímulos positivos y el cuidado de la autoestima de los estudiantes ha sido analizado por Wiggins (1998), quien señala que el estímulo positivo es importante ('te

*mantiene en el juego*) pero insuficiente: si el alumno no recibe devolución específica en relación a sus errores y carencias, difícilmente pueda mejorar su desempeño.

### El concepto de evaluación formativa

Con el fin de profundizar la temática, durante las entrevistas se preguntó expresamente a los docentes acerca de sus ideas sobre la evaluación formativa. A partir de sus respuestas es posible identificar tres conceptos principales.

Para muchos docentes la evaluación formativa es aquella que se enfoca en el proceso por el cual el estudiante avanza en sus aprendizajes. Esta definición tiene dos variantes vinculadas a la conceptualización del término 'proceso'. Algunos docentes ponen el foco en los procesos cognitivos de los estudiantes –*"estoy permanentemente preguntando qué entendiste por esto, no con ánimo de caerle con el mazo sino de ver si entendió, y si no entendió de explicárselo o que otro se lo explique"* (URY-6)–, en tanto que para otros el término refiere al avance de los alumnos hacia lo que se les pedirá en la evaluación sumativa –*"es un tipo de evaluación para saber ellos... qué les falta para repasar antes de la prueba"* (CHI-7)–.

Una definición complementaria, que aparece con fuerza en Chile, es que la evaluación formativa es la evaluación sin calificación. *"Es una evaluación que no tiene calificación y que ayuda a los niños y a mí para ir viendo el avance que tienen con respecto a un contenido"* (CHI-3).

Una tercera acepción para el término es la relativa a la formación integral y actitudinal. Evaluación formativa es aquella que se enfoca en los hábitos y actitudes de los estudiantes. Esta visión está presente fuerte y principalmente en Perú y Colombia. En Colombia esta perspectiva fue prácticamente unánime. Consultados sobre la evaluación formativa, la mayoría de los entrevistados se refirió a la evaluación de aspectos actitudinales del alumno, en contraposición al conocimiento de la disciplina. *"Su actitud en clase: su posición, cómo se debe sentar o caminar, y las normas sobre todo de convivencia... Eso lo evaluó formativamente: cuando se presenta un libro cómo lo tratan, cómo lo debe de cuidar, los pupitres, el mobiliario del salón. Por ejemplo, dejar el salón arreglado como lo encontramos, que las basuras no me las arrojen, que no se tiren objetos..."* (COL-3). *"En cuanto a lo formativo está el cuaderno, porque les ayuda a llevar un orden, a que sean niñas organizadas, la estética, a que no hagan las cosas por hacerlas, se maneja mucho la parte de valores, la responsabilidad, por ahí le veo la parte formativa"* (COL-14).

Una concepción similar de evaluación formativa predomina entre los docentes peruanos. Lo formativo aparece más ligado a la formación integral del estudiante que al fomento de competencias para aprender. *"La evaluación formativa se ve en la puntualidad en entrega del trabajo, el orden en esos trabajos, la presentación también dentro del aula [...] su orden y limpieza... ahí va evaluación formativa..."* (PER-10). *"La presentación del educando en la parte formativa, la forma como se sienta, la práctica de valores, la parte de la formación [...] También el cuidado, la higiene de los estudiantes... La higiene es muy importante"* (PER-3).

En resumen, la conceptualización de la evaluación formativa tiene semejanzas y matices en los distintos países. En Chile y Uruguay está vinculada al proceso de aprendizaje de los estudiantes. En el primer caso, con una orientación más marcada hacia lo que se espera de los estudiantes en las pruebas de calificación. En el segundo caso, más orientada hacia la discusión de los procesos de aprendizaje y hacia el registro de aspectos "calificables" del desempeño de los estudiantes. En Colombia y Perú la evaluación formativa es asociada por los docentes con la formación de hábitos y actitudes en los estudiantes.

### ¿Qué se valora al calificar a los estudiantes?

La asignación de calificaciones a los estudiantes es una actividad cotidiana en todos los países. A diario los docentes asignan notas o toman registros de las actividades de los alumnos, que luego serán utilizados para calificarlos.

Según se explicó en el apartado conceptual, la calificación tiene dos significados principales. Por un lado, es la expresión de una función social propia del trabajo docente: dar cuenta en forma pública del grado en que los estudiantes han aprendido lo que se esperaba de ellos. Desde esta perspectiva, se supone que la calificación es una expresión más o menos 'objetiva' de lo aprendido por los estudiantes.

Por otro lado, la calificación es una forma de exigencia y motivación extrínseca, para que los estudiantes realicen su 'trabajo' de aprender: dado que necesitan obtener una calificación aceptable para aprobar el curso, la misma se convierte en un instrumento que el docente utiliza para exigir esfuerzo y estudio, comportamiento apropiado en el aula y realización de tareas<sup>11</sup>.

Esta función está vinculada a una tensión propia de la condición humana: los estudiantes, especialmente en la adolescencia, actúan tanto por motivaciones internas -el gusto e interés por aprender- como por exigencias externas -la necesidad de aprobar los cursos-. Esto convierte a la calificación en un arma de doble filo, dado que puede transformarse en el objetivo central de los alumnos, desplazando al aprendizaje. Esto ocurre cuando aquella puede lograrse por otros caminos como, por ejemplo, demostrando buen comportamiento en clase. También puede desplazar al aprendizaje reflexivo, cuando las pruebas se enfocan excesivamente en saberes que hay que recordar y/o enunciar.

La construcción de la calificación es compleja y encierra múltiples dilemas para los profesores. ¿Qué y cómo se califica? ¿La calificación debe vincularse exclusivamente a los aprendizajes logrados y demostrados en pruebas? ¿Se debe calificar otros aspectos como el esfuerzo, la participación o las actitudes de los estudiantes? ¿Cómo resolver si un alumno aprueba o reprueba un curso? ¿Qué hacer con los alumnos que están en el límite de lo aceptable? Una postura demasiado exigente por parte del docente puede llevar al estudiante al fracaso y a la deserción, en tanto que una postura demasiado benigna puede llevar a que los estudiantes avancen en el sistema educativo sin haber logrado los aprendizajes esperados. Estas tensiones son más difíciles de resolver en la educación media que en la primaria, porque los profesores tienen a su cargo varios cientos de alumnos y pocas horas de clase a la semana con cada grupo.

### Criterios para asignar las calificaciones

Con el fin de caracterizar los criterios que los profesores priorizan en cada país a la hora de calificar a los estudiantes, en el marco de la entrevista en profundidad se les entregó una tabla con ocho afirmaciones y la siguiente consigna: "*Por favor, marque con una cruz en cada línea, según sean sus prácticas habituales para calificar a los alumnos*". Las afirmaciones fueron las siguientes:

1. Me baso principalmente en el resultado de pruebas escritas.
2. Hago un promedio de todas las notas de cada alumno por distintas tareas y conceptos y me atengo a dicho resultado.
3. Tengo en cuenta el esfuerzo realizado por el estudiante durante el año y le mejoro la nota a quienes se han esforzado.
4. Tengo en cuenta la participación del estudiante en las clases y le mejoro la nota a quienes participan regularmente.
5. Tengo en cuenta la realización de trabajos domiciliarios y otro tipo de proyectos o actividades extra clase.
6. Tengo en cuenta cómo le fue a cada estudiante en comparación con los otros.
7. Tengo en cuenta el comportamiento y le bajo la nota a los estudiantes que molestan en clase.
8. Asigno tareas complementarias a los estudiantes que están en el límite de la suficiencia, para que puedan subir su nota.

Para cada una de estas afirmaciones los entrevistados debían marcar una de las siguientes opciones: '*casi siempre*', '*frecuentemente*', '*a veces*', '*nunca*'. En la tabla 11 se presenta el promedio para cada afirmación y país, a partir de asignar los códigos 4, 3, 2 y 1, respectivamente, a las respuestas mencionadas. Promedios cercanos a 1 significan que mayoritariamente los docentes no tienen en cuenta la práctica mencionada en la afirmación en cuestión, en tanto que promedios de 3 o superiores significan que lo hacen frecuentemente o casi siempre.

El panorama que ofrece la Tabla 11 es muy elocuente para caracterizar las prácticas comunes y las diferencias entre países. Si se pone la atención en los dos aspectos más priorizados en cada país, se observa que en Chile son las pruebas y los promedios de notas, en Colombia la participación en clase y los trabajos extra clase, en Perú los promedios de notas y el esfuerzo y, en Uruguay, los trabajos extra clase y el esfuerzo.

**Tabla 11 - Valoración promedio por país de distintos elementos utilizados por los profesores para definir las calificaciones finales de los estudiantes (escala de 1 a 4)**

PAÍS	Pruebas escritas	Promedio de notas	Esfuerzo	Participación en clase	Trabajos y proyectos extra clase	Comparación entre alumnos	Comportamiento en el aula	Tareas extra para alumnos insuficientes
CHI	3,00	3,42	2,33	2,33	2,67	2,83	1,50	1,83
COL	2,33	3,04	3,50	3,75	3,58	2,54	1,79	3,29
PER	2,61	3,67	3,50	3,22	3,06	2,94	2,00	2,78
URY	1,95	2,23	3,32	3,05	3,45	1,18	2,45	2,59

En Uruguay aparece como llamativamente baja la importancia otorgada a las pruebas escritas. En Chile, en términos comparativos con los demás países, es más baja la importancia dada a la participación en clase y al esfuerzo. Colombia y Perú se caracterizan por tener la mayor cantidad de aspectos a los que se otorga alta importancia: 5 de los 8 ítemes en el primero y 4 en el segundo, tienen promedios superiores a 3. En todos los países la comparación entre alumnos y el comportamiento en el aula son los aspectos que tienen menos importancia en la calificación.

Las percepciones de los estudiantes son consistentes con lo anterior. Uruguay aparece como el país en que menos importancia se otorga a las pruebas escritas en la determinación de las calificaciones. Lo más valorado sería el promedio por distintas tareas, la participación en clase y la realización de trabajos extra. En Chile es dónde más se valora el resultado en la prueba escrita. En Colombia y Perú el énfasis principal está puesto en el promedio por distintas tareas y conceptos, junto con el comportamiento. Es interesante notar que, en la percepción de los estudiantes, los docentes califican el comportamiento y la disciplina más de lo que ellos mismos reconocen.

**Tabla 12 - Percepciones de los estudiantes acerca de qué tienen en cuenta los profesores al asignar calificaciones (porcentajes de respuestas frecuentemente y casi siempre por país)**

	CHI	COL	PER	URY
1. Se basan principalmente en el resultado de pruebas escritas	71,8	73,1	74,7	55,0
2. Hacen un promedio de todas las notas que tuviste en el período por distintas tareas y conceptos	66,1	90,6	89,7	85,0
3. Tienen en cuenta el esfuerzo que has realizado	53,6	66,0	72,2	69,6
4. Tienen en cuenta tu participación en las clases	52,8	77,4	80,2	83,7
5. Tienen en cuenta si realizaste de trabajos domiciliarios o actividades extra clase	56,5	55,9	65,4	83,0
6. Tienen en cuenta tu comportamiento (conducta, disciplina) en clase	64,2	89,9	90,0	86,0

### Evaluación integral y actitudes: “Yo les califico hasta los respiros”

Los docentes entrevistados en Colombia asignan gran importancia a lo personal y social a la hora de calificar a los estudiantes, en detrimento de lo cognitivo. Esto es conceptualizado como ‘evaluación integral’: “Es posible que el estudiante no sea el mejor a nivel cognitivo pero si es un estudiante que lo ves atento, que es responsable y que es puntual también tiene un peso [...] Esto nos va a determinar que el estudiante va a ser buena persona en un futuro y que se va a enfrentar a la vida diaria que es tan dura [...] ahí es donde está la evaluación formativa integral” (COL-13). “Hago mucho énfasis en lo personal y social porque son factores importantes para el buen desarrollo de la



clase. Si no ponen cuidado, disciplina y orden, tampoco [lo harán] en la toma de conocimiento" recalca otra docente, que asigna 50% de la calificación a estos aspectos (COL-3).

"La parte humana, forma de expresarse del alumno, una persona respetuosa, disciplinada, pulcra. Todo eso para mí es de evaluación. Para mí es más importante la persona, el ser humano que el 5 que sacan en un examen" (COL-11).

"Ellas conmigo tienen ocho, nueve o diez notas en un período, porque yo les califico hasta los respiros" (COL-21). "Hago un promedio de todas las notas: disposición, memoria, redacción, habilidad para cortar, para pegar, evaluación escrita, participación, actitud, llegada temprano, comportamiento, trabajo en grupo. Para mí es un trabajo dispendioso porque hay alumnos que llegan a tener en un periodo 13 notas" (COL-17).

Algo similar se verifica entre los docentes peruanos, que enfatizan el concepto de evaluación permanente, concebida como calificación de toda actividad realizada y toda actitud demostrada por el alumno, su minucioso registro, sumatoria y promedio matemático al final de período. "Todo es evaluable. Y la evaluación es constante [...] constantemente estoy observando. Todo no se anota, es bien difícil, lo ideal sería que se pudiese anotar todo, pero hay una ficha que me ayuda a anotar algunas cosas." (PER-9). "La calificación es permanente. La prueba objetiva es un insumo más. Por ejemplo, en el laboratorio los criterios y materiales que trabajan, el trabajo activo entre ellos, eso ya también es acumulación. Entonces el alumno así tiene notas" (PER-3).

Varios docentes mencionan criterios de calificación que tienen que ver con la presentación estética de los trabajos, la vestimenta, el uso de colores o el pegado de materiales en el cuaderno. "Un dibujo a veces lo llevan a la casa, lo traen más arregladito y obtiene mejor nota que los trabajos de clase" (PER-14). "Uno observa, por decir algo, el color azul. Bueno sería que no tuviera azul sino lapicero negro y empiezan a colorear de negro la sustancia que habíamos observado y tengo que corregir [...] en las ciencias exactas... tienes que pintar del color con que lo mides" (PER-1).

En ambos países los docentes expresan que los estudiantes pueden aprobar el curso antes de llegar a la instancia de evaluación formal de aprendizajes: "mucha gente cuando llega el examen del período, ya la materia la tiene ganada, por todas las notas que tienen de todas las evaluaciones, si yo le dijera a usted que sólo evaluó con el examen sería otra cosa" (COL-7). "A veces hasta no hacemos evaluación bimestral, solo evaluación de prácticas si al alumno le va bien" (PER-5).

### **El esfuerzo: "el que hizo el trabajo de hormiga, pasa"**

Entre los profesores uruguayos se observó un énfasis especial en la valoración del esfuerzo y la auto-superación individual, más que en el logro de metas preestablecidas. Se valora especialmente "el proceso que tuvieron en el año, cómo comenzaron, qué logros se tuvo al final" (URY-2). "Hay que hacer un seguimiento del estudiante del comienzo del año al final, no pretendiendo que todos lleguen al mismo punto, pero sí que haya habido una evolución desde marzo a noviembre" (URY-17). "No me guió por el que le fue mejor, sino el cómo le fue a cada uno en relación a antes, yo creo que la mejor medida es uno mismo" (URY-12). Se valora "el esfuerzo que hacen por tratar de llegar a un mínimo aceptable" (URY-14). "Tiene que haber tenido un trabajo continuo... El que hizo el trabajo hormiga, que se mantuvo todo el año, con altibajos pero todo el año trabajó, ese pasa" (URY-1). "El grado de cumplimiento de tareas es evaluado, más allá de si está bien o mal la tarea, si intentó hacer algo, la información, si hizo una búsqueda" (URY-16).

Como consecuencia de este énfasis, los docentes uruguayos otorgan una importancia menor a las pruebas. "Los escritos son un porcentaje importante, pero no son la nota final ni el 50%, sobre todo considerando el nivel de estrés que genera, el escrito no refleja siempre el trabajo cotidiano y las herramientas que adquirió" (URY-15).

También en Chile varios docentes expresaron tener en cuenta el esfuerzo a la hora de calificar. "Tengo un tic [marca] de quien participó en clase o no. Si tengo alguien que nunca participó en clase ni hizo actividades y después tiene mala calificación, es una cosa de que no está poniendo su esfuer-



zo" (CHI-6). Este aspecto es tenido en cuenta para que los alumnos puedan realizar tareas adicionales para mejorar su calificación. *"Vienen y te dicen que tienen mala nota, si podrían hacer un trabajo o podría yo remediar. Conversas con ellos y transas, con algunos puedes hacer excepciones y con otros no, depende del desempeño que tengan en la asignatura y, sobre todo, del esfuerzo que ponen. Hay niños que les cuesta mucho cognitivamente pero se esfuerzan un montón, entonces tú a ellos también los ayudas"* (CHI-3).

### La calificación como motivación: "la nota es como la plata"

A lo largo de las entrevistas aparecieron, en todos los países, múltiples referencias al uso de la calificación como instrumento de disciplinamiento y de motivación extrínseca para los estudiantes.

En Chile una docente reconoce la utilidad de la evaluación como herramienta de presión *"que obliga a las niñas a estudiar en determinados momentos, porque si no, no estudian, se quedarían con lo de la clase nomás"* (CHI-9). Otro docente explica cómo utiliza 'décimas' de puntos de calificación para motivar a los estudiantes: *"en todas las clases estaba preguntando para que se ganaran ellos sus décimas. Los chicos funcionan mucho con estímulo positivo de notas. Ojalá no fuera así pero en la casa si te portas bien tienes permiso para el cine y acá décimas de nota [...] tiene que haber una presión, algo escrito para que ellos se motiven a hacer algo..."* (CHI-12).

Los alumnos sólo toman en serio aquello que se califica: *"uno lo quiere hacer pero si esto no tiene notas, se bajan al tiro y hacen los trabajos a medias. Es como por plata, la nota es la plata. Y todo es para lograr algo, nada se trabaja porque sí"* (CHI-8).

En Colombia los entrevistados estuvieron de acuerdo en que el aprendizaje es el fin último de la evaluación, pero la mayoría reconoció, además, la utilidad de emplear la evaluación y el miedo que esta produce en los estudiantes, como herramienta para lograr buena conducta, responsabilidad y compromiso con el estudio. Una docente lo resume en estos términos: *"eso hace que ellos trabajen y se afanen, porque saben que se van a ganar una buena nota"* (COL-22). Otra docente expresa: *"Cuando a un estudiante lo evalúan se compromete más. Hace un tiempo el gobierno decretó que los estudiantes no podían perder el año y ellas habían perdido el interés, entonces todo les daba lo mismo."* (COL-14). *"Dicen que la evaluación no es la parte numérica pero a ellas le gusta que las estimulen de esa manera, dando una nota"* (COL-15).

### La construcción de la calificación: promedios vs. Conceptos

¿Cómo se llega a asignar una calificación en un boletín y a la decisión consiguiente respecto a la aprobación o reprobación de un curso por un estudiante? En este punto es posible identificar dos posturas extremas, una en la que predomina la "objetividad" matemática de los promedios y otra en la que predomina la percepción o "concepto" que el docente se forma con relación a cada estudiante.

#### El promedio matemático

Entre los docentes chilenos predomina la 'objetividad' y el uso de la tecnología para promediar las diferentes notas: *"aquí una pasa al computador las notas que uno lleva, el computador hace el promedio y una no hace nada"* (CHI-8). *"En el software que trabajamos, ingresas todas las notas de todas las asignaturas y te entrega el informe por niño"* (CHI-3). La mayoría de los docentes entrevistados parecen estrictos respecto a la objetividad, que vinculan con la igualdad de oportunidades: *"si la nota de aprobación es 4 y su nota es 3.9, esa queda como 3.9 y no como 4 porque todos han tenido los mismos trabajos anteriores. Entonces dar una posibilidad para que suba me parece injusto para el resto, tendría una posibilidad que los demás no tuvieron antes"* (CHI-10).

En Colombia la cuestión de los promedios está fuertemente establecida, con ponderaciones específicas para la amplia diversidad de aspectos que son evaluados y que fueron presentados antes.

"En lo que se llama valoración académica, que es la parte de conocimientos en sí [...] todo eso tiene un valor del 30%. La valoración actitudinal vale también el 30%, el examen parcial que vale 20% y el examen final que vale 20%. Son muchas cosas que se le evalúa al niño, ellos no pueden decir 'ah yo perdí porque nunca me recogió el cuaderno'; uno tiene muchas maneras de evaluar a estos niños" (COL-22).

"Por período hay una evaluación parcial y una final, eso tiene un valor del 50% de la nota final. El otro 50% está repartido, un 65% que es la parte cognitiva, hablamos de los trabajos, actividades, talleres, todo lo que se trabaja en la clase, laboratorios; y el 35% es la parte afectivo-espiritual que maneja el colegio [...] son los porcentajes que maneja la institución y eso no se puede cambiar" (COL-4). "Hay 15 y hasta 18 notas por período [bimestre]. Porque se manejan muchas tareítas de calificación y en eso va acumulando el niño notas" (COL-18).

En la Figura 6 se puede observar una planilla de registro de calificaciones de Colombia. Las columnas numeradas del 1 al 9 en la parte superior, corresponden a las siguientes dimensiones: 1) evaluación escrita; 2) participación en clase; 3) disposición actitudinal; 4) asistencia; 5) sustentación salida al tablero; 6) disciplina; 7) trabajo en equipo; 8) coevaluación; 9) trabajo en clase.

Figura 9

También en Perú los comportamientos son minuciosamente registrados y calificados. "Hay una ficha o lista de cotejo por alumno, dentro de los criterios para esta actitud ante el área y los valores. Por ejemplo 'levanta la mano para opinar' y tiene sus puntajes. Si el alumno recoge un papelito del suelo, se tiene en cuenta un tema transversal, la conciencia ambiental. Si participan en aula, en lluvia de ideas, también se le pone check. Al final tienen un valor. 5 puntos: siempre participa; 3 puntos: a veces participa; o 0: nunca participa" (PER-8).

"Criterio 1, los conocimientos: salen de las tres pruebas trimestrales y mensuales, y la prueba de recuperación... ese es el promedio del que sale el criterio 1. Criterio 2, trabajos de investigación o indagación, guías de prácticas, proyectos y las fichas prácticas del aula. Criterio 3, las actitudes: el cuaderno, el comportamiento, las pruebas orales y la asistencia. Así sale el promedio final, del promedio de los tres criterios" (PER-13).

### El 'concepto' del profesor: "Yo pienso que le tengo que poner tanto"

En Uruguay la calificación no se construye como promedio matemático sino de una manera mucho más subjetiva. "Tomás en cuenta todo para la calificación final, pero no es un promedio. Uno no le pone al esfuerzo una nota, hay cosas que no podés hacer un promedio porque no son un número. Esfuerzo es un concepto del estudiante, más allá de los contenidos. Incluye todo lo que es el estudiante, no es un número, no podés hacer promedio con cosas que son subjetivas. Las tenés que evaluar y no entra en el promedio matemático" (URY-2).

"Lo hago subjetivamente porque además depende para quien: un 6 para vos no es lo mismo que un 6 para mí. Hay un grado de aceptabilidad en función de características de cada uno: no puedo hacer pesar más el escrito a alguien que tiene problemas para expresarse por escrito, sino que le hago pesar más lo oral de todos los días. Es más, lo hago al revés, no sé si esto es correcto o no, tengo la visión global de 'yo pienso que le tengo que poner tanto', y cambio la nota matemática en el oral para que después le de eso y no me joroben con lo matemático. Si le va bien en el escrito y no hace nada, le amarreo en el oral para que haga, su aceptabilidad está en un nivel diferente" (URY-9).

*"En ciertas situaciones límite también lo conversamos con los coordinadores viendo el panorama general del alumno en el resto de asignaturas. Si puede terminar condicionado a repetir o no el curso, vemos qué es lo más conveniente. Si estoy en la duda y evaluamos que es mejor que no pierda el año, capaz le doy aprobado el curso para que rinda examen de las demás materias" (URY-16).*

También en Colombia algunos docentes pusieron en tela de juicio la utilidad de los promedios y expresaron que priorizan su propio conocimiento y opinión sobre cada alumno: *"El sistema me exige evaluaciones y notas, aunque yo soy muy de que a veces evalúo al ser humano sin tener en cuenta las planillas, sé quién pasa y quién no, a quien dejamos un tiempo o ponemos a que se esfuerce más" (COL-11).* *"A veces lo hago y a veces no lo hago. Yo no necesito llevar una planilla con tantos números... yo te puedo decir a quien le puedo poner un 5 y a quién le puedo poner un 2 sin que ellos me respondan" (COL-17).*

En resumen y en forma esquemática, en Chile las calificaciones están fuertemente enfocadas en los resultados de pruebas escritas y se construyen a través de promedios. En Uruguay lo más característico es la valoración del esfuerzo y de trabajos extra. Se llega a la calificación principalmente a través de un juicio valorativo global del profesor (según emerge, además, del marco normativo). Colombia y Perú se destacan por la importancia otorgada a los aspectos actitudinales, conceptualizados como evaluación integral o permanente. Esto incluye la asignación de puntos por todo tipo de aspectos, incluida la prolijidad en los trabajos y conductas tales como levantar un papel del suelo. La utilización de las calificaciones con fines de disciplinamiento existe. En la visión de los profesores este uso es limitado. En la percepción de los estudiantes, en cambio, es alto en todos los países, salvo en Chile.

## Desafíos para la formación y el desarrollo profesional de los docentes

El trabajo realizado y la información presentada constituyen una primera aproximación descriptiva al objeto de estudio. Esta descripción es relevante dado el carácter cerrado que tienen las actividades que se realizan en el aula.

La descripción realizada permitió identificar, a la vez, problemáticas comunes y modos distintos de realizar la evaluación en los distintos países. En Chile y Uruguay se enfatiza más la evaluación de saberes, en tanto que en Colombia y Perú es notorio el énfasis en aspectos actitudinales y de hábitos de orden, prolijidad e higiene, así como la preocupación por la autoestima de los alumnos, con su correspondiente riesgo, el de dejarlos en la ignorancia de sus propias limitaciones. En Chile y Colombia, seguidos por Perú, es notorio el uso de pruebas de opción múltiple, en tanto que en Uruguay predominan las tareas de respuesta construida. Los docentes uruguayos otorgan, además, especial importancia al esfuerzo de los estudiantes. Asignan las calificaciones a partir de un 'concepto' que se forman de cada estudiante, mientras que en los demás países se prioriza el promedio matemático de notas asignadas por una amplia variedad de aspectos.

En todos los países las tareas de evaluación requieren de los estudiantes, principalmente, que recuerden o enuncien contenidos. En Uruguay, de la mano de las preguntas de respuesta abierta, aparecen tareas que requieren procesos cognitivos un poco más complejos. La falta de complejidad cognitiva de las propuestas de evaluación, así como la ausencia de información para analizar o situaciones reales para resolver, determina que exista poco espacio para la realización de devoluciones propiamente formativas, en el sentido de propiciar la reflexión de los alumnos sobre sus procesos de aprendizaje. La devolución en las aulas suele enfocarse en el análisis de los errores en las respuestas a las preguntas.

La calificación es un elemento omnipresente. Si bien los docentes la experimentan como una obligación administrativa, al mismo tiempo la utilizan como una forma eficaz de motivar y "formar". La tensión entre el peso relativo a asignar a lo cognitivo y a lo actitudinal, así como a los resultados tangibles y al esfuerzo, están presentes en forma notoria en todos los países.

Escapa al alcance de este estudio el análisis de las razones que están detrás de estas diferencias. Algunas explicaciones son bastante obvias como, por ejemplo, la influencia de SIMCE en Chile y de las pruebas SABER en Colombia, en el tipo de pruebas que los profesores utilizan dentro del aula.

Otras explicaciones seguramente tienen que ver con la historia de la formación docente en cada país.

Más allá de las diferencias, hay un conjunto de desafíos comunes, en mayor o menor medida, para las prácticas de evaluación en todos los países. Algunas claves que deberían ser tenidas en cuenta para mejorar las habilidades de evaluación de los futuros profesores son las siguientes:

Un primer aspecto clave es comprender el carácter cultural de las prácticas de enseñanza y de evaluación. A evaluar de manera diferente no se aprende en forma teórica, sino práctica, comenzando por la propia experiencia de ser evaluado. Los docentes han aprendido qué es la evaluación desde los seis años de edad, como alumnos. Por ello, una primera cuestión clave es modificar el modo en que los futuros docentes son evaluados durante el período de su formación. No tendría mayor efecto 'enseñarles' a evaluar de un modo diferente y continuar evaluándolos a ellos con tareas memorísticas, promedios de notas sin un significado claro y sin instancias en las que puedan experimentar lo que es la devolución.

Experimentar la devolución es una cuestión clave para los docentes. Experimentarla como evaluados, antes que como evaluadores. Algo de esto seguramente ocurre en las instancias de práctica profesional durante la formación docente, aunque es probable que reciban más valoraciones y orientaciones, que devoluciones propiamente dichas (en el sentido que Wiggins da a estos términos). Es probable que la mayoría de los docentes, tanto durante su formación como durante el ejercicio profesional, reciban juicios de valor sobre lo bien o mal que han hecho las cosas, así como indicaciones de lo que deberán cambiar. Probablemente son pocos los que han tenido oportunidad de reflexionar por sí mismos acerca de sus propios modos de dar clase y de evaluar a sus alumnos.

De la misma manera, las instancias de discusión de las tareas que se propone a los estudiantes y de los procesos cognitivos que requiere su resolución, es otra vía privilegiada para generar un laboratorio de experimentación y cambio. De allí la importancia de construir bases de registros como el de este estudio. La discusión y análisis de pruebas y preguntas de evaluación debería ser parte central de los programas de formación docente.

Los procedimientos y criterios para calificar a los estudiantes también deberían ser objeto de discusión y análisis cuidadoso durante la formación docente. Para evitar que las calificaciones sean un 'número mágico', resultado de un promedio matemático supuestamente 'objetivo', o del 'conocimiento directo' que el docente tiene de sus alumnos, es necesario establecer claramente qué cosas serán tenidas en cuenta, qué peso tendrá cada una, y, probablemente reportar por separado, cuestiones actitudinales y cuestiones cognitivas. Hace falta mucho trabajo de explicitación de modos de proceder y de construcción de herramientas que permitan expresar qué es lo que se espera de los estudiantes en los distintos ámbitos de su desempeño. Es necesario también analizar las ventajas y los problemas del uso de la calificación como herramienta de motivación y disciplinamiento.

Mejorar las evaluaciones y la enseñanza requiere tiempo. Una condición de posibilidad para cualquier cambio es avanzar hacia una nueva concepción del trabajo docente, que no se limita a 'dar' clases. Los países de América Latina deberían avanzar hacia modelos de contratación por jornada completa, con permanencia del docente en la institución, tanto para dar sus clases como para realizar todo el conjunto de actividades adicionales vinculadas a la docencia.

Más allá de la formación docente, es preciso señalar que la amplia mayoría de los profesores entrevistados son conscientes de buena parte de las debilidades que tienen sus propias evaluaciones y de las tensiones que implica la armonización de diversos criterios importantes a la hora de calificar a los estudiantes.

En el marco de la entrevista se les preguntó por sus dificultades con la evaluación, y los profesores destacaron las siguientes:

La falta de hábitos de estudio y trabajo en los estudiantes. "Un poco difícil porque a veces los alumnos no estudian, no repasan y aparte que se les explica no toman interés, aunque a veces sí tienen buenas calificaciones. Y ellos mismos reconocen cuando salen mal 'ay que yo no estudié', que verdaderamente no han estudiado". (PER-18).

Las limitaciones de los estudiantes para expresarse por escrito. "Ellos no son de estar mucho tiempo tranquilos, entonces tienes que hacer pruebas cortas porque muy largas los pierdes a los diez minutos" (CHI-3). "Las propuestas escritas son cada vez más difíciles a nivel de tercero. Hay muchos alumnos que tienen tolerancias diversas, porque son hiperactivos, no se pueden enfrentar a un escrito entero sino de a partes, pregunta a pregunta, para poderse concentrar. Los alumnos de tercero son cada vez más dependientes del profesor, hacen preguntas, acabás casi que haciendo escrito con ellos" (URY-3).

La complejidad y tiempo que requiere la elaboración de una buena prueba, que tome en cuenta todos los contenidos de manera balanceada y que se adapte a las diferentes habilidades de los alumnos. Muchos docentes son conscientes de la dificultad para evaluar habilidades o pensamiento crítico, y de que sus propuestas suelen limitarse a los contenidos memorísticos. "Me resulta difícil porque tenemos que pasar demasiado contenido en poco tiempo, entonces uno deja de lado la evaluación de habilidades y se concentra mucho en el contenido" (CHI-4). "Si queremos elevar la calidad de la evaluación, eso implica dedicar tiempo suficiente para elaborar el instrumento, pensarlo bien, no hacer instrumentos que sean muy simples, por decirlo así... toma su tiempo preparar los instrumentos" (PER-9).

El tiempo que requiere la corrección de trabajos y pruebas. "Una carga la cantidad, el tiempo que te lleva corregir 90 escritos cada vez que los pones. Tengo 90 alumnos, y no es nada al lado de otros colegas con 60 horas de clase. Tengo siempre mi casa llena de cosas para corregir, tapada de informes de escritos, pruebas, tareas. Cansa, se vuelve tedioso, se vuelve una carga, te baja el nivel [...] Me da mucho trabajo y tengo que corregir mucho para eso" (URY-11). "Una carga cuando son muchas evaluaciones que tú tienes que hacer... ¿En qué momento reviso, saco la nota y los promedios? El problema es que me lleva todo el fin de semana" (CHI-7). "Tengo 11 grupos con un promedio de 40 alumnos por grupo, es decir que estoy administrando 440 o 450 alumnos. Si me pongo todos los días a hacer evaluaciones individuales con los muchachos, en la clase siguiente no he tenido todavía el tiempo suficiente para haber revisado esas 450 evaluaciones" (COL-24).

La frustración por los malos resultados, porque la evaluación es el momento en que los profesores constatan que sus alumnos no aprenden lo que ellos querían. "Los resultados poco satisfactorios que tienen los estudiantes, eso genera en uno como cierto sentimiento de tristeza, de melancolía que lo pone a pensar a uno qué es lo que está pasando aquí" (COL-1). "La evaluación a mí me deja mucho sinsabor porque veo que mi trabajo no tiene el fruto que yo quiero encontrar y hay mucha gente que pierde" (COL-23). "Con colegas hablamos mucho y hacemos terapias grupales por el fracaso permanente [...] hacés todo esfuerzo de plantear actividades diferentes, que atiendan a las capacidades diferentes que pueden tener ellos, pensás y pensás [...] y de repente el resultado no es el mejor. Genera un sentimiento de frustración, de 'nado, nado y no llego a la orilla'" (URY-1).

Estos aspectos deberían ser contemplados en un programa de formación en servicio orientado al desarrollo de habilidades para la evaluación. Debe tenerse en cuenta, asimismo, que la enorme mayoría de los profesores, en todos los países, manifestó que en general trabaja solo en el diseño y corrección de sus evaluaciones.

A modo de cierre del capítulo transcribimos el testimonio de una joven profesora de Uruguay, con tan solo tres años de experiencia, que realiza una fuerte crítica a la formación docente, así como a las prácticas de enseñanza y evaluación de aprendizajes que los docentes utilizan, a las que identifica con la adhesión a un paradigma obsoleto:

*"Venimos con una herencia de años de otro paradigma. ... Hemos heredado ese paradigma y nos arraigamos a ese modelo... que es el más fácil y cómodo, que simplemente implica que si al alumno le va mal es porque no estudia, no entiende, viene mal de la escuela, y no porque estamos haciendo algo mal. Los cambios son difíciles... La formación docente ahí entra en un rol primordial, si eso no se lo mostramos en formación docente, siguen saliendo docentes con el mismo modelo y prácticas de hace 45 años. ¿Cómo se nos pasa eso no puede ser! ... Sé que hay casos [excepcionales], pero en los hechos de la vida cotidiana de las asignaturas en el aula no ha cambiado nada de manera significativa" (URY-13).*

## Bibliografía

- Administración Nacional de Educación Pública (2008). Plan Nacional Integrado de Formación Docente - 2008. Acceso web el 31 de julio de 2011.  
[http://www.cfe.edu.uy/index.php?option=com\\_content&view=article&id=116&Itemid=91](http://www.cfe.edu.uy/index.php?option=com_content&view=article&id=116&Itemid=91)
- LOUREIRO, G. (2009), "Evaluación en el aula, currículo y evaluaciones externas", en *GTEE-PREAL*, descargado el 28/09/2009 de: <http://www.grade.org.pe/gteepreal/evaluacion/>
- National Center for Education Statistics (2003), *Teaching Mathematics in Seven Countries. Results from the TIMSS Video Study*, Washington D.C.: Autor.
- National Center for Education Statistics (2006), *Teaching Science in Five Countries. Results from the TIMSS 1999 Video Study*, Washington D.C.: Autor.
- Organización para la Cooperación y el Desarrollo Económico (2006), *PISA 2006. Marco de la evaluación. Conocimientos y habilidades en Ciencias, Matemática y Lectura*, Madrid, OCDE – Santillana.
- Organización para la Cooperación y el Desarrollo Económico (2009), *Creating Effective Teaching and Learning Environments: First Results from TALIS*, París, Autor.
- Organización para la Cooperación y el Desarrollo Económico (2010), *PISA 2009 Results: What Makes a School Successful? Resources, Policies and Practices. Volume IV*, París, Autor.
- PICARONI, B. (2009), *La evaluación en las aulas de primaria: usos formativos, calificaciones y comunicación con los padres*, en *GTEE-PREAL*, descargado el 28/09/2009 de:  
<http://www.grade.org.pe/gteepreal/evaluacion/>
- RAVELA, P. (2009), "Consignas, devoluciones y calificaciones: los problemas de la evaluación en las aulas de educación primaria en América Latina", en *Páginas de Educación*, v. 2, Montevideo, Universidad Católica del Uruguay, pp. 49-89.
- RAVELA, P. (2010a), *¿Qué pueden aportar las evaluaciones estandarizadas a la evaluación en el aula?*, en *Documento de Trabajo N° 47*, Santiago de Chile, Programa de Promoción de la Reforma Educativa en América Latina (PREAL).
- RAVELA, P. (2010b), "La evaluación de aprendizajes en las aulas de primaria en América Latina. Enfoques y prácticas", en C. WAINERMAN & M. DE VIRGILIO (comps.), *El Quehacer de la Investigación en Educación*, Buenos Aires, Manantial - Universidad de San Andrés.
- STIGLER, J. & HIEBERT, J. (2009). *The Teaching Gap*. New York: Free Press.
- UNESCO (2007). *Enhancing learning. From access to success. Report of the first experts' meeting: Defining areas of action*. París: autor.
- WIGGINS, G. (1998). *Educative Assessment. Designing Assessments to Inform and Improve Student Performance*. San Francisco: Jossey-Bass.

## Notas

- 1 El estudio se realizó en el marco y gracias al apoyo del Fondo de Investigaciones Educativas (FIE) de PREAL.
- 2 Proyecto conjunto IEE-UCU y OEI (Oficina de Buenos Aires); "Prácticas de Enseñanza en las Aulas de Primaria en América Latina. Banco de Registro Digitales". Incluyó la filmación de clases en aulas de Primaria en Santiago de Chile, Córdoba, México y Cali. Se encuentra en su fase final y dará lugar a un Informe de Investigación y a un sitio WEB con video clips de clases de Lenguaje y Matemática.
- 3 *Impactos del Plan CEIBAL en las Prácticas de Enseñanza en las Aulas de Primaria* (2012). Incluyó la observación de 200 períodos de clase de 100 maestras (2 por maestra), en Uruguay. Informe disponible en el sitio web de IEE-UCU (<http://www.ucu.edu.uy/Default.aspx?tabid=1164>).
- 4 Ref. Estudio Evaluación de Aprendizajes en las aulas de América Latina.
- 5 Ref. documentos, sitio web y revista Páginas.
- 6 Un informe completo y más detallado sobre la metodología y hallazgos del estudio será publicada en formato digital en el sitio web del IEE-UCU.



- <sup>7</sup> En rigor el estudio se desarrolló en cuatro ciudades (Santiago, Cali, Lima y Montevideo).
- <sup>8</sup> Agradecemos especialmente la colaboración de la Fundación Chile en Santiago; del Grupo de Investigación sobre Evaluación de la Calidad de la Educación -GIECE- de la Universidad Sanbuenaventura en Cali; y de la Unidad de Medición de la Calidad -UMC- del Ministerio de Educación de Perú.
- <sup>9</sup> En el estudio de video de TIMSS, a pesar de que la muestra original aleatoria incluía 100 profesores de Matemática, se concluyó que con alrededor de 50 era más que suficiente para caracterizar la enseñanza en un país, dada la redundancia de las prácticas observadas.
- <sup>10</sup> En rigor, debemos decir que podemos caracterizar las prácticas predominantes en los establecimientos urbanos de las ciudades estudiadas. Por cierto, las prácticas pueden variar en zonas rurales o en regiones particularmente diferentes.
- <sup>11</sup> Esta función presupone que a todos los estudiantes les importa aprobar sus cursos. Esto es algo que podía darse por supuesto hace algunas décadas, pero que no necesariamente ocurre en el presente.

### Resumen

El estudio realizado se propuso analizar las prácticas de evaluación de los profesores de ciencias naturales de educación secundaria básica en Chile, Colombia, Perú y Uruguay. La evidencia empírica recogida muestra que las tareas que los estudiantes deben responder cuando son evaluados por sus profesores requieren, principalmente, que recuerden o enuncien conceptos, o que apliquen fórmulas para resolver ejercicios. Las devoluciones de carácter formativo suelen estar enfocadas en el análisis de los errores en las respuestas. Como éstas son poco complejas, también lo es el tipo de reflexión que puede darse en estos análisis. En cuanto a la asignación de calificaciones, predominan dos modalidades: el cálculo de un promedio matemático a partir de notas asignadas por distintos aspectos y la asignación subjetiva de una calificación por parte del profesor, a partir del 'concepto' que se ha formado sobre el estudiante. Los aspectos tenidos en cuenta en cualquiera de estas modalidades incluyen el esfuerzo del estudiante, su participación en clase y la realización de tareas extra escolares. La calificación es utilizada con frecuencia como un instrumento para motivar y exigir esfuerzo y dedicación a los estudiantes. La tarea de evaluar es percibida como una pesada carga por parte de los docentes, en especial por el tiempo que requiere, tanto la preparación de propuestas como su corrección. Los docentes son conscientes de que no logran plasmar sus objetivos educativos -formar estudiantes reflexivos- en sus propuestas de evaluación. Pero no encuentran los espacios y tiempos para el desarrollo de nuevas propuestas. En todos los países la mayoría de los profesores prepara sus evaluaciones en forma individual, con escasa interacción con otros colegas.

### Palabras clave

Evaluación Educativa - Prácticas Docentes - Nivel Secundario - Ciencias Naturales - Pruebas

### Abstract

*This study aims to analyze the evaluation practices of the natural science teachers in basic secondary education in Chile, Colombia, Peru and Uruguay. The empirical evidence gathered shows the tasks students need to complete when they are evaluated by their teachers mainly require them to remember or state concepts, as well as to apply formulas to solve exercises. Feedback often focuses on analyzing the mistakes in the responses given. As the tasks are of low complexity, so are the considerations that can be derived from analyzing student responses. Two grading methods prevail: (1) the calculation of the mathematical average of all the grades given in that period to the student, and (2) the subjective grade assignment- based on the concept the teacher has of the student. Both methods take into account the student's efforts, in-class participation and extra-credit homework completion. Grades are often used to foster student motivation, effort and dedication. Teachers perceive evaluation as a heavy burden, especially given the time required to both prepare and correct the completed tasks. Teachers are aware of not being able to capture their educational objectives -forming thoughtful students- in their evaluation proposals but they do not have the space and time to develop new and more suitable proposals. Most teachers in all the participating countries prepare their evaluations individually. Interaction with colleagues is very limited.*

### Key words

Educational Evaluation - Practice Teaching - Secondary Level - Natural Sciences - Tests