COLECCIÓN EDUCACIÓN UC

El poder de la evaluación en el aula

Mejores decisiones para promover aprendizajes

CARLA E. FÖRSTER (editora)

EDICIONES UC

El poder de la evaluación en el aula

Mejores decisiones para promover aprendizajes

CARLA E. FÖRSTER (editora)

EDICIONES UC

EDICIONES UNIVERSIDAD CATÓLICA DE CHILE
Vicerrectoría de Comunicaciones
Av. Libertador Bernardo O'Higgins 390, Santiago, Chile

editorialedicionesuc@uc.cl
www.ediciones.uc.cl

EL PODER DE LA EVALUACIÓN EN EL AULA
Mejores decisiones para promover aprendizajes
Carla E. Förster (editora)

© Inscripción Nº 285.251
Derechos reservados
Diciembre 2017
ISBN Edición Impresa N° 978-956-14-2170-7
ISBN Edición Digital N° 978-956-14-2244-5

Corrección de estilo: Tiarella Moreira Muñoz
Diseño: Francisca Galilea

Diagramación digital: ebooks Patagonia
www.ebookspatagonia.com
info@ebookspatagonia.com

CIP-Pontificia Universidad Católica de Chile
El poder de la evaluación en el aula : mejores decisiones para promover aprendizajes / Carla E. Förster, editora.
Incluye notas bibliográficas.
1. Evaluación educacional.
2. Evaluación en educación.
3. Calidad de la educación.
I. Förster Marín, Carla E., editora.
2017 371.26 + DDC23 RDA

Índice

PRESENTACIÓN
Lorena Medina Morales, decana Facultad de Educación UC

AUTORES

Capítulo 1
ALFABETIZACIÓN EVALUATIVA DE LOS DOCENTES
Carla E. Förster

Capítulo 2
APRENDIZAJE Y EVALUACIÓN: LO QUE NO SE EVALÚA, NO SE APRENDE
Carla E. Förster y Cristian A. Rojas-Barahona

Capítulo 3
PLANIFICANDO INTEGRADAMENTE LA ENSEÑANZA Y LA EVALUACIÓN
Sandra Zepeda Aguirre y Carla E. Förster

Capítulo 4
EL FIN JUSTIFICA LOS MEDIOS: INTENCIONALIDADES DE LA EVALUACIÓN
Sandra Zepeda Aguirre

Capítulo 5
LA RETROALIMENTACIÓN EFECTIVA Y SU POTENCIAL PARA MEJORAR EL APRENDIZAJE
Sandra Zepeda Aguirre

Capítulo 6
LOS ESTUDIANTES COMO AGENTES DE EVALUACIÓN

Carla E. Förster

Capítulo 7
INSTRUMENTOS PARA LA EVALUACIÓN DE APRENDIZAJES, ¿CON QUÉ EVALUAR?
Carla E. Förster, Sandra Zepeda Aguirre y Claudio Núñez Vega

Capítulo 8
LA EVALUACIÓN CON ORGANIZADORES GRÁFICOS: CÓMO Y CUÁNDO UTILIZARLOS
Paola Marchant-Araya

Capítulo 9
CRITERIOS DE CALIDAD PARA LA CONSTRUCCIÓN DE INSTRUMENTOS DE EVALUACIÓN
Carla E. Förster y Cristian A. Rojas-Barahona

Capítulo 10
UNA METODOLOGÍA DE ANÁLISIS DE CASOS PARA ORIENTAR LA TOMA DE DECISIONES EN EL CONTEXTO ESCOLAR
Paola Marchant-Araya y Carla E. Förster

Presentación

La Colección Educación UC es una contribución de la Facultad de Educación de la Pontificia Universidad Católica de Chile y Ediciones UC a la labor cotidiana de los profesores en ejercicio, que nace del convencimiento sobre la necesidad de contar con textos que orienten el trabajo práctico de los educadores en terreno. Cada uno de estos títulos es fruto de investigación interdisciplinaria actualizada y de la puesta en práctica de las propuestas concretas que ellos ofrecen.

La evaluación es sin duda uno de los ámbitos más relevantes y necesarios en los que la formación de profesores, tanto inicial como continua, parece seguir en deuda. Se discute mucho acerca de grandes evaluaciones, estandarizadas, censales o muestrales, pero se orienta poco la evaluación en el aula vista como parte de los procesos de enseñanza y aprendizaje. Por ello, este libro se centra en relevar el rol crucial que tiene la retroalimentación en el proceso de aprendizaje, las consideraciones necesarias para los diferentes formatos en los que se presenta, así como en el involucramiento de los estudiantes en la comprensión de lo que ella implica, con miras a reducir la brecha entre su estado actual y la meta a la que se espera que lleguen.

Se presenta aquí un libro pensado como apoyo a los profesores de aula y equipos directivos, con estrategias e ideas concretas para implementar. En él se plantea una actualización de lo que se entiende por evaluación en el contexto actual, lo que en inglés se denomina *assessment literacy* y que supone un docente alfabetizado en evaluación, poniendo el foco en la complementariedad de la evaluación formativa y sumativa de los aprendizajes. Al mismo tiempo, asigna a los estudiantes el rol de agentes evaluativos y no

solo el de participantes pasivos, lo que implica darles responsabilidades en su propio proceso de aprendizaje y en el de sus pares a través de la evaluación.

Esta propuesta pone énfasis en las consecuencias que las prácticas evaluativas que los profesores realizan tienen en las trayectorias escolares y universitarias de los estudiantes, pues es clave que los docentes sean conscientes de esto y sepan resguardar los principios éticos asociados a la evaluación, los criterios de validez y de confiabilidad de las situaciones evaluativas que aplican.

Por último, este texto aporta con una herramienta metodológica de análisis de casos que permite a los docentes y equipos directivos abordar necesidades o casos críticos que pueda tener un establecimiento y trabajar los antecedentes de forma sistemática para levantar una propuesta de mejora.

A través de esta colección, la UC realiza un nuevo aporte a la comunidad escolar, como parte de su compromiso público con la educación del país, que esperamos los docentes puedan aprovechar para potenciar estas temáticas y sus prácticas pedagógicas en las salas de clases.

Lorena Medina Morales, decana Facultad de Educación UC.

Autores

Editora

CARLA E. FÖRSTER

Bióloga marina. Magíster en Evaluación Educacional. Doctora en Ciencias de la Educación, Pontificia Universidad Católica de Chile.

Profesora asociada de la Facultad de Educación UC. Especialista en evaluación para el aprendizaje, metacognición y formación inicial docente.

ceforste@uc.cl

Colaboradores

PAOLA MARCHANT-ARAYA

Asistente social. Magíster en Psicología, mención Psicología Educacional. Doctora en Ciencias de la Educación, Pontificia Universidad Católica de Chile.

Profesora asistente de la Escuela de Trabajo Social, Facultad de Ciencias Sociales UC. Especialista en diseño y evaluación curricular y en formación en Educación Superior.

mpmarcha@uc.cl

CLAUDIO NÚÑEZ VEGA

Profesor de Ciencias Naturales y Biología, Pontificia Universidad Católica de Chile. Doctor en Filosofía y Ciencias de la Educación de la Universidad Complutense de Madrid, España.

Profesor asociado, Facultad de Educación UC. Especialista en evaluación de aprendizajes, evaluación de la formación inicial de docentes y formación práctica de profesores.

cnunezv@uc.cl

CRISTIAN A. ROJAS-BARAHONA

Psicólogo. Doctor en Psicología, Universidad de Granada, España. Posdoctorado en el Developmental Brain Behaviour Laboratory, Academic Unit of Psychology, University of Southampton, UK.

Profesor asociado, Departamento de Aprendizaje y Desarrollo, Facultad de Educación UC. Especialista en funciones ejecutivas y desarrollo cognitivo.

crojash@uc.cl

SANDRA ZEPEDA AGUIRRE

Asistente social. Magíster en Ciencias de la Educación con mención en Evaluación Pontificia Universidad Católica de Chile. Doctora (c) en Estudios Americanos (IDEA-Usach).

Docente de la Facultad de Educación UC en programas de formación de pre y posgrado. Especialista en desarrollo curricular y evaluación para el aprendizaje en educación escolar y superior.

sczepeda@uc.cl

CAPÍTULO 1

Alfabetización evaluativa de los docentes*

Carla E. Förster

INTRODUCCIÓN

Existe consenso sobre que el qué y cómo se evalúa condiciona los aprendizajes de los estudiantes y las decisiones pedagógicas que los profesores toman (Brown, 2011; Himmel, 2003; Salinas, 2002; Shepard, 2006; Xu & Brown, 2016, entre otros). Para ningún profesor es novedad las preguntas clásicas que hacen los estudiantes frente a una evaluación escrita: ¿profesora, qué entra?, ¿cómo va a ser la prueba?, ¿es de desarrollo o de alternativas? Estas preguntas conducen al estudiante en la forma de aproximarse a la evaluación: según la cantidad de "materia que entra" y "el tipo de preguntas" será la estrategia que usarán para estudiar. Ellos saben que si la prueba es de selección múltiple, entonces deben estudiar sin aprender los conceptos de memoria porque la respuesta está en una de las opciones, pero si es de desarrollo tienen que saber algunas palabras claves de memoria para poder explicar o argumentar en sus respuestas, y comprender mejor el contenido, aunque hay mayor flexibilidad en la forma de responder; en palabras de los estudiantes: "hay que saber más cuando la prueba es de desarrollo", "pero se puede chamullar un poco y salvar".

También sabemos que cuando una tarea evaluativa implica habilidades más complejas e integradas, como la elaboración de un proyecto, una investigación o un portafolio, si son bien llevados como un proceso evaluativo, generan aprendizajes más profundos en los estudiantes (Wiggins & McTighe, 2005), pero requieren de mayor dominio evaluativo del docente para llevarlos a cabo, ya que debe retroalimentar el proceso (con poco tiempo y mucho contenido que abordar) y elaborar rúbricas para revisar esos trabajos.

Por otra parte, el sistema escolar está constantemente pidiendo al docente que gestione los resultados de sus evaluaciones, y que tome decisiones respecto de la efectividad de las prácticas de enseñanza desde las evidencias de aprendizaje de sus estudiantes, pero el análisis siempre se realiza a partir de las calificaciones obtenidas y no

desde el avance o logro en los aprendizajes esperados. Es habitual ver en los colegios prácticas asociadas al porcentaje de estudiantes que tienen una nota inferior a 4,0 en una evaluación, situación en la que se espera que el docente aplique una acción remedial, que puede ir desde repetir dicha evaluación hasta hacer una nueva y promediar las notas para quienes tuvieron “un rojo”, o conversar con los estudiantes y reforzar el aprendizaje menos logrado sin modificar la calificación.

La literatura especializada señala que los profesores que tienen un conocimiento inadecuado en evaluación a nivel del aula o en medición para la rendición de cuentas al sistema tienen a su vez menor eficacia en sus prácticas de enseñanza, lo que redunda en aprendizajes de menor calidad en sus estudiantes (Popham, 2009; Rogers & Swanson, 2006; Stiggins, 2004). A estas competencias en evaluación se les ha denominado internacionalmente *assessment literacy* o alfabetización evaluativa.

ALFABETIZACIÓN EN EVALUACIÓN

La alfabetización en evaluación se ha definido como “la comprensión de un docente respecto de los conceptos fundamentales y procedimientos de evaluación, los que probablemente considerará en las decisiones pedagógicas que tome” (Popham, 2011, p. 267). El concepto de “alfabetización” se adoptó desde la idea de que una persona “alfabetizada”, en este caso un docente, tiene conocimientos sobre una temática (la evaluación educacional), los comprende y posee las habilidades requeridas para ponerlos en funcionamiento efectivo en su quehacer diario (Ray & Margaret, 2003). Es por esto que la alfabetización en evaluación implica un docente que es capaz de construir evaluaciones confiables y luego gestionarlas y calificarlas para facilitar decisiones válidas respecto de su enseñanza, coherentes con los estándares educativos a los que el colegio adscribe (DeLuca, LaPointe-McEwan & Luhanga, 2016; Popham, 2004, 2013;
Stiggins, 2002, 2004).

Para comprender mejor por qué hoy es tan importante un profesor “alfabetizado” en evaluación es necesario entender la evolución histórica de la concepción de evaluación de los docentes. Deneen y Brown (2016) identifican tres etapas temporales: la primera, anterior a los años 90, muestra que los profesores veían la evaluación en aula desde una lógica influenciada fuertemente por las teorías y prácticas de la medición educacional y psicológica. La formación de los docentes hacía énfasis en la evaluación de los aprendizajes y en cómo generar instrumentos con altos estándares de calidad que permitieran discriminar entre los estudiantes que “sabían la materia” y los que no, pero les entregaban escasas herramientas para que realizaran evaluación formativa y dieran a sus estudiantes una retroalimentación de calidad. La segunda etapa se reconoce entre los ’90 y principios del siglo XXI, se caracteriza porque los profesores se centran en la evaluación para el aprendizaje y se fortalece el balance entre la evaluación formativa y la evaluación sumativa; no obstante, los docentes demuestran una predisposición negativa hacia la medición, produciendo un quiebre entre la formación docente y los requerimientos del sistema relacionados con la rendición de cuentas y las pruebas estandarizadas. La tercera etapa, que representa la actualidad, muestra que hoy día la concepción de evaluación que tienen los docentes rescata las prioridades de las otras dos etapas, esto es, entenderla como evaluación del y para el aprendizaje y que, por tanto, se requiere dominio de la evaluación formativa, la retroalimentación efectiva y la evaluación sumativa como componentes claves en el aula para monitorear y certificar aprendizajes, pero también hacer uso y estar dispuestos a trabajar con las pruebas estandarizadas, negociando así las tensiones y demandas en términos evaluativos, que son cada vez más complejas (figura 1).

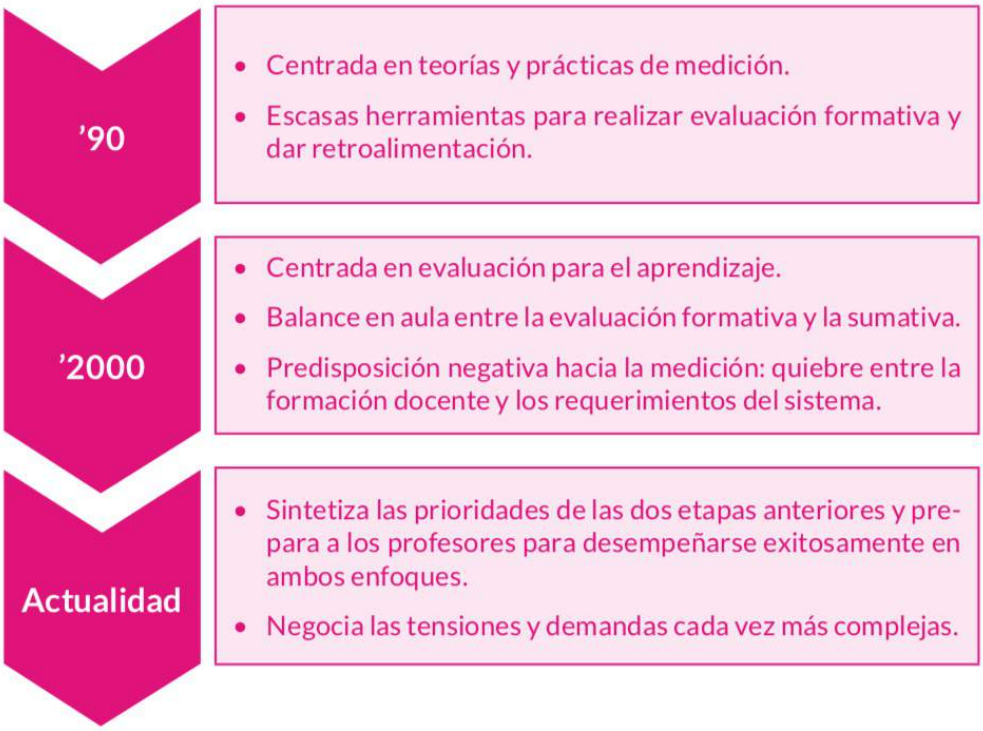


FIGURA 1: Evolución de la concepción de evaluación en aula en la formación de profesores

Un docente con un nivel de alfabetización en evaluación adecuado requiere dominio en distintos ámbitos para que su práctica sea eficiente: a) usar múltiples evaluaciones de alta calidad, alineadas con sus objetivos de aprendizaje, los cuales deben estar definidos con precisión; b) interpretar el desempeño de los estudiantes a la luz de formas específicas de evaluación e hipotetizar sobre errores comunes propios de la disciplina; c) gestionar y calificar las evaluaciones de manera adecuada; d) comunicar con precisión los resultados a las partes interesadas, y e) llevar a cabo todas las responsabilidades de la evaluación, legal y éticamente (Popham, 2009; Stiggins, 1991; Stiggins, Arter, Chappuis & Chappuis, 2007). Profundizaremos cada uno a continuación.

a) Evaluaciones válidas y confiables

Es común encontrar lineamientos, tanto en los decretos de evaluación ministeriales como en los reglamentos de evaluación de los colegios,

que ponen énfasis en la idea de utilizar múltiples herramientas de evaluación, sin embargo esta diversidad se reduce considerablemente en la práctica, ya que solo se considera su utilización asociada a calificación (Ravela, Leymonié, Viñas & Haretche, 2014) y, por tanto, muchas de estas herramientas evaluativas no son utilizadas por su complejidad para calificarlas. Si entendemos la evaluación como un proceso constante de monitoreo de los aprendizajes, se podrían intencionar otras situaciones evaluativas que son de gran efectividad para recoger evidencias sobre los logros de los estudiantes, como contestar por escrito una pregunta de síntesis al final de una clase o armar en conjunto un organizador gráfico de cierre de la temática, las cuales requieren poco tiempo de revisión y no necesitan calificarse.

Por otra parte, es necesario que estos instrumentos o situaciones evaluativas cumplan con criterios de calidad para que la información que recojamos con ellos sea confiable y las decisiones que tomemos sean pertinentes al aprendizaje de los estudiantes. Un criterio fundamental de calidad supone que dichas evaluaciones estén alineadas con los objetivos de aprendizaje que se espera medir con la tarea. Muchas veces estos objetivos pueden estar planteados en forma amplia o no ser directamente observables, por lo que se necesita reelaborarlos para que den cuenta de un desempeño concreto por parte del estudiante (ver ejemplo 1).

Como se observa en el ejemplo 1, la precisión de un objetivo tiene relación con su operacionalización concreta para poder levantar evidencia de dicho aprendizaje y emitir un juicio correcto de su nivel de desarrollo. Estos indicadores de evaluación no representan el objetivo en su totalidad, sino que son acciones que el estudiante puede hacer para demostrar su aprendizaje. Es importante que los alumnos conozcan estos indicadores antes de ser evaluados, de esta forma hacemos explícitas nuestras expectativas como docentes y los estudiantes (y sus padres, si son pequeños) sabrán cómo orientar el estudio o refuerzo de dichos aprendizajes de acuerdo con lo esperado.

EJEMPLO 1:

Un objetivo planteado para la asignatura de Lenguaje en 4° Básico es "Comprender y disfrutar versiones completas de obras de la literatura, narradas o leídas por un adulto, como: cuentos folclóricos y de autor, poemas, fábulas y leyendas, capítulos de novelas". ¿Cómo puedo saber yo como profesora que uno de mis estudiantes comprende las obras y las disfruta? Ambas acciones son internas de cada alumno y, además, de naturaleza diferente (comprenderlas es cognitivo y disfrutarlas es actitudinal), por lo que para evaluarlas necesitamos que este objetivo sea más preciso. Esto implica desglosarlo en aspectos más específicos que requerirán distintas tareas de evaluación. Así, para el objetivo asociado a comprensión de las obras, una propuesta de indicadores de evaluación puede ser:

- Explica las consecuencias que tienen las acciones de ciertos personajes de los textos.
- Fundamenta su postura con ejemplos a partir de las obras leídas.

Y para el objetivo asociado al disfrute de la lectura:

- Formula preguntas que demuestran su interés por la obra leída.
- Lee diversas obras en sus versiones completas sin que esté programado evaluarlas.
- Hace comentarios que demuestran disposición positiva hacia la lectura.
- Comenta aspectos específicos de obras complementarias leídas fuera del contexto de la clase.

También hay que considerar que aunque un objetivo de aprendizaje sea preciso, puede tener particularidades que requieran más de una instancia de evaluación o más de un instrumento para certificar que efectivamente se está logrando. Veamos el ejemplo 2:

EJEMPLO 2:

Si mi objetivo de aprendizaje es "Identificar y comparar, por medio de la exploración, los estados sólido, líquido y gaseoso del agua", está implícito que los estudiantes no deberían ser evaluados solo con una prueba sobre los estados de la materia, específicamente, del agua. Con ese instrumento podremos certificar si identifican y comparan los estados del agua, pero no estaremos evaluando el desarrollo de la habilidad de explorar para comprender el fenómeno, habilidad clave en Ciencias. De este modo, si incluimos en nuestra estrategia evaluativa una actividad en la que los estudiantes observen concretamente los estados del agua y, utilizando preguntas claves, ellos expliquen qué criterios usan para distinguir entre los estados (dureza, flexibilidad frente al contenedor, temperatura, entre otros) podremos verificar su proceso de exploración del fenómeno y su capacidad de observación. Esto supone entender la enseñanza y la evaluación como una misma actividad que tiene dos momentos y no como dos acciones separadas (Montes, 2014).

A esta coherencia entre lo que enseñamos y lo que evaluamos se le denomina "validez instruccional", por tanto, cuando hablamos de un instrumento que debe ser válido y confiable, estamos suponiendo que este cumple con los criterios de calidad técnica en su elaboración y que está alineado con los objetivos de aprendizaje que se pretenden evaluar con él.

El *Joint Committee on Standards for Education Evaluation* (2003) define unos estándares asociados a la ética evaluativa que dan cuenta de los principales puntos que hay que considerar para asegurar una práctica de calidad en este ámbito, los cuales serán discutidos en el último apartado de este capítulo.

b) Interpretación del desempeño de los estudiantes

Interpretar el desempeño de los estudiantes a la luz de formas específicas de evaluación e hipotetizar sobre errores comunes propios de la disciplina supone comprender la manera en que cada instrumento se puntúa, las limitaciones y fortalezas que tiene cada

tarea evaluativa para valorar distintos tipos de aprendizaje y cómo construirlos para evidenciar errores comunes en el contenido disciplinar de las asignaturas.

Respecto de las fortalezas y limitaciones de cada instrumento o tarea evaluativa, es importante considerar que en un formato de prueba escrita, por ejemplo, lo que se levanta como evidencia es conocimiento declarativo. Esto es muy eficiente para evaluar habilidades cognitivas, pero no para medir aspectos actitudinales o valóricos ni para procedimientos prácticos que requieren manipulación de material o implementación de procedimientos o técnicas. Del mismo modo, una escala de valoración autoevaluativa es de gran utilidad para medir la percepción de quien responde sobre un tema, pero no necesariamente para evaluar su conducta real.

También es fundamental el dominio del contenido pedagógico que tenga el docente de la disciplina que enseña, ya que para abordar las ideas claves con sus estudiantes necesita saber cuáles son esas ideas (Shulman, 2005) e identificar las dificultades que los estudiantes tienen en su razonamiento para poder afrontarlas (Zohar & Schwartzer, 2005). Esto es válido no solo al momento de enseñar sino también al de evaluar (Carr et al., 2000; Mishra & Koehler, 2006), pues se ha visto que un docente que tiene un buen dominio del contenido pedagógico interactúa con sus estudiantes potenciando preguntas en la clase para recoger evidencias de cómo están pensando y aprendiendo sobre el tema que enseña y, además, tiene mejores prácticas evaluativas, las que incorporan tanto lo formativo como lo sumativo (Jones & Moreland, 2005). Un ejemplo de esto es el siguiente:

EJEMPLO 3:

Un error común en Matemáticas es que los estudiantes asumen reglas a partir de sus experiencias previas. En fracciones por ejemplo, cuando escribimos 3¼, hay algunos alumnos que asumen que como no hay un símbolo matemático entre el 3 y el ¼, entonces va un signo de multiplicación, pensando que eso es lo típico si no hay nada escrito sin comprender que en realidad los 3 enteros están sumados con ¼, con lo

que queda en evidencia que no comprenden lo que pasa en el número mixto. Al final lo resuelven correctamente de forma mecánica multiplicando el entero con el denominador y sumando el numerador, sin tener idea de qué significa lo que están haciendo, por tanto llegan a la solución, pero cuando lo deben explicar, lo hacen de manera incorrecta porque no comprenden lo que está a la base.

Como podemos apreciar en el ejemplo 3, como docentes debemos, por una parte, intencionar acciones en el aula que nos permitan identificar este tipo de errores, de los cuales no nos daremos cuenta solo con una guía de ejercicios porque tendrán el resultado correcto y, por otra, debemos manejar el contenido para poder identificar qué tipo de error es el que están cometiendo los estudiantes para abordarlo de forma adecuada.

c) Gestionar y calificar las evaluaciones de manera adecuada

Uno de los problemas más comunes que se evidencian en las prácticas evaluativas de los docentes tiene relación con la forma en que se analizan los resultados de una evaluación. La mayoría de las veces corresponde a la suma de los puntajes parciales obtenidos en cada parte del instrumento evaluativo, sin considerar el peso que tiene cada uno ni a qué objetivo de aprendizaje responde. Se ha visto que los juicios evaluativos que hacen los profesores sobre el rendimiento de sus estudiantes (notas, clasificaciones, *ranking*) tienen un impacto considerable en sus experiencias de aprendizaje y en sus trayectorias educativas (Südkamp, Kaiser & Möller, 2012). Estas experiencias negativas generan en los estudiantes, por una parte, desmotivación por aprender la disciplina y, por otra, tomar decisiones en base a supuestos erróneos como asumir que “soy malo para las Matemáticas” o que “las mujeres son humanistas y los hombres, científicos”.

Estos juicios errados o poco confiables que hacemos como docentes sobre el rendimiento de nuestros alumnos pasan por la forma en que

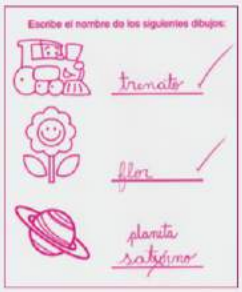
gestionamos y calificamos las evidencias de aprendizaje que recogemos. Muchas veces sucede que elaboramos una pauta para un trabajo que tiene una serie de indicadores de evaluación que constituyen aspectos formales o de la estructura del trabajo, y que terminan sumando más puntaje que los aspectos asociados al contenido de la disciplina; entonces los estudiantes que son ordenados y siguen instrucciones tendrán una buena calificación aun cuando no hayan profundizado tanto en los contenidos que debían incluir, pero un estudiante que falla en la estructura, aunque tenga mayor dominio de los contenidos, obtendrá una nota más baja.

También los profesores tienden a asociar las calificaciones con la conducta de un estudiante o con sus características personales y familiares, emitiendo juicios subjetivos respecto de su rendimiento (Brookhart, 1994; Kaiser, Retelsdorf, Südkamp & Möller, ٢٠١٣; Stiggins, Frisbie & Griswold, ١٩٨٩; Südkamp et al., ٢٠١٢). Así, por ejemplo, un profesor que corrige una prueba de un estudiante que presenta conductas cooperativas en el aula y que en general tiene buen rendimiento tiende a ser más permisivo con los errores o imprecisiones en comparación con un estudiante que tiene conductas disruptivas en el aula o cuyo rendimiento habitual es bajo, es decir, la aplicación de la pauta de corrección es inconsistente (Dompnier, Pansu & Bressoux, ٢٠٠٦). El siguiente caso ejemplifica una situación típica en la cual el estudiante queda en desventaja frente al juicio del profesor, ya que su respuesta no se ajusta literalmente a la pauta de corrección, aun cuando su conocimiento es igual o superior al esperado.

EJEMPLO 4:

Esta es parte de la respuesta de una estudiante de bajo desempeño en 1° Básico que tiende a conversar mucho y a ser muy inquieta durante la clase. El propósito de la guía es evaluar vocabulario a partir de la interpretación gráfica. Como se puede ver, se marca como errónea la palabra “Saturno” y se corrige señalando que la palabra correcta es “planeta”, sin considerar que el conocimiento

de la niña es más profundo, ya que el planeta que tiene anillos efectivamente es el que ella señala. Con esta corrección, el mensaje que se le da es que su conocimiento es erróneo, obligándola a expresarse de acuerdo con lo que la profesora cree y sin dejar opción de desplegar la riqueza del aprendizaje propio, obtenido, tal vez, fuera del aula.



Pauta de respuestas:

1. Tren
2. Flor
3. Planeta

También se ha visto que los docentes tienen un dominio limitado del uso y manejo de datos cuantitativos (Mandinach & Gummer, 2016). En los últimos años se ha enfatizado la importancia que tiene para el mejoramiento de las prácticas docentes –y por tanto para los aprendizajes de sus estudiantes– el que los profesores estén preparados para usar los datos de sus evaluaciones como parte del "arsenal de herramientas" que manejan durante su ejercicio profesional (*National Council for Accreditation of Teacher Education*, 2010). Esta competencia a menudo se confunde con la capacidad del docente para evaluar (Greenberg & Walsh, 2012), no obstante, son aspectos diferentes. En la práctica, un profesor puede hacer muy buenas tareas evaluativas para monitorear el aprendizaje de sus estudiantes, cumpliendo con los criterios de calidad en su construcción, pero si no analiza correctamente la evidencia que recoge, el proceso queda trunco, es decir, sus decisiones serán intuitivas, basadas en una percepción global o incluso sesgada de los resultados que obtuvieron sus alumnos (Feldman & Tung, 2001; Hamilton et al., 2009; Mandinach & Gummer, 2012; Mandinach & Jackson, 2012).

Del mismo modo, se evidencian prácticas erróneas en el cálculo de calificaciones como sacar conclusiones del desempeño de un estudiante o curso a partir de las notas obtenidas, sin considerar qué aprendizajes se estaban evaluando, el nivel de dificultad con que se

evaluaron, la forma en que se trabajaron en la clase, entre otros aspectos.

d) Comunicar con precisión los resultados a las partes interesadas

La comunicación de los resultados se puede dividir en tres ámbitos: 1) la comunicación asociada a la evaluación formativa, que requiere una retroalimentación que permita al estudiante saber cómo fue su desempeño y los aspectos que debe mejorar, tema que abordaremos con detalle en el capítulo 5; 2) la comunicación en el contexto de una evaluación sumativa que, igual que en la evaluación formativa, requiere de retroalimentación, pero esta tiene otras particularidades, ya que la mejora de los aprendizajes de los estudiantes se produce fuera del proceso de enseñanza planificado y el estudiante tiene mayor responsabilidad en ello, y 3) las características para que dicha comunicación sea efectiva, considerando la forma en que estos resultados son comunicados a los interesados, resguardando la confidencialidad de los mismos y que estos sean comprensibles según el destinatario (estudiantes, apoderados, otros profesores o directivos) (Stiggins et al., 2007).

La comunicación de resultados en el contexto de la evaluación sumativa en general se traduce en una calificación, por lo que debemos asegurarnos de que nuestros estudiantes entiendan de manera unívoca qué significa ese número en relación con sus aprendizajes. Esta comunicación tiene lugar al final de un período y consiste en dar a conocer un juicio sobre el desempeño del alumno con respecto a los objetivos de aprendizaje. Esta información se traduce en una calificación que se comunica al propio alumno, a los padres y a las autoridades externas. Sería deseable que dicha comunicación se adaptara al destinatario en cuanto al formato en que se presentan los resultados y, además, en cuanto a las evidencias entregadas. Por ejemplo, los padres requieren la calificación y evidencias que la sustentan, mientras que las autoridades educativas solo necesitan la valoración final. También es necesario definir con

claridad los símbolos utilizados en los reportes escritos o verbales, si se emplean números o letras para calificar y emitir un juicio global, los usuarios deben entender de manera unívoca su significado y que los resultados estén fundamentados en un proceso de evaluación válido y confiable (Airasian, 2002) y como señalan García, Aguilera, Pérez y Muñoz (2011), es importante que “el reporte de las evaluaciones separe los resultados relacionados con los conocimientos y habilidades de la participación, el esfuerzo u otros comportamientos” (p. 99).

Esta comunicación de los resultados, para que sea efectiva, requiere que el docente comparta con sus estudiantes y apoderados información sobre la forma en que se evaluará un aprendizaje determinado y los fundamentos evaluativos que están tras esos métodos, de manera tal que todos los actores involucrados entiendan el propósito a la base de dicho proceso evaluativo (Rogers & Swanson, 2006). Un ejemplo de mala comunicación se detalla a continuación:

EJEMPLO 5:

En un colegio decidieron implementar una tarea integrada en I° Medio entre las asignaturas de Lenguaje y Artes Visuales, para lo cual se trabajaría en la elaboración y puesta en escena de una obra de teatro. Esta obra implicaba trabajar en Lenguaje los OA asociados a la Unidad 3: Relaciones humanas en el teatro y la literatura, y en Artes Visuales, en la Unidad 2: Arquitectura. Tanto profesores como estudiantes estaban muy entusiasmados, trabajaron durante ocho semanas en el proyecto y fueron entregando avances, los cuales se retroalimentaron. Cuando se presentó la obra y los estudiantes fueron calificados, la expectativa de los padres era que todos tuvieran un 7, ya que era una “actividad entretenida”. Frente a algunos resultados diferentes, como el de un estudiante que tuvo un 5,0 en Lenguaje y un 4,8 en Artes, su padre fue a reclamar al colegio insistiendo en que “esto afectaría su promedio para entrar a la universidad” y que encontraba “insólita esta calificación”, en Lenguaje lo podía entender porque su hijo era matemático, pero en Artes no “porque a todos les va bien y el colegio no está formando actores ni diseñadores de ambiente”.

La situación presentada en el ejemplo da cuenta de una escasa o nula comunicación efectiva de los resultados de aprendizaje del estudiante, a él y a su familia, ya que solo se entregó la calificación sin una explicación de las razones de dicha nota y de los aprendizajes asociados a ella. Tampoco se hizo explícito el proceso de recogida de evidencias que estuvo detrás; de ser así, tanto el estudiante como sus padres estarían en conocimiento de lo que faltó mejorar durante la preparación y entenderían que la puesta en escena tenía un ponderación muy menor, ya que efectivamente no era el propósito de la actividad desarrollar un talento para ser actores o diseñadores de ambiente.

En relación con esto, el *Joint Committee on Standars for Education Evaluation* (2007) define la comunicación de resultados como un punto crítico desde la ética evaluativa y plantea un estándar que da cuenta de los principales puntos que hay que considerar para asegurar una práctica evaluativa de calidad en este ámbito. Dicho estándar se presenta en el siguiente punto.

e) Llevar a cabo todas las responsabilidades de la evaluación legal y éticamente

El último aspecto del cual se debe hacer cargo un docente que está “alfabetizado” en evaluación es considerar y aplicar los estándares legales y éticos asociados a cualquier proceso evaluativo. Los estándares legales están dados por los decretos de evaluación vigentes, los reglamentos internos de evaluación de los colegios y los Marcos y Bases Curriculares; no obstante, las responsabilidades éticas asociadas a la evaluación son más difusas, ya que abordan diferentes ámbitos del quehacer docente.

Como plantea Moreno (2011), toda evaluación tiene siempre una dimensión ética relacionada en gran medida con la pregunta ¿por qué evaluar?, pero que en la práctica evaluativa del docente se desdibuja, ya que pasan a ser más importantes las preguntas relacionadas con qué y cómo evaluar, asumiendo implícitamente una concepción de la evaluación como un proceso objetivo y técnico en el que el estudiante

y el propio docente no tienen injerencia y los resultados obtenidos a través de un riguroso “método” de recogida de información, son irrefutables respecto del logro de aprendizajes. Considerando que las prácticas evaluativas son subjetivas en esencia, desde que seleccionamos un contenido por sobre otro para incluir en una prueba o el tipo de formato que definimos para una evaluación, debemos tener en cuenta las implicancias éticas asociadas a nuestras decisiones.

La subjetividad no significa que la evaluación sea de menor calidad o arbitraria, sino que en la evaluación hay un componente propio de cada persona, es decir, en el proceso evaluativo se conjugan las concepciones, disposiciones y decisiones tomadas por el docente, por tanto, es necesario considerarlo como un proceso que no es aséptico y que siempre tiene consecuencias.

Son numerosos los textos que han abordado la ética en las prácticas asociadas a la evaluación educacional (ver, por ejemplo, Covacevich, 2014; Guidelines for Educational and Psychological Testing, 1989), sin embargo, en la mayoría de los casos, el enfoque ha estado puesto en un paradigma psicométrico que sitúa la evaluación en mediciones de gran escala, como pruebas nacionales o mediciones individuales asociadas a investigaciones realizadas en la escuela, pero pocos se refieren a los principios éticos de la evaluación en el aula, que es donde el docente realiza sus acciones día a día. A continuación se presenta una síntesis de los principios éticos asociados a la evaluación de los aprendizajes en el contexto aula, tomando como base los documentos *Principles for fair student assessment practices for education in Canada* (1993) y *Joint Committee on Standars for Education Evaluation* (2003). Hay que destacar que estos principios y directrices, si bien son exhaustivos y detallados, no se agotan aquí y puede haber otros no explícitos que cada actor podría incluir. Además se reconoce que no todas las directrices son igualmente aplicables en todas las circunstancias. No obstante, la consideración del conjunto de principios y orientaciones asociadas a la evaluación en aula que presentan estos estándares debería ayudar a lograr justicia y equidad para los estudiantes al ser evaluados. Son cinco:

I. Desarrollo y elección de métodos de evaluación: los métodos de evaluación[1] deberían ser apropiados y compatibles con el propósito y contexto de la evaluación.

II. Recopilación de información en las evaluaciones: los estudiantes deben tener diversas y suficientes oportunidades para demostrar los conocimientos, habilidades, actitudes o conductas que se les están evaluando.

III. Puntuación y calificación del desempeño estudiantil: los procedimientos para puntuar o juzgar el desempeño del estudiante deben ser apropiados al método de evaluación utilizado y ser aplicados y monitoreados constantemente.

IV. Síntesis e interpretación de resultados: los procedimientos para resumir e interpretar resultados de las evaluaciones deben corresponder a representaciones precisas e informativas del rendimiento de un estudiante en relación con las metas y objetivos de aprendizaje para el período del informe.

V. Comunicación de los resultados de la evaluación: los informes o reportes de evaluación deben ser claros, precisos y de valor práctico para las audiencias a las que van destinados.

I. DESARROLLO Y ELECCIÓN DE MÉTODOS DE EVALUACIÓN: los métodos de evaluación deberían ser apropiados y compatibles con el propósito y contexto de la evaluación

1. Los métodos de evaluación deben ser desarrollados o seleccionados de modo que las inferencias que se realicen respecto de los conocimientos, habilidades, actitudes y conductas que posee cada estudiante sean válidas[2] y no den lugar a malentendidos.

2. Los métodos de evaluación deben estar claramente relacionados con las metas y objetivos de aprendizaje, y ser compatibles con los métodos de enseñanza utilizados.

Planificar el diseño de la evaluación al mismo tiempo que se planifica la enseñanza ayudará a integrar ambos de manera significativa. Dicha planificación conjunta provee una perspectiva general sobre los conocimientos, habilidades, actitudes y comportamientos que se espera sean aprendidos y evaluados, y los contextos en los que se aprenderán y evaluarán.

3. Al desarrollar o elegir métodos de evaluación se debe considerar las consecuencias de las decisiones que se tomarán con la información obtenida. Así, los resultados de algunas evaluaciones pueden ser más críticos que otros. Por ejemplo, la mala interpretación del nivel de rendimiento en una prueba puede permitir a un estudiante pasar de curso sin tener las competencias o aprendizajes mínimos, lo que será perjudicial para su progreso posterior.

4. Se debe utilizar más de un método de evaluación para garantizar la obtención de una imagen más completa o de un perfil de los conocimientos, habilidades, actitudes o comportamientos de un estudiante, que permita distinguir patrones y tendencias consistentes de su rendimiento. El uso de más de un método también ayudará a minimizar la inconsistencia provocada por diferentes fuentes de error de medición. Por ejemplo, bajo rendimiento debido a un “día ausente”; falta de concordancia entre los ítemes incluidos en una prueba y la escala de calificación; inestabilidad emocional del estudiante a través del tiempo.

5. Los métodos de evaluación tienen que ser adecuados a los antecedentes y experiencias previas de los estudiantes. Deben estar libres de sesgo provocado por factores estudiantiles externos al propósito de la evaluación. Estos posibles factores incluyen la cultura, la etapa de desarrollo, la etnia, el género, la situación socioeconómica, el idioma, los intereses especiales y

necesidades educativas especiales. Por ejemplo, el éxito de los estudiantes al responder a las preguntas en un examen o en un cuestionario oral no debe depender del conocimiento cultural previo, como comprender la alusión a una tradición cultural o valor, a menos que tal conocimiento esté dentro del dominio del contenido que se evalúa. Todos los estudiantes deben tener las mismas oportunidades de mostrar sus fortalezas.

6. El contenido y el lenguaje que generalmente se considera como sensible, sexista u ofensivo debe ser evitado. Ejemplos clásicos en este sentido son la asociación de las mujeres (mamás, tías, abuelas) a labores domésticas o compras en el *mall*, y de los hombres a trabajos fuera de la casa, que marcan estereotipos obsoletos.

7. La traducción o adaptación de instrumentos de evaluación a una segunda lengua o transferidos desde otro contexto o lugar debe ir acompañada de pruebas que aseguren que estos instrumentos son válidos para los fines previstos. Muchas veces se cree que como la prueba original era en español no es necesario hacer adecuaciones de lenguaje, pero encontramos palabras como “alberca” o “bolígrafo” que no son de uso cotidiano para nuestra población, y que configuran un problema de validez semántica o pueden inducir a error en la respuesta de los estudiantes.

II. RECOPILACIÓN DE INFORMACIÓN EN LAS EVALUACIONES: los estudiantes deben tener diversas y suficientes oportunidades para demostrar los conocimientos, habilidades, actitudes o conductas que se les están evaluando

1. Los estudiantes deben saber cuál es el propósito, qué información se está recopilando y cómo se utilizará dicha información. Por ejemplo, si ellos saben que el propósito de la evaluación es diagnosticar las fortalezas y debilidades en lugar de asignarles una calificación, estarán más dispuestos a revelar debilidades y no solo exponer sus fortalezas.

2. Un procedimiento de evaluación debe aplicarse en condiciones y formas adecuadas al propósito para el cual fue construido. Por ejemplo, se deben considerar las condiciones ambientales de la sala (luz, ventilación, temperatura y ruido ambiental), el tiempo estimado para la evaluación, el formato del instrumento (tamaño de la letra, calidad de las imágenes, distribución de los ítemes), etc.

3. En las evaluaciones que involucran observaciones y que utilizan instrumentos como listas de cotejo o escalas de valoración, el número de características incluidas en el instrumento que deben ser evaluadas en paralelo debe ser lo suficientemente limitado y concreto de modo que las observaciones pueden realizarse con precisión y no se genere un efecto halo (asumir que la característica estuvo presente aunque no la veamos porque otras de la misma naturaleza sí fueron observadas).

4. Las instrucciones que se entregan a los estudiantes deben ser claras, completas y adecuadas a su capacidad, edad y curso.

5. En las evaluaciones que involucran elementos de selección (p. ej. verdadero-falso, selección múltiple), las instrucciones deben indicar a los estudiantes que respondan a todos los ítemes y hacer explícito que no hay penalización por respuestas incorrectas. En ocasiones se utiliza una fórmula de corrección (descuento de buenas/malas) para disuadir las respuestas por “adivinación”. Esta práctica tiene la intención de animar a los estudiantes a omitir ítemes si no conocen la respuesta, en

lugar de adivinarla. Pero la evidencia científica indica que los beneficios esperados de esta corrección no son tales y no se recomienda el uso de la fórmula, por el contrario, se debería incentivar a los estudiantes a utilizar cualquier conocimiento parcial que tengan al momento de elegir sus respuestas y responder a todos los ítemes.

6. Al recoger información de las evaluaciones, las interacciones con los estudiantes deben ser apropiadas, justas y consistentes. Por ejemplo, cuando se evalúan las presentaciones orales en grupos es necesario distribuir las preguntas entre todos para que tengan la misma oportunidad de demostrar sus conocimientos. Si es una prueba escrita, la aclaración de un elemento ambiguo debe realizarse para todo el curso y no solo para el estudiante que tiene la duda.

7. Circunstancias imprevistas que interfieren con la recogida de información de la evaluación se deben registrar y posteriormente tomar en cuenta al interpretar la información obtenida. Por ejemplo, eventos como un simulacro de incendio, una asamblea no programada, materiales insuficientes, un sismo o corte de luz.

8. En las instituciones debe haber una política escrita, desarrollada por los profesores y directivos, que guíe las decisiones sobre el uso de procedimientos de evaluación alternativos para la recogida de información de los alumnos con necesidades educativas especiales y para los estudiantes cuyo dominio del idioma es insuficiente como para que respondan de la manera prevista.

III. PUNTUACIÓN Y CALIFICACIÓN DEL DESEMPEÑO ESTUDIANTIL: los procedimientos para puntuar o juzgar el

desempeño del estudiante deben ser apropiados al método de evaluación utilizado y ser aplicados y monitoreados constantemente

1. Antes de utilizar un método de evaluación se debe tener claro el procedimiento de puntuación que se hará al juzgar la calidad de una actuación o producto, la adecuación de una actitud o conducta, o la regularidad de una respuesta. Por ejemplo, en pruebas de respuesta cerrada se debe tener previamente la pauta de corrección de los ítemes, y si la respuesta es abierta o lo que se está evaluando es un desempeño práctico, se debería incluir la especificación de lo que se espera como desempeño del estudiante y, en la medida de lo posible, las descripciones de los diferentes niveles de rendimiento o de la calidad de un producto.

2. Antes de utilizar un método de evaluación, a los estudiantes se les debe decir cómo serán juzgados o indicar cuál es el desempeño esperado para cada criterio. Esta práctica ayuda a asegurar que las expectativas del docente y de los alumnos sean similares respecto del desempeño deseado en la actividad evaluativa. Por ejemplo, si al explicar una guía para elaborar un ensayo les presentamos a los estudiantes algunos ejemplos de ensayos (con diferentes niveles de calidad) y analizamos su nivel de desempeño explicando por qué el ensayo A está en el nivel más alto y cuáles son los errores cometidos en el ensayo B que lo sitúan en un nivel de desempeño inicial, los alumnos tendrán más clara la tarea y lo que se espera que realicen.

3. Se debe cuidar que los resultados de una evaluación no estén influenciados por factores que no son relevantes para el propósito de esa evaluación específica. Por ejemplo, bajar puntos por ortografía si eso no es parte de los objetivos de aprendizaje evaluados, o los sesgos personales asociados con

la conducta en la sala de clases de determinados estudiantes, que distorsionan el juicio sobre su logro alcanzado.

4. Entregar una retroalimentación a través de comentarios en forma oral y escrita, que proporcione información respecto de los errores o inconsistencias que el estudiante puede corregir, debería formar parte de la calificación. Esta retroalimentación tiene que basarse en las respuestas dadas por los alumnos y los comentarios deben ser presentados de una manera que los estudiantes puedan comprenderlos y utilizarlos.

5. Cualquier cambio que se realice durante el proceso de puntuación debe basarse en un problema del procedimiento inicial de puntuación. Por ejemplo, si en la pauta de corrección de una prueba se incluía un puntaje asociado a una pregunta que está mal hecha y debe ajustarse la pauta, es necesario volver a revisar las respuestas de todos los estudiantes.

6. Se debe contemplar un proceso de apelación de los resultados de una evaluación para los estudiantes al comienzo de cada año escolar o curso. Por ejemplo, pueden surgir situaciones en las que un estudiante crea que un resultado refleja incorrectamente su nivel de rendimiento y en el colegio debe haber un protocolo establecido y socializado con los estudiantes y docentes para resolverlas. Este procedimiento puede ir desde la comprobación de la suma de puntajes u otros errores de graduación hasta solicitar que la evaluación sea revisada y puntuada por una segunda persona cualificada.

IV. SÍNTESIS E INTERPRETACIÓN DE RESULTADOS: los procedimientos para resumir e interpretar resultados de las evaluaciones deben corresponder a representaciones precisas

e informativas del rendimiento de un estudiante en relación con las metas y objetivos de aprendizaje para el período del informe

1. Los procedimientos para resumir e interpretar los resultados de las evaluaciones de un período (trimestre, semestre, año) deben guiarse por una política escrita. Las calificaciones e informes de otros aspectos del desempeño de un estudiante sirven para una variedad de funciones, como informar a los estudiantes y sus padres/apoderados del progreso en el aprendizaje, y a profesores y directivos para guiar las estrategias de enseñanza, determinar los criterios de promoción, identificar a los estudiantes que requieren una atención especial y para ayudar a los alumnos a desarrollar planes a futuro.

2. Se debe explicar a los estudiantes y sus padres/apoderados la forma en que son formulados e interpretados los informes de otros aspectos del desempeño y las calificaciones.

3. Los resultados individuales y el proceso seguido para el cálculo de las notas y para la realización de los informes de otros aspectos del desempeño del estudiante deben describirse con el suficiente detalle para que el significado sea claro para los distintos actores, indicando el énfasis relativo dado a cada resultado y el proceso seguido para combinar los resultados.

4. Combinar tipos de información de fuentes dispares en un solo resumen del proceso debe hacerse con cautela. En la medida de lo posible, el rendimiento académico, el esfuerzo, la participación y otros comportamientos deben ser clasificados por separado. De esta forma, el estudiante puede saber su desempeño en aspectos específicos y mejorar lo que sea necesario.

5. La base para la interpretación de los resultados debe ser conocida por la comunidad educativa. En este sentido, la contrastación de la información de las evaluaciones con un estándar o aprendizaje esperado es lo recomendable, ya que tener como referencia un estándar externo permite asegurar el logro de las metas curriculares establecidas y no basarse en el desempeño de un grupo de estudiantes en particular.

6. Las interpretaciones de los resultados de la evaluación deben tener en cuenta los conocimientos y experiencias de aprendizaje de los estudiantes. Esto se relaciona con el levantamiento de conocimientos e ideas previas de los alumnos para entender cuál es su punto de partida y así analizar con evidencias su progreso y, además, con la validez instruccional, ya que los resultados de una evaluación de aula estarán condicionados por la forma y actividades que los estudiantes llevaron a cabo para aprender lo evaluado. El bajo rendimiento en una evaluación puede ser atribuible a la falta de oportunidades para aprender.

7. Los resultados de las evaluaciones deben interpretarse en relación con el contexto personal y social del estudiante. Entre los factores a tener en cuenta están la edad, el sexo, el idioma, la motivación, la oportunidad de aprender, la autoestima, la situación socioeconómica, los intereses particulares, las necesidades educativas especiales y las habilidades “para rendir exámenes”. Por ejemplo, la motivación para hacer las tareas escolares, la capacidad de lenguaje o el entorno familiar pueden influir en el aprendizaje de los conceptos evaluados. Esto supone que el docente conoce a sus estudiantes y familias, que es consciente de las dinámicas sociales del curso y que no solo se preocupa del rendimiento de los estudiantes en un área disciplinar específica.

8. Los resultados de las evaluaciones que combinan comentarios y calificaciones deben ser almacenados en una forma que asegure su exactitud en el momento que se resumen e interpretan.

9. Las interpretaciones de los resultados de las evaluaciones deben hacerse teniendo en cuenta las restricciones de los métodos de evaluación utilizados, los problemas encontrados en la recopilación de la información y las limitaciones que puede haber en su análisis.

V. COMUNICACIÓN DE LOS RESULTADOS DE LA EVALUACIÓN: los informes o reportes de evaluación deben ser claros, precisos y de valor práctico para las audiencias a las que van destinados

1. El sistema de gestión de la información de una escuela debe estar guiado por una política escrita. Entre los elementos a tener en cuenta se incluyen aspectos como audiencias, formato, contenido, nivel de detalle, frecuencia y confidencialidad. La política para orientar la preparación de informes en la escuela debe ser desarrollada por los profesores o equipos directivos, en consulta con los representantes de las audiencias que tienen derecho a recibir dicho informe, como son los centros de padres y de estudiantes (p. ej., los resultados de una evaluación; informes trimestrales o semestrales). La participación cooperativa no solo conduce a lograr una información más adecuada y útil, sino también aumenta la probabilidad de que los informes y resultados sean comprendidos y utilizados por los destinatarios.

2. Los informes escritos y orales deben contener una descripción de los objetivos y metas de aprendizaje a los que hacen referencia las evaluaciones. Un informe estará limitado por una serie de consideraciones prácticas, pero el foco central debe estar en los objetivos de aprendizaje y los tipos de desempeños que representan el logro de estos objetivos.

3. Los informes deben incluir en las áreas problemáticas abordadas la descripción de fortalezas y debilidades de los estudiantes. La precisión al informar fortalezas y debilidades ayuda a reducir el error sistemático[3] y es esencial para estimular y reforzar un mejor rendimiento. Los informes deben contener la información que ayudará y guiará a los estudiantes, sus padres/apoderados y los profesores a tomar acciones de seguimiento pertinentes. No se deben reportar en estos informes aspectos del estudiante que no son susceptibles de cambiar, si son condiciones estructurales o de base hay que trabajar con ellas y no considerarlas una debilidad en sí mismas.

4. Los informes de las evaluaciones deben proporcionar a los profesores información para preparar las reuniones con los padres/apoderados. Cuando sea apropiado, los estudiantes también deben participar en estas reuniones. Ellas deben ser programadas a intervalos regulares y, en caso necesario, a petición de los padres/apoderados, como una oportunidad para discutir los procedimientos de evaluación usados, aclarar y comprender los resultados de las evaluaciones, calificaciones e informes, o para trabajar en planes de seguimiento o de acciones pertinentes a las necesidades del estudiante.

5. El acceso a la información de una evaluación debe regirse por una política escrita que sea consistente con las leyes y con los principios básicos de la justicia y los derechos humanos. Para orientar las decisiones en cuanto a la liberación de información de las evaluaciones de los estudiantes se debe utilizar una

política por escrito, desarrollada por la comunidad escolar. La información de una evaluación debe estar a disposición de aquellas personas a las que se aplica (estudiantes y sus padres/apoderados y los profesores) y la información debe ser utilizada de manera constructiva a favor de los alumnos. Asimismo, la información de una evaluación podría ponerse a disposición de otros que justifiquen su necesidad de acceso, previo consentimiento del evaluado y/o su apoderado (p. ej., instituciones postsecundarias, empleadores potenciales, investigadores). El tema de los consentimientos informados también debe abordarse en esta política escrita.

La literatura plantea que un evaluador no debe ceñirse solo a un código ético (Moreno, 2011), ya que hay situaciones específicas que pueden no estar cubiertas, y tener distintos puntos de vista ayuda a complementar el conocimiento de las implicancias éticas de la evaluación. Como señala Stake (2006), es la interpretación personal y contextual la que condiciona la toma de decisiones.

Por la razón antes expuesta rescatamos el aporte de Beauchamp y Childress (1994), quienes han definido seis principios éticos generales asociados a la ética en Medicina, pero que hacen mucho sentido en Educación y han sido adaptados al contexto escuela. Si vemos las consecuencias que tiene la evaluación en la vida de nuestros estudiantes, debemos considerar que la responsabilidad de nuestras acciones es tan importante y delicada como la de un médico; en este sentido, destacamos estos principios como complementarios a los anteriores, que se adaptaron para abordar la ética de la evaluación desde una mirada profesional. Estos principios son:

1) Beneficencia: se basa en la norma moral de que siempre debe promoverse el bien de los estudiantes (o evaluados) y en ella existe una obligación para el docente (o evaluador) de brindar un servicio de calidad, tener un buen trato hacia sus educandos y respetar condiciones, credos e ideologías. Para que esto ocurra en el aula,

los docentes deben estar preparados y actualizados para enseñar el contenido y las habilidades, ser competentes para realizar la enseñanza y la evaluación, y conocer las características de los estudiantes a los que atienden. Esto implica que el estudiante debe beneficiarse con cualquier acto evaluativo y la evaluación debe ser una oportunidad más de aprendizaje.

2) No maleficencia: este principio se refiere a no dañar al estudiante, es la formulación en negativo del principio de beneficencia. El docente debe analizar los riesgos/beneficios ante la toma de decisiones, respetando la integridad física y psicológica de los alumnos. Se diferencia del principio de beneficencia en que en ocasiones se puede tener la intención de hacer el bien al estudiante, pero que ocurran situaciones que eventualmente lo perjudiquen, directa o indirectamente. Este principio debe considerarse integrado con otros principios, ya que a veces se puede privilegiar un beneficio global que puede implicar en algún momento del proceso un malestar para el estudiante. Por ejemplo, cuando un alumno se pone muy nervioso al presentar en forma oral, se debe analizar cómo lograr que desarrolle la habilidad de exponer sus ideas en público, sin que esta tensión implique un daño a su autoestima y la evaluación acuse un perjuicio para él.

3) Autonomía: este principio hace referencia a la libertad que tienen los estudiantes para aceptar o no participar en una evaluación y la obligación de los docentes de hacer explícitos los objetivos y propósitos de las evaluaciones y solicitar un consentimiento informado a los participantes. Este principio aplica principalmente a las evaluaciones externas al aula, producto de investigaciones; sin embargo, a nivel aula se puede entender la autonomía en un sentido más concreto, esto es, garantizar a los estudiantes el acceso a la información de manera oportuna y veraz respecto del proceso evaluativo, lo que se evidencia en situaciones como, por ejemplo, pautas de corrección autoexplicativas que no necesiten del docente para entender la puntuación obtenida.

4) Justicia: este principio se basa en la igualdad de oportunidades sin exclusión ni privilegios para algunos. Se basa en la norma moral

de dar a cada quien lo que necesita, de la cual se derivan diversas obligaciones, como realizar una adecuada distribución de los recursos, proveer a cada estudiante un nivel adecuado de atención y disponer de los recursos indispensables para garantizar una enseñanza apropiada. La justicia y equidad se relacionan con la educación entendida como un derecho fundamental de todo ser humano, que debe ser garantizado por la sociedad y, por tanto, los docentes tienen la obligación de proveer a sus estudiantes oportunidades de acuerdo con las necesidades de cada uno (evaluación diversificada o diferenciada, cuando sea el caso), entregar retroalimentación oportuna y de calidad respecto de las fortalezas y limitaciones en su aprendizaje, y no beneficiar o perjudicar a algunos por sobre otros. Es común que los docentes tengamos una actitud diferente con un alumno que tiene buena disciplina y que participa activamente en nuestro curso, que con otro estudiante que se ubica en el polo opuesto en estas conductas. Pero en la medida que somos conscientes de nuestra limitación, podemos corregir nuestro accionar y cumplir con el principio de justicia.

5) Privacidad: este principio hace referencia a la confidencialidad en el manejo de los resultados de la evaluación de cada estudiante (informes, pruebas, certificaciones, etc.). Si bien se asume que los principios de beneficencia y no maleficencia se cumplieron, los resultados de una evaluación pueden ser complejos o difíciles de asumir para un estudiante y tiene el derecho de querer mantenerlos en reserva. Prácticas habituales en el aula como leer las notas de los estudiantes en voz alta mientras se entrega un trabajo, pegarlas en la puerta de la sala para que ellos se busquen y vean cómo les fue, u ordenar las pruebas desde el que sacó mejor nota y hacer comentarios como “hasta aquí están los azules” o “solo quedan dos con nota Ɛ” atentan significativamente contra este principio. También es importante considerar que los comentarios que podamos hacer con otros profesores o apoderados cuando no sea necesario divulgar el resultado de la evaluación de un estudiante

particular es a lo menos inadecuado, ya que puede llevar a la estigmatización de dicho estudiante.

6) Integridad: se refiere a la rectitud, honradez y veracidad que debe impregnar el proceso evaluativo a sus participantes. Este principio es particularmente importante no solo por el cumplimiento de una norma moral básica para la sana convivencia en la sociedad, sino porque la evaluación recoge evidencias del aprendizaje de los estudiantes y con esa información se toman decisiones pedagógicas. Si, por ejemplo, la información que recopilamos no es confiable porque fue producto de plagio, tendremos información de mala calidad y asumiremos erróneamente que nuestros estudiantes aprendieron, lo que indudablemente afectará su aprendizaje futuro. Esto aplica también para los docentes, ya que alude a las prácticas que tenemos al momento de utilizar textos, imágenes o ideas de otros y las reconocemos como propias, con lo que además estamos modelando en nuestros estudiantes una mala práctica.

EN SÍNTESIS, para que un docente lleve a cabo prácticas evaluativas de calidad debe ser competente no solo en la construcción de los instrumentos que aplica, sino que debe estar informado y ser consciente de una serie de otros elementos que permiten llevar a cabo una evaluación de calidad que cumpla con principios éticos, lo cual constituye una tarea compleja en la que se evidencia su profesionalismo docente y su responsabilidad para enseñar una disciplina y monitorear cómo van aprendiendo los estudiantes. Así, es necesario intencionar en el aula prácticas evaluativas que favorezcan el aprendizaje de los alumnos y que mejoren nuestras prácticas pedagógicas.

BIBLIOGRAFÍA

Airasian, P. (٢٠٠٢). La evaluación en el salón de clases. *Biblioteca para la Actualización del Maestro*. México DF: Mc Graw-Hill Interamericana Editores.

Beauchamp, T. & Childress, J. (1994). *Principles of biomedical ethics* (2nd ed.). New York: Oxford University Press.

Brookhart, S. (1994). Teachers' grading: practice and theory. *Applied Measurement in Education*, *7*(4), 279-301.

Brown, G. (2011). Teachers' conceptions of assessment: comparing primary and secondary teachers in New Zealand. *Assessment Matters*, *3*, 45-70.

Carr, M., McGee, C., Jones, A., McKinley, E., Bell, B., Barr, H. & Simpson, T. (2000). *Strategic research initiatives: the effects of curricula and assessment on pedagogical approaches and on educational outcomes*. Wellington, New Zealand: Ministry of Education.

Covacevich, C. (2014). *Cómo seleccionar un instrumento para evaluar aprendizajes estudiantiles*. Nota técnica 738, Banco Interamericano de Desarrollo.

DeLuca, C., LaPointe-McEwan, D. & Luhanga, U. (2016). Teacher assessment literacy: a review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, *28*(3), 251-272.

Deneen, C. C. & Brown, G. (2016). The impact of conceptions of assessment on assessment literacy in a teacher education program. *Cogent Education*, *3*, 1225380, 1-14.

Dompnier, B., Pansu, P. & Bressoux, P. (2006). An integrative model of scholastic judgments: pupils' characteristics, class context, halo effect and internal attributions. *European Journal of Psychology of Education*, *21*(2). Recuperado de http://dx.doi.org/10.1007/BF03173572

Feldman, J. & Tung, R. (2001). Using data-based inquiry and decision making to improve instruction. *ERS Spectrum*, *19*(3), 10-19.

García, A., Aguilera, M., Pérez, M. & Muñoz, G. (2011). Evaluación de los aprendizajes en el aula. *Opiniones y prácticas de docentes de primaria en México.* México: Instituto Nacional para la Evaluación de la Educación.

Greenberg, J. & Walsh, K. (2012). *What teacher preparation programs teach about K-12 assessment: a review*. Washington DC: National Council on Teacher Quality.

Guidelines for Educational and Psychological Testing (1986). Ottawa: Canadian Psychological Association.

Hamilton, L., Halverson, R., Jackson, S., Mandinach, E., Supovitz, J. & Wayman, J. (2009). *Using student achievement data to support instructional decision making.* IES Practice Guide, NCEE, 4067. Recuperado de https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/dddm_pg_092909.pdf

Himmel, E. (2003). Evaluación de aprendizajes en la educación superior: una reflexión necesaria. *Revista Pensamiento Educativo*, *33*(2), 199-211.

Joint Committee on Standards for Educational Evaluation (2003). *The student evaluation standards: how to improve evaluations of students*. Thousand Oaks, CA: Corwin.

Jones, A. & Moreland, J. (2005). The importance of pedagogical content knowledge in assessment for learning practices: a case-study of a whole-school approach. *The Curriculum Journal*, *16*(2), 193-206.

Kaiser, J., Retelsdorf, J., Südkamp, A. & Moller, J. (2013). Achievement and engagement: how student characteristics influence teacher judgments. *Learning and Instruction*, *28*, 73-84. Recuperado de http://dx.doi.org/10.1016/j.learninstruc.2013.06.001

Mandinach, E. & Gummer, E. (2012). *Navigating the landscape of data literacy: it IS complex.* Washington DC & Portland OR: WestEd and Education Northwest.

Mandinach E. & Gummer, E. (2016). What does it mean for teachers to be data literate: laying out the skills, knowledge and dispositions. *Teaching and Teacher Education*, *60*, 366-376.

Mandinach, E. & Jackson, S. (2012). *Transforming teaching and learning through data driven decision making*. Thousand Oaks, CA: Corwin.

Mishra, P. & Koehler, M. (2006). Technological pedagogical content knowledge: a framework for teacher knowledge. *Teachers College Record*, *108*(6), 1017-1054.

Montes, N. (2014). Introducción. Enseñanza y evaluación: dos caras de la misma moneda. *Propuesta Educativa*, *41*(1), 6-8.

Moreno, T. (2011). Consideraciones éticas en la evaluación educativa. REICE, *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, ١٤٤-١٣٠ ,(٣)٩.

National Council for Accreditation of Teacher Education, NCATE (2010). *Transforming teacher education through clinical practice: a national strategy to prepare effective teachers*. Washington DC: Autor.

Principles for fair student assessment practices for education in Canada (1993). Edmonton, Alberta: Joint Advisory Committee.

Popham, W. J. (2004). Why assessment illiteracy is professional suicide. *Educational Leadership*, *62*, 82-83.

Popham, W. J. (2009). *Unlearned lessons: six stumbling blocks to our schools' success*. Cambridge MA: Harvard Education Press.

Popham, W. J. (2011). Assessment literacy overlooked: a teacher educator's confession. *Teacher Educator*, *46*, 265-273.

Popham, W. J. (2013). *Evaluación trans-formativa. El poder transformador de la evaluación formativa*. Madrid: Narcea.

Ravela, P., Leymonié, J., Viñas, J. & Haretche, C. (2014). La evaluación en las aulas de secundaria básica en cuatro países de América Latina. *Propuesta Educativa*, *41*(1), 20-45.

Ray, A. & Margaret, W. (Eds.) (2003). *Programme for international student assessment (PISA)*. PISA 2000 technical report. Paris: OECD.

Rogers, T. & Swanson, M. (2006). *Effective student assessment and evaluation in the classroom: knowledge + skills + attributes*. Alberta, Canada: Alberta Education.

Salinas, D. (2002). ¡Mañana *examen! La evaluación entre la teoría y la realidad*. Barcelona: Graó.

Shepard, L. (2006). La evaluación en el aula. En R. Brennan (Ed.), *Educational Measurement* (pp. 623-646). Westport: ACE Praeger.

Shulman, L. S. (2005). Conocimiento y enseñanza: fundamentos de la nueva reforma. *Profesorado. Revista de Currículum y Formación de Profesorado*, *9*(2), 1-30.

Stake, R. (٢٠٠٦). *Evaluación comprensiva y evaluación basada en estándares*. Barcelona: Graó.

Stiggins, R. (1991). Assessment literacy. *Phi Delta Kappan*, *72*(7), 534-39.

Stiggins, R. (2002). Assessment crisis: the absence of assessment for learning. *Phi Delta Kappan*, *83*, 758-765.

Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, *86*, 22-27.

Stiggins, R., Frisbie, D. & Griswold, P. (1989). Inside high school grading practices: building a research agenda. *Educational Measurement: Issues and Practice*, *8*(2), 5-14.

Stiggins, R., Arter, J., Chappuis, J. & Chappuis, S. (2007). *Classroom assessment for student learning: doing it right - using it well*. Columbus, Ohio: Pearson Merrill Prentice Hall.

Südkamp, A., Kaiser, J. & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: a meta-analysis. *Journal of Educational Psychology*, *104*(3), 743-762. doi: 10.1037/a0027627

Wiggins, G. & McTighe, J. (2005). *Understanding by design*. Virginia: Association for Supervision and Curriculum Development.

Xu, Y. & Brown, G. (2016). Teacher assessment literacy in practice: a reconceptualization. *Teaching and Teacher Education*, *58*, 149-162. doi: 10.1016/j.tate.2016.05.010

Zohar, A. & Schwartzer, N. (2005). Assessing teachers' pedagogical knowledge in the context of teaching higher-order thinking. *Journal of Science Education*, *27*(13), 1595-1620.

* Este capítulo está basado en los resultados del proyecto Fondecyt de Iniciación Nº 11140713, financiado por Conicyt, Chile.

[1] El término “método de evaluación” refiere a las diversas estrategias y técnicas que los profesores pueden utilizar para obtener información de los estudiantes sobre su progreso hacia la consecución de los conocimientos, habilidades, actitudes o comportamientos que se espera que aprendan.

[2] La validez se refiere al grado en que las inferencias extraídas de los resultados de las evaluaciones son significativas. Por lo tanto, el desarrollo o la selección de métodos de evaluación para la recopilación de deben estar claramente vinculados con los fines para los cuales se van a utilizar dichos resultados.

[3] “Error sistemático” es un tipo de error de medición o recogida de información que afecta a todos los estudiantes, por lo que analizar los resultados individuales ayudará a reducir este error al incluir elementos cualitativos asociados a su desempeño.

CAPÍTULO 2

Aprendizaje y evaluación: lo que no se evalúa, no se aprende*

Carla E. Förster
Cristian A. Rojas-Barahona

INTRODUCCIÓN

Como señalamos en el capítulo anterior, la evaluación condiciona el aprendizaje de los estudiantes, ya que ellos estudian de acuerdo al contenido y formato de la evaluación. Es por esto que definir qué evaluar es clave para lograr aprendizajes de calidad. Siempre la base está dada por los aprendizajes establecidos en el currículum, pero alguna veces estos lineamientos no son tan claros y no aportan a la definición de qué es lo relevante y necesario monitorear. Lo que aportará a mejorar la claridad será el conocimiento de cómo aprenden los estudiantes y, luego, considerando esta información junto al conocimiento del desarrollo del alumno, generar indicadores lo más concretos posible. En lo específico, como parte de lo que se desarrollará en el presente capítulo, los docentes cuentan con taxonomías que facilitan dicho proceso.

Antes de presentar las distintas taxonomías es importante resaltar la necesidad de que el profesor debe conocer ambos temas mencionados: el aprendizaje y desarrollo de sus estudiantes. Ambos se influyen mutuamente, el aprendizaje impactará en el desarrollo del niño y su desarrollo influirá en su aprendizaje. Esto implica que si el profesor quiere promover el aprendizaje y el desarrollo de sus estudiantes, debe ser capaz de identificar los procesos involucrados en el aprendizaje y el desarrollo del estudiante en todas sus dimensiones: física, afectiva, cognitiva, social, cultural, moral y espiritual. Así, corresponde a los docentes, a través de la evaluación, monitorear que su enseñanza esté impactando o promoviendo el aprendizaje y desarrollo de sus estudiantes.

COMPRENDIENDO EL APRENDIZAJE

Creemos necesario, para una mejor comprensión de las taxonomías, recordar elementos centrales de las distintas formas de aprendizaje, en coherencia con las grandes teorías sobre este tema. Con un fin práctico, entenderemos por aprendizaje lo propuesto por Shunk

(2012): un cambio duradero del comportamiento o en la capacidad de comportarse, resultado de la práctica u otras formas de experiencia. El problema no sería la definición de aprendizaje, sino la interpretación de sus elementos constitutivos. En ese sentido se hace necesario presentar, aunque de manera breve, algunos aspectos clave de las tres grandes visiones que dan respuesta a la pregunta de cómo se aprende: el conductismo, el cognitivismo y el constructivismo (para una mayor revisión, ver, por ejemplo, Shunk, 2012).

La primera visión que abordaremos será el conductismo, que plantea que el aprendizaje sucede por una asociación. Desde la mirada del conductismo clásico, la clave está en la asociación entre un estímulo incondicionado y uno neutro, la que sucederá si se presentan varias veces de manera continua (primero el neutro), logrando una conducta similar a la del estímulo incondicionado. Desde el conductismo operante, el foco estará en un estímulo reforzador que fortalecerá la asociación entre un estímulo discriminante y su respuesta o conducta. En otras palabras, el docente debe atender no solo a la conducta esperada del estudiante, sino también a los estímulos ambientales, discriminando cuál es el causante de ese comportamiento, para asociarlo con otro o para que cuando se genere la conducta esperada, se produzca el estímulo reforzador adecuado para fortalecer la asociación. Para un profesor, la tarea determinante es ser consciente de las asociaciones que están sucediendo en el aula, y conocer a sus estudiantes para entregar los estímulos reforzadores adecuados. Basado en el asociacionismo, Bandura propone una teoría (llamada “cognitivista social” o “socialcognitiva”) en la que no solo reconoce la importancia del ambiente y la conducta (y su determinación recíproca), sino de la persona que observa el ambiente (puede aprender sin necesidad de recibir el estímulo reforzador o castigo de manera directa) y que es capaz de autorregular su conducta. De esta forma, resalta la importancia no solo de los procesos sociales, sino también de la cognición, destacando el rol activo del estudiante. Al mismo tiempo, reconoce la relevancia, en el proceso de aprendizaje por observación, de la atención, la retención, la reproducción motora y la motivación, este último punto es

fuertemente desarrollado tanto desde lo emocional como desde lo cognitivo. Además destaca la importancia de la autoeficacia del estudiante (juicio construido a partir de sus experiencias previas en relación con su desempeño) en el logro o no logro del aprendizaje. Finalmente, el autor plantea que es fundamental tener clara la meta, y llegar a ella de manera progresiva, especialmente cuando es de alta complejidad. Sin duda Bandura es el puente entre el conductismo y el cognitivismo.

El cognitivismo, basado principalmente en las teorías del procesamiento de la información, se focaliza en la importancia de esta, de su codificación adecuada, clasificación y organización, clave para guardar la información en la memoria de largo plazo, para luego rescatarla. Se identifican procesos relevantes en el aprendizaje, como la atención, el uso efectivo de la memoria de trabajo, de las estrategias de incorporación y rescate de la información. Se logran identificar las diferencias de procesamiento de un novato versus un experto, del uso del espacio mental y de cómo un experto ocupará menos espacio en su memoria de trabajo, dejando mayor espacio para procesar información. El profesor tiene el desafío de trabajar con conceptos definidos de manera clara, considerar el orden del nuevo conocimiento buscando conectarlo con el conocimiento que ya tiene el estudiante. Desde este enfoque teórico emerge el conocido y necesario concepto de metacognición, destacándose la importancia de la autorregulación en el aprendizaje, en que el monitoreo y la evaluación de los nuevos conocimientos son determinantes para consolidar los nuevos aprendizajes.

Por último, el constructivismo, enfoque representado principalmente por Piaget (una mirada más individual) y Vygotski (una mirada más social), focaliza el aprendizaje en el estudiante que construye su conocimiento de manera activa, en el que el ambiente seguirá jugando un papel relevante, tanto para que el estudiante vivencie directamente un conflicto cognitivo (conocimiento nuevo con conocimiento o esquemas previos) como para que actúen sus propios mediadores a través de pares o un profesor, facilitando el movimiento en la zona de desarrollo próximo (entre el desarrollo real y potencial

que tiene cada estudiante), es decir, estimulando o permitiendo el aprendizaje. Aquí el rol del profesor es clave para mediar la interacción del alumno con el ambiente y con las personas. También se reconoce la importancia del lenguaje como una herramienta que facilita la transformación del mundo interno. El profesor debe tomar conciencia de que cada alumno construye el nuevo conocimiento a partir de sus experiencias o conocimientos previos. Incluso debe saber que los conocimientos previos (difíciles de mover) influirán en la construcción de los conocimientos nuevos.

Como se dijo al principio del capítulo, teniendo como base las teorías expuestas, la consideración del desarrollo y la necesidad de generar indicadores claros para evaluar el progreso de los aprendizajes ha surgido una serie de taxonomías y clasificaciones que ayudan a posicionar los aprendizajes en un continuo de complejidad y, de esta forma, definir si lo que estamos enseñando y evaluando es consistente con la edad y el desarrollo de nuestros estudiantes.

TAXONOMÍAS DE PROGRESO

Una taxonomía es un marco de clasificación que refleja los patrones comunes de uno o varios aspectos de un organismo, sujeto u objeto y que se construye a través de la sistematización de múltiples observaciones que permiten generar tipologías a partir de dichos patrones o características (Enghoff, 2009).

En el caso de las taxonomías de progreso, describen tipos de conductas, de aprendizaje y desempeños que deseamos que los estudiantes desarrollen o realicen; enumeran niveles genéricos de complejidad que pueden ser usados para clasificar e interpretar tanto los requerimientos de la tarea que se le solicitará al estudiante como la respuesta de este en dicha tarea (Hutchinson, Francis & Griffin, 2014). Tienen como función determinar de forma precisa aprendizajes que deben ser logrados por los estudiantes, y son usadas para identificar diferentes etapas del desarrollo y del aprendizaje para así entregar una herramienta que permita distinguir resultados de aprendizajes particulares (O’Neill & Murphy, 2010). La principal

ventaja de estas taxonomías es que por su carácter genérico pueden ser utilizadas en diferentes situaciones, sin embargo, tienen la de que pueden ser usadas de forma inapropiada malinterpretando la progresión de desarrollo y con esto, limitando el crecimiento de los estudiantes. Haremos énfasis en este aspecto más adelante.

Se han definido distintas taxonomías en el ámbito de la educación escolar y universitaria. Las más conocidas y usadas en evaluación son la de Bloom, Englehart, Furst, Hill y Krathwohl (1956), que aborda tres dominios (cognitivo, afectivo y psicomotor); la revisión del dominio cognitivo realizada por Anderson et al. (2001); la de Marzano y Kendall (2007, 2008), que considera tres dominios (de conocimiento, de procesos mentales y de procesos psicomotores), y la S.O.L.O. (Biggs & Collins, 1982), que considera un escalamiento cuantitativo y cualitativo de los aprendizajes y se ha utilizado más para evaluar competencias en el contexto universitario.

TAXONOMÍA DE BLOOM ET AL. (1956) REVISADA POR ANDERSON ET AL. (2001)

Esta taxonomía se basa en la propuesta original de Bloom et al. (1956) y aborda el dominio cognitivo dividiéndolo en dos dimensiones (procesos cognitivos y conocimientos); por una parte presenta seis niveles asociados a los procesos cognitivos que un estudiante ejecuta en una tarea evaluativa (ver tabla 1) y, por otra, desagrega el tipo de conocimiento que está en juego asociado a dicha tarea. Esta distinción permite elaborar objetivos de aprendizaje y de evaluación más precisos y, de esta manera, diseñar tareas evaluativas que evidencien con mayor cercanía los aprendizajes de los estudiantes. Si bien ambas dimensiones se presentan en un continuo de complejidad, los desempeños de los alumnos pueden no ser lineales, es decir, un estudiante que hace una tarea asociada a análisis no necesariamente realiza de forma previa procesos de aplicación, lo mismo con el tipo de conocimientos: un estudiante que utiliza conocimiento metacognitivo no necesariamente tiene dominio de los conocimientos procedimentales asociados a la tarea.

En la revisión realizada por Anderson et al. (2001) se señala que la desagregación en la dimensión de procesos cognitivos no es suficiente para clasificar correctamente los desempeños de los estudiantes y se agrega la dimensión de conocimiento, que presenta cuatro categorías: conocimiento factual, conceptual, procedimental y metacognitivo (tabla 2).

Esta definición de procesos cognitivos y tipos de conocimientos permite una formulación más precisa de los objetivos o metas de aprendizaje que se espera que los estudiantes logren y, con esto, la elaboración de tareas evaluativas también se facilita, logrando así un alineamiento curricular (Anderson, 2002) o alineamiento constructivo (Biggs & Tang, 2011), que consisten en que haya coherencia entre los objetivos establecidos, lo que se enseña y lo que se evalúa. En la tabla 3 se presentan ejemplos de desempeños asociados a cada cruce de dimensión cognitiva y tipo de contenido. Resaltamos esto porque se tiende a pensar que los conocimientos factuales solo se asocian a la habilidad cognitiva de recordar, no obstante, se pueden aplicar a procesos cognitivos más complejos.

Para utilizar correctamente los dominios, Hutchinson et al. (2014) señalan que el docente debe definir una o varias hipótesis de progresión respecto de cómo el estudiante irá avanzando en la meta de aprendizaje. Estas hipótesis corresponden a la definición que hace un profesor de cómo debería ser la progresión del aprendizaje o desempeño en una tarea determinada a partir de su conocimiento sobre el contenido, adaptando alguna taxonomía genérica a la tarea específica. Se puede hacer la progresión utilizando la taxonomía completa o solo lo que el docente considera que efectivamente se podrá observar en el desempeño de los estudiantes. Un ejemplo de hipótesis de progresión para la escritura de una carta utilizando la taxonomía de Bloom et al. (1956) revisada por Anderson et al. (2001) se presenta en la tabla 4.

TABLA 1: Descripción de los niveles de la dimensión de procesos cognitivos

Nivel cognitivo	Descripción	Categorías	Ejemplo de desempeño	Sugerencias de verbos
Conocer	Los estudiantes recuperan información relevante desde la memoria de largo plazo. Reconocen o evocan información, ideas y principios en forma similar a la que ellos aprendieron.	Reconoce	Reconocer las fechas de acontecimientos importantes en la historia de su país.	Distingue, identifica, subraya, señala.
		Evoca	Evocar las fechas de acontecimientos importantes en la historia de su país.	Recita, recuerda, rememora, reproduce.
		Lista	Nombra los alimentos de un grupo en la pirámide alimentaria.	Anota, cita, registra, enumera, rotula, nombra, ordena.
Comprender	Los estudiantes construyen significados a partir de lo que se les enseñanza, mediante la comunicación oral, escrita y gráfica. Traducen, comprenden e interpretan información basados en aprendizajes previos.	Interpreta	Parafrasear discursos y documentos importantes.	Clarifica, parafrasea, representa, traduce.
		Ejemplifica	Dar ejemplos de varios estilos pictóricos.	Ilustra, ejemplifica.
		Clasifica	Clasificar tipos de metales.	Categoriza, subsume.
		Resume	Escribir un breve resumen de los hechos retratados en un documental.	Extracta, generaliza.
		Infiere	En el aprendizaje de una lengua extranjera, inferir principios gramaticales a partir de ejemplos.	Concluye, extrapola, interpola, predice.
		Compara	Comparar hechos históricos con situaciones contemporáneas.	Contrasta, mapea, establece correspondencias.
		Explica	Explicar las causas de determinados hechos importantes del siglo XVIII en Francia.	Explica, relata, formula, establece, manifiesta.
Aplicar	Los estudiantes llevan a cabo o usan un procedimiento en una situación dada.	Ejecuta	Dividir un número entero por otro, ambos de varias	Aplica, lleva a cabo.

	Seleccionan, transfieren y usan datos y principios para completar un problema o tarea.		cifras.	
		Implementa	Determinar en qué situaciones es aplicable la segunda ley de Newton.	Aplica, utiliza.
Analizar	Los estudiantes descomponen el material en sus partes constituyentes determinando de qué manera se relacionan unas con otras y con una estructura o propósito general. Distinguen, clasifican y relacionan los supuestos, hipótesis, la evidencia o la estructura de un estamento o pregunta.	Diferencia	Distinguir entre los números relevantes e irrelevantes en un problema matemático expresado con palabras.	Diferencia, discrimina, distingue, enfoca, selecciona, separa.
		Organiza	Estructurar los hechos de una descripción histórica en acontecimientos a favor y en contra de una determinada explicación histórica.	Integra, esquematiza, estructura, analiza sintácticamente, ordena.
		Atribuye	Determinar el punto de vista del autor de un ensayo en función de su perspectiva política.	Deconstruye, distingue, determina.
Evaluar	Los estudiantes hacen juicios, aprueban o critican basados en criterios y estándares específicos.	Comprueba	Determinar si las conclusiones de un científico se deducen de hechos observados.	Chequea, coordina, detecta, monitorea, verifica, valida, diagnostica.
		Critica	Juzgar cuál de los métodos presentados es el mejor para resolver un problema dado.	Juzga, califica, atribuye, argumenta, critica, convence.
Crear	Los estudiantes reúnen elementos para formar un todo coherente o funcional; reorganizan elementos en un nuevo modelo o estructura. Originan, integran y combinan ideas en un producto, plan o propuesta que son nuevos para ellos.	Genera	Generar hipótesis que expliquen un determinado fenómeno.	Hipotetiza, integra, combina, reorganiza.
		Planifica	Planificar un documento de investigación sobre un tema histórico dado.	Planea, diseña, planifica, prepara.
		Produce	Construir hábitats para unas especies y propósitos determinados.	Construye, inventa, elabora, produce.

Adaptado de Anderson et al. (2001)

TABLA 2: Descripción de las categorías de la dimensión de conocimiento

Conocimiento	Descripción	Categorías	Ejemplos
A. Factual	Conocimiento de terminología, detalles y elementos específicos. Son fragmentos de información aislados que no necesariamente "encajan" dentro de una perspectiva disciplinar más amplia o más sistemática.	A1. Conocimiento de terminología.	• Vocabulario técnico. • Símbolos musicales. • Símbolos matemáticos. • Estructura química de un compuesto.
		A2. Conocimiento de detalles y elementos específicos.	• Principales recursos naturales. • Fuentes fiables de información. • Autor de una obra. • Nombre de un personaje.
B. Conceptual	Refiere al "qué" de un contenido. Conocimiento de clasificaciones, categorías, principios, generalizaciones, teorías, modelos y estructuras. Son formas de conocimiento organizadas y complejas; los hechos o fragmentos de información se integran y organizan sistemáticamente de forma disciplinar o útil. Supone la interrelación de elementos básicos en una estructura mayor que los habilita para que actúen funcionalmente.	B1. Conocimiento de clasificaciones y categorías.	• Períodos geológicos. • Clasificación de los seres vivos. • Tipos de textos.
		B2. Conocimiento de principios y generalizaciones.	• Teorema de Pitágoras. • Ley de la oferta y la demanda. • Ley de conservación de la materia. • Reglas ortográficas.
		B3. Conocimiento de teorías, modelos y estructuras.	• Teoría de la evolución de Darwin. • Estructura del Congreso. • Modelos atómicos. • Estructura de un cuento o una noticia.
C. Procedimental	Conocimiento de serie o secuencia de pasos a seguir para realizar algo, supone el "cómo" hacer	C1. Conocimiento de	• Pasos para pintar con

	un determinado proceso, discriminando las condiciones más apropiadas para llevarlo a cabo. Representa solo el conocimiento de los procesos, no su puesta en práctica real.	habilidades y algoritmos específicos de la asignatura.	acuarela. • Secuencia para realizar un salto alto. • Algoritmo de la división por números enteros.
		C2. Conocimiento de técnicas y métodos específicos de la asignatura.	• Técnicas de la entrevista. • Método científico. • Ubicación de coordenadas en un mapa. • Análisis de fuentes históricas. • Producción de un texto argumentativo.
		C3. Conocimiento de criterios para determinar cuándo utilizar los procedimientos apropiados.	• Criterios usados para determinar cuándo aplicar un procedimiento relacionado con la segunda ley de Newton. • Criterios para determinar qué escala o tipo de gráfico emplear según los datos.
D. Metacognitivo	Conocimiento de cognición en general, así como la conciencia y conocimiento de su propia cognición. Incluye el conocimiento de las estrategias generales para aprender, pensar y resolver problemas (conocimiento estratégico), el conocimiento de la dificultad de una tarea cognitiva y de las demandas que puede requerir, y el conocimiento de los propios puntos fuertes y débiles en relación con la cognición y el aprendizaje (autoconocimiento).	D1. Conocimiento estratégico.	Conocimiento de: • Organizadores gráficos para ordenar la información. • Varias estrategias mnemotécnicas para memorizar información. • Estrategias de planificación y establecimiento de metas para la lectura de un libro.

	D2. Conocimiento sobre las tareas cognitivas, incluido el conocimiento contextual y el conocimiento condicional adecuados.	Conocimiento sobre: • La memoria que requieren los distintos tipos de ítemes de prueba. • Las demandas cognitivas que exige un ensayo o un resumen. • Las normas sociales, convencionales y culturales de ámbito local y general.
	D3. Autocono-cimiento	• Conocimiento de las capacidades propias para realizar una tarea específica acorde a la realidad. • Conciencia del nivel de conocimiento propio. • Conocimiento del interés personal que se tiene en una tarea.

Adaptado de Anderson et al. (2001)

TABLA 3: Ejemplos de desempeño o tarea que puede realizar un estudiante integrando un proceso cognitivo y una dimensión de conocimiento

Dimensión de conocimiento	Dimensión del proceso cognitivo					
	Recordar	Comprender	Aplicar	Analizar	Evaluar	Crear
Factual	Lista colores primarios y secundarios.	Asocia a un autor con una obra.	Responde según normas de cortesía.	Selecciona la lista más completa de materiales para realizar una actividad.	Chequea consistencia entre fuentes.	Genera un registro de sus actividades diarias.
Conceptual	Reconoce las características de un mamífero.	Clasifica los animales según su hábitat.	Asesora a un compañero que sabe menos sobre fracciones.	Diferencia un cuento de una noticia.	Determina la relevancia de los datos en un problema.	Diseña un sistema solar con material reciclado.
Procedimental	Recuerda cómo hacer una carta.	Explica las instrucciones de un juego.	Calcula una multiplicación de fracciones.	Compara un cuento con una fábula.	Juzga una carta elaborada por un compañero.	Diseña un plan de trabajo para hacer una clasificación.
Metacognitivo	Identifica las estrategias que usa para memorizar información.	Predice su propia respuesta frente a un problema ecológico.	Usa las técnicas de estudio que domina mejor.	Deconstruye sus propios prejuicios sobre participación ciudadana.	Reflexiona sobre su propio progreso en un aprendizaje.	Crea un portafolio de su aprendizaje.

Adaptado de Heer (2012)

TABLA 4: Hipótesis de progresión de la escritura de una carta usando la taxonomía de Bloom et al. (1956), revisada por Anderson et al. (2001)

NIVEL	DESEMPEÑO QUE DEMUESTRA EL CONOCIMIENTO

Crear	Escribe una carta con un tema libre.
Evaluar	Chequea la calidad de su carta a partir de la pauta dada en clases, identificando fortalezas y debilidades.
Analizar	Revisa una carta y reconoce los elementos de su estructura.
Aplicar	Escribe una carta con un tema dado por la profesora al curso.
Comprender	Describe los elementos centrales de la estructura de una carta.
Recordar	Lista las partes de una carta.

DOMINIO AFECTIVO DE LA TAXONOMÍA DE KRATHWOHL, BLOOM Y MASIA (1964)

Este dominio fue elaborado por Krathwohl et al. (1964) y, a pesar del tiempo transcurrido, no ha sido modificado. Aborda los aprendizajes socioafectivos y valóricos de los estudiantes, describiendo etapas actitudinales a través de las cuales estos se mueven cuando están adquiriendo nuevas ideas, hábitos o conductas (Hutchinson et al., 2014). En esta taxonomía se plantea una graduación de la interiorización de una actitud, valor o apreciación que se evidencia en el cambio de conducta de un mismo estudiante. Esta taxonomía supone que una persona actúa consistentemente con los valores internalizados; si bien sabemos que no siempre ocurre así en las etapas intermedias de adquisición de un valor, esta clasificación ayuda a generar tareas evaluativas en las que los estudiantes demuestren su nivel de apropiación. Se ha descrito ampliamente en la literatura que los componentes afectivos y motivacionales son claves en el aprendizaje de los estudiantes (p. ej. Bandura & Cervone, 1983; Morgan, Bourke & Thompson, 2001), no obstante, este dominio ha sido puesto en un lugar secundario en las distintas asignaturas, relegado a un conjunto de habilidades genéricas también llamadas relacionales o “blandas”, cuyo desarrollo no es evaluado. El currículum escolar plantea en las distintas asignaturas un desglose de actitudes y habilidades afectivas que se espera desarrollen los estudiantes y esta taxonomía aporta a la operacionalización de estas para evaluarlas. Los autores reconocen cinco niveles de progresión, los cuales se presentan en la tabla 5.

TABLA 5: Dimensiones del dominio afectivo de Krathwohl et al. (1964)

Dimensión	Descripción	Categoría	Descripción	Ejemplos	Sugerencia de verbos
Recepción	El aprendiz está dispuesto a recibir, escuchar y obtener información e ideas pero no toma decisiones sobre el valor de la información. Toma conciencia de ideas y fenómenos.	Toma de conciencia.	Conocer la importancia del tema o situación, capacidad de tenerlo en cuenta, noción simple.	• Escucha pasivamente al profesor y a sus pares. • Muestra sensibilidad o empatía frente a un problema.	Pregunta, sigue, repite, acepta, prefiere, escucha, recibe, favorece.
		Disposición a recibir.	Tolerar o aceptar los estímulos, disposición a participar.		
		Atención controlada o selectiva.	Preferencia del alumno por ciertos aspectos del estímulo, como estéticos, humorísticos, críticas.		
Respuesta	El aprendiz participa activamente e interactúa con la nueva información o procedimientos sin estar de acuerdo o hacer propias estas ideas. Demuestra complacencia, deseos y satisfacción con las ideas.	Consentimiento para responder.	Primer nivel de respuesta activa del alumno.	• Participa en las discusiones de clases. • Pregunta nuevas ideas y conceptos. • Completa un set de tareas diseñadas para desarrollar una actitud específica. • Concuerda con una idea respondiendo a ella.	Responde, narra, ejecuta, reporta, selecciona, sigue, explora, exhibe, conforma, completa, escribe, registra, aprueba, ofrece voluntariamente, dedica tiempo a.
		Disposición a responder.	Voluntad de responder, disposición del alumno.		
		Satisfacción por responder.	Agrado o entusiasmo por responder.		

TABLA 5: Dimensiones del dominio afectivo de Krathwohl et al. (1964) (cont.)

Dimensión	Descripción	Categoría	Descripción	Ejemplos	Sugerencia de verbos
Valoración	El aprendiz es capaz de ver el valor de la nueva información y procedimientos. Varía desde la aceptación y preferencia por un valor hasta la aceptación de un compromiso. Reconocimiento y actitud motivada.	Aceptación de un valor.	Atribuir valor a un fenómeno, cosa o comportamiento.	• Comparte perspectivas respetando las diversas opiniones del grupo y es capaz de solucionar problemas. • Está dispuesto a ser visto como "valorador" de una idea o material.	Argumenta, confía, renuncia, abandona, acepta, reconoce, participa, indica, decide, ayuda, asiste, niega, protesta, provee, defiende, inicia.
		Preferencia por un valor.	Disposición a buscar, desear o procurar un valor.		
		Compromiso.	Convicción absoluta del valor del fenómeno o cosa.		
Organización y conceptualización	El aprendiz está incorporando nueva información y procedimientos en sus esquemas existentes. Proceso por el cual el aprendiz soluciona conflictos y comienza a interiorizar valores.	Conceptualización de un valor.	Relacionar valores adquiridos con los en adquisición.	• Equilibra libertad y responsabilidad, acepta diferentes puntos de vista y construye con ellos para desarrollar nuevas perspectivas y comprensión de ideas. • Organiza de forma independiente el tiempo y materiales para implementar una rutina o hábito.	Organiza, asocia, determina, compara, define, formula, abstrae, encuentra, correlaciona, integra, dispone, hace un balance, teoriza.
		Organización de un sistema de valores.	Interrelación coherente de un conjunto de valores.		
Caracterización e internalización de valores	El aprendiz comienza a ser partidario y defiende la nueva información e ideas, posee un sistema de valores relacionado con las creencias, ideas y actitudes que controlan su comportamiento de una manera previsible y consistente.	Conjunto generalizado.	Coherencia interna con el sistema de actitudes y valores.	• Muestra autoconfianza, colabora positivamente con otros miembros del grupo y actúa éticamente, valorando a las personas por lo que son y no por cómo lucen. • Actúa consistentemente con los valores internalizados. • Lidera voluntariamente una actividad asociada a la temática trabajada.	Es consecuente con, se revela por, da el ejemplo, revisa, cambia, evita, resuelve, resiste, discrimina, cuestiona, influencia.
		Caracterización.	Cosmovisión, filosofía de vida, actitud fundamental.		

La evaluación del dominio afectivo no es fácil, ya que siempre va unida al ámbito cognitivo a través de la forma en que el estudiante

procesa la información que recibe, y la única manera en que el profesor puede evaluarlo como agente externo es mediante la observación de sus conductas. No obstante, el uso de esta taxonomía permite definir la progresión de conductas que evidencien la adquisición y desarrollo de valores, hábitos y actitudes que están aprendiendo los estudiantes y que guían sus acciones. Analicemos un ejemplo usando el objetivo actitudinal de Ciencias Naturales: “Demostrar curiosidad e interés por conocer seres vivos, objetos y/o eventos que conforman el entorno natural” (tabla 6). En él podemos ver que la taxonomía nos ayuda a definir de forma más clara una progresión en cuanto a complejidad o interiorización de la actitud, y los verbos propuestos apoyan la definición de las conductas observables mediante las cuales vamos a monitorear el desarrollo de dicha actitud. Hay que recalcar que el desarrollo puede tener distintos tiempos, por lo que algunos estudiantes (o todos) podrían no alcanzar el nivel más alto. Aquí es el docente quien, en función de la complejidad de la actitud y las características e intereses de los alumnos, debe definir cuál es su expectativa de desarrollo para un período de tiempo delimitado y comunicarles a sus estudiantes estas metas. Una actividad que podría motivarlos es pegar la progresión en la pared de la sala y que cada alumno se vaya autoevaluando respecto de su progreso, usando algún marcador con su nombre.

TABLA 6: Ejemplo de hipótesis de progresión del dominio afectivo para la actitud “Demostrar curiosidad e interés por conocer seres vivos, objetos y/o eventos que conforman el entorno natural”

NIVEL	DESEMPEÑO QUE DEMUESTRA EL CONOCIMIENTO
Recepción	Escucha a su profesor/a cuando da información de los hábitats y su relación con la vida de los animales.
Respuesta	Realiza las guías de clases asociadas a los hábitats y pregunta elementos específicos para profundizar en la información.
Valoración	Reconoce ante sus compañeros que no le gusta que el ser humano esté destruyendo el hábitat de los animales y que quisiera poder hacer algo.
Organización y conceptualización	Discute con sus compañeros sobre formas de conocer más sobre el hábitat de animales en peligro de extinción y todos acuerdan investigar sobre uno de ellos.
Caracterización e internalización de	Voluntariamente busca información sobre fundaciones de protección del animal definido e invita a sus compañeros a realizar una actividad de apoyo (colecta,

valores	carta, etc.).

Dominio psicomotor de la taxonomía de Krathwohl et al. (1964)

DOMINIO PSICOMOTOR DE LA TAXONOMÍA DE SIMPSON (1972)

La taxonomía de Simpson (1972) fue creada para abordar aspectos educativos del currículum que no estaban cubiertos en los dominios cognitivos y afectivos, y ha sido utilizada ampliamente en educación física y en educación técnica y profesional (Seidel, Perencevich & Kett, 2005). Si bien no existe una definición única respecto de lo que se entiende por aprendizaje psicomotor o habilidad perceptivo-motora, hay un consenso en que se refiere a acciones que implican que el estudiante debe manipular, mover objetos o controlar partes del cuerpo y cuyo desempeño se refleja en movimientos (Adams, 1987; Rosenbaum, Carlson & Gilmore, 2001; Seidel et al., 2005; Singer, 1980). Estas acciones y movimientos pueden implicar motricidad gruesa, es decir, movimientos amplios que involucran desplazamiento por parte del estudiante, como saltar en un pie, tocarse los hombros con las manos, correr de un lugar a otro, o pueden ser de motricidad fina, lo que implica coordinación y precisión en el movimiento, como abrochar un botón, pintar sin salirse del borde, poner una cantidad específica de líquido usando una pipeta o trazar una línea en un papel milimetrado.

En cualquier caso, las habilidades motoras requieren una alta demanda cognitiva en el procesamiento de la información mientras el estudiante aprende y las automatiza, por lo que el dominio cognitivo siempre está presente en una tarea psicomotora (Ackerman & Cianciolo, 2000; Adams, 1987). A medida que el alumno va automatizando la habilidad, la demanda cognitiva decrece, ya que utilizará los heurísticos[1] pertinentes a la situación; no obstante, la transferencia a largo plazo requiere que el estudiante procese cantidades significativas de información para construir patrones a partir de su experiencia en diversas situaciones y en múltiples contextos (Seidel et al., 2005). El ejemplo más común en que se visualiza esta transferencia en el largo plazo es andar en bicicleta: en general uno aprende cuando es niño y aunque haya pasado mucho

tiempo, nos subimos a una bicicleta y comenzamos a pedalear sin mayores complicaciones, pero para lograr esto se requirió muchas horas de práctica y que usáramos distintas bicicletas a lo largo de nuestra vida y que anduviéramos por terrenos de diferentes tipos.

Así, para automatizar un aprendizaje psicomotor se requiere que el estudiante tenga múltiples oportunidades para practicar y avanzar en la adquisición de la habilidad, ya que cada nueva situación implica que vuelva al nivel inicial en una lógica circular de progresión. Por ejemplo, cuando un alumno practica un deporte en equipo, cada partido implica un nuevo aprendizaje, en los entrenamientos se practican jugadas predefinidas para hacerlas cada vez más automáticas, pero al enfrentarse a un oponente, hay elementos que son contingentes, impredecibles con anterioridad y, por lo tanto, se puede volver a un dominio inicial de la habilidad en esa situación específica. La experiencia de aprendizaje de las habilidades motoras va mucho más allá del contexto escolar y la vivimos en forma cotidiana, como cuando escribimos en un computador, nuestra velocidad para mecanografiar puede ser cada vez mayor mientras más practicamos, llegando incluso a tener memorizado el teclado y no necesitar mirar o pensar dónde presionar al escribir, pero si nos cambian el equipo y el teclado tiene otra configuración, cometeremos muchos errores y necesitaremos mayor concentración (demanda cognitiva) para recordar las teclas que ahora son diferentes y poder escribir sin mayores problemas.

La propuesta de Simpson (1972) con las dimensiones de la taxonomía del dominio psicomotor ordena el desarrollo de estos aprendizajes psicomotores en una progresión de siete niveles asociada a la automatización y complejización de los procesos motores (tabla 7).

TABLA 7: Descripción de los niveles de la taxonomía de Simpson (1972) para el dominio psicomotor

Nivel	Características	Ejemplos de desempeño	Sugerencias de verbos
		• Observa un	

Percepción **Observación**	El estudiante usa sus sentidos para obtener claves que guíen la acción: va desde la conciencia del estímulo hasta la traducción de la percepción de la señal a la acción.	procedimiento. • Identifica una herramienta o material. • Selecciona un material entre un grupo.	Observa, escucha, elije, describe, detecta, diferencia, distingue, identifica, aísla, relaciona, selecciona, separa.
Preparación	El estudiante está mental, física y emocionalmente preparado para realizar la acción.	• Manipula el material o herramienta sin usarlo. • Realiza partes de la acción.	Comienza, muestra, explica, mueve, procede, reacciona, responde.
Respuesta guiada **Manipulación**	Conocimiento de los pasos requeridos para realizar una tarea: incluye imitación y ensayo y error.	• Realiza la acción siguiendo instrucciones, pero necesita una demostración o guía.	Copia, sigue, reproduce, evidencia, reacciona, responde.
Mecanización	El estudiante realiza tareas y acciones de manera cotidiana. Es una etapa intermedia en que va creciendo la eficiencia, la habilidad y la confianza.	• Opera herramientas o materiales de forma competente. • Ensambla un material de laboratorio. • Organiza los materiales para hacer un grabado. • Realiza un salto en caballete con imprecisiones en la carrera.	Organiza, manipula, ensambla, calibra, construye, desmantela, despliega, fija, muele, calienta, mide, repara, mezcla, organiza, esboza.
Respuesta compleja * La diferencia con el nivel anterior es precisión y rapidez.	El estudiante ha desarrollado la habilidad y la ejecución es automática, rápida y precisa, o con pequeñas vacilaciones. Las acciones involucran patrones complejos de movimiento.	• Opera herramientas y materiales con precisión. • Realiza un salto en caballete correctamente.	Organiza, manipula, ensambla, construye, calibra, desmantela, exhibe, fija, muele, calienta, mide, mezcla, dibuja.
		• Mezcla colores para generar el que no tiene en una clase de arte.	

Adaptación	El estudiante tiene desarrollada la habilidad y puede modificar patrones de movimiento para dar cuenta de situaciones problemáticas o nuevas sin guía.	• Cambia o agrega un ingrediente de una receta para mejorar el sabor. • Planifica los trabajos en una obra de teatro escolar.	Reorganiza, modifica, reordena, internaliza, adapta, altera, revisa, varía, planifica.
Creación	El estudiante ha internalizado el dominio automático de la habilidad, crea nuevos patrones de movimiento para dar cuenta de situaciones problemáticas o nuevas, crea nuevas tareas que incorporan lo aprendido.	• Evalúa resultados de un procedimiento. • Incorpora los errores detectados en una acción en la vez siguiente. • Diseña una coreografía. • Improvisa para solucionar un problema.	Compone, construye, diseña, inicia, crea, arregla, combina, compone, diseña, origina, realiza, improvisa.

Adaptado de Ferris y Aziz (2005); Kasilingam, Ramalingam y Chinnavan (2014)

A partir de esta taxonomía, Dave (1975) realizó una propuesta, reduciéndola a cinco niveles:

Imitación: fusiona las dos primeras dimensiones de la original (observación y preparación) y agrega que el desempeño puede ser de baja calidad, con lo que se amplía de una manipulación parcial a una realización completa, pero cuyo resultado da cuenta de un desarrollo deficiente, que lo sitúa en un nivel inicial (ej. copiar una obra de arte).

Manipulación: es similar a la tercera dimensión (respuesta guiada), sin hacer modificaciones; supone la realización de la tarea utilizando algún apoyo (ej. crear una obra después de tomar clases o leer instrucciones).

Precisión: fusiona la cuarta y quinta dimensión (mecanización y respuesta compleja) haciendo referencia a la adquisición gradual

de la exactitud y disminución de los errores (ej. tocar y ajustar una canción en guitarra hasta que esté “a punto”).

Articulación: es similar a la sexta dimensión de la original (adaptación); refiere a la coordinación de una serie de acciones que funcionan armónicamente y con consistencia interna (ej. producir un cortometraje que incorpore imágenes como fotos y videos, que tenga música y cambios de color).

Naturalización: es similar a la última dimensión original (creación); supone que el desempeño comienza a ser natural y la persona no necesita pensar en la tarea motora que está realizando (ej. un bailarín profesional, un concertista de guitarrista).

Cualquiera de estas taxonomías que se utilice apunta a describir una progresión del dominio de una habilidad psicomotora, que permite al docente graduar una meta de aprendizaje en desempeños intermedios y monitorear así a sus estudiantes desde que están comenzando a adquirir la habilidad hasta que son expertos o la han automatizado. Veamos ahora un ejemplo de hipótesis de progresión para el objetivo de aprendizaje de Música “Cantar al unísono y tocar instrumentos de percusión convencionales y no convencionales” (tabla 8).

TABLA 8: Ejemplo de hipótesis de progresión del dominio psicomotor para Música

NIVEL	DESEMPEÑO QUE DEMUESTRA EL DESARROLLO DE LA HABILIDAD
Percepción Observación	Percibe visual y auditivamente cómo se tocan distintos instrumentos de percusión como forma de acompañamiento vocal.
Preparación	Manipula instrumentos de percusión, convencionales y no convencionales, sin realizar una ejecución rítmica.
Respuesta guiada Manipulación	Sigue instrucciones para tocar un instrumento de percusión, convencional o no convencional, y ensaya el uso del instrumento según las indicaciones recibidas.
Mecanización	Toca el instrumento de percusión y canta al mismo tiempo, cometiendo errores en su ejecución o dejando de realizar una de las acciones en algunos momentos.
Respuesta compleja	Toca el instrumento de percusión y canta al mismo tiempo, sin errores en su ejecución.

Adaptación	Realiza pequeños arreglos con el instrumento de percusión y/o la voz, modificando uno de los medios expresivos.
Creación	Crea una línea melódica y su respectivo acompañamiento rítmico para tocar un instrumento de percusión y cantar al mismo tiempo.

INTEGRANDO LOS TRES DOMINIOS TAXONÓMICOS EN UNA TAREA

Si miramos los objetivos de aprendizaje de Ciencias Naturales, podemos ver que se espera el desarrollo de habilidades científicas que implican poner en juego los tres dominios descritos anteriormente, pero no todos se presentan con el mismo peso ni la misma temporalidad dentro de una actividad concreta. Como los conocimientos en la realidad no están separados y como los docentes esperamos que nuestros estudiantes aprendan de forma integrada los contenidos, lo ideal es diseñar tareas evaluativas que efectivamente permitan que los alumnos demuestren su nivel de desarrollo o aprendizaje desde una mirada multidimensional. En la figura 1 se puede observar la estructura de una actividad que implica enseñar primero el contenido disciplinar que está a la base (conceptos científicos por trabajar) y explicar la actividad de laboratorio que realizarán, en la cual los estudiantes deben trabajar en duplas. Puesto que está intencionado como aprendizaje también el trabajo en equipo, se deben explicar los aspectos que se espera observar en la dinámica de trabajo (dominio afectivo). El docente debe monitorear que ambos aprendizajes sean comprendidos por los alumnos, de lo contrario la actividad puede fracasar. Posteriormente realizan las acciones de preparación del material de laboratorio y la ejecución del experimento (dominio psicomotor), y sintetizan su trabajo en un informe que luego presentan al curso. El dominio cognitivo está en todas las etapas del proceso de aprendizaje, ya que los estudiantes solo serán capaces de aplicar bien sus habilidades manuales (dominio psicomotor) después de haber adquirido el conocimiento e información conceptual (dominio cognitivo). Del mismo modo, el dominio afectivo asociado al objetivo de aprendizaje de trabajo en equipo estará presente en toda la actividad. Respecto de la evaluación, según sea la etapa de la

actividad, se deben monitorear o certificar los aprendizajes esperados. En la figura 1 se muestra la estructura y los posibles momentos de evaluación.

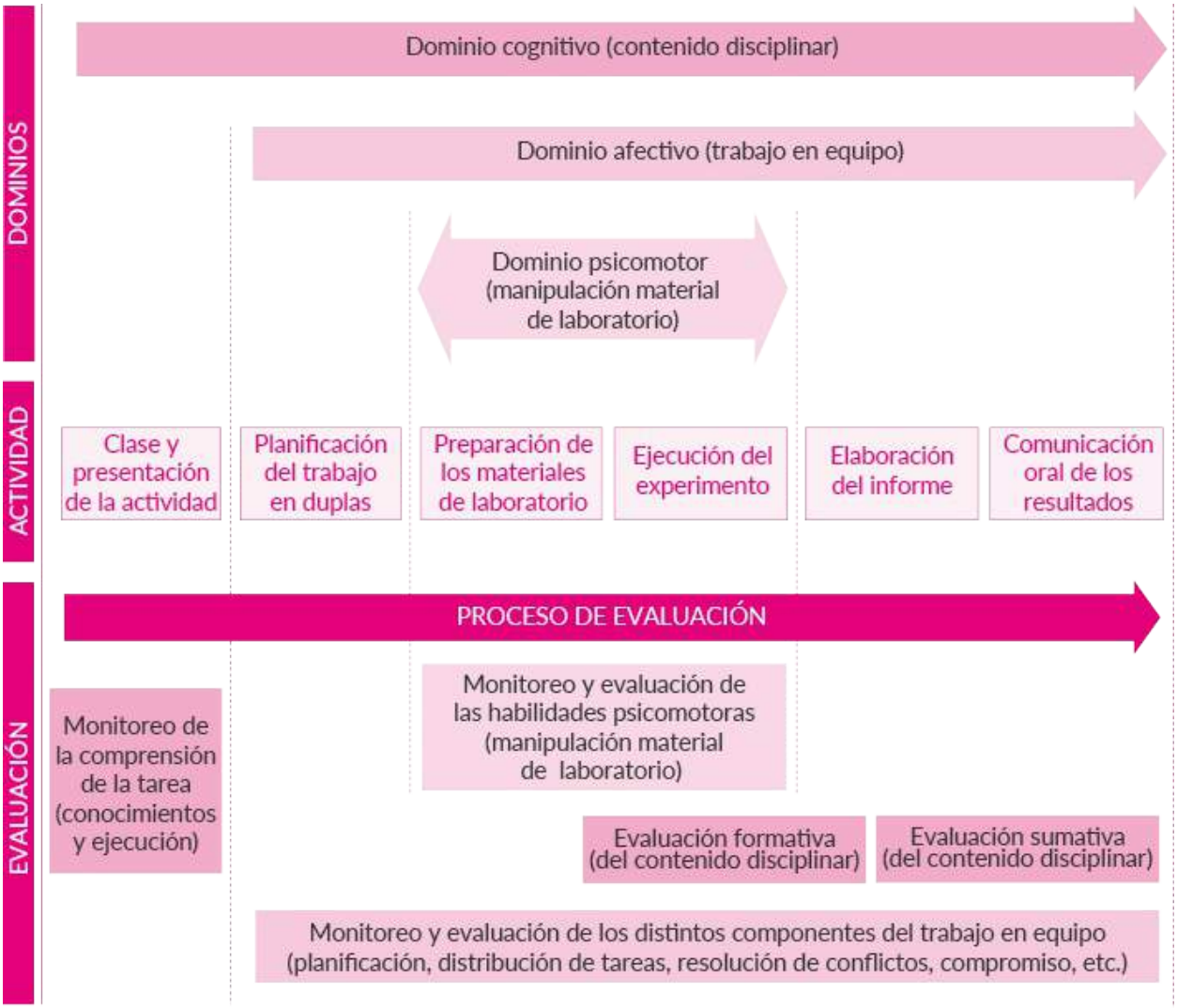


FIGURA 1: Ejemplo de integración de los tres dominios de la taxonomía de Bloom et al. (1956) en una actividad de laboratorio de Ciencias

TAXONOMÍA *STRUCTURE OF OBSERVED LEARNING OUTCOMES* (S.O.L.O.)

Esta taxonomía fue elaborada por Biggs y Collins (1982) y ha sido utilizada principalmente en educación superior, aun cuando hoy en día la educación escolar supone el desarrollo de competencias o

aprendizajes complejos, esta taxonomía es una buena opción para ordenar los desempeños de los estudiantes. Los autores plantean que a medida que los alumnos van consolidando sus aprendizajes, van avanzando en cinco niveles de complejidad estructural crecientes. Biggs (1992) señala además que hay dos tipos de cambios que se pueden observar en el desarrollo de un aprendizaje: los cuantitativos, referidos al aumento en la cantidad de detalles que un estudiante incluye en su respuesta, y los cualitativos, que refieren a la forma en que el estudiante va integrando estos detalles en un modelo estructural. Los niveles cuantitativos se producen primero y luego el aprendizaje va cambiando cualitativamente.

Bajo esta premisa se plantea la taxonomía S.O.L.O. (por su sigla en inglés, *Structural Outcomes Learning Observed*), que proporciona, por una parte, una forma sistemática de jerarquizar en modo creciente el desempeño de un alumno a medida que avanza en el desarrollo de un aprendizaje y, por otra, puede orientar la formulación de objetivos de aprendizaje para su enseñanza y su posterior evaluación. Se presentan cinco niveles que describen el desempeño desde un novato hasta un experto (figura 2). La tabla 9 muestra la descripción de los niveles, ejemplos de una progresión de desempeños y una propuesta de verbos para guiar la formulación de objetivos o indicadores de evaluación.

Como podemos ver en el ejemplo de la tabla 9, la progresión de la complejidad en el aprendizaje está referida a la respuesta que da un estudiante, en este caso, a una pregunta. Es importante tener en cuenta que dicha respuesta puede parecer, en estructura, de nivel más complejo de lo que en realidad es si durante la clase se le enseñaron las relaciones entre los conceptos de forma explícita o se hizo una abstracción desde elementos específicos hasta generalizaciones teóricas o predicciones, entonces debemos tener especial cuidado con la tarea evaluativa, ya que los estudiantes estarán repitiendo una respuesta que ya fue elaborada previamente en la clase y, por tanto, el nivel de complejidad será más bajo de lo que se esperaba.

TABLA 9: Taxonomía S.O.L.O. (Biggs & Collins, 1982)

Cambio	Niveles	Descripción	Ejemplo progresión de desempeño	Verbos
Cuantitativo	Preestructural	El estudiante no ha entendido realmente el tema o tarea, utiliza una manera demasiado simple de hacerla, o la realiza de forma inapropiada, sin conocimientos en el área.	Frente a la pregunta ¿Cuál es la diferencia entre los distintos géneros narrativos?, la respuesta es un vaciado de ideas inconexas en las que utiliza lenguaje disciplinar de forma inapropiada o básica.	(No hay verbos porque no se considera un nivel de desempeño).
	Uniestructural	El estudiante solo conoce un aspecto relevante del área o temática abordada.	Frente a la pregunta ¿Cuál es la diferencia entre los distintos géneros narrativos? solo define qué son los géneros narrativos o solo hace referencia a uno de ellos (ej. la novela).	Lista, nombra, memoriza, identifica.
	Multiestructural	El estudiante conoce y utiliza aspectos relevantes de la tarea, pero los trata de forma independiente y aditiva.	Frente a la pregunta ¿Cuál es la diferencia entre los distintos géneros narrativos? define cada uno, evidenciando dominio conceptual, pero no los compara.	Describe, enumera, clasifica, combina.
Cualitativo	Relacional	El estudiante integra los aspectos relevantes del tema o tarea en una estructura coherente. Este nivel es lo que normalmente se entiende por una adecuada comprensión del contenido.	Frente a la pregunta ¿Cuál es la diferencia entre los distintos géneros narrativos? presenta características que los diferencian, utilizando criterios de comparación y conectores como “mientras que”, “en cambio”, que denotan relación entre los conceptos.	Analiza, explica, integra.
	Abstracción ampliada	El estudiante integra los aspectos relevantes de la tarea en una estructura y los conceptualiza en	Frente a la pregunta ¿Cuál es la diferencia entre los distintos géneros narrativos?, el estudiante da una respuesta relacional pero plantea posibles explicaciones contextuales para ciertos subtipos de géneros narrativos. Ej., señala que las novelas	Predice, refleja,

	un nivel superior de abstracción y generalización hacia un nuevo tema o área de contenido.	de ciencia ficción surgen masivamente en Estados Unidos asociadas al desarrollo de la carrera espacial y argumenta que un nuevo boom se podría dar nuevamente asociado al desarrollo tecnológico actual.	teoriza.

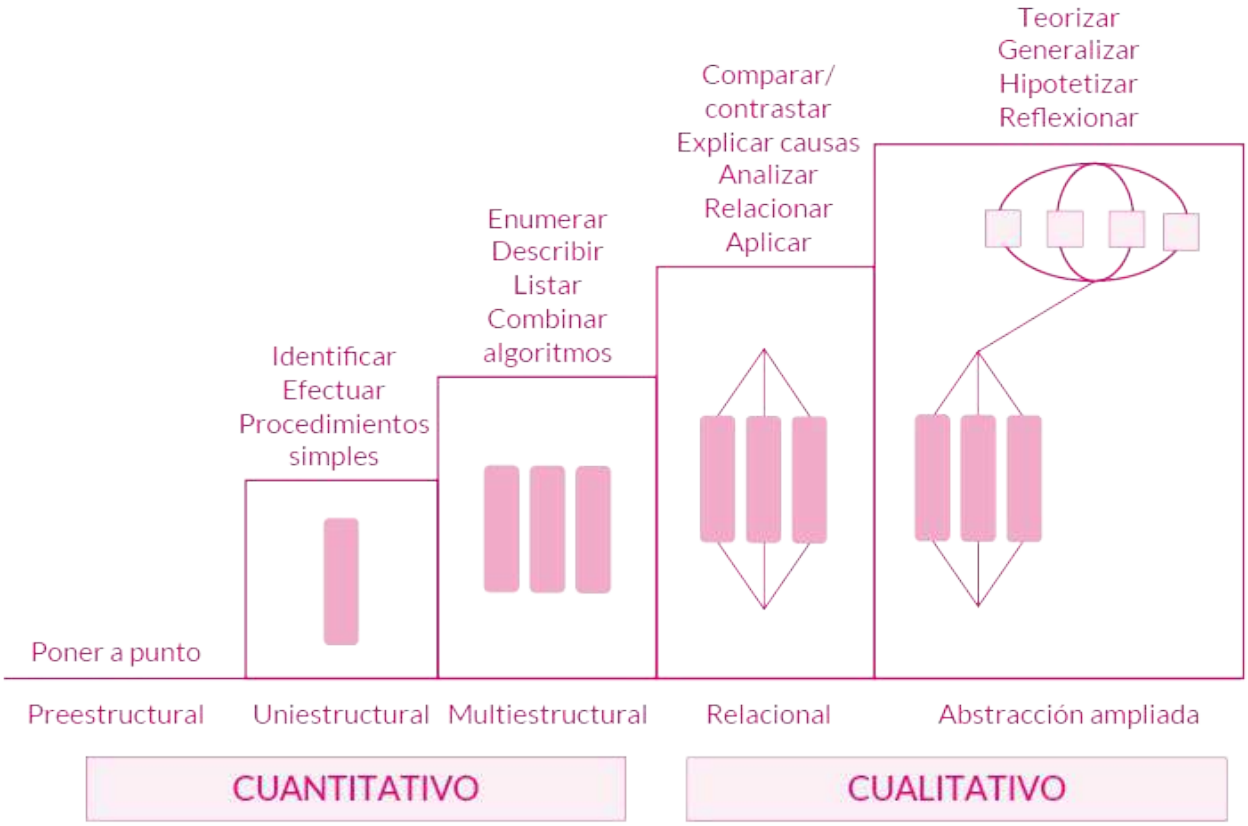


FIGURA 2: Niveles de organización de la información según la taxonomía S.O.L.O.

INDICADORES DE EVALUACIÓN Y DEFINICIÓN DE TAREAS EVALUATIVAS

Las taxonomías no tienen un sentido en sí mismas, fueron diseñadas como una forma de apoyar la formulación de objetivos de aprendizaje y el alineamiento con las tareas que se utilicen para enseñar y evaluar el logro de dichos aprendizajes (Anderson, 2005; Biggs & Tang, 2011). Así, contar con niveles que permitan desarrollar progresiones de aprendizaje es de gran ayuda para la formulación de los objetivos, pues permite al docente centrarse en qué y cómo deben aprender los

estudiantes y no solo en los temas que tiene que enseñar. Este giro hacia el resultado del aprendizaje esperado es crucial hoy en Educación, ya que obliga al docente a identificar la meta a la cual quiere llevar a sus estudiantes, explicitando el contenido y en qué nivel espera que este contenido se aprenda. Los objetivos o resultados de aprendizaje "son declaraciones, escritas desde la perspectiva de los estudiantes, que indican el nivel de comprensión y desempeño que se espera lograr como resultado de participar en la experiencia de enseñanza y aprendizaje" (Biggs & Tang, 2011, p. 100).

La estructura de redacción de un objetivo se basa en un "verbo + un contenido + una condición o contexto" (Anderson, 2005; Goff et al., 2015) (se ha omitido en la estructura de la oración el sujeto, ya que siempre es "el estudiante"). El verbo debe ser una acción observable que represente un conocimiento, una habilidad o un valor (Goff et al., 2015) y tiene como función informar al alumno qué es lo que se espera que realice con ese contenido disciplinar (Biggs & Tang, 2011). El contenido y la condición constituyen una declaración específica del aprendizaje que debe ser demostrado (Goff et al., 2015). Por ejemplo, el objetivo "Explica que una fracción representa la parte de un todo de manera concreta, pictórica y simbólica" supone que no solo el contenido de fracciones debe ser enseñado, también se tienen que dar oportunidades al estudiante de aprender a explicar qué representa una fracción, por lo que esta debe ser enseñada de distintas maneras (concreta, pictórica y simbólica) y luego evaluar dicho objetivo utilizando tareas que permitan al estudiante demostrar que lo aprendió. Además, el verbo "explica" se sitúa según los procesos cognitivos planteados por Anderson et al. (2001) en la categoría "comprender" y el contenido "fracción" asociado a las condiciones lo ubica en un conocimiento "conceptual", lo que nos permite visualizar la complejidad cognitiva del aprendizaje que queremos potenciar.

Goff et al. (2015, p. 8) plantean que hay cuatro elementos que constituyen el ABCD (por su sigla en inglés, Audience Behavior Conditions Degree) a considerar en la formulación de un resultado de

aprendizaje: a) audiencia: ¿quién es el que aprende?, b) conducta: ¿qué será lo que ellos serán capaces de saber, valorar o hacer?, c) condición: ¿bajo qué circunstancias/contexto ocurrirá el aprendizaje? y d) grado: ¿cómo será el desempeño y en qué nivel?

PROBLEMAS COMUNES EN LA FORMULACIÓN DE OBJETIVOS Y RESULTADOS DE APRENDIZAJE

Los objetivos o resultados de aprendizaje esperados, y el desglose que se haga de ellos en indicadores de logro, deben ser escritos como conductas observables, medibles y posibles de alcanzar durante un período de tiempo establecido. Formularlos no es fácil, ya que se requiere una estructura específica y unos verbos que cumplan con ciertos criterios, que sean suficientemente genéricos para dar flexibilidad al docente en su enseñanza y posterior evaluación, pero que no sean tan amplios que resulten ambiguos. A continuación se analizan los tres principales problemas que se observan en su formulación:

PROBLEMA 1: Utilización de verbos ambiguos. Los verbos con que se inician los objetivos o resultados de aprendizaje deben ser precisos. Utilizar verbos como “comprende” o “reflexiona” no orienta al estudiante respecto de la actividad que debe realizar ni en qué nivel se espera que se desarrolle un aprendizaje específico; además, no son directamente observables, ya que aluden a comportamientos internos y se necesitaría leer la mente del estudiante para poder emitir un juicio sobre su aprendizaje. Cuando se le pregunta a un docente ¿cómo sabes que tu alumno “comprendió” un contenido?, la respuesta más común es “me explica, hace comparaciones, infiere, describe…”, aquí es donde aparecen de forma natural los verbos que reflejan conductas observables.

Potter y Kustra (2012) han definido un grupo de verbos y frases que se utilizan habitualmente en la formulación de objetivos de aprendizaje, pero que son ambiguos y no observables, y que deberíamos erradicar de nuestro repertorio:

1) Comprende	7) Acepta	12) Percibe
2) Aprecia	8) Conoce	13) Valora
3) Entiende	9) Es consciente de	14) Consigue
4) Capta	10) Tiene en cuenta	15) Aprehende
5) Sabe	11) Aprende	16) Está familiarizado con
6) Ve		

Cada taxonomía propone una lista de verbos que pueden ser utilizados para formular objetivos de aprendizaje e indicadores de logro, que constituyen un buen recurso, sin embargo, hay algunos verbos que se repiten en más de un nivel, y es aquí donde el docente debe discriminar a qué nivel corresponderán de acuerdo con la complejidad del contenido que acompañe al verbo en la frase (Anderson et al., 2001; Biggs & Collins, 1982).

PROBLEMA 2: En la formulación de los objetivos se confunden los medios con los resultados que se espera alcancen los estudiantes. En este caso, se formulan objetivos de aprendizaje que son actividades de enseñanza o tareas evaluativas. Una característica de un buen objetivo o resultado de aprendizaje es que sea suficientemente genérico y, por lo tanto, el docente tenga múltiples formas de enseñarlo y evaluarlo.

En la tabla 10 se presentan algunos ejemplos de resultados de aprendizaje formulados como actividades y su propuesta como objetivo de aprendizaje o indicador de logro. En el ejemplo 1 se puede ver que el resultado de aprendizaje de Ciencias Naturales “Obtiene información en internet sobre el aporte de algunos científicos relevantes para la comprensión de los fenómenos del entorno” es una actividad porque el foco no está puesto en que demuestre lo que aprendió sobre el aporte de ciertos personajes al desarrollo de la ciencia, sino en la búsqueda de información en internet; de esta forma, el medio o estrategia de enseñanza pasa a ser el fin, con lo que el propósito formativo curricular se desdibuja, además restringe las opciones de búsqueda para los estudiantes solo a un recurso digital y reduce las opciones de tareas evaluativas. Plantear el resultado de aprendizaje como “Explica el carácter provisorio del

conocimiento científico, en base a investigaciones clásicas y contemporáneas" sitúa el desempeño del estudiante en el foco central de la asignatura y es suficientemente genérico como para que el docente pueda utilizar diversas actividades para enseñarlo y también múltiples tareas evaluativas.

TABLA 10: Características y ejemplos de una actividad y de un objetivo de aprendizaje

Actividad	**Objetivo de aprendizaje o indicador de logro**
• Es específica. • Tiene un momento puntual. • Es el medio para evidenciar el indicador.	• Es genérico. • Se evidencia en diferentes momentos. • Es la conducta o desempeño que se va a observar.
1) Obtiene información en internet sobre el aporte de algunos científicos relevantes para la comprensión de los fenómenos del entorno.	1) Explica el carácter provisorio del conocimiento científico, en base a investigaciones clásicas y contemporáneas.
2) Responde la pauta de observación diseñada previamente.	2) Utiliza distintos sistemas de registro para observar un fenómeno natural.
3) Participa en una conversación en pequeños grupos basada en información recabada sobre la contaminación ambiental durante la clase.	3) Interactúa de acuerdo con las convenciones sociales en diferentes situaciones utilizando contenido disciplinar en su argumentación.
4) Formula preguntas para aclarar algunos elementos de la civilización griega de la Antigüedad.	4) Reconoce aspectos cotidianos de la civilización griega en la Antigüedad.

PROBLEMA 3: Elaboración de objetivos o resultados de aprendizaje demasiado específicos. Cuando el objetivo de aprendizaje es muy específico o restringido, puede cumplir con los criterios de ser preciso, observable y medible, pero además constituirse en un arma de doble filo por la poca flexibilidad que da al docente para enseñarlo y luego evaluarlo, perdiendo la relevancia del aprendizaje para la asignatura.

Si el objetivo de aprendizaje lo formulamos tan específico como "Reconoce aspectos de la vestimenta usada por las mujeres de la civilización griega de la Antigüedad en Corinto" se focaliza tanto en un aspecto puntual como es la vestimenta en un lugar de Grecia que se

constituye en un hecho factual, poco relevante, que complica innecesariamente al docente para enseñarlo y luego evaluarlo, con lo cual además se descarta cualquier posibilidad para el estudiante de tener aprendizajes de mayor complejidad cognitiva o desarrollar su creatividad en el proceso de aprendizaje (Potter & Kustra, 2012).

Para evitar cometer estos errores u otros al formular los objetivos o resultados de aprendizaje y los indicadores de evaluación, se presentan cinco preguntas con las que podemos chequear si nuestra formulación está bien hecha (Potter & Kustra, 2012):

1. ¿Este objetivo o resultado de aprendizaje refleja una conducta o desempeño externo, observable y medible?
2. ¿El objetivo o resultado de aprendizaje es suficientemente genérico para ser evaluado con diferentes tareas?
3. ¿Cómo sabrán el docente y los estudiantes cuándo se ha logrado este objetivo o resultado de aprendizaje?
4. ¿Qué evidencia dada por el estudiante consideraré mínima para decir que se ha alcanzado este objetivo o resultado de aprendizaje?
5. ¿Qué conducta o desempeños asociaría con alguien que ha llegado a este estado?

EN SÍNTESIS, en este capítulo se deja clara la importancia de que para realizar una adecuada evaluación se deben conocer los elementos clave del aprendizaje y del desarrollo de los estudiantes. Para ello se resumen algunos elementos que hoy se consideran en el aprendizaje, elementos basales para construir las diversas taxonomías de progreso. Estas clasificaciones fueron creadas para realizar una operacionalización ordenada y progresiva de los aprendizajes. Finalizamos destacando el rol que tienen las taxonomías en Educación y algunos de los problemas que pueden emerger en la formulación de objetivos y resultados de aprendizaje.

En el próximo capítulo veremos cómo aplicar las taxonomías en la planificación de la enseñanza y la evaluación y la relevancia que tiene para el aprendizaje que esos procesos pedagógicos estén integrados desde un inicio.

BIBLIOGRAFÍA

Ackerman, P. L. & Cianciolo, A. T. (2000). Cognitive, perceptual-speed and psychomotor determinants of individual differences during skill acquisition. J*ournal of Experimental Psychology: Applied, 6*(4), 259-290.

Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention and transfer of human motor skills. *Psychological Bulletin, 101*(1), 41-74.

Anderson, L. W. (2002). Curricular alignment: a re-examination. *Theory into Practice*, *41*, 255-260.

Anderson, L. W. (2005). Objectives, evaluation and the improvement of education. *Studies in Educational Evaluation, 31*, 102-113.

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J. & Wittrock, M. C. (Eds.) (2001). *A revision of Bloom's taxonomy of educational objectives*. Boston MA: Pearson.

Bandura, A. & Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. *Journal of Personality and Social Psychology, 45*, 1017-1028.

Biggs, J. & Tang, K. (2011). *Teaching for quality learning at university*. 4th ed. Maidenhead: Oxford University Press.

Biggs, J. B. & Collins, K. F. (1982). *Evaluating the quality of learning: the SOLO taxonomy*. New York: Academic Press.

Biggs, J. B. (1992). A qualitative approach to grading students. *HERDSA News, 14*(3), 3-6.

Bloom, B., Englehart, M., Furst, E., Hill, W. & Krathwohl, D. (1956). *Taxonomy of educational objectives: the classification of educational goals. Handbook I: cognitive domain*. New York, Toronto: Longmans, Green.

Dave, R. H. (1975). *Developing and writing behavioural objectives*. In R. J. Armstrong (Ed.). Tuckson: Educational Innovators Press.

Enghoff, H. (2009). What is taxonomy? An overview with myriapodological examples. *Soil organisms, 81*(3), 441-451.

Ferris, T. & Aziz, S. M. (2005). A psychomotor skills extension to Bloom's taxonomy of education objectives for engineering education. *Exploring Innovation in Education and Research*, 1-5.

Goff, L., Potter, M. K., Pierre, E., Carey, T., Gullage, A., Kustra, E., Lee, R., Lopes, V., Marshall, L., Martin, L., Raffoul, J., Siddiqui, A. & Van Gastel, G. (2015). L*earning outcomes assessment, a practitioner's handbook*. Ontario: Higher Education Quality Council of Ontario.

Heer, R. (2012). A model of learning objectives. Iowa State University Center for Excellence in Learning and Teaching. Recuperado de

www.celt.iastate.edu/teaching/RevisedBlooms1.html

Hutchinson, D., Francis, M. & Griffin, P. (2014). *Developmental teaching and assessment*. In P. Griffin (Ed.), *Assessment for teaching* (pp. 26-57). New York: Cambridge University Press.

Kasilingam, G., Ramalingam, M. & Chinnavan, E. (2014). Assessment of learning domains to improve student's learning in higher education. J*ournal of Young Pharmacists, 6*(4), 27-33.

Krathwohl, D. R., Bloom, B. S. & Masia, B. B. (1964). *Taxonomy of educational objectives the classification of educational goals. Handbook II: affective domain*. London: Longman Group Ltd.

Marzano, R. J. & Kendall, J. S. (2007). *The new taxonomy of educational objectives.* California: Corwin Press.

Marzano, R. J. & Kendall, J. S. (2008). *Designing and assessing educational objectives: applying the new taxonomy*. California: Corwin Press.

Morgan, P., Bourke, S. & Thompson, K. (2001). *The influences of personal school physical education experiences on non-specialist teachers attitudes.* Article presented in The Annual Conference of the Australian Association for Research in Education in Fremantle. Faculty of Education, University of Newcastle.

O'Neill, G. & Murphy, F. (2010). *Guide to taxonomies of learning*. UCD Teaching and Learning Assessment Resources. Recuperado de http://www.ucd.ie/t4cms/ucdtla0034.pdf

Potter, M. K., & Kustra, E. (2012). *A primer on learning outcomes and the SOLO taxonomy.* Centre for Teaching and Learning, University of Windsor. Recuperado de http://www1.uwindsor.ca/ctl/system/files/PRIMER-on-Learning-Outcomes.pdf

Rosenbaum, D. A., Carlson, R. A. & Gilmore, R. O. (2001). Acquisition of intellectual and perceptual-motor skills. *Annual Review of Psychology, 52*, 453-470.

Seidel, R. J., Perencevich, K. C. & Kett, A. L. (2005). *From principles of learning to strategies for instruction*. Empirically Based Ingredients to Guide Instructional Development. New York: Springer Science + Business Media.

Shunk, D. H. (2012). *Teorías del aprendizaje. Una perspectiva educativa*. Naucalpan de Juárez, México: Pearson.

Simpson, E. J. (1972). *The classification of educational objectives in the psychomotor domain*. Washington DC: Gryphon House.

Singer, R. N. (1980). *Motor learning and human performance: an application to motor skills and movement behaviors*. New York: Macmillan Publishing.

* Este capítulo está basado en los resultados del proyecto Fondecyt de Iniciación Nº 11140713, financiado por Conicyt, Chile.

[1] Un heurístico es una regla mental que se utiliza para simplificar la realidad, lo que en la práctica se convierte en un atajo mental que permite un ahorro de procesos cognitivos.

CAPÍTULO 3

Planificando integradamente la enseñanza y la evaluación

Sandra Zepeda Aguirre
Carla E. Förster

INTRODUCCIÓN

La evaluación educacional no es algo reciente, la tarea del profesor en su interacción con los estudiantes en el aula ha incluido desde siempre, como una dimensión fundamental, evaluar los avances en los aprendizajes de cada uno (Shepard, 2006). Lo que sí es reciente es la preocupación por enriquecer esta dimensión, pues hay vasta bibliografía que sostiene que para enriquecer las prácticas pedagógicas de los profesores es necesario mejorar sus prácticas evaluativas, dado que la forma como se evalúa condiciona lo que aprenden los estudiantes (Broadfoot et al., 2002; Deneen & Brown, 2016; Gibbs, 2010; Himmel, 2003; Looney, 2011; Salinas, 2002; Stiggins, 2006).

Diversos autores, entre ellos Gibbs (2010), sostienen que el diseño de las estrategias de evaluación y las situaciones evaluativas tienen una fuerte influencia sobre la cantidad de tiempo que los alumnos estudian, lo que estudian y la calidad de su compromiso con el aprendizaje.

A partir de lo anterior surgen preguntas sobre cómo lograr este compromiso o involucramiento del estudiante, qué estrategias facilitan el aprendizaje, cómo se prepara la enseñanza y cómo se recogen evidencias acerca del aprendizaje, en qué momento evaluar y qué rol juegan los estudiantes en la evaluación. Todas estas preguntas serán parte de los aspectos que examinaremos en el presente capítulo, en el que analizaremos el papel de la planificación de la estrategia de evaluación y los componentes centrales que deben planearse para desarrollar la enseñanza y asegurar el logro del aprendizaje de los estudiantes.

PLANIFICACIÓN PARA EL APRENDIZAJE

Según las orientaciones técnicas del Ministerio de Educación acerca de la planificación, ella puede ser entendida como “un proceso

sistémico y flexible en que se organizan y anticipan los procesos de enseñanza y aprendizaje, con el propósito de orientar la práctica pedagógica en función de apoyar a los y las estudiantes a avanzar hacia el logro de los aprendizajes esperados u objetivos de aprendizaje, propuestos en el currículum nacional" (Mineduc, 2016a, p. 3).

Por su parte, los desempeños asociados al Dominio A del Marco para la Buena Enseñanza se demuestran principalmente a través de las planificaciones, y en los efectos de estas en el desarrollo del proceso de enseñanza y de aprendizaje en el aula. Este Dominio A indica que "el docente, basándose en sus competencias pedagógicas, en el conocimiento de sus alumnos y en el dominio de los contenidos que enseña, diseña, selecciona y organiza estrategias de enseñanza que otorgan sentido a los contenidos presentados; y, estrategias de evaluación que permitan apreciar el logro de los aprendizajes de los alumnos y retroalimentar sus propias prácticas" (Mineduc, 2003, p. 9).

Considerando los antecedentes del contexto institucional, la planificación representa una herramienta profesional clave al servicio de los aprendizajes de los estudiantes. Los profesores debieran analizar profundamente los aprendizajes del currículum nacional y local a la luz de las características de sus estudiantes y, a partir de todos estos elementos, tomar un conjunto de decisiones para diseñar una propuesta didáctica fundamentada y pertinente que organice "temporalmente" tanto el proceso de enseñanza como las estrategias para monitorear los avances en los aprendizajes de sus alumnos. Esta visión reconoce en la planificación una tarea profesional inherente al quehacer pedagógico, pues está vinculada al trabajo reflexivo de los profesores al ir levantando propuestas de diseño de estrategias para desarrollar y hacer seguimiento del aprendizaje de los estudiantes.

La visión de la planificación a la que adherimos considera que la evaluación debe ser parte integral de la planificación de la enseñanza. Ella debe tener como foco los criterios de evaluación desde los cuales se planean las oportunidades de aprendizaje y las estrategias para obtener información acerca del progreso de las metas propuestas.

Además se debe planificar cómo los estudiantes comprenderán los criterios de evaluación, cómo recibirán la retroalimentación, cómo autoevaluarán sus aprendizajes y cómo se les ayudará a progresar aun más (Assessment Reform Group, 2002).

La finalidad principal de la planificación del proceso de enseñanza y de evaluación en conjunto es organizar de forma coherente las prácticas de aula con el propósito de tener una meta de aprendizaje clara, junto con estrategias para lograrla y para monitorear su logro (Mineduc, 2006). Por lo tanto, debiéramos considerar la planificación como una formulación propositiva o tentativa de aquello que se realizará en el aula. Como toda propuesta, puede ser tensionada por la práctica misma, tanto durante su implementación como con posterioridad a ella, lo que implica que podemos (y debemos) modificarla si vemos que algo que planificamos no está dando los resultados que esperamos.

Dados los antecedentes expuestos, es evidente la importancia de la planificación para desarrollar y evaluar el logro de los aprendizajes de nuestros estudiantes, bajo la lógica de que la planificación no solo debe considerar estrategias didácticas de enseñanza, sino que también las estrategias de evaluación. La planificación debiera tener en cuenta de un modo integrado tanto el diseño de la enseñanza como el de las estrategias de evaluación. La figura 1 representa los pasos que deberíamos realizar al planificar la enseñanza y la evaluación integradas, propuestos por Wiggins y McTighe (1998) en su modelo *backward design*, en el que se plantea la idea de hacer un diseño en reversa, primero ver adónde queremos llegar y luego diseñar el cómo llegar ahí (estrategias de enseñanza). En este modelo, el primer paso es identificar el o los aprendizajes que se desea que los estudiantes logren y esto está directamente relacionado con el currículum; el segundo paso es definir cómo ellos van a demostrar que lograron esos aprendizajes, que apunta al modo como se evaluará ese logro, y el tercer paso es diseñar las experiencias u oportunidades de aprendizaje que ayudarán a que los estudiantes desarrollen el o los aprendizajes definidos.

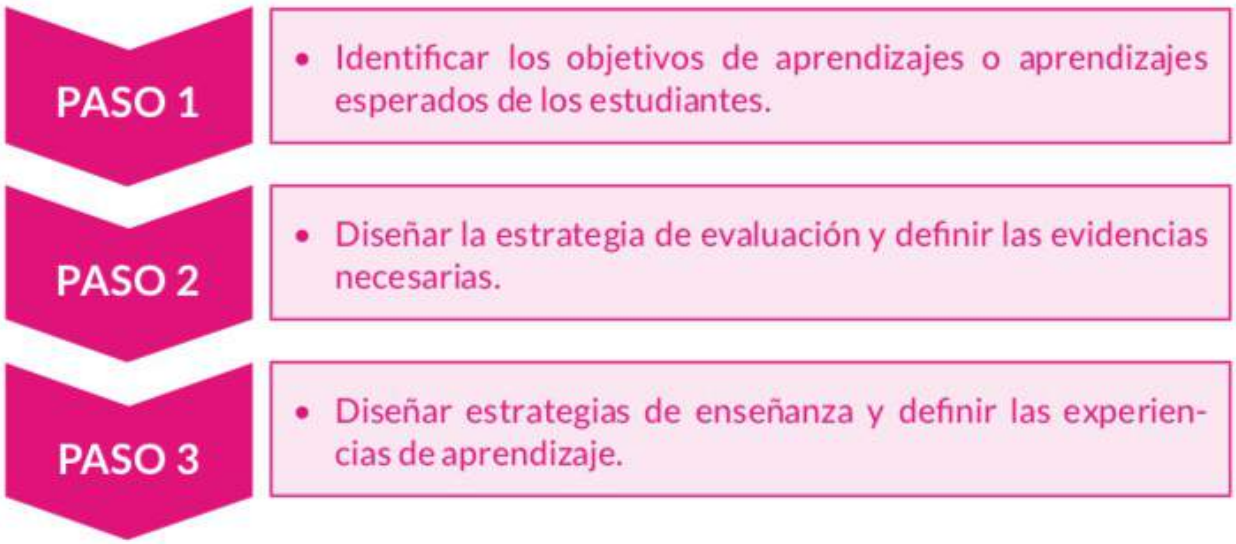


FIGURA 1: Pasos para realizar la planificación de la enseñanza y de la evaluación integradas (adaptado de Wiggins & McTighe, 1998)

COMPONENTES DE LA ESTRATEGIA DE EVALUACIÓN

En términos generales, una estrategia representa un conjunto de decisiones que orientan el logro de un objetivo o propósito, y supone la coordinación de acciones para alcanzar dicho objetivo. Es un proceso regulable, que corresponde a criterios para asegurar una serie de decisiones, coherentes en cada momento del diseño de los pasos y procedimientos, para lograr un propósito que en el ámbito educativo es que nuestros estudiantes logren el objetivo de aprendizaje.

En términos específicos, una estrategia de evaluación representa un conjunto de criterios para tomar decisiones asociadas al proceso de evaluación del aprendizaje de los estudiantes. Considera la definición de indicadores de evaluación, la selección o diseño de los procedimientos para obtener información acerca del progreso hacia los objetivos de aprendizaje, la definición de procedimientos y herramientas para que los estudiantes conozcan y comprendan los indicadores y criterios de evaluación, la definición de instancias y espacios para que los estudiantes reciban retroalimentación oportuna

orientada a la tarea, tengan oportunidades de autoevaluar sus aprendizajes y reciban apoyo para progresar aun más.

Aunque hay múltiples decisiones involucradas en un proceso de planificación, a continuación se describirán en detalle las que nos parecen esenciales para diseñar la estrategia de evaluación:

PRIMERA DECISIÓN: definición de objetivos de aprendizaje y criterios o indicadores de evaluación

La primera y más importante decisión en el diseño de la estrategia de evaluación supone la definición de los objetivos de aprendizaje y los criterios o indicadores de evaluación que serán el foco de la estrategia.

Los objetivos de aprendizaje son parte de la definición curricular nacional y pueden ser contextualizados y enriquecidos a nivel local con las precisiones que se plantean en el proyecto educativo de cada establecimiento. Dada esta condición prescriptiva de los aprendizajes, se plantea que el referente respecto del cual se diseña la estrategia evaluativa es criterial. Esto supone que, para determinar el nivel de logro obtenido, las evidencias que recogemos durante la evaluación de cada estudiante las debemos comparar con los objetivos de aprendizaje e indicadores de evaluación, y el juicio no debiera depender del resultado del resto de los alumnos.

Se espera que los indicadores de evaluación guarden coherencia con los objetivos de aprendizaje. Los indicadores se “desprenden” de los objetivos y corresponden a descripciones que explicitan en términos evidenciables u observables la o las acciones que deberá realizar un estudiante para considerar que ha alcanzado un objetivo de aprendizaje determinado. Como vimos en el capítulo anterior, la formulación de los indicadores de evaluación, al igual que los objetivos y los resultados de aprendizaje, supone una estructura sintáctica que contiene un verbo + un contenido + una condición. Así, esta primera decisión requiere tener claros los aprendizajes (cognitivos, procedimentales y afectivos) que se quieren relevar para definir los indicadores de evaluación que permitirán recoger las evidencias de su logro.

En el currículum escolar posiblemente encontremos propuestas de objetivos de aprendizaje que cumplan con esta estructura y otros que no se ajusten totalmente, también hallaremos algunos que son muy amplios y, por lo tanto, es posible desagregarlos en múltiples indicadores de evaluación, y otros que serán tan acotados como un indicador, en el que su desagregación no será posible. Lo importante como profesores es tener la capacidad para visualizar esta diversidad y distinguir lo que está bien planteado y modificar aquello que nos parezca mejorable, para utilizarlo en el ámbito de la planificación de la estrategia. Se sugiere partir trabajando desde una base, más que elaborar los indicadores desde cero, por lo que la invitación en este punto es a tomar la decisión con argumentos técnicos que permitan justificarla ante otros, y no de forma intuitiva. Debemos recordar nuestra responsabilidad como docentes, ya que lo que está en juego en esta primera decisión es a qué aprendizajes estamos dando énfasis para que los desarrollen nuestros estudiantes.

Este primer componente no solo supone una decisión sobre los objetivos de aprendizaje y sus criterios de evaluación, sino que también se debe diseñar cómo se logrará que los estudiantes los comprendan y orienten sus actuaciones hacia ese logro. Parte importante del éxito de la planificación y en especial de la estrategia de evaluación es que los alumnos comprendan y hagan suyas las metas de aprendizaje propuestas.

SEGUNDA DECISIÓN: definición de la finalidad o intencionalidad de la evaluación

Se debe discernir en qué momento del proceso de enseñanza la evaluación tendrá un foco formativo, con énfasis en la detección de conocimientos iniciales, en el monitoreo, mejoramiento y retroalimentación del aprendizaje de los estudiantes; o bien, una intencionalidad sumativa, cuyo foco será la determinación del nivel de logro de los aprendizajes obtenidos por los estudiantes al finalizar el proceso de enseñanza de un objetivo de aprendizaje (Hattie, 2003; Shepard, 2006; Wiliam, 2000).

La evaluación tiene una intencionalidad formativa inicial (o diagnóstica) cuando el profesor usa las evidencias que recoge para detectar conocimientos previos de los estudiantes, errores conceptuales, valoraciones, creencias y experiencias que pueden ser necesarias de rescatar para considerarlas en el inicio del proceso de enseñanza que se ha planificado (Shepard, 2006).

La evaluación tiene intencionalidad formativa (de proceso) cuando el profesor utiliza las evidencias de aprendizaje de sus estudiantes para monitorear el proceso de aprendizaje, retroalimentarlos respecto de sus avances y espacios de mejora, y considerar ajustes en las estrategias de enseñanza para acompañar sus logros (Black & Wiliam, 1998; Wiliam, 2006).

La evaluación tiene intencionalidad sumativa cuando el profesor usa las evidencias de aprendizaje para dar cuenta del logro alcanzado por los estudiantes en los objetivos de aprendizaje desarrollados durante un período (una unidad, por ejemplo). De esta manera, el docente deberá recoger evidencias suficientes para determinar el nivel de logro de cada uno de los aprendizajes definidos y certificarlos.

Es importante señalar que una misma estrategia de evaluación debiera definir más de una intencionalidad o finalidad, pues ellas ocurren en diversos momentos del proceso de enseñanza (Looney, 2011). Por ejemplo, al iniciar una unidad de aprendizaje, el profesor puede definir realizar una evaluación formativa inicial, pues su finalidad es detectar conocimientos previos de los estudiantes en un determinado contenido asociado a un objetivo de aprendizaje, a través de un instrumento de evaluación como una prueba. Una vez que se avanza en el desarrollo de la unidad, el mismo profesor puede definir hacer una evaluación con carácter formativo, recurriendo a la elaboración de un organizador gráfico como instrumento de evaluación para detectar a través de él cuánto ha mejorado el aprendizaje en desarrollo. En un tercer momento, el profesor puede definir realizar una evaluación con finalidad sumativa para determinar el nivel de logro del aprendizaje definido para la unidad, recurriendo a la elaboración de un proyecto como instrumento de evaluación.

Esta temporalidad se puede dar también asociada a una clase, en la

que los instrumentos o situaciones de evaluación debieran ser menos identificables para un estudiante que una evaluación formal, pero que nos permiten como docentes recoger evidencias de aprendizaje de nuestros alumnos. Para una clase establecemos unas metas de aprendizaje u objetivo a lograr, debemos realizar una evaluación formativa inicial, que puede ser un par de preguntas de discusión que permitan al docente dar cuenta de los aprendizajes de la clase anterior (si es una clase intermedia de la unidad) y así reforzar si algo no quedó claro. Luego, durante la clase, realizamos distintas actividades para enseñar el contenido de la asignatura; ahí deberíamos ir monitoreando cómo están desarrollando estas actividades y si están comprendiendo. Por ejemplo, se podrían revisar los resultados de algunos ejercicios de la guía, que demuestren distintos niveles de comprensión por parte de los estudiantes (ejercicios que representen distintos niveles de complejidad) y, finalmente, certificamos el logro de esta meta específica a través de la elaboración de un mapa conceptual que recogemos al finalizar la clase.

TERCERA DECISIÓN: definición de los procedimientos evaluativos

La tercera decisión apela a la definición de los procedimientos y técnicas evaluativas con las que se recogerán las evidencias acerca de los aprendizajes de los estudiantes.

El profesor debe asegurar que los procedimientos de evaluación seleccionados sean válidos para dar cuenta de los indicadores o criterios de evaluación definidos. Esto significa que deben entregar evidencia central acerca de los aprendizajes que fueron definidos. Junto con ello, el profesor debe verificar que los procedimientos evaluativos seleccionados entreguen evidencia relevante, suficiente y variada sobre los aprendizajes de todos los estudiantes.

Finalmente, se debe cautelar la precisión de la información generada desde la aplicación de los procedimientos e instrumentos, pues su uso es central para la viabilidad en la implementación de la estrategia.

CUARTA DECISIÓN: agentes evaluadores

La cuarta decisión supone discernir sobre quién será el evaluador, si es el profesor, cada estudiante o sus pares. Para responder al agente evaluador o a la pregunta ¿quién evalúa? se han identificado en el aula básicamente dos actores implicados en el proceso: estudiantes y profesores.

La forma más tradicional es la evaluación que realiza el profesor al estudiante, que se denomina heteroevaluación. Este tipo de evaluación tiene a la base el criterio del nivel jerárquico dentro de la sala de clases; en ella el evaluador posee un nivel jerárquico o autoridad superior al evaluado. En cambio, la evaluación de pares apela a un mismo nivel jerárquico en la sala de clases, pues quienes evalúan serían los mismos estudiantes a sus pares. Cuando el alumno se evalúa a sí mismo se denomina autoevaluación.

Incorporar a los estudiantes como agentes evaluativos representa un recurso muy relevante para el aprendizaje por cuanto los sitúa como protagonistas no solo al desarrollar las oportunidades de aprendizaje, sino que además participan en el proceso de evaluación y monitoreo de las mismas.

El profesor puede diseñar la estrategia evaluativa considerando un único agente, o variedad de ellos, lo importante es que su decisión tome en cuenta como criterio qué actor o agente puede otorgar evidencias suficientes y variadas acerca de el o los aprendizajes descritos en los indicadores de evaluación definidos para la estrategia.

ORGANIZACIÓN DE LA ESTRATEGIA DE EVALUACIÓN

Como la estrategia de evaluación es parte integral de la planificación de las acciones para desarrollar los aprendizajes, se propone organizarla en conjunto con la estrategia de enseñanza, para lo cual se podría utilizar una matriz de planificación que tome como foco curricular una unidad (o subunidad) y aborde aspectos clave en el proceso de planificación de la enseñanza y la evaluación para el aprendizaje de los estudiantes. En la tabla 1 se presenta una matriz

de planificación que contiene el diseño de una estrategia evaluativa. En esta tabla se describen los elementos que debemos considerar en cada uno de los componentes.

Luego se presenta un ejemplo de matriz de planificación de 5° Básico para la asignatura de Ciencias Naturales (tabla 2). Esta matriz se centra en la Unidad 2 y considera el desarrollo de un objetivo de aprendizaje clave de la asignatura, junto con objetivos actitudinales que enriquecen la estrategia evaluativa. La unidad se desarrollará en 4 semanas, durante las cuales la estrategia de evaluación considera una evaluación formativa inicial para detectar los conocimientos previos de los estudiantes acerca de los indicadores del objetivo de aprendizaje. La estrategia considera además dos instancias de evaluación formativa durante el desarrollo de la unidad:

- mediante una prueba recoge evidencia acerca de la comprensión de los conceptos clave, previo a la elaboración del proyecto, y
- una evaluación formativa del proyecto propiamente tal.

Finalmente se diseñaron dos instancias de evaluación sumativa:

- se recurre a la prueba para dar cuenta del nivel de aprendizaje de los aspectos conceptuales del objetivo, y
- al proyecto y producto elaborado (prototipo de juego), para dar cuenta del nivel de aprendizaje de los aspectos conceptuales, procedimentales y actitudinales de la unidad.

En esta estrategia se recurre a instrumentos de evaluación como pruebas, elaboración de proyecto, escala y rúbrica. Además el profesor evaluará los indicadores centrales del objetivo de aprendizaje y los pares, el proyecto y los aspectos actitudinales desplegados durante el desarrollo de este.

Se presenta además un ejemplo de matriz de planificación para Enseñanza Media, de Iº Medio, para la asignatura de Historia, Geografía y Ciencias Sociales (tabla 3). Esta matriz de planificación

se centra en la Unidad 3 de aprendizaje, considera tanto el desarrollo del objetivo de aprendizaje de un eje temático como el desarrollo de la habilidad de analizar y evaluar críticamente las fuentes, junto con demostrar actitudes como la convivencia pacífica, propia de la asignatura.

La unidad se desarrollará en 8 a 12 horas pedagógicas, que aproximadamente corresponden a dos o tres semanas, durante las cuales la estrategia de evaluación considera una evaluación formativa inicial que recogerá evidencias sobre el nivel de conocimiento de los estudiantes acerca de la Guerra del Pacífico. Junto con ello, la estrategia propone una evaluación formativa durante el desarrollo de la unidad, que aborda –a través del trabajo grupal– aspectos centrales del análisis crítico de las fuentes y del contraste de visiones sobre los factores, causas y efectos de la guerra. Esta evaluación considera la retroalimentación de profesor, y una autoevaluación del estudiante acerca de los aspectos consignados en el análisis de las fuentes y la dimensión actitudinal de las relaciones pacíficas entre países.

Finalmente, la unidad tiene diseñada una evaluación sumativa que se desarrollará en duplas y que recurre al análisis de fuentes para evaluar los indicadores de la unidad. El profesor evaluará esta prueba de aplicación a través de rúbrica.

Adicionalmente se desarrollará una autoevaluación cuyo centro son la habilidad de análisis crítico de fuentes y la dimensión actitudinal de valorar la importancia de las relaciones pacíficas.

Esta estrategia recurre a la evaluación del profesor y del alumno, y utiliza situaciones tipo prueba en diversas modalidades de aplicación: una pregunta abierta para la evaluación formativa inicial y una prueba de aplicación desarrollada en duplas para la evaluación final. Además elabora una guía para la evaluación formativa, se diseñan una escala para la autoevaluación intermedia y una rúbrica para la autoevaluación final.

TABLA 1: Matriz de planificación de la enseñanza y la evaluación para el aprendizaje

Unidad: indicar el nombre y número de la unidad curricular. Puede considerar la misma unidad planteada en los programas de estudio del Mineduc o una unidad (o subunidad) organizada por el propio docente.
Temporalidad: indicar desde y hasta qué fecha se desarrollará la unidad que se planifica.

Objetivos de aprendizaje	Contenidos (qué se va a enseñar)	Indicadores de logro o criterios de evaluación	Estrategias de enseñanza/ aprendizaje	Recursos	Estrategia de evaluación
Son descripciones de lo que deben aprender los alumnos en relación con los contenidos que se desarrollarán en la unidad. Representan aquellos conocimientos, habilidades, actitudes y formas de comportamiento que se espera que los alumnos logren durante un período de trabajo.	Son la desagregación de los conocimientos expresados en los objetivos de aprendizaje. Se expresan como sustantivos. Por lo general se desagregan en tres ámbitos asociados a los tres dominios descritos en la taxonomía de Bloom: Conceptuales: se refieren a conceptos, principios o hechos que los estudiantes deben saber y que están contenidos en los objetivos de aprendizaje. Procedimentales: corresponden a "secuencias de pasos que pueden ser llevados a cabo para alcanzar una meta" o "técnicas que implican un conjunto de acciones ordenadas, orientadas a la consecución de una meta". Pueden ser hábitos, estrategias (algoritmos y heurísticos), métodos y técnicas.	Son descripciones que explicitan en términos evidenciables u observables la o las acciones que deberá realizar un estudiante para considerar que ha alcanzado un objetivo de aprendizaje. Permiten delimitar lo expresado en el objetivo de aprendizaje y en los contenidos, orientar las conductas que deben observarse en los estudiantes para determinar el nivel de logro que ellos han alcanzado. Estos "aprendizajes específicos" deben ser considerados tanto en la evaluación procesual como en la final.	Son las estrategias, mecanismos y actividades que se desarrollarán para trabajar los contenidos y sus objetivos de aprendizaje. Pueden organizarse por clase o en forma global por unidad.	Corresponden a los materiales que se ocuparán o los textos (libros y documentos, láminas, etc.) que se trabajarán, entre otros.	En esta columna se describirán los aspectos centrales de la estrategia de evaluación de la unidad. Se presentarán los indicadores de evaluación, la finalidad, los procedimientos y el agente de evaluación. Se asume que el referente siempre es criterial, ya que está asociado al currículum vigente y no al rendimiento del curso. • **Indicadores de evaluación:** señalan el o los indicadores que serán evaluados.

TABLA 1: Matriz de planificación de la enseñanza y la evaluación para el aprendizaje (cont.)

Objetivos de aprendizaje	Contenidos (qué se va a enseñar)	Indicadores de logro o criterios de evaluación	Estrategias de enseñanza/ aprendizaje	Recursos	Estrategia de evaluación
Están definidos en los Programas de Estudio, en las Bases Curriculares de Mineduc, bajo el formato de Objetivos de aprendizaje de cada asignatura y Objetivos de aprendizaje transversales que pueden abarcar todas las asignaturas.	Responden a contenidos que implican un saber hacer en contextos determinados, y están presentes en algunos de los objetivos de aprendizaje. Actitudinales: se refieren a valores, actitudes y normas que se van a trabajar conjuntamente con los otros dos tipos de contenidos. Responden a la pregunta ¿además de trabajar contenidos conceptuales y procedimentales, qué aspectos formativos de mis alumnos voy a tratar de fomentar y desarrollar en mi clase? Pueden estar asociados a un objetivo de aprendizaje transversal o a actitudes de la propia asignatura.	Se enuncian de la misma forma que los objetivos de aprendizaje (verbo + contenido + contexto o condición). Deben tener coherencia con los objetivos de aprendizaje y con los contenidos. En el caso de los contenidos conceptuales y procedimentales, se expresan a través de verbos relacionados con las habilidades cognitivas (Anderson et al., 2001) o con las habilidades psicomotoras (Simpson, 1972) que se espera desarrollen.			• **Finalidad** (formativa o sumativa). Explicita la o las finalidades de las evaluaciones de la unidad, indicando en qué momento se realizarán. • **Tipo de procedimientos:** indica el o los instrumentos que se van a utilizar para las evaluaciones de la unidad. • **Agente:** indica quién va a hacer las evaluaciones de la unidad, el profesor o los estudiantes (heteroevaluación, autoevaluación, evaluación de pares).

TABLA 2: Ejemplo de planificación de la evaluación en enseñanza básica

MATRIZ DE PLANIFICACIÓN PARA EL APRENDIZAJE

Unidad: 5° Básico, Ciencias Naturales, Unidad 2: Sistemas del Cuerpo Humano. Subunidad: Sistema Circulatorio.
Temporalidad: 4 semanas.

Objetivos de aprendizaje	Contenidos (qué se va a enseñar)	Indicadores de logro o criterios de evaluación	Estrategias de enseñanza/ aprendizaje	Recursos	Estrategia de evaluación
OA 4: "Explicar la función de transporte del sistema circulatorio (sustancias alimenticias y oxígeno y dióxido de carbono), identificando sus estructuras básicas (corazón, vasos sanguíneos y sangre)" (Mineduc, 2016b). **OA actitudinal** Manifiesta un trabajo riguroso y colaborativo.	**Conceptuales:** • Componentes: corazón, vasos sanguíneos y sangre. • Características y componentes del corazón, vasos sanguíneos, sangre. • Componentes sanguíneos y su función. • Grupos sanguíneos. • Presión sanguínea. • Función: transporte de sangre, sustancias alimenticias, oxígeno y dióxido de carbono. **Procedimentales:** • Interpretación de esquemas y tablas. • Planteamiento de hipótesis sobre causa-efecto en el funcionamiento anómalo de estructuras del sistema circulatorio.	1. Identifica las funciones, estructura y características de los componentes del sistema circulatorio. 2. Relaciona estructuras musculares del corazón con su continuo trabajo de impulsar la sangre a todo el organismo. 3. Infiere la relación de los anticuerpos y los grupos sanguíneos. 4. Reconoce los distintos grupos sanguíneos y su condición de donantes y receptores. 5. Compara los componentes sanguíneos de acuerdo a su función. 6. Interpreta esquemas sobre los niveles de componentes sanguíneos o de la presión sanguínea. 7. Predice problemas que produciría la falta de sangre oxigenada o de nutrientes en algún órgano.	Se realizarán clases expositivas sobre el sistema circulatorio con la finalidad de informarles la creación del proyecto. Luego, se evaluarán conocimientos mediante una prueba (con carácter formativo) que examinará la comprensión de los conceptos clave para el desarrollo posterior del proyecto. **Proyecto:** construir juegos para explicar las funciones de los distintos sistemas.	- Libro de texto otorgado por Mineduc. - Guías con actividades y contenido. - Ppt o láminas con imágenes. - Pizarrón. - Cartulinas. - Tijeras. - Pegamento. - Cartón, etc.	• Referente: indicadores de logro: 1, 2, 4 y 5. • Intencionalidad: diagnóstica (formativa inicial). Se realizará al comienzo de la unidad. • Tipo de procedimientos evaluativos: prueba de selección múltiple e identificación. • Agente: heteroevaluación. • Referente: indicadores de logro: 1, 2, 3, 4, 5, 8, 9. • Intencionalidad: formativa. Se realizará en medio de la unidad (semana 2). • Tipo de procedimiento evaluativo: prueba de ítem V o F con justificación. • Agente: heteroevaluación. • Referente: indicadores de logro: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. • Intencionalidad: sumativa. Se realizará al finalizar la unidad. • Tipo de procedimiento evaluativo: prueba de selección múltiple y preguntas abiertas de aplicación. • Agente: heteroevaluación.

TABLA 2: Ejemplo de planificación de la evaluación en enseñanza básica (cont.)

Objetivos de aprendizaje	Contenidos (qué se va a enseñar)	Indicadores de logro o criterios de evaluación	Estrategias de enseñanza/ aprendizaje	Recursos	Estrategia de evaluación
	• Investigación bibliográfica sobre el sistema circulatorio. **Actitudinales:** • Respeto por el trabajo y opiniones propias y de otros. • Trabajo colaborativo, interacciones efectivas para el trabajo en equipo. • Importancia de seguir normas y procedimientos que resguarden y promuevan la seguridad personal y colectiva.	8. Ilustra el movimiento de la sangre en el corazón y en todo el organismo. 9. Distingue la circulación sistémica de la pulmonar, con su respectiva función. 10. Plantea hipótesis sobre el aumento o el descenso de niveles normales de componentes sanguíneos o de la presión sanguínea. 11. Diseña un proyecto donde se apliquen los conocimientos sobre el sistema circulatorio. 12. Construye un proyecto (aplicado a un juego) con los conocimientos sobre el sistema circulatorio. 13. Aplica normas y procedimientos para resguardar su seguridad y la del equipo de trabajo en el desarrollo del proyecto. 14. Trabaja colaborativa y responsablemente junto a su equipo en el desarrollo del proyecto.	Los estudiantes divididos en grupos deberán crear un proyecto o prototipo de juego que represente la estructura, función e importancia del sistema circulatorio, con preguntas, adivinanzas, predicciones, etc. O bien, de alguna de sus estructuras u órganos. Para ello se define un trabajo de 3 semanas.		• Referente: indicadores de logro: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14. • Intencionalidad: formativa. Se realizará luego de la entrega del proyecto (prototipo del juego). • Tipo de procedimiento evaluativo: guía de elaboración de un proyecto (juego, prototipo) y rúbrica. • Agente: heteroevaluación. • Referente: indicadores de logro: 11, 13, 14. • Intencionalidad: formativa. Se realizará luego de la entrega del proyecto (prototipo del juego) corregido. • Tipo de procedimiento evaluativo: escala de valoración. • Agente: evaluación de pares. • Referente: indicadores de logro: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14. • Intencionalidad: sumativa. Entrega final del proyecto y su producto. • Tipo de procedimiento evaluativo: elaboración del producto, su proyecto y rúbrica. • Agente: heteroevaluación.

TABLA 3: Ejemplo de planificación de la estrategia evaluativa en Educación Media

MATRIZ DE PLANIFICACIÓN PARA EL APRENDIZAJE

Unidad: 1° Medio, Historia Geografía y Ciencias Sociales, Unidad 3: La conformación del territorio chileno y de sus dinámicas geográficas: caracterización e impactos de las políticas estatales de expansión (la planificación aborda uno de sus Objetivos de aprendizaje).
Temporalidad: 8 a 12 horas pedagógicas.

Objetivos de aprendizaje	Contenidos (qué se va a enseñar)	Indicadores de logro o criterios de evaluación	Estrategias de enseñanza/ aprendizaje	Recursos	Estrategia de evaluación
OA 15: Analizar la Guerra del Pacífico considerando el conflicto económico en torno al salitre, el impacto de la guerra en múltiples ámbitos de la sociedad chilena y la ampliación del territorio nacional, y evaluar su proyección en las relaciones con los países vecinos.	**Conceptuales:** consecuencias económicas, sociales, políticas y territoriales de la Guerra del Pacífico. **Procedimentales:** análisis de diversas fuentes. **Actitudinales:** importancia de las relaciones pacíficas.	1. Explica los factores asociados a una guerra y los relaciona con el conflicto chileno. 2. Identifica las consecuencias que tuvo la Guerra del Pacífico en cada ámbito del país. 3. Compara las consecuencias que tuvo la Guerra del Pacífico para los países involucrados. 4. Analiza a través de diversas perspectivas las causas y consecuencias que incidieron en la Guerra del Pacífico, tanto en Chile como en los demás países involucrados.	La unidad se inicia con la recogida y devolución de la información inicial acerca de la Guerra del Pacífico. Posteriormente los estudiantes desarrollan un trabajo grupal en clases recurriendo a guías de análisis de fuentes de diversos historiadores. Explicación de las causas y consecuencias que trajo al país el desarrollo de la Guerra del Pacífico de manera grupal y plenario, luego del análisis de fuentes.	- Diapositivas. - Fuentes (primarias y secundarias). - Imágenes y mapas. - Texto escolar. - Guías con actividades.	• Referente: indicadores de logro 1 y 2. Intencionalidad formativa inicial. • Tipo de procedimiento evaluativo: prueba con preguntas abiertas de respuesta intermedia. • Agente: heteroevaluación. • Referente: indicadores de logro 1, 2, 3, 4, 5, 6 y 7. Intencionalidad formativa. Se realizará en medio del desarrollo de la unidad. • Tipo de procedimiento evaluativo: guía de análisis de fuentes con rúbrica. Escala para la autoevaluación.

TABLA 3: Ejemplo de planificación de la estrategia evaluativa en Educación Media (cont.)

Objetivos de aprendizaje	Contenidos (qué se va a enseñar)	Indicadores de logro o criterios de evaluación	Estrategias de enseñanza/ aprendizaje	Recursos	Estrategia de evaluación
Objetivos de aprendizaje: Habilidades G. Analizar y evaluar críticamente la información de diversas fuentes para utilizarla como evidencia en argumentaciones sobre temas del nivel. Actitudes: F. Demostrar valoración por la vida en sociedad, a través del compromiso activo con la convivencia pacífica, el bien común, la igualdad de hombres y mujeres y el respeto a los derechos fundamentales de todas las personas (Mineduc, 2016c).		5. Contrasta las distintas visiones de las fuentes en cuanto a causas y consecuencias de la Guerra del Pacífico. 6. Evalúa lo que significa la guerra en la relación económica, social, cultural entre los países involucrados (Perú y Bolivia). 7. Reconoce la importancia de las relaciones pacíficas entre los países.	En la siguiente clase se culmina con el análisis de fuentes y se sintetizan las discusiones acerca de los factores, causas y consecuencias de la Guerra del Pacífico, además de examinar los tratados firmados luego de la guerra y los cambios que significaron en los países hasta la actualidad. Discusión en torno a la importancia que tienen las relaciones internacionales del país con el exterior, en especial con los países vecinos, con énfasis en las relaciones pacíficas.		• Agente: heteroevaluación. Autoevaluación (especialmente para indicadores 5, 6 y 7). • Referente: indicadores de logro de toda la unidad (1, 2, 3, 4, 5, 6 y 7). Intencionalidad sumativa. Se realizará al finalizar la unidad. • Tipo de procedimiento evaluativo: prueba de aplicación desarrollada en duplas, que contenga análisis de fuentes y una pregunta argumentativa. Rúbrica para prueba y para autoevaluación, junto con preguntas abiertas. • Agente: heteroevaluación y autoevaluación (evaluación en dupla).

EN SÍNTESIS, a lo largo de este capítulo hemos examinado el peso que tiene la planificación de la estrategia de evaluación en el diseño de las oportunidades de aprendizaje que damos a los estudiantes, y

que ella se potencia y orienta hacia los aprendizajes cuando se planifica de manera integrada con la estrategia de enseñanza. De esta manera, constatamos el papel central que le otorgamos a la evaluación como un proceso permanente y estrechamente vinculado a la enseñanza. Para que esta visión o enfoque de la evaluación pueda desplegarse debemos asumir un enfoque evaluativo que conciba la evaluación formativa y sumativa como parte de un mismo ciclo evaluativo, y no como evaluaciones contrapuestas u opuestas, pues como pudimos revisar en los ejemplos que presentamos en el capítulo, ambas son parte de un mismo ciclo evaluativo y son necesarias para asegurar el aprendizaje de los estudiantes.

En el siguiente capítulo examinaremos las intencionalidades de la evaluación, sus fundamentos, características y cómo ellas pueden implementarse en el aula de manera integrada. Así tendremos más herramientas para incluir en la planificación estas actividades de forma natural.

BIBLIOGRAFÍA

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J. & Wittrock, M. C. (Eds.) (2001). *A revision of Bloom's taxonomy of educational objectives*. New York: Longman Publishing.

Assessment Reform Group (2002). *Assessment for learning: 10 principles research-based principles to guide classroom practice*. London: Assessment Reform Group.

Black, P. & Wiliam, D. (1998). Assessment and classroom learning education. *Principles, Policy and Practice, 5*(1), 7-74.

Broadfoot, P. M., Daugherty, R., Gardner, J., Harlen, W., James, M. & Stobart, G. (2002). *Assessment for learning: 10 principles.* Cambridge: University of Cambridge School of Education.

Deneen, C. C. & Brown, G. (2016). The impact of conceptions of assessment on assessment literacy in a teacher education program. *Cogent Education, 3*, 1225380, 1-14.

Gibbs, G. (2010). *Using assessment to support student learning*. East Anglia, England: Leeds Met Press.

Hattie, J. (2003). *Teachers make a difference: what is the research evidence?* Proceedings of the ACER Research Conference, Melbourne.

Himmel, E. (2003). Evaluación de aprendizajes en la Educación Superior: una reflexión necesaria. *Pensamiento Educativo, 33,* 199-211.

Looney, J. W. (2011). *Integrating formative and summative assessment: progress toward a seamless system?* OECD Education Working Papers, 58, OECD Publishing. Recuperado de http://dx.doi.org/10.1787/5kghx3kbl734-en

Mineduc (2003). *Marco para la Buena Enseñanza*. Santiago de Chile: Ministerio de Educación.

Mineduc (2006). *Evaluación para el aprendizaje.* Santiago de Chile: Ministerio de Educación.

Mineduc (2016a). *Orientaciones técnicas: la planificación como un proceso sistémico y flexible*. Santiago de Chile: Ministerio de Educación.

Mineduc (2016b). *Ciencias Naturales. Programa de Estudio Séptimo Básico*. Santiago de Chile: Ministerio de Educación.

Mineduc (2016c). *Historia, Geografía y Ciencias Sociales. Programa de Estudio Primero Medio*. Santiago de Chile: Ministerio de Educación.

Salinas, D. (2002). *¡Mañana examen! La evaluación entre la teoría y la realidad*. Barcelona: Graó.

Shepard, L. A. (2006). *La evaluación en aula*. México DF: Instituto Nacional para la Evaluación de la Educación.

Simpson, E. J. (1972). *The classification of educational objectives in the psychomotor domain*. Washington DC: Gryphon House.

Stiggins, R. (2006). Assessment for learning - a key to motivation and achievement. *Edge, 2*(2), 1-19.

Wiggins, G. & McTighe, J. (1998). *Understanding by design*. Alexandria VA: Association for Supervision and Curriculum Development.

Wiliam, D. (2000). Assessment: social justice and social consequences: review essay. *British Educational Research Journal, 26*(5), 661-663.

Wiliam, D. (2006). Assessment: learning communities can use it to engineer a bridge connecting teaching and learning. *Journal of Staff Development, 27*(1), 16-20.

CAPÍTULO 4

El fin justifica los medios: intencionalidades de la evaluación

Sandra Zepeda Aguirre

INTRODUCCIÓN

La evaluación puede servir para muchos propósitos y a menudo se distingue entre fines formativos y sumativos. Por una parte, la evaluación sumativa se centra principalmente en la evaluación de los resultados del aprendizaje y, por otra, la evaluación formativa asume como objetivo apoyar el aprendizaje a través de la enseñanza y la retroalimentación específica (Stobart, 2008). Se distinguen en lo formativo una evaluación inicial (diagnóstico) y evaluaciones durante el proceso (continuas). Estos propósitos de la evaluación se presentan en un ciclo, en el que una instancia sumativa puede ser considerada inicial en otro ciclo de aprendizaje (figura 1).

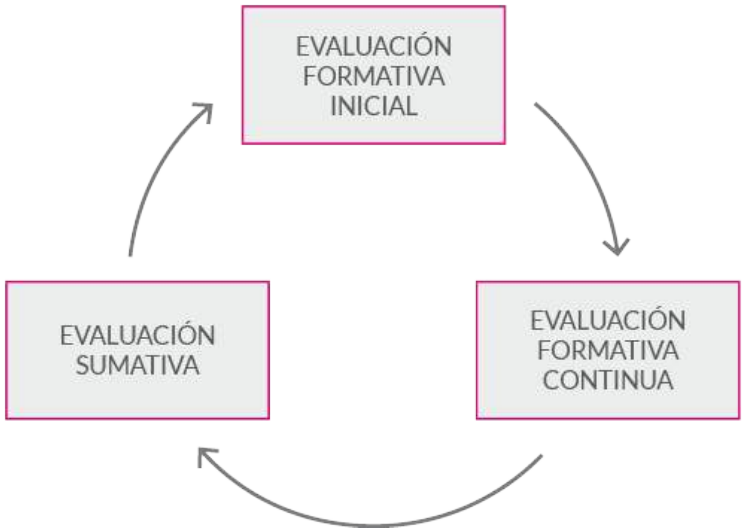


FIGURA 1: Ciclo de la evaluación según intencionalidad

Los profesores están más familiarizados con la forma tradicional de evaluación de los estudiantes y su carácter sumativo, el cual se vincula con pruebas, exámenes finales y calificaciones, y se asocia a una evaluación centrada en la certificación del estudiante. No obstante, la evaluación formativa en el aula también ha adquirido un papel cada vez más visible en los últimos años, por su potencial para

desarrollar aprendizajes en los estudiantes y orientar la enseñanza (Looney, 2011; OCDE, 2005).

A lo largo de este capítulo examinaremos las conceptualizaciones a la base de la evaluación formativa y sumativa, sus aspectos comunes y aquellos que las distinguen. También analizaremos algunas tensiones en torno a ambos propósitos y plantearemos visiones que los integran, como una manera de reconocer el papel que cada una juega en la evaluación de los aprendizajes en aula. Por último, presentaremos los principales usos de la evaluación formativa y sumativa y recomendaciones para implementar estrategias en el aula.

CONCEPTUALIZACIONES DE LA EVALUACIÓN SUMATIVA Y FORMATIVA

En términos generales, se recurre al concepto de “evaluación sumativa” para describir las evaluaciones usadas para certificar el logro de aprendizaje de los estudiantes, en cambio la evaluación formativa se entiende como aquella que proporciona retroalimentación a los estudiantes respecto de cómo mejorar sus aprendizajes, y el concepto de “diagnóstica” se utiliza para aquellas evaluaciones que proporcionan información sobre las ideas y conocimientos previos que tienen los alumnos. Si bien los términos “diagnóstica”, “formativa” y “sumativa” se usan para describir distintos tipos de evaluaciones, los resultados de una misma evaluación podrían servir para más de una función o finalidad en distintos momentos (Wiliam & Black, ١٩٩٦).

Como podemos apreciar, las evaluaciones formativas y sumativas tienen finalidades diferentes. Una hace posible monitorear el proceso de enseñanza y mejorar el aprendizaje y, la otra, ilustra la realización y los logros alcanzados respecto de una meta de aprendizaje. Al rastrear los orígenes de ambos conceptos encontramos que en 1967 Scriven distinguió entre evaluaciones formativas y sumativas, pero Bloom, Hastings y Madaus (1971) fueron los primeros en definir su acepción más aceptada en Educación. Estos autores definieron como evaluación sumativa aquella que se da al final de las unidades, o de

un curso, para juzgar el grado de aprendizaje de los estudiantes, con el propósito de clasificar, certificar y evaluar el progreso o incluso para investigar la efectividad de un currículum.

EVALUACIÓN FORMATIVA

Entendiendo la evaluación formativa como el conjunto de instancias en las cuales se recogen evidencias de los aprendizajes de los estudiantes, para potenciar su desarrollo y logro durante el proceso de enseñanza/aprendizaje, podemos identificar dos momentos: la evaluación formativa inicial, conocida tradicionalmente como diagnóstica, y la evaluación formativa continua. Preferimos hablar de evaluación formativa inicial porque la información que se recoge constituye la primera instancia de un proceso formativo más amplio, mientras que el diagnóstico se asocia en el ámbito escolar a una única instancia, generalmente al inicio del año, restringiendo el potencial del ciclo evaluativo en el acompañamiento que podemos hacer a los estudiantes.

EVALUACIÓN FORMATIVA INICIAL

La evaluación formativa inicial parte desde la convicción de que el conocimiento previo es esencial para el desarrollo de nuevos aprendizajes, pues aprender supone establecer conexiones e integrar una nueva comprensión al conocimiento ya existente. Esta evaluación se desarrolla con el fin de obtener un panorama general de lo que un estudiante sabe, piensa o cree. Levantar evidencias acerca del conocimiento previo de los alumnos incluye el aprendizaje formal adquirido a través de la escolarización, además de las explicaciones y teorías implícitas que ellos tienen sobre cómo funciona el mundo. Ambos conocimientos pueden facilitar u obstaculizar la integración del nuevo aprendizaje, como ocurre con los errores conceptuales (Shepard, 2006). Los profesores pueden recurrir a una variedad de estrategias para recoger evidencias sobre el aprendizaje. Cuanto más información podamos reunir, más precisa será la imagen inicial que tendremos de un estudiante y su aprendizaje en esa unidad de contenido de la asignatura.

Las estrategias eficaces de enseñanza consideran al conocimiento previo de los estudiantes como recurso. Por ello, el rol de la evaluación formativa inicial apunta a recoger dicho conocimiento al principio de nuevas sesiones o unidades curriculares, desarrollando explícitamente en los estudiantes el hábito de preguntarse acerca de lo que saben, reconocen y han hecho respecto de un determinado aprendizaje. Este tipo de evaluación nos entrega valiosa evidencia para adecuar las estrategias de enseñanza y de acompañamiento a nuestros estudiantes, pues permite evidenciar los errores conceptuales, los aprendizajes adquiridos en forma completa, parcial, incompleta o incorrecta, junto con sus creencias y visiones sobre el tema.

Las evidencias acerca del conocimiento previo no solo corresponden a los conocimientos formales que han adquirido los estudiantes, también incluyen patrones de lenguaje y formas de pensar que estos desarrollan por medio de sus roles sociales y sus experiencias culturales. Hay muchas estrategias que pueden usarse para determinar dónde están los estudiantes al principio de una sesión o unidad y qué necesita suceder en el aula para alcanzar con éxito las metas de aprendizaje; lo importante para que este tipo de evaluación cumpla con su propósito es que la información recogida permita la toma de decisiones oportuna por parte del docente y sus alumnos.

EVALUACIÓN FORMATIVA CONTINUA

Black y Wiliam (2009) aportan la siguiente definición de la evaluación formativa: “La práctica en una clase es formativa en la medida en que la evidencia acerca de los logros de los estudiantes es obtenida, interpretada y usada por docentes, aprendices o sus pares, para tomar decisiones sobre sus próximos pasos en la enseñanza, que tengan probabilidades de ser mejores, o mejor fundadas, que las decisiones que ellos hubieran tomado en la ausencia de la evidencia que fue obtenida” (p. 6). Esta definición pone el acento en la toma de decisiones para apoyar o mejorar el aprendizaje; se considera por ello que la evaluación formativa debiera ser continua y estar centrada en

el progreso de los estudiantes para identificar las necesidades de aprendizaje y dar forma a la enseñanza (OCDE, 2005).

Adicionalmente, la evaluación formativa se enmarca en un proceso de aprendizaje activo e intencional que une a profesores y estudiantes para recoger, constante y sistemáticamente, evidencias del aprendizaje con el objetivo explícito de mejorarlo (Moss & Brookhart, 2009). Esta nueva mirada pone énfasis en que la evaluación formativa es considerada como parte integral del proceso de enseñanza/aprendizaje, y no un complemento de él (Deneen & Brown, 2016; Shepard, 2006).

Así, el propósito primordial de la evaluación formativa continua es mejorar el aprendizaje mientras el estudiante está aprendiendo y no auditarlo, entendiéndola como una evaluación para el aprendizaje en lugar de una evaluación del aprendizaje; por ello, la evaluación formativa continua se utiliza principalmente para el beneficio de los estudiantes como apoyo a sus aprendizajes (Moss & Brookhart, 2009). La distinción crucial de esta evaluación es que solo podría ser catalogada de formativa si ella apoya o estructura aprendizajes posteriores y cuando la evidencia recogida se utiliza realmente para adaptar el trabajo docente atendiendo a las necesidades de estos estudiantes (Black & Wiliam, 1998; Wiliam, 2009).

Para que la evaluación formativa continua se despliegue en el aula de manera efectiva debe: 1) estar intrínsecamente conectada con el ciclo de enseñanza y aprendizaje, permitiendo el monitoreo y ajuste frecuente (Sadler, 1989; Shepard, 2000, 2006); 2) compartir con los estudiantes los objetivos y criterios de calidad esperados, y 3) entregar retroalimentación que le permita al estudiante mejorar su aprendizaje. Este tercer punto será tratado con detalle en el capítulo siguiente.

El enfoque de evaluación formativa fue explicado más ampliamente por Atkin, Black y Coffey (2001) a partir de tres preguntas clave (figura 2). La primera apunta a identificar el aprendizaje o meta a lograr, la segunda identifica el lugar en el que se encuentra el aprendizaje del estudiante en relación con la meta, y la tercera se hace cargo de la brecha entre la meta y la situación actual del

aprendizaje, focalizándose específicamente en lo que se necesita para alcanzar el objetivo.

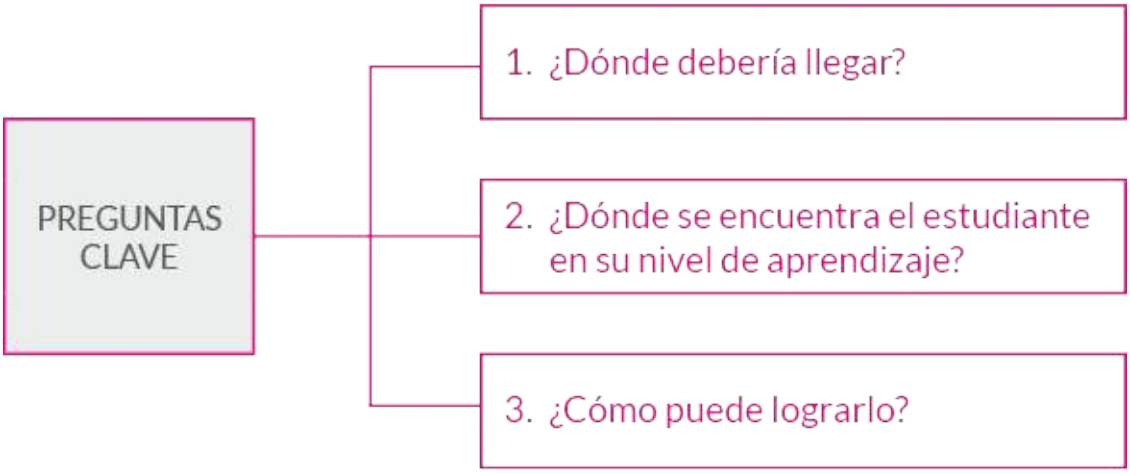


FIGURA 2: Preguntas clave de la evaluación formativa (Atkin et al., 2001)

La evaluación formativa continua recurre a una variedad de estrategias para revelar la comprensión de los estudiantes respecto de una temática, permitiendo que los profesores identifiquen y aborden las dificultades detectadas en el progreso de aprendizaje de sus alumnos. A partir de los datos formativos, los docentes deben decidir cuánto y qué tipo de apoyo y práctica requiere un estudiante para alcanzar la meta. Cuando la evaluación formativa se emplea al inicio, durante y al finalizar la enseñanza, tanto los profesores como los estudiantes pueden orientarse hacia la mejora, pues tienen una medida de progreso tangible a la cual recurrir (Rasmussen, 2017).

Sin duda la evaluación formativa inicial es el primer paso de la evaluación formativa continua para mejorar el logro de los aprendizajes (Moss & Brookhart, 2009). A modo de síntesis se presenta el siguiente esquema, que contiene los aspectos clave de la evaluación formativa descritos en este apartado (figura 3).

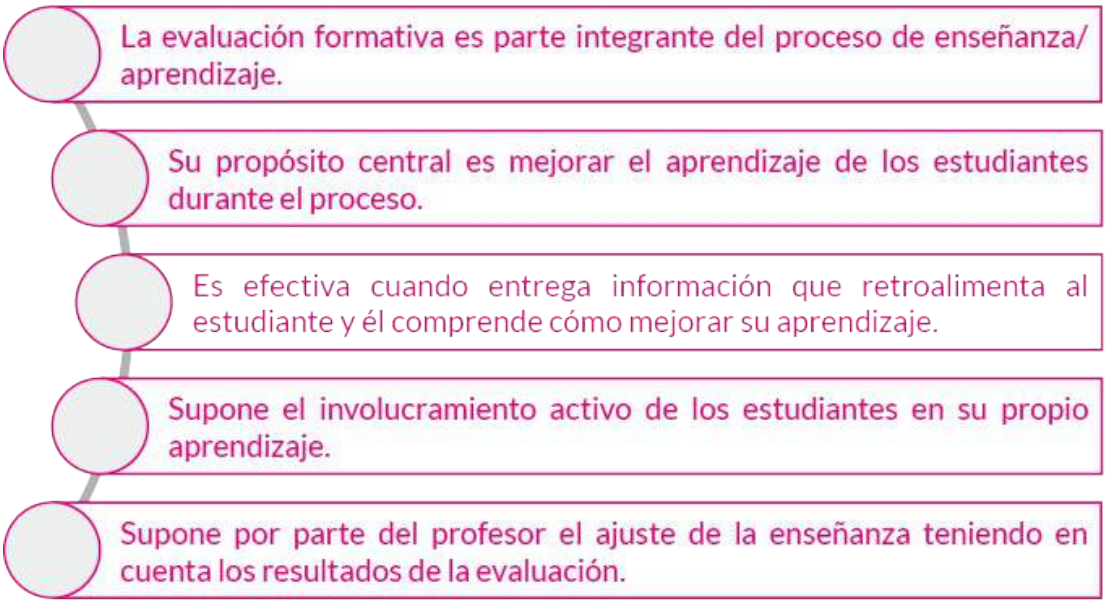


FIGURA 3: Elementos clave de la evaluación formativa

EVALUACIÓN SUMATIVA

Como ya se ha descrito, las evaluaciones sumativas pueden concebirse como parte de un ciclo, o como hitos continuos importantes de aprendizaje que fueron andamiados en las evaluaciones formativas. Desde esta visión, las evaluaciones sumativas ilustran la realización y los logros alcanzados por un estudiante en una meta de aprendizaje (Shepard, 2006).

Las evaluaciones sumativas tienen el propósito de dar cuenta de lo que los estudiantes han aprendido al final de una unidad o período lectivo, para asegurarse de que cumplieron con los estándares requeridos y de esta manera obtener la certificación. Dicha evaluación permite legitimar un proceso de aprendizaje, asignando una valoración o juicio al finalizar un período, con el propósito de promover el avance de los estudiantes al siguiente curso o acceder a otros sistemas educativos o profesionales (OCDE, 2005).

La evaluación sumativa posee más de un uso, ya que hay una variedad de maneras en que se utiliza la información sobre el logro del estudiante en un momento determinado. Estos se pueden agrupar en dos principales: internos y externos a la comunidad escolar. Entre

los usos se incluyen el seguimiento del progreso de los estudiantes en el establecimiento educacional; informar a los padres, alumnos y al próximo profesor lo que se ha logrado; certificar o acreditar el aprendizaje mediante el uso de calificaciones regulares para el mantenimiento de registros, etc.

Se ha visto que la evaluación sumativa realizada por los profesores tiene un impacto más positivo en el aprendizaje cuando sus resultados se integran en la práctica pedagógica diaria del docente que cuando se concentran en ocasiones puntuales (Harlen & Winter, 2004).

Uno de los aspectos que comúnmente se vinculan con la evaluación sumativa son las calificaciones, pero relacionarlas solo a la evaluación sumativa representa una visión restringida del papel que ocupa la calificación en el proceso de evaluación. Diversos autores sostienen que la calificación debe ser coherente con el proceso de evaluación completo, es decir, ella debe representar el aprendizaje demostrado por el estudiante. Esto implica que para calificar con una nota que se reconozca como final, el alumno debe haber sido monitoreado considerando una evaluación que cumpla con las siguientes características: 1) estar vinculada con los objetivos de aprendizaje (que deben ser conocidos y comprendidos por los estudiantes); 2) se apliquen procedimientos de evaluación válidos para recoger evidencias (alineados con las metas de aprendizaje); 3) se otorgue retroalimentación acerca del desempeño y que el estudiante tenga oportunidades para reelaborar o rehacer su tarea evaluativa, y 4) que culmina con la certificación del aprendizaje (Dyer, 2014; Shepard, 2006).

Los resultados de la evaluación sumativa de aula normalmente implican algún tipo de reducción de datos (p. ej. promedio de un período académico; asignación de una calificación basada en notas ponderadas), pero que exige que dicha reducción mantenga la determinación del nivel de logro de cada objetivo de aprendizaje evaluado. Por lo anterior se debe destinar tiempo al diseño de reportes útiles para la enseñanza, que incluyan información cualitativa junto a las puntuaciones obtenidas en cada objetivo de aprendizaje o

indicadores de evaluación, considerando, por ejemplo, información sobre respuestas correctas/incorrectas para los objetivos de aprendizaje o sus indicadores de evaluación, y comentarios sobre lo que implica una respuesta incorrecta complementados con sugerencias para abordar los próximos pasos (Marion, 2007).

DISTINCIONES Y ARTICULACIONES ENTRE EVALUACIÓN FORMATIVA Y SUMATIVA

En la actualidad coexisten en los docentes creencias y prácticas contradictorias. Por una parte está instalada en el discurso la creencia de que la evaluación formativa y sumativa deben ser parte de un solo sistema de evaluación pero, por otra, lo que ocurre en el aula es la supremacía de la evaluación sumativa, invisibilizando la finalidad formativa (Deneen & Brown, 2016; Wiliam, 2000).

Varios autores (ver p. ej. Hattie, 2003; Rea-Dickins, 2006; Shepard, 2006; Wiliam, 2000) cuestionan la visión dicotómica entre evaluación formativa y sumativa y proponen una mirada complementaria e integrada de los dos procesos. La evaluación en el aula debe ser sumativa y formativa, pues establece lo que los estudiantes pueden hacer en un momento determinado, recurre al monitoreo de su progreso y retroalimenta la enseñanza como un medio para apoyar el aprendizaje. Bajo este punto de vista, la evaluación no debiera ser solo formativa, como una herramienta de enseñanza, o solo sumativa, como una herramienta para documentar el logro alcanzado por el estudiante, sino que ambas debieran ser utilizadas conjuntamente para reorientar tanto la enseñanza como el aprendizaje (Wolfel, 2009).

El desafío que plantean diversos autores en cuanto a las finalidades y usos de la evaluación (ver p. ej. Gibbs, 2010; Shepard, 2006; Wiliam, 2000) es que las prácticas formativas y sumativas de evaluación en aula sean coherentes y se respalden mutuamente, que las dos tengan como foco el aprendizaje de los estudiantes y la mejora del proceso de enseñanza. No se espera que compitan, ni menos que una invisibilice a la otra, lo que profesores y alumnos necesitan son

evaluaciones formativas que monitoreen los progresos parciales o continuos del aprendizaje y evaluaciones sumativas que utilicen los mismos criterios para verificar la consecución de logros importantes en la adquisición de competencias por parte de los estudiantes (Shepard, 2006).

Para Hattie (2003), la distinción entre evaluación formativa y sumativa se relaciona más con el momento de las interpretaciones y las decisiones que se adoptan que con la forma o procedimientos con los que se recogen las evidencias. Este autor advierte que si realizamos evaluaciones formativas menos rigurosas que la evaluación sumativa, tanto al inicio como durante el proceso de aprendizaje, es más probable que encaminemos a nuestros estudiantes en una dirección equivocada, o que les demos recomendaciones que no sean lo suficientemente precisas como para acortar la brecha de aprendizaje, socavando la precisión de las correcciones o ajustes que el alumno pueda hacer. Por lo anterior, hacemos un llamado a no minimizar la exactitud en la evaluación formativa inicial y de proceso, porque justamente la retroalimentación entregada en esos momentos tiene un alto potencial para asegurar aprendizajes de calidad.

La lógica propuesta en este apartado de integración de la evaluación formativa y sumativa en un mismo ciclo evaluativo busca que tanto profesores como alumnos se unan con el propósito de crear una forma diferente de aprendizaje, pues, a partir de la experiencia evaluativa y en especial de sus resultados, ambos cambian: los estudiantes van haciendo propios los aprendizajes y los profesores pueden ir ajustando sus estrategias de enseñanza para mejorar las oportunidades de aprendizaje (Wiliam, 2000).

TENSIONES ENTRE LA EVALUACIÓN SUMATIVA Y FORMATIVA

En este apartado describiremos las principales confusiones, malentendidos y tensiones que se reconocen en las prácticas evaluativas, asociadas tanto a la evaluación sumativa como a la evaluación formativa:

1. Generalmente se vincula la evaluación sumativa solo con la calificación, ocultando en primer lugar la tarea central de este tipo de evaluación, que es la determinación y certificación del nivel de logro de los aprendizajes y, en segundo lugar, separando este tipo de evaluación del ciclo evaluativo completo, en que también la evaluación formativa es parte integral.
2. En la práctica evaluativa se reconocen distorsiones respecto del vínculo entre la evaluación sumativa y la calificación, lo que constituye una seria amenaza hacia el enfoque y la convicción de que la evaluación aporta significativamente al logro de los objetivos de aprendizaje.
    - En primer lugar, las pruebas y las tareas calificadas comunican lo que es relevante aprender; si ellas no son coherentes con las metas del aprendizaje que se valoran, entonces los estudiantes concentrarán su atención y esfuerzo solo en las tareas calificadas, pudiendo generar lo que se ha llamado “estrechamiento curricular” (Wiliam, 2000).
    - Una segunda práctica es el uso de las calificaciones como premio o castigo, desvinculándolas completamente de los aprendizajes (Shepard, 2006). Por ejemplo, en situaciones de mal comportamiento se amenaza a los estudiantes con hacer una prueba sorpresa o se ofrece asignar un punto extra a quienes se portan bien.
3. Respecto de la evaluación formativa, muchas veces los docentes realizan prácticas que se asocian a esta intencionalidad, pero que se distancian de sus premisas centrales:
    - Bajo la convicción de que la retroalimentación puede tener un impacto sustancial en el aprendizaje, una serie de estudios muestra prácticas de retroalimentación enjuiciadoras del estudiante, excesivamente directivas, deshonestas, desprolijas, vagas, sin dirección o centradas en aspectos no mejorables, las cuales son percibidas como irrelevantes y frustrantes por los estudiantes y que

finalmente no conducen a mejorar el aprendizaje (Hargreaves, 2013; Kay & Knaack, 2009).

- Un aspecto muy relevante de la evaluación formativa corresponde al uso del error como recurso para el aprendizaje; a pesar de ello, hay estudios que indican que los profesores no abordan explícitamente en la retroalimentación los errores conceptuales comunes a una cierta disciplina (Heitink, Van der Kleij, Veldkamp, Schildkamp & Kippers, 2016). Además se reconocen prácticas en las que el error es penalizado, asignándole valor solo al final del proceso de enseñanza, cuando se están evaluando los aprendizajes para certificarlos (Moss, 2013).
- Respecto de la posibilidad de mejorar el aprendizaje, considerado uno de los aspectos clave en la evaluación formativa, se observan prácticas en que los profesores analizan con sus estudiantes los resultados de una prueba, revisando especialmente las respuestas incorrectas, una vez que ya los han calificado. Con esta acción no sería posible aplicar la premisa de que el aprendizaje puede mejorar, pues no se les ofrecería a los estudiantes la opción de re-corregir o re-elaborar (Shepard, 2006).
- Respecto del rol central del estudiante en la evaluación formativa, estudios recogen prácticas en las que los estudiantes “autocalifican” y no autoevalúan sus aprendizajes, prácticas en las que los estudiantes evalúan entre pares y a sí mismos sin haber comprendido completamente los criterios de evaluación ni haber sido preparados para dar y recibir una retroalimentación acerca del aprendizaje propio y de sus pares (Heitink et al., 2016).
- Por último y en relación al diseño de situaciones de evaluación, uno de los elementos que se espera a través de las instancias de evaluación formativa es que el estudiante asuma un rol activo en su proceso de

aprendizaje. No obstante lo anterior, estudios consignan prácticas de evaluación que no consideran la motivación o involucramiento del estudiante en ellas, que recurren frecuentemente a la aplicación de pruebas descontextualizadas y a tareas de evaluación rutinarias, poco significativas y no auténticas que generan escasa motivación y bajo involucramiento en los estudiantes (Heitink et al., 2016).

BUENAS PRÁCTICAS EVALUATIVAS FORMATIVAS Y SUMATIVAS EN AULA

Asumiendo que la evaluación sumativa y formativa constituyen partes de un mismo ciclo y, por lo tanto, se complementan mutuamente, en este apartado examinaremos cómo pueden desarrollarse prácticas evaluativas en aula en las que se integren ambas intencionalidades:

PRIMERA PRÁCTICA. Metas claras: los objetivos de aprendizaje y sus criterios de evaluación

Tanto la evaluación sumativa como la formativa se guían por los objetivos de aprendizaje planteados en el currículum. Una buena práctica evaluativa debe considerar, en primer lugar, la definición clara de los objetivos de aprendizaje y sus criterios de evaluación por parte del profesor. Se espera que esta definición se plasme en la planificación de su estrategia evaluativa (que fue examinada en el capítulo anterior) y se mantenga coherentemente durante todo el ciclo evaluativo.

En segundo lugar, se espera que los estudiantes logren una comprensión clara de los objetivos de aprendizaje y sus criterios de evaluación; en este sentido, se recomienda al profesor dedicar un espacio en la clase para que los alumnos puedan comprender y asimilar cuál es la meta de aprendizaje y sus criterios de calidad.

SEGUNDA PRÁCTICA. Uso de los conocimientos previos

Para que el ciclo de evaluación se dé virtuosamente se necesita implementar la evaluación formativa inicial para recoger evidencias de

los conocimientos previos de los estudiantes, tanto formales como aquellos asociados a explicaciones implícitas acerca de los nuevos aprendizajes. Estas evidencias resultan clave para que el profesor valide o replantee las estrategias de enseñanza que ha planificado para abordar la unidad o clase (Shepard, 2006). Se puede recurrir a diversas estrategias y procedimientos de evaluación, pero lo más relevante es que permitan recoger evidencias suficientes y variadas de los aprendizajes previos y que se tomen decisiones oportunas a partir de estas evidencias.

TERCERA PRÁCTICA. Uso del error para aprender

Durante todo el proceso de enseñanza, los estudiantes pueden cometer errores y aprender de ellos sin penalización. Los errores solo se convierten en oportunidades de aprendizaje cuando se proporciona retroalimentación que acorte la brecha hacia la meta, pues los estudiantes debieran tener la opción de
re-escribir o re-elaborar su aprendizaje. Los errores además deben ser analizados acuciosamente por el profesor para detectar la presencia de errores conceptuales, de organización o ejecución, y a partir de este análisis tomar decisiones pedagógicas adecuadas. Esta práctica evaluativa es básicamente formativa y prepara al estudiante para certificar su aprendizaje (Dyer, 2014).

CUARTA PRÁCTICA. Evaluación diversificada

El ciclo debe ser diseñado considerando que la evaluación es un proceso continuo, que recoge evidencias permanentemente sobre el avance del aprendizaje de los estudiantes hasta su certificación. Por lo anterior, se debieran diseñar e implementar estrategias evaluativas que consideren contextos auténticos, desafíos variados, con distintos formatos, de modo que se reúnan evidencias válidas y suficientes para emitir juicios sobre el aprendizaje individual de los estudiantes, y que permitan tanto ajustarlo como certificarlo.

QUINTA PRÁCTICA. Retroalimentación efectiva

Se espera que la retroalimentación sea efectiva, específica y oportuna, es decir, que permita que el alumno ajuste o reelabore su aprendizaje, por lo tanto, esta debiera desarrollarse antes de su

certificación y calificación. La retroalimentación debe proporcionar detalles sobre cómo mejorar en lugar de simplemente indicar si el trabajo de un estudiante es correcto o no (Shute, 2008). Los elogios generales o los comentarios personales comúnmente no son útiles para mejorar el aprendizaje (Rasmussen, 2017).

SEXTA PRÁCTICA. El estudiante como protagonista

La práctica evaluativa en aula debe ser diseñada e implementada para que los estudiantes posean algún control sobre su proceso de aprendizaje y que participen activamente en dicho proceso. Además de que las metas sean explícitas, los estudiantes deben conocer lo que están aprendiendo, saber qué productos o
desempeños se espera que elaboren o demuestren, y cómo serán evaluados (Dyer, 2014). Se recomienda recurrir a la autoevaluación y evaluación de pares especialmente en el momento formativo de este ciclo, y emplear autoevaluaciones o registros reflexivos acerca del aprendizaje en el momento sumativo del ciclo. Este punto se profundizará en los capítulos siguientes.

SÉPTIMA PRÁCTICA. Coherencia entre retroalimentación y evaluación sumativa

La práctica evaluativa en aula debe asegurar coherencia entre la retroalimentación entregada en el momento formativo y la determinación del nivel de logro del aprendizaje, informado en el momento sumativo. En esta práctica en aula debemos asegurar que la certificación del nivel de logro del aprendizaje considere como insumo principal la retroalimentación entregada al estudiante, teniendo como base las mismas metas de aprendizaje y criterios de evaluación de todo el ciclo evaluativo.

RECURSOS E INSTRUMENTOS PARA UTILIZAR EN LAS PRÁCTICAS EVALUATIVAS EN AULA

A continuación presentamos una selección de recursos e instrumentos que se podrían utilizar en las prácticas evaluativas recomendadas en el apartado anterior.

PRIMERA PRÁCTICA. Metas claras: los objetivos de aprendizaje y sus criterios de evaluación

Sadler (1989) propone que al inicio de una situación evaluativa, el docente trabaje con los estudiantes mostrando productos terminados del trabajo o tarea evaluativa que hayan sido realizados por alumnos en años anteriores.

- En primer lugar, el profesor debe entregar a los estudiantes esos productos terminados para que los conozcan y no solo se imaginen cuál es el producto de la tarea que se les está solicitando.
- En segundo lugar, el profesor debe pedir a los estudiantes que evalúen cuáles consideran que son productos de buena calidad y distingan lo que son de una calidad deficiente. Así son los propios estudiantes quienes levantan criterios de calidad emergentes.
- En tercer lugar, el profesor les presenta los criterios de calidad que serán utilizados en esta tarea evaluativa para que ellos puedan comprenderlos y visualizarlos en una situación concreta.
- En cuarto lugar, los alumnos solo examinan aquellos productos terminados que hayan obtenido el logro máximo. El profesor debe resguardar que los productos que seleccionó expresen diversas formas de abordar el trabajo. El propósito de este ejercicio es que los estudiantes comprendan que hay variadas maneras de demostrar el logro del aprendizaje y que no se espera una única forma de obtener un logro máximo en esta situación evaluativa.

Este recurso permite que comprendan las metas de aprendizaje y sus criterios de evaluación antes de iniciar una situación evaluativa.

SEGUNDA PRÁCTICA. Uso de los conocimientos previos

a. La primera tarea evaluativa corresponde a la elaboración de un glosario de conceptos en torno al aprendizaje que se quiere indagar. Un glosario corresponde a un grupo de palabras o conceptos de una misma disciplina o de un campo de estudio que aparecen definidas, explicadas o comentadas. Para el caso de esta tarea evaluativa, el profesor deberá solicitar a cada estudiante que explique con sus propias palabras qué entiende por un conjunto acotado de conceptos asociados al tema seleccionado. Se puede además pedir al estudiante que, junto con explicar el concepto, entregue un ejemplo que lo grafique.

   Este ejercicio o tarea evaluativa es individual y se desarrolla por escrito. Se entrega al profesor para que este pueda analizar los conocimientos previos y tomar decisiones al respecto. Este mismo glosario inicial puede ser retomado por el estudiante en un momento más avanzado o al final, para que lo reelabore en función de los aprendizajes logrados durante el proceso de enseñanza.

b. La segunda tarea evaluativa corresponde a la elaboración de un organizador gráfico en torno al aprendizaje que se quiere indagar. Un organizador gráfico es una representación visual de conocimientos que muestra información respecto de una temática específica rescatando aspectos importantes de un concepto o materia dentro de un esquema.

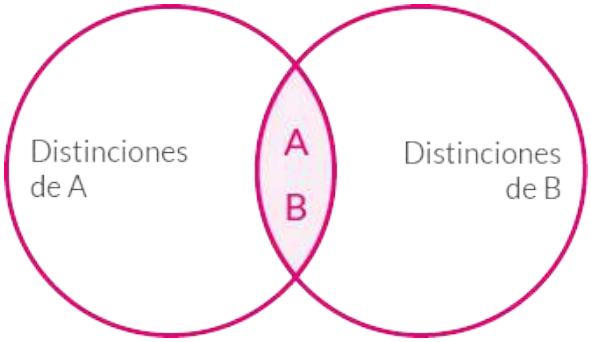
Esta tarea evaluativa es individual y se solicita que el estudiante escoja un tipo de entre algunos organizadores gráficos preseleccionados por el profesor (aquellos que mejor permitan graficar el aprendizaje de que se trata). Profundizaremos este recurso en el capítulo 8.

TERCERA PRÁCTICA. Uso del error para el aprendizaje

Esta tarea evaluativa considera solicitar que el alumno elabore un organizador gráfico específico que compara conceptos, teorías o enfoques, llamado “diagrama de Venn”. Se sugiere aplicar esta tarea evaluativa con finalidad formativa durante la enseñanza, con el objeto

de monitorear el aprendizaje y detectar posibles errores conceptuales o aprendizajes incompletos. El diagrama de Venn es un tipo de organizador conceptual que consta de dos círculos que se unen en el centro. Sirve para contrastar información o puntos de vista, para comparar enfoques, distinguiendo los aspectos comunes y las diferencias.

Modelo



Instrucciones

Se solicita al estudiante que elabore un diagrama de Venn a partir de la discusión y exposición de la unidad, específicamente sobre las diferencias y similitudes entre el enfoque A y B (a definir por el profesor). El alumno debe completar el diagrama de Venn, escribiendo las características distintivas de cada enfoque por separado e incluyendo lo común al centro.

CUARTA PRÁCTICA. Evaluación diversificada

A continuación se expone una estrategia evaluativa completa que ejemplifica la práctica de una evaluación diversificada.

Presentación: los profesores de Historia y Ciencias Sociales y de Matemáticas han diseñado una estrategia evaluativa en la que los alumnos de 6° Básico elaborarán un proyecto grupal de servicio a la comunidad, en el contexto de una evaluación de proceso con intencionalidad formativa y sumativa. Como agentes de la evaluación se han considerado el profesor y los estudiantes, por lo

tanto se desarrollará heteroevaluación, evaluación de pares y autoevaluación.

Fueron definidos objetivos de aprendizaje tanto de la asignatura de Historia y Ciencias Sociales como de Matemáticas para esta estrategia evaluativa, convirtiéndola en un proyecto integrado.

HI06 OA 21: trabajar en equipo de manera efectiva para llevar a cabo una investigación u otro proyecto, asignando y asumiendo roles, cumpliendo las responsabilidades asignadas y los tiempos acordados, escuchando los argumentos de los demás, manifestando opiniones fundamentadas y llegando a un punto de vista común.

HI06 OA 23: participar, mediante acciones concretas, en proyectos que impliquen aportes dentro de la escuela, la comunidad y la sociedad, como voluntariado y ayuda social, ateniéndose a un plan y su presupuesto.

MOA I: extraer información del entorno y representarla matemáticamente en diagramas, tablas y gráficos, interpretando los datos.

Instrucciones: el trabajo se desarrollará grupalmente, la composición de los equipos será asignada por los estudiantes.

El proyecto se desarrollará en tres etapas y cada una de ellas tendrá una evaluación formativa y una sumativa. Se recurrirá a la elaboración de un portafolio, que tendrá dos entregas parciales y una entrega completa.

Los equipos podrán seleccionar un tema a desarrollar entre una nómina propuesta por el profesor, los que se relacionan con conceptos trabajados en Ciencias Naturales, Artes u otras asignaturas.

Los criterios de evaluación serán:

- Argumentación del tema escogido.
- Organización de tareas, responsables y plazos.
- Investigación del tema escogido: personas afectadas, problema o necesidad, recopilación de información,

representación y sistematización de la información.

- Cumplimiento de responsabilidades y tareas grupales.
- Evaluación de pares reflexiva en todas las etapas.
- Propuesta de soluciones efectivas y factibles al problema identificado.
- Organización y ejecución del proyecto.
- Evaluación de los resultados del proyecto y del trabajo grupal.
- Incorporación de las retroalimentaciones al trabajo grupal.
- Desempeño grupal: participación equitativa, responsabilidad, interacciones en el grupo.
- Evaluación del proceso de aprendizaje del estudiante.

Etapa 1

- El grupo selecciona el tema sobre el que va a desarrollar el proyecto de servicio y argumenta su elección.
- El grupo organiza y planifica sus tareas, define responsables, plazos y actividades a realizar, de acuerdo con la pauta entregada por el profesor.
- Los estudiantes del grupo indagan acerca de la necesidad o problema que responderán a través del proyecto: investigan sobre el tema, identifican qué personas están siendo afectadas por el problema o necesidad, generan pautas de entrevistas o encuestas y las aplican, sistematizan la información en gráficos, diagramas o tablas y generan un informe breve acerca de la necesidad que abordarán.
- Se registra la asistencia a las reuniones del grupo, sus principales acuerdos y temas tratados.
- Cada miembro del grupo responde a las siguientes preguntas: ¿cuán efectiva ha sido la planificación que elaboramos en el desarrollo de esta etapa?, ¿qué dificultades hemos tenido?, ¿cómo las hemos enfrentado?

- El grupo hace entrega formativa parcial del portafolio con las evidencias de la etapa 1.
- El grupo recibe la retroalimentación y entrega una semana después el portafolio con evidencias de la etapa 1 para su evaluación sumativa y calificación.

Etapa 2

- El grupo desarrolla una lluvia de ideas para levantar posibles soluciones al problema o necesidad detectada en la etapa 1. El equipo organiza las soluciones, clasificándolas según dos criterios: efectividad y factibilidad. La efectividad considera que respondan a la necesidad y/o solucionen el problema diagnosticado. La factibilidad considera que sean viables, dado el tiempo y recursos con que cuentan.
- El grupo presenta ante el curso tanto el diagnóstico (etapa 1) como las ideas de solución seleccionadas para recibir retroalimentación. Los estudiantes se evalúan entre pares con una escala de valoración preparada por el profesor.
- El grupo selecciona la estrategia o solución y elabora el proyecto recurriendo a una pauta con lineamientos entregada por el profesor.
- Se registra la asistencia a las reuniones del grupo, sus principales acuerdos y temas tratados.
- Cada miembro del grupo responde a las siguientes preguntas: ¿cuál fue la principal dificultad que tuvimos en esta etapa?, ¿cómo la hemos enfrentado?
- El grupo hace entrega formativa parcial del portafolio con las evidencias de la etapa 2.
- El grupo recibe la retroalimentación y entrega una semana después el portafolio con evidencias de la etapa 2 para su evaluación sumativa y posterior calificación.

Etapa 3

- El grupo presenta el proyecto ante sus destinatarios para recibir su retroalimentación. Cuida que la forma y fondo de la presentación sean adecuados a las características de la audiencia.
- El grupo ajusta el proyecto conforme a las recomendaciones de sus destinatarios.
- El grupo organiza la ejecución del proyecto recurriendo a las directrices del profesor.
- El grupo ejecuta el proyecto que diseñó.
- El grupo evalúa los resultados del proyecto.
- El equipo evalúa su trabajo grupal y la ejecución del proyecto con un instrumento dado previamente por el profesor.
- El grupo hace entrega formativa del portafolio completo con las evidencias de las etapas 1, 2 y 3.
- El grupo recibe la retroalimentación y entrega una semana después el portafolio completo para su evaluación sumativa y posterior calificación.

QUINTA PRÁCTICA. Retroalimentación efectiva

Se presenta una serie de recomendaciones para realizar una retroalimentación efectiva frente a una tarea, producto parcial entregado o un desempeño desplegado por el estudiante.

Las recomendaciones se basan en los cuatro principios de la retroalimentación efectiva. Se refieren a:

- que sea descriptiva y no enjuiciadora del aprendizaje
- que considere aspectos logrados y por mejorar
- que sea específica respecto de la tarea o producto que revisa y que oriente al estudiante a reducir su brecha entre el nivel actual del aprendizaje y la meta
- que se centre en los aspectos cambiables por el alumno

Recomendaciones para el profesor:

- Realice la retroalimentación antes de calificar la tarea evaluativa del estudiante.
- En sus comentarios escritos evite incluir *tickets*, calificaciones o puntajes asociados.
- Una vez entregada la retroalimentación, ofrezca al estudiante la oportunidad de mejorar algunos de los aspectos del aprendizaje que usted revisó, de modo que él efectivamente pueda reelaborar o reescribir el producto antes de la evaluación sumativa.
- Céntrese en la descripción hecha por el estudiante o en cómo se llevó a cabo el trabajo analizado evitando mencionar juicios como “incompleto” o “incorrecto” sin describir qué es lo que está incompleto o incorrecto.
- Considere tanto los aspectos logrados por el estudiante en el trabajo o producto revisado como los aspectos por mejorar.
- La retroalimentación debe ser precisa, al estudiante debe quedarle claro cómo mejorar su producto o trabajo, por lo que el profesor tiene que concentrarse en las ideas específicas que el alumno puede trabajar para mejorar.
- Céntrese en los aspectos que el estudiante puede cambiar de su producto o trabajo, y no en aspectos que no dependen de él y que no se pueden modificar.

SEXTA PRÁCTICA. El estudiante como protagonista

1. El primer recurso sugerido es aplicar la técnica CQA (conozco, quiero, aprendí) durante una sesión de clases. Esta técnica surge como una herramienta para propiciar el aprendizaje en la lectura de textos expositivos (Ogle, 1986), pero también puede ser adaptada para usarse en otras situaciones de aprendizaje.
   C: ¿Qué conozco? En esta parte se pide al estudiante que indique y demuestre lo que sabe o conoce acerca de la meta de aprendizaje de la sesión. Se aplica al inicio de la clase.

Q: ¿Qué quiero conocer? En esta parte se le pide que determine lo que quiere aprender, definiendo con sus palabras su propia meta de aprendizaje. Se aplica al inicio de la clase luego de responder a la primera pregunta.

A: ¿Qué aprendí? Al finalizar la clase se le pide al estudiante que identifique lo que ha aprendido durante la sesión en relación con la meta que él mismo definió.

2. El segundo recurso sugerido es aplicar con finalidad formativa "pruebas breves o *quizzes* constructivos". Las pruebas escritas breves son comúnmente usadas para evaluaciones formativas, en las que los estudiantes pueden recibir retroalimentación inmediata. Para que ello ocurra, el profesor debe entregarles la hoja de claves, o bien, recurrir a un recurso digital que autocorrija la prueba breve. Y que a partir de ese momento se comiencen a desarrollar en la clase el análisis y revisión de las respuestas incorrectas y correctas.

Una primera variante de este recurso es pedir a los estudiantes que respondan al inicio de la clase el *quizz*, y luego de desarrollarla, se les ofrezca revisarlo y darles la opción de mejorarlo.

Una segunda variante de este recurso es pedir que los estudiantes respondan la prueba breve individualmente y luego la aborden en duplas, lo que implica negociar, argumentar y analizar los ítemes de la prueba a la luz de los aprendizajes de cada miembro de la dupla.

El tercer recurso sugerido es aplicar con finalidad formativa una autoevaluación. Esta evaluación consiste en que el estudiante reflexione acerca de las estrategias metacognitivas que usa para abordar una situación evaluativa propuesta por el profesor.

Por ejemplo, si se quiere que el estudiante autoevalúe sus estrategias metacognitivas de planificación, la técnica se debe aplicar al inicio, cuando se le presenta la situación evaluativa que deberá desarrollar. En ese contexto, el profesor le pide al alumno que planifique cómo abordará el trabajo y que luego autoevalúe sus estrategias de planificación a través de preguntas que el estudiante debe responder

y analizar antes de que el profesor desarrolle la retroalimentación. Algunas de ellas son:

- ¿Me di cuenta de qué clase de trabajo es este?
- ¿Comprendí cuál es el objetivo de este trabajo?
- ¿Identifiqué qué información necesito para desarrollarlo?
- ¿Reconozco los problemas que podrían surgir mientras esté trabajando y sé cómo podría manejarlos?
- ¿Identifiqué cuáles estrategias pueden ayudarme a desarrollar el trabajo?
- ¿Identifiqué los recursos de que dispongo para él?
- ¿Definí cuánto tiempo me tomará realizar este trabajo?

SÉPTIMA PRÁCTICA. Coherencia entre retroalimentación y evaluación sumativa

La principal recomendación para implementar esta práctica evaluativa la plantea Hattie (2003) y es asegurar, en primer lugar, la precisión y validez de la evaluación formativa otorgando una retroalimentación válida, esto es, alineada con los objetivos de aprendizaje y sus criterios de evaluación, y precisa en las recomendaciones de mejora para que el estudiante reduzca la brecha respecto de su meta de aprendizaje. En segundo lugar, el profesor debe asegurar la coherencia entre la retroalimentación entregada y la posterior evaluación sumativa que aplica, pues ambas se basan en la misma meta de aprendizaje y mantienen los criterios de evaluación.

EN SÍNTESIS, a lo largo del capítulo hemos desarrollado la propuesta de integración de la evaluación formativa y sumativa en un mismo ciclo evaluativo porque ellas se alimentan mutuamente y constituyen dos caras de la misma moneda en la práctica evaluativa en el aula.

Sostenemos que la visión integrada genera que profesores y estudiantes se unan con el propósito de crear una forma diferente de

desarrollar aprendizajes. Bajo este enfoque, tanto profesores como estudiantes cambian. Los alumnos van haciendo propios los aprendizajes y los docentes pueden ir ajustando sus estrategias de enseñanza para mejorar las oportunidades de aprendizaje (Wiliam, 2000).

Finalmente, consideramos necesario describir un conjunto de prácticas evaluativas que afiancen esta alianza, junto con compartir una serie de recursos evaluativos que permiten y posibilitan la implementación del enfoque integrado de la evaluación formativa y sumativa en el aula.

En el siguiente capítulo ahondaremos en un proceso central para desarrollar aprendizajes en los estudiantes, que está estrechamente vinculado con la evaluación formativa que hemos examinado en este capítulo, y que corresponde a la retroalimentación.

BIBLIOGRAFÍA

Atkin, J. M., Black, P. & Coffey, J. (2001). *Classroom assessment and the national science standards*. Washington DC: National Academies Press.

Black, P. J. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles Policy and Practice, 5*(1), 7-73.

Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5.

Bloom, B. S., Hastings, J. T. & Madaus, G. F. (Eds.). (1971). *Handbook on the formative and summative evaluation of student learning*. New York: McGraw-Hill.

Deneen, C. C. & Brown, G. (2016). The impact of conceptions of assessment on assessment literacy in a teacher education program. *Cogent Education, 3*, 1225380, 1-14.

Dyer, K. (2014). *Formative, summative, interim: putting assessment in context*. Recuperado de https://www.nwea.org/blog/2014/formative-summative-interim-putting-assessment-context/

Gibbs, G. (2010). *Using assessment to support student learning*. Leeds: Leeds Met Press.

Hargreaves, E. (2013). Inquiring into children's experiences of teacher feedback: reconceptualising assessment for learning. *Oxford Review of Education, 39*(2), 229-246.

Harlen, W. & Winter, J. (2004). The development of assessment for learning: learning from the case of science and mathematics. *Language Testing, 21*(3), 390-408.

Hattie, J. (2003). *Formative and summative interpretations of assessment information.* Auckland: University of Auckland. Recuperado de www.excellencegateway.org.uk/page.aspx?o=128831

Heitink, M. C., Van der Kleij, F. M., Veldkamp, B. P., Schildkamp, K. & Kippers, W. B. (2016). A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational Research Review, 17*, 50-62.

Kay, R. & Knaack, L. (2009). Exploring the use of audience response systems in secondary school science classrooms. *Journal of Science Education and Technology, 18*(5), 382-392.

Looney, J. W. (2011). *Integrating formative and summative assessment: progress toward a seamless system?* OECD Education Working Papers, 58, OECD Publishing. doi: http://dx.doi.org/10.1787/5kghx3kbl734-en

Moss, C. M. & Brookhart, S. M. (2009). *Advancing formative assessment in every classroom: a guide for the instructional leader*. Alexandria VA: ASCD.

Moss, C. M. (2013). Research on classroom summative assessment. In J. McMillan (Ed.). *SAGE handbook of research on classroom assessment* (pp. 235-255). Virginia: SAGE.

OECD. (2005). Formative assessment: improving learning in secondary classrooms. *The OECD Policy Briefs*. Paris: OECD.

Ogle, D. (1986). K-W-L: A teaching model that develops active reading of expository text. *The*

Reading Teacher, 39(6), 564-570.

Rasmussen, J. (2017). *Formative assessment strategies*. Recuperado de http://scholar.googleusercontent.com/scholar?q=cache:NA5I7rIhKzQJ:scholar.google.com/&hl=es&as_sdt=0,5

Rea-Dickins, P. (2006). Currents and eddies in the discourse of assessment: a learning-focused interpretation. *International Journal of Applied Linguistics, 16*(2), 163-188.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 145-165.

Scriven, M. (1967). *The methodology of evaluation*. Washington DC: American Educational Research Association.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(4), 4-14.

Shepard L. A. (2006). Classroom assessment. In R. L. Brennan (Ed.), *Educational measurement* (pp. 623-646). 4th ed. Westport CT: Praeger.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189.

Stobart, G. (2008). *Testing times: the uses and abuses of assessment*. New York: Routledge.

Wiliam, D. (2000). Integrating formative and summative functions of assessment. *Working group, 10*, 1-25.

Wiliam, D. (2009). Una síntesis integradora de la investigación e implicancias para una nueva teoría de la evaluación formativa. *Archivos de Ciencias de la Educación* (4a. época), *3*(3), 15-44.

Wiliam, D. & Black, P. (1996). Meanings and consequences: a basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal, 22*(5), 537-548.

Wolfel, R. (2009). *Classroom assessment: the confusion of many voices*. Recuperado de http://www.usma.edu/cfe/Literature/Wolfelre/Wolfel_09.pdf

CAPÍTULO 5

La retroalimentación efectiva y su potencial para mejorar el aprendizaje

Sandra Zepeda Aguirre

INTRODUCCIÓN

Diversos investigadores sostienen que la retroalimentación es una poderosa influencia en la mejora del aprendizaje de los estudiantes (Black & Wiliam, 1998; Hattie, Biggs & Purdie, 1996; Hattie & Jaeger, 1998). La conclusión de estudios de metaanálisis es categórica respecto del efecto que tiene la retroalimentación, pues ella se sitúa entre las diez principales acciones que influyen en el logro de aprendizajes (Hattie & Jaeger, 1998; Kluger & DeNisi, 1996). Sin embargo, se ha observado que la varianza de los efectos encontrados es considerable, lo que indica que algunos tipos de retroalimentación son más potentes que otros (Kluger & DeNisi, 1996). De esta forma, pareciera ser que la sola prescripción de un gran número de comentarios o la frecuencia de ellos no aseguran que el aprendizaje se desarrolle y que hay elementos asociados a esos comentarios que condicionan su efectividad en los estudiantes.

Hay evidencia que indica que aunque la retroalimentación sea frecuente en las salas de clases, la mayor parte de esta información es desestimada por los estudiantes y el uso en la revisión de sus trabajos es escaso, ya que la mayoría de los comentarios que los profesores realizan al curso no son considerados como relevantes por sus estudiantes (Carless, 2006). Se ha visto que los profesores encuentran su retroalimentación mucho más valiosa que lo que opinan realmente los alumnos acerca de ella, quienes señalan que la retroalimentación recibida es confusa y no los orienta en qué deben hacer para mejorar su desempeño (Goldstein, 2006; Nuthall, 2007).

A partir de lo planteado surge la duda respecto de cómo se debe dar retroalimentación al estudiante para lograr aprendizajes, en qué momento es más efectiva, de quién depende, qué debe decírsele y qué no debe decírsele a él o ella. Al parecer sabemos mucho sobre el poder de la retroalimentación para generar efectos en el aprendizaje de los estudiantes, pero desconocemos cómo aprovechar dicho poder

y hacer que funcione más eficazmente en el aula de manera sustentable en el tiempo.

A lo largo de este capítulo examinaremos las preguntas anteriores y profundizaremos en los principios y características de una retroalimentación efectiva y los distintos niveles en que se da, analizando críticamente su contribución al aprendizaje. Revisaremos ejemplos para identificar empíricamente prácticas de retroalimentación efectivas y no efectivas.

LA RETROALIMENTACIÓN

A continuación abordaremos los elementos nucleares de la definición de retroalimentación y, a partir de ella, examinaremos la evolución de este concepto.

Ramaprasad (1983) sostiene que el rasgo distintivo de la retroalimentación es que la información generada dentro de un sistema debe tener algún efecto sobre este. El autor plantea que “la retroalimentación es información acerca de la distancia entre el nivel actual y el nivel de referencia de un parámetro del sistema que es utilizada para modificar esa brecha en alguna forma” (p. 4).

Sadler (1989), por su parte, sostiene que un rasgo importante de la definición de Ramaprasad es que la información acerca de la distancia o brecha entre el nivel actual y el de referencia es considerada como retroalimentación siempre y cuando sea usada para reducir esa distancia. Si la información solo es registrada o traspasada a otra persona que no tiene ni el conocimiento ni el poder para cambiar el resultado, o si la información está muy profundamente codificada como para conducir a la acción apropiada (p. ej. un promedio de calificaciones), la información implícita no puede ser reconstruida para que la retroalimentación sea efectiva.

Ambos autores muestran la definición histórica e inicial de retroalimentación, asociada a los sistemas ingenieriles. La idea fuerza que se recoge de estos planteamientos, para el contexto educativo, es que la retroalimentación no puede separarse de sus consecuencias en el aprendizaje. Por eso resulta preocupante que el

aspecto que ha primado en ella en el lenguaje cotidiano haya sido la información que se entrega, y no se haya considerado la segunda parte de esta definición histórica, que nos habla del uso de esa información para la mejora
del sistema.

DESDE UNA VISIÓN CENTRADA EN EL PROFESOR HACIA UNA VISIÓN CENTRADA EN QUIEN APRENDE

Los primeros estudios y teorías sobre retroalimentación tienen casi 100 años y surgieron desde la perspectiva psicológica conductista (Thorndike, 1913 citado en Brookhart, 2008). En dicho contexto, la retroalimentación positiva era concebida como “refuerzo positivo” y la negativa, como “castigo”, y se consideraba que afectaba el aprendizaje. Por lo mismo, la preocupación se centró en estudiar si constituía o no un recurso eficaz. A partir de esa tradición, algunas investigaciones han abordado las diferentes aplicaciones de la retroalimentación, centrando su atención en lo que hace el profesor, esto es, cuándo los comentarios que se entregan son efectivos y cuándo, ineficaces (Bangert-Drowns, Kulik, Kulik & Morgan, 1991; Butler & Winne, 1995; Hattie & Timperley, 2007; Kluger & DeNisi, 1996). Otras investigaciones se han focalizado en la caracterización de una retroalimentación eficaz, preguntándose por el tipo de retroalimentación entregada y las consecuencias en el aprendizaje (Johnston, 2004; Tunstall & Gipps, 1996).

Las bases teóricas de la educación han cambiado, ya no tienen una visión conductista del aprendizaje, asociada a conexiones estímulo-respuesta. Los estudios más recientes reconocen la centralidad del estudiante en el proceso de retroalimentación y se han orientado a conocer el tipo de *feedback* dado, el contexto en el que se desarrolla, y en indagar cómo llega la información de la retroalimentación al estudiante, cómo la interpreta y la “filtra” a través de su percepción, mediada por los conocimientos previos, las experiencias y la motivación. El desafío para el estudiante es entender la tarea evaluativa, comprender y adherir a su meta de aprendizaje, recibir e

interpretar la información y encaminarse para mejorar su aprendizaje, más que responder mecánicamente a determinados estímulos. Este proceso se identifica con la autorregulación del aprendizaje.

La investigación de Butler y Winne (1995) muestra que tanto la retroalimentación externa (entregada por el profesor, por ejemplo) como la retroalimentación interna (autoevaluación del propio estudiante) influyen fuertemente en el conocimiento y en las creencias y habilidades metacognitivas del estudiante. Ambas fuentes de retroalimentación permiten que desarrolle la autorregulación de su propio proceso de aprendizaje y, por ello, que pueda crear o seleccionar las tácticas y estrategias para alcanzar y producir la tarea o desafío requerido.

Uno de los aspectos más relevantes del enfoque centrado en quien aprende es reconocer que los comentarios o información que entrega el profesor no aseguran que el estudiante haga lo necesario para mejorar su aprendizaje, pues es él quien decide y quien se debe movilizar para acortar la brecha. El profesor entrega un comentario, que se convierte en un *input* y, vista así, la retroalimentación depende centralmente de lo que haga el estudiante con ese *input* para el logro de su aprendizaje.

Retomando el concepto original de retroalimentación, observamos que se vincula explícitamente con el enfoque centrado en quien aprende, pues exige asegurar que el ciclo de retroalimentación se haya completado (Boud & Molloy, 2012). Como sostiene Sadler (1989), sin proporcionar estrategias para mejorar el aprendizaje y sin un seguimiento de cómo influye posteriormente la información entregada en el desempeño del estudiante, la retroalimentación puede verse simplemente como “datos inconexos” o cometarios que no conducen u orientan a la acción.

A modo de síntesis, del recorrido por el concepto de retroalimentación podemos observar que su origen proviene de la ingeniería y posteriormente, al llegar al campo de la Educación, se le añade una cualidad propia y única de la retroalimentación educativa: el papel del estudiante, pues no lo considera un objeto pasivo que responde a un estímulo, sino una persona que piensa, toma decisiones y actúa. En

este enfoque, la retroalimentación se posiciona no como un acto episódico vinculado a corregir las tareas evaluadas, sino como un aspecto clave que integra el diseño de la enseñanza y el proceso de enseñanza/aprendizaje desarrollado en el aula (Molloy & Boud, 2013).

EL PODER DE LA RETROALIMENTACIÓN

Ramsden (2003) sostiene que los comentarios efectivos sobre el trabajo de los estudiantes representan una de las características clave de la enseñanza de calidad. Hounsell (2003, p. 67) señala que “desde hace mucho tiempo, tanto los investigadores como los profesionales, reconocen que la retroalimentación juega un papel decisivo en el aprendizaje y el desarrollo dentro y fuera de los entornos educativos formales. Aprendemos más rápido, y mucho más efectivamente, cuando tenemos una idea clara de cómo lo estamos haciendo y lo que podríamos necesitar hacer para mejorar”.

En una discusión sobre las condiciones bajo las cuales la evaluación apoya el aprendizaje, Gibbs & Simpson (2004) destacan la importancia de que la retroalimentación sea comprensible, oportuna y que oriente a los estudiantes a la acción. Yorke (2003), por su parte, indica que a pesar de su importancia, la literatura revela que los estudiantes están a menudo insatisfechos con la retroalimentación que reciben, en términos de que les falta un asesoramiento específico para mejorar (p. ej., ver Higgins, Hartley & Skelton, 2001), que consideran que ella es difícil de interpretar (Chanock, 2000) o reciben a partir de ella un impacto potencialmente negativo en su autopercepción y confianza (James, 2000).

Shute (2008) describe un aspecto controversial de la retroalimentación, pues ella constituye uno de los rasgos más poderosos instrumentalmente del diseño instruccional, pero es a la vez el menos comprendido. La retroalimentación proporcionada durante el proceso de enseñanza/aprendizaje, en el marco de la evaluación formativa, es clave y Wiggins (2012) sostiene que esta retroalimentación continua aumentará en última instancia el logro estudiantil, pues permite oportunidades para que los alumnos

remodelen su desempeño para lograr mejor los objetivos de aprendizaje.

LA RETROALIMENTACIÓN EFECTIVA

En este apartado intentaremos responder a la pregunta ¿cuándo estamos frente a una retroalimentación efectiva?, pues estudios constatan tanto efectos positivos como negativos cuando ella se utiliza en el ámbito educativo.

La retroalimentación efectiva se define como la información comunicada al estudiante con la intención de modificar su pensamiento con el propósito de mejorar el aprendizaje (Shute, 2008), ella cierra la brecha entre el nivel actual de comprensión de los estudiantes y la meta de aprendizaje deseada. Ayuda a los alumnos a entender la relación entre un conjunto claramente definido de criterios o estándares y su nivel actual de desempeño (Clark, 2011).

A pesar de que el profesor también puede recibir información que le permita modificar su proceso de enseñanza, el centro de esta retroalimentación es el estudiante, por lo que es clave que ella sea correcta y posibilite mejorar significativamente los procesos de aprendizaje (Boud & Molloy, 2012). Wiliam (2012) utiliza una analogía para explicar la retroalimentación: al igual que un termostato ajusta la temperatura ambiente, la retroalimentación efectiva ayuda a mantener un ambiente de apoyo para el aprendizaje del estudiante.

Winne y Butler (1994) sostienen que la retroalimentación es información con la que el estudiante puede confirmar, añadir, sobrescribir, afinar o reestructurar la información en su trabajo, ya sea que esa información corresponda a conocimiento disciplinar, conocimiento metacognitivo, creencias sobre sí mismo y acerca de las tareas o estrategias cognitivas e innovadoras.

Para que el efecto de la retroalimentación sea poderoso tiene que desarrollarse en un contexto de aprendizaje, pues ella debe ser considerada parte integrante del proceso de enseñanza, en que temporalmente se despliega luego de que un estudiante haya

respondido una tarea inicial durante la enseñanza (Hattie & Timperley, 2007).

EL ROL DEL CONTEXTO

Una retroalimentación efectiva debe formar parte de un clima o contexto de aula donde los estudiantes consideren y reciban el comentario, la crítica o recomendación como algo positivo y entiendan que el aprendizaje no puede ocurrir sin que ejerciten o actúen para mejorarlo. Cuando estamos frente a una cultura de aula que no da espacio para cometer errores y solo enfatiza hacer las cosas bien, entonces se juzga como algo incorrecto cuando alguna tarea necesita mejoras. Por el contrario, en una cultura de aula que valora la búsqueda, el trabajo persistente y el uso del error para mejorar el aprendizaje, los alumnos podrán recurrir a la retroalimentación para planificar y ejecutar las tareas necesarias para mejorar sus aprendizajes (Brookhart, 2008).

En esta cultura de la retroalimentación efectiva no parece adecuado que se les entregue a los estudiantes una retroalimentación y no tengan oportunidades para usarla. Tampoco es adecuado que se les dé a los estudiantes una retroalimentación constructiva y luego el profesor la utilice como medida para la evaluación sumativa al asignarles una baja calificación final.

COMPONENTES ESENCIALES DE LA RETROALIMENTACIÓN EFECTIVA

Sadler (1989) describió tres componentes esenciales de la retroalimentación, que incluyen:

1. En primer lugar, información sobre el objetivo de aprendizaje, sus estándares y criterios de evaluación.
2. En segundo lugar, información sobre el nivel alcanzado en el producto, desempeño o tarea en ejecución.

3. Por último, las estrategias para abordar la brecha entre el objetivo de aprendizaje y la tarea evaluativa en desarrollo.

La retroalimentación es eficiente cuando se enfoca en la naturaleza cognitiva de la tarea y sus procesos metacognitivos asociados, y se organiza para dar respuesta a tres preguntas principales:

1. ¿Adónde voy?: asociada a comprender cuál es la meta de aprendizaje.
2. ¿Cómo voy?, ¿dónde estoy?: vinculadas al nivel de avance en que se encuentra en la actualidad la tarea evaluativa y, por consiguiente, el aprendizaje del estudiante.
3. ¿Cómo sigo avanzando?: asociada a las estrategias o recomendaciones entregadas al alumno para acortar la brecha de aprendizaje.

Los poderes de la retroalimentación se despliegan más bien desde quien la recibe que desde quien la entrega. Por ello es muy relevante examinar cuándo, cómo y qué contenidos debe tener para que se desarrolle el ciclo efectivo del aprendizaje desde la perspectiva del estudiante (Hattie & Gan, 2011).

TIPOS DE RETROALIMENTACIÓN Y SU EFECTO EN EL APRENDIZAJE

Los efectos de la retroalimentación pueden también analizarse desde su naturaleza. Hattie y Timperley (2007) revisaron múltiples investigaciones y lograron sintetizar un modelo de retroalimentación con cuatro niveles:

1. El primer nivel se asocia a la retroalimentación centrada en la tarea, entregando información sobre aspectos de ella como, por ejemplo, si las respuestas fueron correctas o incorrectas, o dando más instrucciones para obtener más información, o precisando los pasos de la tarea en evaluación.

2. El segundo nivel se asocia a la retroalimentación sobre el procesamiento de la tarea, y se vincula con el proceso que se desarrolla para generar e implementar dicha tarea. Por ejemplo, se retroalimentan las estrategias utilizadas para abordarla si ellas fueron seleccionadas adecuadamente y si fueron aplicadas con precisión.
3. El tercer nivel se asocia a la retroalimentación de la autorregulación del estudiante, abordando las estrategias metacognitivas utilizadas por él para enfrentar la tarea y apelando además al autoconcepto y confianza del estudiante en su capacidad de aprendizaje y motivación por alcanzarlo.
4. El cuarto nivel se asocia a los comentarios o refuerzos realizados por el profesor al estudiante como persona. Considera frases que lo aprueban o desaprueban, como: “eres bueno”, “eres inteligente”, “eres descuidado”.

Los autores afirman que los niveles en los que se enfoca la retroalimentación efectiva corresponden al primero, segundo y tercero, pues los tres permiten acortar la brecha de aprendizaje. En cuanto al cuarto nivel, son categóricos al sostener que este no permite mejorar el aprendizaje ni moviliza a los estudiantes para acortar la distancia hacia la meta, pues se centra en cualidades de la persona que no tienen relación con el nivel del aprendizaje actual, futuro y su brecha.

A continuación, la tabla 1 resume estos cuatro niveles de retroalimentación:

TABLA 1: Niveles de retroalimentación (Hattie & Timperley, 2007)

Primer nivel	Segundo nivel	Tercer nivel	Cuarto nivel
• CENTRADO EN LA TAREA • Efectiva para apoyar el aprendizaje	• CENTRADO EN EL PROCESO • Efectiva para apoyar el aprendizaje	• CENTRADO EN LA AUTORREGULACIÓN DEL APRENDIZAJE • Efectiva para regular el aprendizaje y mantener la motivación hacia él	• CENTRADO EN LA PERSONA • No es efectiva para orientar el aprendizaje

Examinaremos una segunda clasificación de retroalimentación, pero en este caso el criterio responde a la pregunta por el contenido de la retroalimentación, no a su nivel, como en el modelo anterior. Esta tipología clasifica todas las formas de retroalimentación, tanto aquellas efectivas como las que no lo son para mejorar el aprendizaje. Fue desarrollada por Tunstall y Gipps (1996) a partir de un estudio que investigó los distintos tipos de retroalimentación que se entregaban a niños y niñas en la Enseñanza Básica en el área de Matemáticas en Reino Unido.

Aparecen en la parte izquierda de la tabla 2 cuatro tipos de retroalimentación (A1, A2, B1, B2) que no resultan ser efectivos para el aprendizaje, pues se centran en el reforzamiento positivo o negativo. Estos cuatro tipos se centran en “enjuiciar” al estudiante y no apuntan a la mejora de la tarea evaluativa.

En la parte derecha de la tabla se presentan cuatro tipos de retroalimentación (C1, C2, D1, D2) que resultan ser efectivos para mejorar el aprendizaje, pues se enfocan en describir los aspectos logrados y por lograr de la tarea evaluativa, junto con proponer estrategias para mejorarla. Estos cuatro tipos se centran en “describir” los logros y la brecha respecto de la tarea evaluativa.

TABLA 2: Tipología de retroalimentación de profesores hacia estudiantes (Tunstall & Gipps, 1996)

Retroalimentación positiva ⟷ Retroalimentación de logros

Retroalimentación enjuiciadora (+)		**Retroalimentación descriptiva (+)**	
A1 Premiar. Refuerzo positivo o recompensa centrada completamente en la persona del estudiante (ej.: entrega de obsequios).	**B1** Aprobar. Aprobación verbal y no verbal del profesor a la persona o trabajo del estudiante (ej.: “excelente”, vistos buenos, “buena niña”).	**C1** Describir los logros. Identifica aspectos específicos de la tarea que están logrados, lo que apoya el aprendizaje (“esto que has hecho… está bien logrado”).	**D1** Generar mejores niveles de logro. Identifica (junto al estudiante) las estrategias metacognitivas usadas para aprender.
A2 Castigar.	**B2** Desaprobar.	**C2** Especificar los logros y	**D2** Diseñar caminos para

Refuerzo o comentarios negativos. Centrada completamente en la persona del estudiante (ej.: expulsar de la sala, destruir el trabajo).	Desaprobación verbal y no verbal del profesor a la persona o trabajo del estudiante (ej.: cruces y rayas en el trabajo, “estoy desilusionado, esperaba más de ti, pero eres un flojo”).	lo que hay que mejorar. Se enfoca en los logros y/o errores del trabajo realizado y su relación con los aprendizajes, más que en las personas.	mejorar. Describe (junto al estudiante) futuras posibilidades en el aprendizaje logrado y sus vínculos con la trasferencia a otras situaciones de la misma índole.
Retroalimentación enjuiciadora (-)		**Retroalimentación descriptiva (-)**	
Retroalimentación negativa ⟷ Retroalimentación para mejorar			

ASPECTOS CLAVE ASOCIADOS A LA RETROALIMENTACIÓN EFECTIVA

Se describirá a continuación un conjunto de aspectos que permitirán abordar consistentemente una retroalimentación efectiva:

¿CUÁNDO ENTREGAR LA RETROALIMENTACIÓN?

Antes de la evaluación sumativa

La retroalimentación proporcionada principalmente después de las evaluaciones sumativas a menudo llega demasiado tarde en el proceso, por lo que los estudiantes pueden encontrarla sin sentido (Huxham, 2007).

De acuerdo con esta recomendación, la retroalimentación formativa tiene que entregarse cuando los estudiantes siguen siendo conscientes de los resultados del aprendizaje y tienen tiempo para actuar sobre la información que se les entrega para mejorar, reelaborar la tarea evaluativa, el producto o el desempeño. Esto puede incluir devolver una prueba o tarea evaluativa al día siguiente o dar respuestas orales inmediatas a los errores o conceptos imprecisos de los estudiantes (Shute, 2008).

Se recomienda (Rasmusen, 2017; Shute, 2008) que la retroalimentación se entregue al estudiante durante la evaluación

formativa. Se sostiene que la motivación y el esfuerzo de los alumnos también aumentan cuando los profesores usan la retroalimentación formativa para resolver esta brecha. El clima en el aula se convierte en un elemento positivo en el aprendizaje en lugar de un aspecto que genera alta ansiedad por la calificación (Fluckiger, Vigil, Tixier, Pasco & Danielson, 2010).

Oportuna

Se ha demostrado que la retroalimentación es más efectiva cuando es oportuna, está ligada a las expectativas e incluye sugerencias específicas sobre cómo mejorar el rendimiento en las metas de aprendizaje (Looney, 2011). Otros autores sostienen que la retroalimentación es más efectiva cuando se proporciona en forma inmediata o a lo sumo dentro de un período de días (Wiliam, 2006). Por otra parte y para algunos tipos de tareas evaluativas, la retroalimentación no debe ser proporcionada con demasiada rapidez, pues se espera que el estudiante tenga la oportunidad de reflexionar sobre su propio proceso de aprendizaje.

Respecto de la oportunidad de la retroalimentación, algunos autores indican que ella debe ser inmediata cuando se trata de tareas complejas, o bien, en el caso de estudiantes que hayan demostrado bajos niveles de desempeño. En cambio, debiera ser más distanciada cuando se trata de tareas que necesiten de la revisión o reflexión previa del estudiante.

En síntesis, se recomienda recurrir a la retroalimentación antes de calificar el aprendizaje, de modo que el estudiante tenga oportunidad para mejorarlo. Se aconseja que la retroalimentación sea inmediata para el caso de conocimientos prácticos, pero se espera que se difiera en el tiempo para conocimientos conceptuales, con el propósito de permitir que el estudiante reflexione sobre su aprendizaje antes de recibir comentarios del profesor.

¿CUÁL DEBE SER EL CONTENIDO DE UNA RETROALIMENTACIÓN EFECTIVA?

Información necesaria

Un concepto clave que permite comprender cuál debe ser el contenido de la retroalimentación es considerar el andamiaje, esto es, se debe proporcionar mucha o poca información solo en la medida de lo que el estudiante necesita para alcanzar el siguiente nivel de su meta de aprendizaje.

La retroalimentación efectiva entrega a los estudiantes lo que ellos necesitan para que puedan entender dónde están en su aprendizaje y los conduzca u oriente sobre qué hacer a continuación: el factor cognitivo. Una vez que sientan que entienden, desarrollan la percepción de que tienen el control sobre su propio proceso aprendizaje, lo que se asocia a un fuerte factor motivacional que los mantiene trabajando en la tarea evaluativa y los hace persistir hasta alcanzar la meta de aprendizaje (Brookhart, 2008).

La retroalimentación efectiva contiene información que un estudiante puede usar, lo que significa que tiene que ser capaz de oírla (o leerla en caso de que sea escrita) y entenderla.

Centrada en la tarea, el proceso y la autorregulación

Desde el punto de vista del modelo de niveles de retroalimentación ya descritos en el apartado anterior (Hattie & Timperley, 2007), la retroalimentación puede centrarse en la tarea que el estudiante está desarrollando (primer nivel), en el proceso o estrategias que permiten desarrollar la tarea (segundo nivel), y en las estrategias metacognitivas que regulan y monitorean la mejora de la tarea y su proceso (tercer nivel).

Dentro de la retroalimentación desarrollada para los tres niveles (tarea, proceso, autorregulación) aparece lo que Narciss (2008) denomina “retroalimentación informativa tutorial” para referirse al desarrollo de estrategias de retroalimentación que proporcionan información elaborada (pautas, rúbricas, comentarios detallados, recomendaciones específicas) para guiar a los estudiantes hacia la finalización exitosa de la tarea evaluativa. La retroalimentación

informativa elaborada puede tomar diversas formas, por ejemplo: a) reglas de la tarea, restricciones de las tareas y sus requerimientos, b) conocimiento conceptual vinculado a la tarea, c) errores o equivocaciones detectados, d) conocimientos procedimentales, e) conocimiento de estrategias metacognitivas.

Esta retroalimentación como tutoría se centra en orientar a los estudiantes en la detección de errores, en superar obstáculos y aplicar estrategias más eficaces para completar la tarea de aprendizaje (Narciss, 2008).

Priorizar información

Recomendaciones sobre qué retroalimentar indican que se deben priorizar los aspectos importantes a destacar, eligiendo aquellos que se relacionan directamente con las metas principales de aprendizaje y, finalmente, se debe considerar el nivel de desarrollo del estudiante (Brookhart, 2008).

¿CÓMO DESARROLLAR UNA RETROALIMENTACIÓN EFECTIVA?

Modalidad

La retroalimentación puede ser entregada en modalidades, escrita, oral o a través de videos. Algunos tipos de tareas evaluativas se prestan mejor para los comentarios por escrito, por ejemplo, la revisión de textos elaborados por los estudiantes, en los que son completamente pertinentes los comentarios escritos. La retroalimentación oral es más adecuada para desempeños observables, como por ejemplo, mirar y comentar cómo los estudiantes resuelven los problemas de Matemáticas a través de un trabajo individual. Algunos de los mejores comentarios pueden ser el resultado de entrevistas breves con el estudiante. Por ejemplo, en lugar de decirle de inmediato los aspectos logrados y aquellos que debe mejorar en su trabajo, puede comenzar haciendo preguntas como: ¿qué fue lo más difícil para ti en esta tarea?, ¿cómo lo

resolviste?, ¿hay algo que te sorprendió? Estas preguntas reflexivas tienen el potencial de colaborar productivamente para que la retroalimentación se constituya en un espacio donde profesor y estudiantes dialogan acerca del aprendizaje (Brookhart, 2008).

Las decisiones sobre la modalidad de retroalimentación deberán considerar primeramente la naturaleza de la tarea evaluativa, además de la capacidad de lectura de los estudiantes y también el momento y oportunidad para entregarla. Generalmente hablar con ellos suele ser mejor, porque se puede desarrollar una conversación, pero esto tiene la dificultad de que no se cuenta con el tiempo para desplegarlo con cada uno. Ante esta restricción, Boud y Molloy (2012) recomiendan grabar audios o videos breves (minivideos) en los que el profesor pueda describir a los estudiantes los logros y brechas acerca de la tarea evaluativa, y ofrecer o recomendar estrategias específicas para superarla.

¿QUIÉN RETROALIMENTA?

Retroalimentación sustentable, una mirada hacia el futuro

El enfoque tradicional de la retroalimentación, basado en una noción unilateral, en la que la información se transmite desde el profesor hacia el estudiante, debiera transitar hacia un enfoque más multilateral. Este sitúa a los estudiantes como sujetos activos que generan sus propios juicios acerca de la tarea evaluativa, recurriendo a información desde varios otros actores (profesores, pares y sí mismos) (Boud & Molloy, 2012).

La retroalimentación multilateral no solo proporciona al estudiante una fuente de datos adicional para complementar su aprendizaje, sino que el hecho de buscar y dar sentido a la retroalimentación desde múltiples fuentes constituye además una práctica clave para el aprendizaje permanente a lo largo de la vida. Es por ello que este tipo de retroalimentación se ha llamado “retroalimentación sostenible”, pues no depende solo de la persona del profesor para apoyar el aprendizaje, sino de la capacidad de autorregulación del propio aprendizaje del sujeto (Boud & Molloy, 2013).

Para que este enfoque multilateral de la retroalimentación sostenible se despliegue en las prácticas de aula, debe ser parte del diseño general del curso y el profesor debe organizarlo de modo que tenga cabida en la planificación de la enseñanza. Si la retroalimentación solo recae a nivel individual en el profesor podría convertirse en un mecanismo episódico en el desarrollo de la enseñanza, por la alta demanda de tiempo que requiere, reduciendo así su potencial. En cambio, si a esta práctica se suman los pares y el propio estudiante, se amplifica su efecto, se desarrolla la autorregulación en ellos y los hace más conscientes de las metas y de su proceso de aprendizaje.

Las consideraciones acerca del momento, contenido, rol de los estudiantes y del profesor deben ser resueltas antes, básicamente durante el diseño y la planificación del proceso de enseñanza/aprendizaje. Por ello, la retroalimentación no debiera ser un complemento de las actividades de enseñanza y las tareas evaluativas, sino ser una característica de la planificación de la enseñanza y su posterior implementación en el aula.

La naturaleza de la información proporcionada necesita centrarse en sus efectos sobre el estudiante y lo que puede hacer con ella, más que analizar la información en sí a partir de cualidades descontextualizadas. La conclusión que se extrae de esta afirmación es que ninguna forma, modo o estrategia de retroalimentación es apropiada en sí misma, sino que ella debe considerar diferentes propósitos en diferentes contextos, las expectativas de mejora, la complejidad de la tarea evaluativa y el papel que cada agente va a jugar en su desarrollo.

PRÁCTICAS DE RETROALIMENTACIÓN EFECTIVA

Sobre la base de una serie de estudios, Nicol y Macfarlane-Dick (2006) y Wiggins (2012) han demostrado que la retroalimentación puede dar lugar a mejoras sustanciales en el aprendizaje. Los autores proponen siete buenas prácticas asociadas a esta retroalimentación efectiva, que examinaremos detenidamente:

PRIMERA BUENA PRÁCTICA: Ayudar a clarificar lo que es un buen rendimiento supone definir explícitamente los objetivos de aprendizaje y sobre todo que los estudiantes comprendan los criterios de evaluación y los estándares esperados para la realización de la tarea evaluativa, producto o desempeño.

Ejemplos para implementar la práctica:

Un ejemplo demostrado por ser particularmente poderoso en los objetivos de aprendizaje y los criterios aclaratorios ha sido proporcionar a los estudiantes “ejemplares” del desempeño (Orsmond et al., 2002). Los ejemplares (productos elaborados por estudiantes en períodos lectivos anteriores) son eficaces porque hacen explícito lo que se requiere y definen un estándar válido con el que los estudiantes pueden comprender más empíricamente qué se espera de su trabajo.

Otras estrategias que han demostrado ser eficaces en la clarificación de criterios, estándares y objetivos de aprendizaje para los estudiantes consideran además el desarrollo creciente de la autorregulación:

- Entregar una mejor definición de los requisitos del trabajo o tarea evaluativa utilizando criterios cuidadosamente construidos, elaborando descripciones escritas y definiciones del nivel de rendimiento.
- Promover la discusión y reflexión sobre los criterios y las reglas de la clase (por ejemplo, antes de la tarea evaluativa).
- Promover la participación de los estudiantes en los ejercicios de evaluación formativa, en la que puedan comentar el trabajo de otros estudiantes en relación con los criterios y estándares definidos.
- Desarrollar talleres en que los estudiantes, en colaboración con el profesor, puedan diseñar o negociar sus propios criterios de evaluación para una parte del trabajo.

SEGUNDA BUENA PRÁCTICA: Facilitar el desarrollo de la autoevaluación como una reflexión interna del estudiante acerca del

nivel que está alcanzando su producto, trabajo o desempeño, teniendo como base los estándares esperados para el cumplimiento del objetivo de aprendizaje.

Ejemplos para implementar la práctica:

Se presentan ejemplos de reflexión estructurada y de autoevaluación de los estudiantes:

- Solicitarles que identifiquen el tipo de retroalimentación que les gustaría recibir sobre su trabajo o tarea.
- Pedir a los alumnos que reconozcan las fortalezas y debilidades en su propio trabajo en relación con los criterios o estándares antes de entregarlo para la retroalimentación del profesor.
- Pedirles que reflexionen sobre sus logros y selección de las evidencias de su trabajo con el fin de incluirlas en un portafolio.
- Pedir al alumno que antes de la retroalimentación de su tarea, reflexione sobre los hitos de rendimiento y sobre los avances hacia la siguiente etapa de la tarea evaluativa (Cowan, 1999).

Un método adicional supone entregarles oportunidades para evaluar y proporcionar retroalimentación sobre el trabajo de sus pares. La evaluación de pares colabora en desplegar oportunidades para que el estudiante aprenda a emitir juicios objetivos basados en los estándares; este aprendizaje puede ser transferido cuando los alumnos vuelven a producir y regular su propio trabajo (Boud et al., 1999; Gibbs, 1999).

TERCERA BUENA PRÁCTICA: Entregar información de alta calidad a los estudiantes sobre su aprendizaje.

La investigación muestra que los profesores cumplen un rol central en el desarrollo de la capacidad de sus estudiantes para la autorregulación y a la vez constituyen una fuente importante de retroalimentación externa. La retroalimentación entregada por los profesores es una fuente con la que los estudiantes pueden evaluar el progreso de su tarea evaluativa, y permite comparar sus propias evaluaciones internas con los objetivos, criterios y estándares. Los profesores son mucho más eficaces en la identificación de errores

conceptuales o equivocaciones en el trabajo o tarea evaluativa que los propios estudiantes o sus pares; de este modo la retroalimentación de los docentes complementa la información obtenida a través de procesos de autoevaluación.

Respecto de lo que se entiende por una retroalimentación externa de calidad, se apela a aquella que entrega la información que ayuda a los estudiantes a solucionar su propio rendimiento y autocorrección, es decir, los ayuda a que tomen decisiones para reducir la brecha entre sus intenciones y los efectos resultantes.

Ejemplos para implementar la práctica:

Se presentan algunas estrategias que incrementan la calidad de la retroalimentación externa provista por los profesores:

- Asegurarse de que la retroalimentación sea válida, lo que implica que tenga a la base criterios predefinidos.
- Proporcionar retroalimentación en tiempo oportuno, es decir, antes de la evaluación sumativa, para que los estudiantes tengan la oportunidad de modificar o reescribir el trabajo o tarea evaluativa.
- Entregar un asesoramiento correctivo, no solo centrar la información en las fortalezas y debilidades, sino que apoyarlos con estrategias para que mejoren las dificultades detectadas.
- Priorizar y ordenar (p. ej. temporalmente) la cantidad de aspectos por mejorar en la tarea evaluativa, considerando prioridades, secuencias y las metas intermedias de aprendizaje.
- Proveer de pruebas en línea a las que los estudiantes puedan acceder a su resultado en cualquier momento, en cualquier lugar y las veces que lo necesiten.

CUARTA BUENA PRÁCTICA: Promover diálogos en torno al aprendizaje con el profesor y compañeros.

Bajo un enfoque de autorregulación, para que la retroalimentación externa sea efectiva debe ser interiorizada por el estudiante antes de que pueda ser utilizada para hacer mejoras productivas en su desempeño o tarea. Sin embargo, en la literatura de investigación

(Chanok, 2000; Hyland, 2000) existe mucha evidencia de que los estudiantes no entienden la información dada por externos (p. ej., “tu argumento es incompleto”, “este ensayo no es suficientemente analítico”) y, por lo tanto, no pueden usarla para tomar medidas de mejora, pues no saben qué hacer para reducir la brecha (respecto de los ejemplos: no saben qué le falta al argumento ni cómo completarlo, no saben cómo se elabora un ensayo más analítico).

Una estrategia para aumentar la eficacia de la retroalimentación externa, y la posibilidad de que la información entregada sea entendida y usada por los estudiantes, es concebir la retroalimentación más como un diálogo que como la transmisión de información. El diálogo significa que el estudiante no solo recibe información de retroalimentación inicial, sino que tiene la oportunidad de participar en la conversación con el profesor acerca del aprendizaje y su tarea evaluativa. Lamentablemente el tamaño de las clases hace difícil implementar esta estrategia, por lo que se recomienda recurrir al uso de tecnologías de la información y comunicación y al diálogo entre pares.

Ejemplos para implementar la práctica:

Se presentan a continuación ejemplos de estrategias de retroalimentación dialógicas que apoyan la autorregulación de los estudiantes:

- Proporcionar retroalimentación usando la técnica de “respuestas de un minuto” en la clase (ver Angelo & Cross, 1993). Esta corresponde a una actividad de escritura muy corta, que se puede desarrollar al finalizar la clase (tarda un minuto o menos en completarse) y en la que el estudiante responde a una pregunta formulada por el profesor. Dicha técnica lleva a los estudiantes a reflexionar sobre la sesión del día y le entrega al profesor información muy relevante acerca de la comprensión, aplicación o valoración del aprendizaje por parte del estudiante.
- Revisar la retroalimentación externa: se les pide a los estudiantes que lean los comentarios de retroalimentación que les han dado antes en una tarea y que hablen de ellos con sus compañeros

(también se les puede pedir que sugieran estrategias para mejorar el rendimiento la próxima vez).

- Pedir a los alumnos que revisen la retroalimentación externa que reciben, y destaquen uno o dos ejemplos de aquellos aspectos que encontraron útiles en ella y que expliquen cómo los ayudaron en la mejora de su tarea evaluativa.
- Solicitar que los estudiantes se entreguen entre pares retroalimentación descriptiva sobre su trabajo, teniendo como base los criterios evaluativos presentados previamente.
- Revisar en grupo un proyecto inicialmente solicitado. Se trata de que los estudiantes puedan discutir tanto los criterios de evaluación como sus estándares antes de que el proyecto comience.

QUINTA BUENA PRÁCTICA: Potenciar y fortalecer las creencias motivacionales positivas y la autoestima.

Como ya hemos señalado, la motivación y la autoestima juegan un papel importante en el aprendizaje y la evaluación. Los estudios de Dweck (1999) muestran que los estudiantes poseen diferentes estructuras de motivación en función de sus creencias sobre el aprendizaje. Estas estructuras o marcos de referencia acerca de la motivación afectan las respuestas de los estudiantes frente a la retroalimentación externa e interna y, por tanto, influyen en su compromiso con la autorregulación del aprendizaje.

La investigación en el ámbito escolar ha demostrado que la evaluación frecuente compuesta solo de marcas y calificaciones tiene un impacto negativo sobre la motivación para el aprendizaje que actúa en contra de la preparación para el aprendizaje permanente (Harlen & Crick, 2003). Dweck (1999) sostiene que estas evaluaciones hacen que los estudiantes se enfoquen solo en el desempeño de la meta específica de la prueba, en lugar de centrarse en los objetivos de aprendizaje.

Los estudiantes deben comprender que la retroalimentación es una evaluación no de la persona sino de su desempeño en un contexto

específico. Esto es cierto tanto en la retroalimentación derivada de fuentes externas como en la generada a través de la autoevaluación.

Ejemplos para implementar la práctica:

Se presentan ejemplos de estrategias que ayudan a fomentar altos niveles de motivación y autoestima en los estudiantes:

- Proporcionar marcas en el trabajo escrito solo después de que los estudiantes hayan respondido a los comentarios de retroalimentación (Gibbs, 1999).
- Dar un tiempo para que puedan reescribir o reelaborar su tarea evaluativa o trabajo, pues esto ayudaría a cambiar las expectativas de los estudiantes sobre el propósito y los objetivos de aprendizaje. Ese tiempo debe ser realista.
- Aplicar pruebas automatizadas que provean retroalimentación inmediata.

SEXTA BUENA PRÁCTICA: Proporcionar oportunidades para acortar la brecha entre el desempeño actual y el deseado.

Esta práctica se vincula directamente con la autorregulación, pues la pregunta que subyace es cómo la retroalimentación influye en el comportamiento y el trabajo escolar elaborado por el estudiante.

De acuerdo con Yorke (2003), se pueden distinguir dos preguntas respecto de la retroalimentación externa. En primer lugar, ¿es la retroalimentación de la mejor calidad? y, en segundo lugar, ¿se realizan los cambios en el aprendizaje de los estudiantes? Muchos autores se han centrado en la primera pregunta, pero la segunda es igualmente importante.

La retroalimentación externa proporciona una oportunidad para cerrar una brecha entre el desempeño actual y el rendimiento esperado por el profesor. Como Boud y Molloy (2012) señalan: la única manera de saber si los resultados de la retroalimentación mejoran el aprendizaje es que los estudiantes desplieguen algún tipo de respuesta para completar el ciclo de retroalimentación. Este es uno de los aspectos olvidados más frecuentemente en la evaluación formativa.

Ejemplos para implementar la práctica:

Se presentan ejemplos de estrategias que ayudan a los estudiantes a usar la retroalimentación externa para regular y cerrar la brecha de desempeño:

- Proporcionar retroalimentación sobre el trabajo en progreso y aumentar las oportunidades para reelaborarlo o reescribirlo.
- Introducir las tareas evaluativas en dos etapas del ciclo de evaluación: una primera etapa formativa, con entrega de retroalimentación, y una segunda etapa en la que incorporan las recomendaciones y mejoran su tarea, trabajo o desempeño (Gibbs & Simpson, 2004).
- Modelar las estrategias que se debieran utilizar para cerrar una brecha de desempeño en la clase (por ejemplo, el docente modela frente a los estudiantes cómo estructurar un ensayo a partir de una pregunta).
- Proporcionar específica y explícitamente algunos aspectos a mejorar o líneas de acción, junto con la entrega de la retroalimentación normal al estudiante.
- Involucrar a los estudiantes en un trabajo grupal orientado a la identificación de sus propios puntos de actuación en clases, después de haber leído los comentarios acerca de sus tareas evaluativas. Esta estrategia permite integrar la retroalimentación a la enseñanza e involucrar a los estudiantes de forma más activa en la generación y el uso previsto de la retroalimentación.

SÉPTIMA BUENA PRÁCTICA: Proporcionar información a los profesores para ayudar a enriquecer la forma y estrategias para la enseñanza.

Esta práctica trata de proporcionar una buena información a los docentes y como señala Yorke (2003), la evaluación tiene un efecto sobre el profesor, quien puede adaptar su enseñanza a partir de la retroalimentación.

Al profesor le permite descubrir dificultades o errores de los estudiantes con los aprendizajes (p. ej., errores conceptuales), analizar la efectividad de los métodos de enseñanza, el desarrollo de

tareas de evaluación frecuentes, especialmente las pruebas formativas iniciales, pues le entregan información acumulada sobre los niveles de conocimiento y habilidad de los estudiantes, de modo que pueda adaptar su enseñanza coherentemente.

Ejemplos para implementar la práctica:

Algunas estrategias para ejemplificar el uso que pueden hacer los profesores de esta práctica y ayudarlos a generar y evaluar la calidad de la información sobre el aprendizaje de sus estudiantes son:

- Hacer que los estudiantes soliciten la retroalimentación que les gustaría recibir cuando hacen la entrega de una tarea evaluativa (p. ej., marcando en una pauta donde estén todos los criterios de evaluación presentados al inicio, solo aquellos en los que quieren que los retroalimente el profesor).
- Solicitar a los estudiantes que identifiquen dónde están teniendo dificultades en el desarrollo de la tarea evaluativa.
- Pedirles que en grupos presenten una pregunta relevante, basada en el estudio previo de la tarea evaluativa, sus estándares y metas de aprendizaje, focalizándose en lo que les gustaría explorar en un corto plazo.

EN SÍNTESIS, a lo largo de este capítulo hemos explorado la trayectoria del concepto de retroalimentación, describimos sus principales características y analizamos cuándo esta es efectiva para garantizar el aprendizaje de los estudiantes.

Además hemos revisado el potencial de la retroalimentación para la autorregulación del aprendizaje de los estudiantes, un aspecto que pone en relieve al propio estudiante y sus pares como actores clave para el desarrollo de procesos metacognitivos y factores motivacionales que facilitan el mejoramiento del aprendizaje. Finalmente, presentamos ejemplos para implementar buenas prácticas de evaluación y retroalimentación en el aula y desplegar así todo su potencial formativo.

En el siguiente capítulo profundizaremos en cómo los estudiantes conforman un agente clave para hacer que la evaluación formativa y la retroalimentación sean prácticas sustentables en el aula, analizando su rol en la autoevaluación y la evaluación de pares y estrategias para incluirlos como agentes evaluadores.

BIBLIOGRAFÍA

Angelo, T. A. & Cross, K. P. (1993). *Classroom assessment techniques: a handbook for college teachers* (2nd ed.). San Francisco: Jossey-Bass.

Bangert-Drowns, R. L., Kulik, C. L. C., Kulik, J. A. & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*(2), 213-238.

Black, P. J. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles Policy and Practice, 5*(1), 7-73.

Boud, D. & Molloy, E. (2012). Decision-making for feedback. In D. Boud & E. Molloy (Eds.), *Feedback in Higher and Professional Education* (pp. 202-217). London: Routledge.

Boud, D. & Molloy, E. (2013). Rethinking models of feedback for learning: the challenge of design. *Assessment & Evaluation in Higher Education, 38*(6), 698-712.

Brookhart, S. M. (2008). *How to give effective feedback to your students.* Vancouver: Association for Supervision and Curriculum Development.

Butler, D. L. & Winne, P. H. (1995). Feedback and self-regulated learning: a theoretical synthesis. *Review of Educational Research, 65*(3), 245-281.

Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education, 31*(2), 219-233.

Chanock, K. (2000). Comments on essays: do students understand what tutors write? *Teaching in Higher Education, 5*(1), 95-105.

Clark, I. (2011). Formative assessment: policy, perspectives and practice. *Florida Journal of Educational Administration & Policy, 4*, 158-180.

Cowan, J. (1999). *Being an innovative university teacher*. Buckingham: Open University Press.

Dweck, C. (1999). *Self-theories: their role in motivation, personality and development.* Philadelphia PA: Psychology Press.

Fluckiger, J., Vigil, Y., Tixier, Y., Pasco, R. & Danielson, K. (2010). Formative feedback: involving students as partners in assessment to enhance learning. *College Teaching, 58*, 136-140.

Gibbs, G. & Simpson, C. (2004). Conditions under which assessment supports students' learning? *Learning and Teaching in Higher Education, 1*, 3-31.

Gibbs, G. (1999). Using assessment strategically to change the way students learn. In S. Brown & A. Glasner (Eds.), *Assessment matters in higher education: choosing and using diverse approaches.* Buckingham: Open University Press.

Goldstein, L. (2006). Feedback and revision in second language writing: contextual, teacher and student variables. In K. Hyland & F. Hyland, *Feedback in Second Language Writing: Contexts and Issues* (pp. 185-205). Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139524742.012

Harlen, W. & Crick, R. D. (2003). Testing and motivation for learning. *Assessment in Education, 10*(2), 169-207.

Hattie, J. & Jaeger, R. (1998). Assessment and classroom learning: a deductive approach. *Assessment in Education, 5*(1), 111-122.

Hattie, J., Biggs, J. & Purdie, N. (1996). Effects of learning skills intervention on student learning: a meta-analysis. *International Journal of Educational Research, 11*, 187-212.

Hattie, J. & Gan, M. (2011). *Instruction based on feedback*. Handbook of research on learning and instruction. USA: Routledge.

Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Higgins, R., Hartley, P. & Skelton, A. (2001). Getting the message across: the problem of communicating assessment feedback. *Teaching in Higher Education, 6*(2), 269-274.

Hounsell, D. (2003). Student feedback, learning and development. In M. Slowey & D. Watson (Eds), *Higher education and the lifecourse*. Maidenhead: Open University Press.

Huxham, M. (2007). Fast and effective feedback: are model answers the answer? *Assessment & Evaluation in Higher Education, 32*(6), 601-611.

Hyland, P. (2000). *Learning from feedback on assessment.* In A. Booth & P. Hyland (Eds.), *The practice of university history teaching*. Manchester: Manchester University Press.

James, M. (1992). *Assessment for learning*. Paper presented at the Annual Conference of the Association for Supervision and Curriculum Development. New Orleans.

Johnston, P. H. (2004). *Choice words: how our language affects children's learning*. Portland: Stenhouse.

Kluger, A. N. & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254-284.

Looney, J. W. (2011). *Integrating formative and summative assessment: progress toward a seamless system?* OECD Education Working Papers, 58, OECD Publishing. doi: http://dx.doi.org/10.1787/5kghx3kbl734-en

Molloy, E. & Boud, D. (2013). *Changing conceptions of feedback. Feedback in higher and professional education: understanding it and doing it well*. London: Routledge.

Narciss, S. (2008). Feedback strategies for interactive learning tasks. *Handbook of research on educational communications and technology, 3*, 125-144.

Nicol, D. J. & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218.

Nuthall, G. (2007). *The hidden lives of learners*. Wellington: NZCER Press.

Orsmond, P., Merry, S. & Reiling, K. (2002). The use of formative feedback when using student derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education, 27*(4), 309-323.

Ramaprasad, A. (1983). On the definition of feedback. *Behavioural Science, 28*(1), 4-13.

Ramsden, P. (2003). *Learning to teach in higher education*. London: Routledge.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119-144.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189.

Tunstall, P. & Gipps, C. V. (1996). Teacher feedback to young children in formative assessment: a typology. *British Educational Research Journal, 22*(4), 389-404.

Wiggins, G. (2012). Seven keys to effective feedback. *Educational Leadership, 70*(1), 10-16.

Wiliam, D. (2012). Feedback: part of a system. *Educational Leadership, 70*(1), 30-34.

Winne, P. H. & Butler, D. L. (1994). Student cognition in learning from teaching. In T. Husen & T. Postlewaite (Eds.), *International encyclopedia of education* (pp. 5738-5745), 2nd ed. Oxford: Pergamon.

Yorke, M. (2003). Formative assessment in higher education: moves towards theory and the enhancement of pedagogic practice. *Higher Education, 45*, 477-501.

CAPÍTULO 6

Los estudiantes como agentes de evaluación*

Carla E. Förster

INTRODUCCIÓN

Muchas veces hemos escuchado a los estudiantes decir en los pasillos: “No sé por qué el profesor me puso esa nota” o “nunca entendí lo que tenía que hacer y obvio que me fue mal” y deseamos que no haya sido una experiencia en nuestra sala de clases a la que están haciendo referencia. Si involucramos a nuestros alumnos en el proceso de evaluación, necesariamente, deberán conocer los criterios y lo que se espera que ellos realicen, con lo que nos acercamos más a dos elementos básicos en un docente competente en evaluación: tener criterios claros y conocidos por los estudiantes y el involucramiento de ellos en su proceso de aprendizaje.

Así, incorporar otros agentes o actores, distintos al profesor, en las evaluaciones que realizamos en el aula se enmarca en un cambio que se está llevando a cabo (lentamente) en las concepciones que tenemos como docentes respecto del proceso de enseñanza/evaluación y de cómo el estudiante logra los aprendizajes que esperamos. Diversos autores (Boud, 1995; Sebba et al., 2008; Noonan & Duncan, 2005; Spiller, 2012) señalan que los nuevos enfoques que se asocian a la evaluación para el aprendizaje ponen énfasis en la participación activa de los estudiantes en su propio aprendizaje, su responsabilidad en el proceso, el desarrollo de habilidades metacognitivas y en la necesidad de implementar un modelo dialógico y colaborativo de enseñanza y aprendizaje. Es cada vez más evidente que la evaluación en la que el profesor tiene todo el poder hace que se limite el potencial de aprendizaje de las acciones realizadas en el aula.

Esta idea de incorporar a los estudiantes en su evaluación no es nueva y se han hecho bastantes revisiones que abordan la efectividad de la autoevaluación y de la evaluación de pares en el aprendizaje de los alumnos (Boud, 1995; Boud & Falchikov, 1989; Dochy, Segers & Sluijsmans, 1999; Falchikov & Goldfinch, 2000; Sebba et al., 2008; Topping, 2009). En ellas se pueden distinguir tres tipos de impacto en

los estudiantes: 1) resultados relacionados con el logro en la asignatura, 2) resultados relacionados con la autoestima, y 3) resultados relacionados con su proceso de aprendizaje (metacognición, autorregulación).

Es interesante conocer estos tres impactos porque muchas veces esperamos que nuestras acciones innovadoras o cambios de estrategias evaluativas tengan una repercusión directa en mejorar el aprendizaje que estamos enseñando en ese momento en la asignatura, lo que no siempre ocurre. Probablemente nos sintamos frustrados porque teníamos expectativas altas de que funcionaría, invertimos tiempo y ganas en hacer un cambio en nuestras prácticas evaluativas y finalmente el resultado fue mínimo y habríamos logrado lo mismo con menos esfuerzo. Lo que nos muestran estos estudios es que en la mayoría de los casos, los resultados son del tipo 2 y 3, es decir, se mejora la autoestima de los estudiantes y sus habilidades de pensamiento para enfrentar futuras tareas, mientras que los avances en la tarea específica de la asignatura no son tan evidentes.

En ese contexto es que debemos ser conscientes de nuestro rol docente en formar personas integrales (aunque parezca un cliché) y no solo esperar resultados inmediatos en la asignatura. Un estudiante con mejor autoestima se sentirá más motivado por aprender, más capaz de aprender y de enfrentarse a tareas desafiantes; si desarrolla sus habilidades de pensamiento tendrá mayor capacidad para planificar, monitorear sus acciones, prestar atención selectiva a lo relevante de la tarea que realiza, es decir, será un “mejor” estudiante y esto, por consecuencia, se verá reflejado en los resultados de aprendizaje de nuestra asignatura. La invitación, entonces, es a no desmotivarse y a seguir involucrando a los estudiantes en los procesos evaluativos (Pintrich & Schrauben, 1992).

Como ya hemos visto en capítulos anteriores, hay dos aspectos de la evaluación que definen las decisiones que se toman en este ámbito. La primera es tener claro cuáles son las metas de aprendizaje, que en el caso de nuestro sistema escolar están definidas en el currículum nacional o local, que explicita lo que se espera que el estudiante aprenda (objetivos de aprendizaje). El segundo aspecto corresponde

a los juicios que emitimos sobre el logro del estudiante, que sean lo más fidedignos posible respecto de la meta (Boud, 1995). En este segundo punto es que la incorporación de otros actores puede marcar esa diferencia y tener una brecha más pequeña entre lo que el estudiante realmente aprendió (interno) y lo que evidenciamos en sus desempeños (externo). En la figura 1 se presentan los distintos agentes evaluadores, los cuales serán analizados luego en detalle.

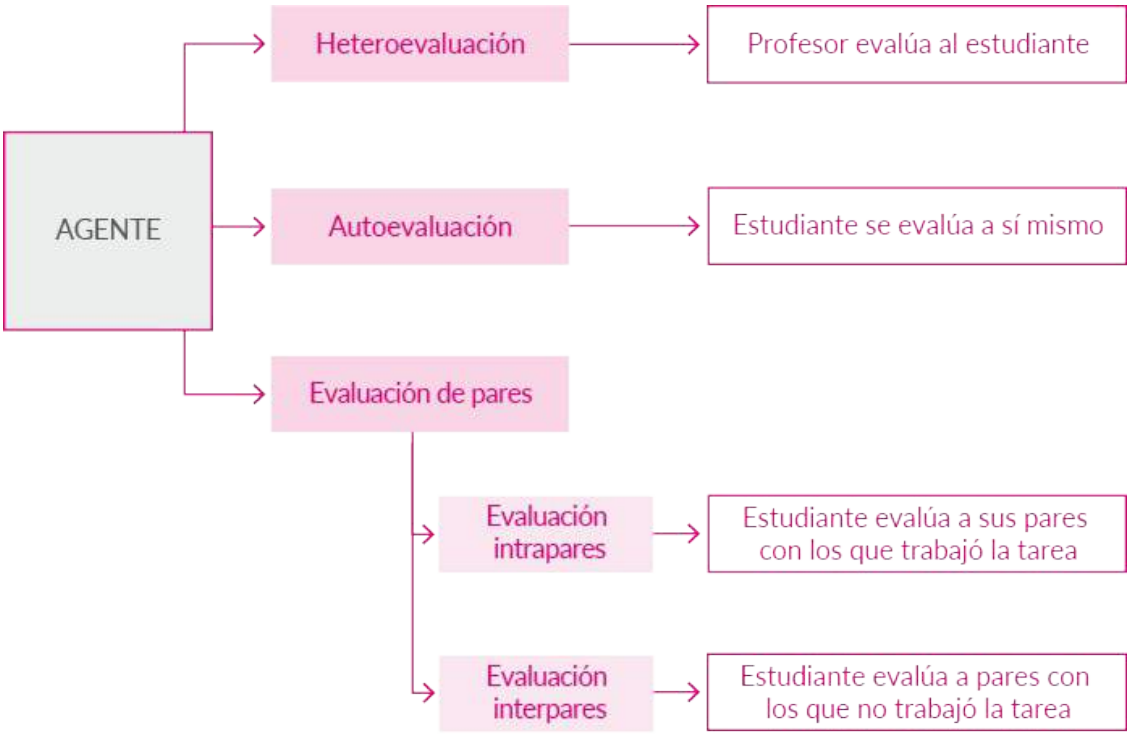


FIGURA 1: Clasificación de los agentes según su rol en la evaluación

Las razones de incorporar a distintos agentes evaluadores en el proceso son varias. Primero, hace que la cultura evaluativa sea mucho más transparente y que los estudiantes adquieran una mejor idea de lo que se espera de ellos y dónde tienen que focalizar sus esfuerzos (Race et al., 2005). Segundo, cuando los alumnos se sienten partícipes de una actividad de enseñanza/evaluación aprenden más profundamente porque esta tiene para ellos un sentido de propiedad sobre las acciones y decisiones asociadas a la tarea. Tercero, el “analizar” y “juzgar” son habilidades cognitivas más

complejas que “leer” o “escuchar”, por lo que una clase donde se incorporan estas habilidades usando como insumo los productos propios o de sus compañeros genera experiencias de aprendizaje más profundos (Race et al., ٢٠٠٥). Por último, pero no menos importante, desde el punto de vista de un docente que incorpora la evaluación formativa como un proceso continuo en su sala de clases, la participación activa de los estudiantes se constituye en algo esencial para hacer sostenible su labor, sin tener una recarga de trabajo producto de múltiples revisiones y de la necesidad de retroalimentación constante (Boud & Falchikov, ٢٠٠٦; Race et al., ٢٠٠٥; Wiliam, ٢٠١٠).

En este capítulo revisaremos las principales características de la autoevaluación y de la evaluación entre pares, qué elementos se deben considerar para que las instancias evaluativas que incorporan a los estudiantes sean buenas prácticas y algunos ejemplos contextualizados que pueden servir a los docentes como ideas a utilizar en sus salas de clases.

LA AUTOEVALUACIÓN

La autoevaluación, en términos generales, se define como el proceso mediante el cual los estudiantes hacen juicios respecto de sus propios logros y procesos de aprendizaje y toman parte en las decisiones sobre las acciones para seguir avanzando (Noonan & Duncan, 2005; Sebba et al., 2008). Andrade y Du (2007) agregan aspectos más precisos señalando que la autoevaluación “es un proceso de evaluación formativa durante el cual los estudiantes reflexionan y evalúan la calidad de su trabajo y su aprendizaje, juzgan el grado en que [su desempeño] refleja metas o criterios declarados explícitamente, identifican fortalezas y debilidades en su trabajo y revisan sus consecuencias” (p. 160).

¿POR QUÉ INCORPORAR LA AUTOEVALUACIÓN COMO PRÁCTICA EVALUATIVA?

La autoevaluación pone el foco en la responsabilidad del estudiante y en la emisión de juicios sobre sus propios desempeños, lo que constituye una habilidad que utilizará durante toda su vida (Boud, 1995) y lo preparará para resolver problemas cuando se enfrente a una tarea (Brew, 1995). También desarrolla en el alumno la conciencia de qué estrategias metacognitivas usar y cuándo utilizarlas (planificación, monitoreo y evaluación de las respuestas) (McMillan & Hearn, 2008). Además se ha visto que hay una estrecha relación entre la capacidad que tienen los estudiantes para autoevaluarse y la mejora en su rendimiento, así como en su motivación y persistencia por terminar una tarea (Kitsantis, Reisner & Doster, 2004).

La autoevaluación también provee información que no siempre es fácil de conseguir de otra manera, como cuánto esfuerzo está invirtiendo un estudiante en una tarea (Ross, 2006); se sabe que la percepción de esfuerzo está asociada a la autoeficacia, por lo que a un profesor, saber esta información, le permite mediar la tarea de forma tal que el estudiante no desista o que invierta menor tiempo y, por tanto, se crea capaz de terminar su tarea con éxito (Pintrich & De Groot, 1990; Pintrich & Schunk, 1996).

Spiller (2012) señala que todos tenemos una tendencia natural a comprobar el progreso de nuestro aprendizaje, y hacer juicios sobre el avance del propio desempeño es parte integral del proceso de aprendizaje, por lo que la autoevaluación fomenta la reflexión sobre el propio aprendizaje y puede promover la responsabilidad y la independencia del alumno; además, fomenta un enfoque de proceso centrado en el estudiante y no en el profesor. De esta forma, las tareas de autoevaluación promueven la apropiación del aprendizaje por parte del estudiante y cambian la idea de que es algo impuesto por otra persona. En este sentido, los alumnos desarrollan habilidades cognitivas complejas, ya que llevar a cabo una autoevaluación requiere habilidades de pensamiento superiores, como interpretar su desempeño usando criterios establecidos (Ross, 2006).

Otro elemento muy importante hoy en nuestras aulas es el abordaje que hacemos de la diversidad, y la autoevaluación se puede

acomodar a un aula heterogénea en cuanto a dominio conceptual, experiencia y antecedentes de los estudiantes.

En la autoevaluación, la distinción entre intencionalidad formativa y sumativa es en general difusa, ya que la retroalimentación sobre su rendimiento es inmediata; así los estudiantes no tienen que esperar a que alguien les diga lo bien o mal que han aprendido, con lo cual una instancia específica de evaluación certifica el logro alcanzado hasta ahí (sumativa), pero como está la opción de seguir mejorando, se transforma en formativa. Sin embargo, la magnitud de la mejora estará condicionada por la comprensión y el compromiso de los alumnos con el objetivo de aprendizaje o las metas a alcanzar y también por su capacidad para identificar sus necesidades y tomar las acciones necesarias para dar el siguiente paso en su aprendizaje. Así, para que la autoevaluación se pueda llevar a cabo de manera efectiva es necesario que tanto el docente como el estudiante tengan una comprensión clara de los objetivos de aprendizaje y de los criterios que deben aplicar para juzgar su grado de cumplimiento de dichos objetivos.

¿CÓMO IMPLEMENTAR LA AUTOEVALUACIÓN?

En la literatura especializada (Boud, 1995; Earl & Katz, 2006; Ross, 2006) se sostiene que la forma en que la autoevaluación se implementa es fundamental para que los estudiantes la acepten como válida y que cumpla con su propósito. Se pueden distinguir cinco elementos comunes, sugeridos por distintos autores, que se deben cumplir para que sea efectiva:

1) Debe haber una razón clara para realizarla. La autoevaluación debe ser una actividad específica que tenga explícitos cuáles son los propósitos, no pueden ser propósitos genéricos, ya que no resultarán significativos para los estudiantes (Boud, 1995). Por ejemplo, la autoevaluación que aparece al final de una lección en los libros de texto donde se plantean las preguntas “¿Qué aprendí hoy?”, “¿Qué me falta por aprender?” se transforma en algo mecánico que el estudiante realizará por cumplir si no hay una acción pedagógica asociada que le dé sentido a la autoevaluación.

2) Los estudiantes deben tener claros los criterios de evaluación y lo que se espera de ellos. Independientemente de la forma en que se establezcan los criterios (dados por el profesor o con participación de los estudiantes), es necesario clarificar estos criterios para que ellos los puedan aplicar correctamente. Los alumnos deben tener un ejemplo que oriente su reflexión respecto de su aprendizaje; para las preguntas “¿Qué aprendí hoy?”, “¿Qué me falta por aprender?”, el estudiante debe saber cómo contestar, no es lo mismo hacer una lista de conceptos que sabe y que le falta por aprender, que señalar algún aspecto específico del tema como, por ejemplo, decir “No entiendo las fracciones” aporta menos que “No sé pasar las fracciones a dibujos”. La calidad y precisión de las respuestas de los estudiantes permite focalizar la forma de lograr el aprendizaje que no ha sido alcanzado aún. En este aspecto es clave hacer un modelaje de lo que se espera que hagan los estudiantes porque no son expertos, necesitan entrenamiento, práctica y apoyo en el desarrollo de habilidades de autoevaluación (Earl & Katz, 2006).

3) Instaurar un entorno seguro en el que los estudiantes puedan ser honestos respecto de su propio desempeño sin temor a exponer información que pueda ser utilizada en su contra. Gran parte de la investigación asociada a autoevaluación sostiene que el aprendizaje puede mejorar de manera más efectiva cuando esta práctica no implica calificación (Kirby & Downs, 2007). Si la autoevaluación es calificada, la tendencia del estudiante será evaluarse sobredimensionando su aprendizaje de manera tal de obtener una buena calificación, especialmente si esta tiene consecuencias como premios, *ranking* o promoción, ¿quién sería honesto sobre su aprendizaje si no tendrá ningún beneficio y por el contrario, lo perjudicará? El incentivo se vuelve perverso y de alto costo para el estudiante, perdiendo todo el propósito formativo de la autoevaluación.

También es interesante la idea planteada por Earl y Katz (2006), quienes señalan que se debe enseñar a los estudiantes que el aprendizaje supone que en algunos momentos nos sintamos

incómodos y que tengamos incertidumbre e inseguridad respecto de lo que estamos haciendo, pero lo importante es resolverlo y salir adelante sin paralizarnos; de esta forma, la autoevaluación honesta no desmotivará al estudiante. Al exponer las fortalezas y aspectos menos logrados, se debe dar el espacio para dialogar y que el alumno pueda tener orientación respecto de cómo abordarlos, andamiando su autonomía.

4) Confianza en que otros estudiantes harán lo mismo, y que la trampa o el engaño serán detectados y desalentados. Este principio parece extraño si se considera la autoevaluación como un proceso individual sin consecuencias para el alumno, pero en el contexto escolar se puede dar una competencia soterrada entre los integrantes de un curso y que puede ser reforzada por el docente (por ignorancia) al poner de ejemplo a un estudiante que ya ha logrado la meta o señalar a quienes les falta lograrla. Además, en la etapa escolar, los alumnos valoran en el profesor el sentido de justicia, y si queremos que la autoevaluación sea una instancia reconocida por su aporte al aprendizaje, no podemos dejar pasar faltas de honestidad (engaño o trampa) en este proceso.

5) La autoevaluación debe ser diseñada para contextos disciplinarios particulares. Muchas veces utilizamos pautas generales que en apariencia sirven para cualquier tarea o producto, sin embargo, si queremos que sea útil al estudiante para mejorar su desempeño en la asignatura, entonces debe ser una evaluación específica y no genérica (Boud, 1995). Esta es una de las principales razones por las que los resultados que se observan en cuanto a la efectividad de la autoevaluación están asociados a autoestima y a habilidades de pensamiento. Si lo que potenciamos que autoevalúen es precisamente eso, no está mal, pero no es lo que se espera para que el estudiante mejore su desempeño en la tarea específica. Por ejemplo, si queremos que nuestros alumnos se autoevalúen en la producción de un texto narrativo, debemos diseñar una pauta que incorpore elementos de la producción de textos, como el tipo de narrador, el tiempo en que ocurre la historia, el ambiente, los personajes. Si mi instrumento de autoevaluación se

sitúa en el proceso general de pensamiento, como por ejemplo, en si tenía los materiales (cuaderno, lápiz de mina, goma) y en si planificó los pasos que haría antes de escribir el texto o si lo leyó después de terminarlo, el foco de la autoevaluación cambia hacia el proceso de aprendizaje general y se sale del objetivo de aprendizaje “producir un texto narrativo”.

También debemos considerar que cuando decidimos incorporar la autoevaluación en nuestras actividades de clases, hay factores como la edad de los estudiantes y su experiencia en prácticas de autoevaluación que pueden influir en el éxito que tengamos con esta acción. Está implícito en estos factores que una autoevaluación efectiva se verá condicionada por las habilidades para comunicarse con otros que tengan los estudiantes, ya que deben expresar verbalmente o por escrito su pensamiento (Towler & Broadfoot, 1992).

Destacamos además las cuatro fases propuestas por Greenway y Crowther (citados en Munby, Phillips & Collins, 1989) en las cuales se puede apreciar cómo abordar estos factores. Se ha visto que en niños menores de 5 años, las habilidades de autoevaluación son difíciles de trabajar, por dos razones: el desarrollo del lenguaje, que influye en su capacidad para comunicar su pensamiento, y la claridad temporal para determinar lo que hicieron y pensaron en retrospectiva. No significa que los niños en edades tempranas no las puedan realizar, pero es más complejo para el docente recoger evidencias de aprendizaje confiables a través de este agente. Las fases son las siguientes:

1) Fase de conocimiento: esta etapa hace énfasis en que el estudiante recuerde las acciones que ha realizado en una tarea evaluativa y principalmente diferencie lo que “hizo” de lo que “le gustó”. Por ejemplo, al resolver una guía de problemas en Matemáticas, un alumno que no está acostumbrado a autoevaluarse puede decir: “Me gustó el problema de los globos porque me recordaron mi cumpleaños”, y uno más experimentado en la autoevaluación dirá: “Lo primero que hice fue leer la guía para saber lo que tenía que hacer”. Con esto podemos ver que el estudiante novato no está autoevaluando su trabajo, por tanto, la

siguiente vez que se enfrente a una tarea similar no tendrá experiencia acumulada para abordarla. Es común que en las preguntas de autoevaluación que damos a nuestros alumnos aparezca alguna asociada a lo que más les gustó, pero ya vemos que no es adecuado para desarrollar pensamiento estratégico.

2) Fase de análisis/comprensión: se refiere a la búsqueda de explicaciones de por qué sucedieron ciertas cosas en la forma en que ocurrieron. Tiene relación con hacer consciente la toma de decisiones en la planificación y ejecución de la tarea. Por ejemplo, en una tarea de Tecnología en la que los estudiantes debían diseñar un robot con ojos de luces que prendían (circuito eléctrico), podemos ver distintas respuestas: a) respuesta sobre lo que el alumno aprendió en función de los aspectos positivos: “aprendí que no es fácil, hay que pensar en pegar los botones de la boca con pegamento, la caja donde van las pilas hay que amarrarla con cuerda, para poner bien la pila con la ampolleta hay que poner los cables separados…”. b) Otro estudiante analiza su desempeño desde los aspectos no logrados: “nosotros necesitamos ayuda de la profesora porque no pudimos cortar el cartón para hacer el robot, también tuvimos problemas porque Oliver estuvo pidiendo pintura todo el rato y nos distraía porque nosotros teníamos lista nuestra pintura, pegamento y cinta adhesiva…” (Towler & Broadfoot, 1992).

3) Fase de evaluación: en esta fase el estudiante debe hacer un juicio sobre la tarea realizada y descomponerla explicitando los aspectos positivos y los que no resultaron, para luego emitir un juicio respecto del resultado. Por ejemplo, “El mapa que hicimos en el colegio antes de la visita al museo no me ayudó porque…”, “Teníamos todos los materiales para construir el puente y el diseño estaba bien, pero no nos resultó tan bien como lo diseñamos porque en el grupo teníamos distintas formas de trabajar y nos costó mucho ponernos de acuerdo y finalmente nos faltó tiempo”. Se ha visto que a medida que los estudiantes tienen más edad son capaces de hacer un mejor análisis de lo que aprendieron y de lo que les faltó, especialmente cuando la evaluación se asocia a una

tarea que implica aplicar conocimiento o crear algo a partir de lo aprendido (Towler & Broadfoot, 1992).

4) Fase de síntesis: esta es la etapa final del proceso de autoevaluación y se espera que el estudiante considere lo que aprendió, hizo o dejó de hacer cuando se enfrente a situaciones futuras o en otros contextos. Ejemplo, “Cuando haga de nuevo un experimento, lo primero que tengo que hacer es saber cuál es la hipótesis e identificar las variables”, “Si tuviera que partir de nuevo con el mismo grupo, además de ver los materiales y el diseño, me pondría de acuerdo antes en quién va a pegar los palitos del puente y quién va a dar las instrucciones que están en el diseño”. También en esta etapa se puede ver que la profundidad de las reflexiones varía según la edad y la experiencia en la realización de este tipo de tareas autoevaluativas (Towler & Broadfoot, 1992), por ejemplo, una estudiante de 7 años, frente a la tarea de diseño, señala como conclusión: “Lo que más me gustó fue cuando finalizó la actividad porque estaba muy cansada” y agrega “para la próxima vez debería hacer un diseño más simple”.

Algunas estrategias concretas para desarrollar pensamiento estratégico y potenciar el aprendizaje a través de la autoevaluación son:

Estrategia 1: Preguntar a los estudiantes qué es lo que aprendieron en esta sesión de clase y que lo relacionen con lo que aprendieron en la última sesión, en la semana anterior o en una unidad pasada. De esta manera a ellos les hará sentido lo que están aprendiendo y no será un aspecto aislado y sin asociación con lo que han visto en clases.

Estrategia 2: Intencionar que los estudiantes parcelen su experiencia de aprendizaje y la descompongan en los aspectos positivos y los que necesitan más apoyo. Para esto se puede preguntar qué parte de la tarea resultó difícil o tediosa, y qué parte fue más fácil o entretenida, siempre solicitando que justifiquen su respuesta y den las razones de su clasificación. Estas razones deberían estar centradas en elementos propios de la tarea y del aprendizaje asociado, si no es así, nosotros como docentes

debemos estar atentos para redireccionar su reflexión hacia este foco.

Estrategia 3: Es importante también intencionar que los estudiantes sean conscientes de lo que hacen para aprender y compartan sus experiencias de aprendizaje con sus pares y profesores, de forma tal que puedan recurrir a otros (personas y recursos) cuando lo necesiten. Así, les podemos preguntar por sus conductas como aprendices: ¿Comparten sus dudas y logros con sus compañeros?, ¿Hablan con sus pares sobre su trabajo?, ¿Leen algo más que lo que se les da en la clase para realizar sus trabajos o complementar lo aprendido o lo que les faltó por aprender? Estas preguntas deben asociarse a un ¿por qué?, ya que lo que marcará la diferencia es la justificación que den de sus acciones.

Estrategia 4: Poder recurrir a experiencias previas que han resultado exitosas o a experiencias fallidas permite al estudiante aprender de sus acciones y aplicar lo aprendido en nuevas situaciones, pero para esto es necesario que como docentes dirijamos a nuestros alumnos a reconocer esas experiencias y los elementos puntuales que las han hecho exitosas o fallidas. Preguntas como ¿Han identificado qué trabajo les resultó bien o no tan bien y cuáles fueron las razones de ese resultado?, ¿Qué podrías mejorar si tuvieras que hacer nuevamente esta actividad? o ¿qué volverías a hacer si te tocara realizar una actividad como esta? podrían ser la base para consolidar estos aprendizajes. Se espera que las respuestas, dependiendo de la tarea, tengan una parte asociada a estrategias genéricas como planificación, trabajo en equipo, etc. y otra parte relacionada con aspectos disciplinares específicos de la asignatura.

Estrategia 5: Pedir a los estudiantes que realicen la tarea y luego pasarles un modelo correcto extraído de un libro (si la tarea es realizar un esquema, por ejemplo) o darles una pauta de revisión, idealmente una rúbrica, para que sean ellos mismos quienes vean en qué nivel quedaron. Esta estrategia se puede realizar previo a la entrega de un trabajo, y pedirles que incluyan su autoevaluación. El

profesor revisa sobre esa evaluación conversando con el estudiante los puntos discrepantes.

Estrategia 6: Incentivar a los alumnos a proyectar sus acciones en futuras tareas. Por ejemplo, al final de sus trabajos o guías, plantear una pregunta como las siguientes: ¿Qué tengo que tener presente cuando haga nuevamente un… (trabajo, guía, experimento, etc.)?, ¿Qué acciones no debo repetir si me vuelvo a enfrentar a una tarea como esta?

Estrategia 7: Utilizar formularios KPSI (*Knowledge and Prior Study Inventory*). Este instrumento consiste en una pequeña escala en la que se incluyen indicadores del aprendizaje esperado o aprendizajes previos que se requieren para aprender el nuevo conocimiento, y en los cuales el estudiante debe autoevaluar su nivel de dominio según unas categorías preestablecidas (Young & Tamir, 1977). Hay que considerar que en algunas ocasiones el alumno cree que sabe algo y el profesor debe ayudarle a reconocer que su autoevaluación no fue adecuada respecto de su conocimiento real (Sanmartí, 2007). En el cuadro 1 se puede ver un ejemplo de KPSI para un aprendizaje de Ciencias Naturales:

CUADRO 1: Ejemplo de formulario KPSI de Ciencias Naturales

Formulario KPSI

Marca con una X en la casilla que corresponda al nivel de conocimiento que tú crees que tienes en cada afirmación.

Afirmaciones	Se lo podría explicar a un compañero	Lo comprendo bien	Lo conozco un poco	No lo sé / no lo entiendo
El agua tiene distintos estados en la naturaleza.				
Hay una relación entre el tiempo atmosférico y el ciclo del agua.				
Las marejadas son producidas por el viento.				
Los pueblos originarios creían que los dioses hacían que lloviera.				

Estrategia 8: ¿Cómo sé que logré el aprendizaje esperado? Esta estrategia se basa en una tabla donde el estudiante va llenando para cada criterio de evaluación las acciones que debe realizar (planificación) y luego poner la evidencia que da cuenta de que realizó todo lo que se indicaba. En la tabla 1 se muestra un ejemplo:

TABLA 1: Ejemplo de matriz de logro del aprendizaje esperado[1]

Criterio	Para alcanzar el objetivo (OA), ¿qué necesito hacer?	¿Cuál es la evidencia de que lo hice?
Comunicación científica	• Mis datos deberán estar en un gráfico o tabla y estarán etiquetados correctamente.	

/ uso de datos	• Mis datos deben probar mi exploración. • Cualquiera debe poder leer mi explicación y entenderla.	
Conceptos científicos y contenidos relacionados	• Los términos que debería usar y comprender son: …….. • Las cosas que debo asegurarme de observar o poner atención son: …….. • Una “idea central” que me ayudará a conectar mi aprendizaje a otras cosas que ya sé o que quiero aprender más es: …….	
Herramientas científicas y tecnología	• Estas son las herramientas que necesito usar para recoger los datos y realizar la tarea (el estudiante debe crear una lista). • Yo necesito revisar si hay errores.	
Procedimientos científicos y estrategias de razonamiento	• Mi hipótesis es: …….. • Para completar la tarea necesito seguir estos pasos: …….. • Necesito registrar estos datos: ……..	

¿QUÉ SE PUEDE EVALUAR A TRAVÉS DE LA AUTOEVALUACIÓN?

Entendiendo que la autoevaluación no es un instrumento en sí mismo sino la incorporación del estudiante como agente evaluador de su propio aprendizaje o desempeño, puede entonces integrarse en la mayoría de las actividades de aprendizaje. Hay que tener en cuenta que se deben generar el espacio y los lineamientos para proporcionar oportunidades para que los estudiantes identifiquen o reflexionen sobre su progreso respecto de las metas de aprendizaje esperadas.

Las instancias evaluativas asociadas a tareas prácticas y desempeños más complejos, como elaboración de informes, desarrollo de investigaciones, análisis de fuentes o diseño de proyectos se prestan con mayor facilidad para la autoevaluación de manera estructurada, sin embargo, durante cualquier actividad de clase, los estudiantes pueden ser invitados a monitorear su progreso en el logro de sus aprendizajes, lo importante es que tengan claros y entiendan bien los criterios de evaluación.

La autoevaluación no es recomendable como única instancia evaluativa en procesos sumativos, no obstante es una muy buena

opción si se utiliza conjuntamente con la evaluación de pares y/o profesores.

VENTAJAS Y LIMITACIONES DE LA AUTOEVALUACIÓN

Ventajas

1. Puede proporcionar al profesor información muy práctica y funcional sobre lo que los estudiantes están aprendiendo, el progreso que han hecho, sus dificultades, las acciones que les han resultado exitosas de forma tal de poder apoyarlos en aspectos específicos, focalizando la retroalimentación.
2. Ayuda a los estudiantes a pensar en su propio aprendizaje, progreso y problemas, y luego encontrar maneras de mejorar. Cuando los alumnos son capaces de analizar su propio progreso pueden encontrar maneras, métodos o estrategias para hacer cambios y ser mejores estudiantes, y los empodera en su rol de agentes evaluadores.
3. Algunos estudiantes tienden a ignorar las correcciones, sugerencias o retroalimentación hechas por los docentes, pero cuando están forzados a corregirse, existe una mayor probabilidad de que esos errores sean analizados y abordados.
4. Existe una amplia variedad de técnicas de autoevaluación, por lo que el profesor puede elegir la que considere mejor para la clase, teniendo en cuenta las características de sus alumnos.
5. Ayuda a los estudiantes a tener una idea más clara de cuáles son los objetivos que están tratando de alcanzar y qué es lo que se espera que realicen.
6. Puede dar retroalimentación al profesor sobre el progreso de los estudiantes sin tener que corregir o revisar cada trabajo o tarea que realizan, por lo que le será posible ahorrar tiempo.
7. Tiene un impacto directo en el proceso de aprendizaje. Cuando está bien diseñada y se focaliza en aspectos específicos de la

tarea, puede tener un efecto directo, ya que el estudiante mejora su desempeño inmediato.

Limitaciones

1. Si los estudiantes no tienen claros los objetivos de la tarea o deben autoevaluar aspectos específicos de la disciplina en los que aún no poseen suficiente dominio, el proceso de autoevaluación puede ser inválido, con lo que tanto la retroalimentación para el estudiante como para el docente entrega información errónea o deficiente.
2. Los estudiantes necesitan tener un grado muy alto de conciencia de lo que se espera como estándar para que puedan ser capaces de analizar los errores que han cometido y su desempeño durante el curso o unidad.
3. La autoevaluación puede requerir mucho tiempo, ya que hay que dar espacio para clarificar los objetivos y que realicen las actividades de autoevaluación. Para contrarrestar esto, los docentes deben planificarla de antemano y diseñar un formato adecuado que no tarde mucho tiempo y que sea fácil de revisar.
4. Dependiendo de la tarea, en ocasiones solo es adecuada cuando los estudiantes alcanzan niveles intermedios a superiores de dominio de la temática, porque tienen las habilidades para analizar su desempeño más fácilmente.
5. Tal vez la desventaja más significativa sea la falta de madurez de los estudiantes para realizar una autoevaluación seriamente. Algunos no son conscientes de la gravedad o importancia del proceso, especialmente si está asociado a una calificación; por lo tanto, tienden a asignarse la nota máxima, que altera el resultado final. También puede suceder al revés, estudiantes muy autocríticos pueden subvalorar su desempeño, pero esta situación es menos común.
6. Es complejo “negociar” una calificación respecto de la autoevaluación. Por una parte, los estudiantes no trabajan a

menos que haya un incentivo y, por otra, se recomienda que la autoevaluación no lleve nota. ¿Cómo hacerlo entonces? Como docentes, si no tenemos más opción que asignar una calificación, deberíamos dejar claro a los estudiantes que la autoevaluación no tendrá un peso importante en su calificación final.

7. Necesita ser integrada con otras actividades del aula, si no se planifica, queda descontextualizada y se pierde el poder de retroalimentación y autorregulación que tiene implícita.
8. La autoevaluación solo funciona si los estudiantes han sido "entrenados" para autoevaluarse. La razón de esto es porque hacerlo de forma rigurosa implica conocimientos sobre la disciplina, claridad sobre el objetivo de aprendizaje y habilidades comunicativas para expresar su reflexión.

LA EVALUACIÓN DE PARES

La evaluación de pares se entiende como la evaluación de productos o resultados de aprendizaje realizada por una persona que tiene el mismo nivel jerárquico que quien está siendo evaluado. Se considera un elemento central de la evaluación formativa, que incorpora a los estudiantes en el proceso evaluativo, dándoles una responsabilidad colaborativa y no individual, como en la autoevaluación (Black & Wiliam, 1998; Panadero & Brown, 2017). Los beneficios en el aprendizaje de los estudiantes cuentan con suficiente evidencia, no obstante, la efectividad de este agente evaluador pasa por la mediación que hagamos como docentes.

Para poder profundizar en los beneficios y dificultades asociadas a la evaluación de pares es necesario tener clara su conceptualización. Puede haber distintas definiciones y también se le denomina "coevaluación", sin embargo, hay algunos que señalan que la coevaluación solo se puede realizar al interior de un grupo de trabajo (Casanova, 1995) y otros que amplían este agente a la idea de que el evaluador y el evaluado tienen el mismo estatus jerárquico en la sala de clases (Sanmartí, 2007). En nuestra propuesta hemos preferido hablar de evaluación de pares y no de coevaluación, de forma tal de

no restringir la participación de los estudiantes solo a un tipo de tarea evaluativa (trabajo en grupo), entendiendo que la idoneidad del evaluador está dada por su capacidad para observar los criterios de evaluación que debe aplicar y, por tanto, es labor del docente resguardar la validez del instrumento de acuerdo a quién evalúa.

Reforzando esta idea, distintos autores (ver, p. ej., Race et al., 2005; Sebba et al., 2008) hacen precisiones respecto de la idoneidad para evaluar a un par, señalando que cuando lo que queremos evaluar es el proceso de elaboración de un producto, solo pueden hacerlo quienes han participado de dicho proceso (evaluación intrapares), en cambio si lo que deseamos es evaluar el producto terminado, pueden hacerlo todos los estudiantes (evaluación interpares), siempre y cuando los criterios de evaluación sean explícitos y haya una evidencia clara de este resultado que ellos puedan observar. La figura 1, presentada al inicio del capítulo, muestra esta distinción.

¿POR QUÉ INCORPORAR LA EVALUACIÓN DE PARES COMO PRÁCTICA EVALUATIVA?

Distintos autores (Race et al., 2005; Panadero & Brown, 2017; Topping, 2013) señalan una lista de razones por las que los docentes deberíamos integrar la evaluación de pares en nuestras prácticas de enseñanza/evaluación. Hemos querido rescatar algunas de ellas que nos parece ilustran bien nuestra realidad y dan sentido a la incorporación de este agente evaluador:

1. Los estudiantes ya lo están haciendo informalmente. Cada vez que un estudiante realiza un trabajo, sus amigos y compañeros de curso le dan retroalimentación y hacen comentarios respecto de la tarea, el problema es que no se consideran voces autorizadas (comparadas con el docente) porque los criterios no siempre son claramente conocidos, por lo que establecer y facilitar instancias formales de evaluación de pares legitima la opinión de los estudiantes, agrega compromiso al aprendizaje propio y de otros y aumenta la retroalimentación que recibe un alumno.

2. No podemos hacer tantas evaluaciones formativas como quisiéramos. El aumento en el número de estudiantes que tenemos en sala, la carga de enseñanza con nuevos desafíos, la preparación de materiales y cada vez menos tiempo disponible hacen que las instancias de evaluación para monitorear el logro de los aprendizajes y apoyar a los estudiantes en su progreso sean cada vez más limitadas, dejando espacio solo para aplicar las evaluaciones calificadas que debemos reportar. Este contexto genera que la capacidad para retroalimentar al estudiante de forma individual sea muy acotada, pero la evaluación de pares, cuando se prepara bien, puede ser una opción para entregarle más retroalimentación.
3. La evaluación de pares permite a los estudiantes aprender de los éxitos y errores de los demás. Los estudiantes, mientras evalúan el trabajo de sus compañeros, no pueden evitar notar los elementos que de alguna manera son mejores que los suyos. Del mismo modo, los alumnos descubren todo tipo de errores en los trabajos de sus compañeros, muchos de estos errores los cometieron ellos también, lo que aumenta su conciencia de lo que no hay que hacer. De esta manera, cuando reciben la retroalimentación, verifican también lo que vieron en los trabajos de sus pares (aspectos positivos y errores) y ajustan el suyo, con beneficios para su propio aprendizaje.
4. La evaluación de pares desarrolla habilidades cognitivas de orden superior y habilidades metacognitivas que le servirán para la vida más allá de la escuela. Cuando un estudiante evalúa el trabajo o desempeño de sus compañeros pone en juego habilidades de análisis y de evaluación que se consideran habilidades cognitivas complejas. Además, para realizar la retroalimentación a su par, debe ordenar la información y chequear su revisión para no cometer errores, acciones que constituyen habilidades metacognitivas. Este desarrollo del pensamiento en los estudiantes genera beneficios para el aprendizaje de la asignatura, pero también crea una estructura

mental que le servirá en su actuar a futuro (Candy, Crebert & O'Leary, 1994).

¿CÓMO IMPLEMENTAR LA EVALUACIÓN DE PARES?

Como ya se mencionó, la efectividad de la implementación de la evaluación de pares requiere nuestra mediación como docentes en la actividad y debemos incluirla en la planificación del proceso de enseñanza como una instancia que tiene un propósito y cuenta con tiempo dentro de la clase. A continuación presentaremos una serie de recomendaciones que se sugiere tener en cuenta cuando se quiere implementar actividades evaluativas entre estudiantes.

1. Proporcionar o acordar criterios de evaluación realmente claros. Si los criterios son inequívocos y explícitos, los estudiantes están obligados a seguirlos al evaluar a sus compañeros, con lo que se reduce considerablemente el sesgo que pueden tener cuando evalúan a sus amigos o cuando los evaluados son compañeros con los que tienen conflictos. Es recomendable también solicitar que cualquier juicio que hagan debe estar justificado con evidencias concretas del trabajo o desempeño del estudiante. Esto cumple una doble función, por una parte resguardar que el juicio que emita el estudiante se base en la realidad y, por otra, que la retroalimentación sea más precisa.
2. Modelar a través del análisis de criterios de evaluación. Para que una evaluación de pares sea efectiva se recomienda hacer un ejercicio en clases en que se clarifiquen los criterios de evaluación que se van a utilizar. Esto no es solo explicar oralmente los criterios, que es lo que la mayoría de las veces, por tiempo, hacemos; este ejercicio supone mostrar a los estudiantes ejemplos de trabajos realizados en años anteriores que sean buenos, malos y regulares, y analizar los criterios de la pauta de revisión (idealmente una rúbrica) juzgando esos ejemplos (Race et al., 2005). Así todos los estudiantes tendrán la posibilidad de ver en la realidad qué se entiende por un desempeño óptimo (y sus categorías descendentes) en cada criterio, y también les

permitirá ajustar sus expectativas en relación al producto. Luego de hacer este ejercicio, ellos estarán mejor preparados para evaluar a sus pares y entregarles retroalimentación que les aporte al avance de su trabajo. Como podemos ver, este ejercicio requiere un espacio de tiempo y recursos, que debe estar planificado como una estrategia de enseñanza y de evaluación, de lo contrario no funcionará y la experiencia de aprendizaje no será exitosa.

3. Moderar los juicios y comentarios de los estudiantes a sus pares. Race et al. (2005) plantean que aunque es importante involucrar a los estudiantes en la revisión de tareas evaluativas, cuando estas son calificadas, el docente debería moderar los juicios y comentarios que ellos realizan a sus pares, ya que su valoración variará según la práctica que tengan en este tipo de evaluación y las consecuencias de la instancia para el alumno evaluado. Sin embargo, estos autores señalan que el mayor beneficio es para el estudiante porque recibirá una retroalimentación más detallada que la que puede hacer el profesor (siempre falto de tiempo) y para el docente, ya que siempre es más fácil revisar sobre los comentarios de otro que partir de cero.
4. Gestionar un clima de confianza y respeto en el aula. Hay que considerar también que debemos gestionar un clima de aula que permita que la evaluación de pares sea una oportunidad de aprendizaje y esto supone minimizar la competencia entre compañeros. Por ejemplo, debemos potenciar que nuestros estudiantes no trabajen por ser “el mejor” del curso, situación en que un desempeño óptimo significa “ganarle” al resto, sino fomentar que el espíritu sea que todos pueden ser “los mejores” si nos apoyamos entre nosotros, lo que estamos aprendiendo no es un “concurso”, por lo que en la sala todos ganamos (Race et al., 2005). Esto puede parecer utópico en un mundo cada vez más competitivo, pero si asumimos que nuestro referente de comparación es un aprendizaje esperado (referente criterial) y no lo personalizamos al desempeño del mejor alumno del curso

(referente normativo), entonces los estudiantes no tendrían por qué sentirse amenazados por apoyar a sus compañeros.

5. Introducirla gradualmente. En algunos colegios, la cultura evaluativa está muy centrada en el profesor, por lo que incorporar la evaluación de pares puede ser difícil de comprender, tanto para colegas y equipos directivos como para los propios estudiantes y sus padres. Por eso es recomendable considerar el contexto e introducirla gradualmente, a pequeña escala, hasta que estemos seguros de cómo funcionará.
6. Comunicar a estudiantes y colegas su propósito con la evaluación de pares. Mantener a los estudiantes y profesores informados de lo que está haciendo y por qué ayuda a que comprendan la estrategia global y su relación con el proceso de aprendizaje y no lo perciban como una acción aislada o improvisada.
7. Incorporar la evaluación de pares en el cálculo de la nota final. Aunque no hay acuerdo en la literatura respecto de la asignación de una calificación a la evaluación de pares, se sugiere que si se va a hacer, debe estar clara la forma en que se asignará. Cuando la evaluación es interpares, se puede calcular un promedio de todas las calificaciones dadas por el curso y que esta nota tenga un porcentaje en la nota final, por ejemplo, 30% y el otro 70% corresponde a la nota puesta por el docente.
8. Debe haber posibilidad de apelar al juicio evaluativo. El estudiante tiene derecho a apelar si considera que la evaluación realizada por uno o más pares fue injusta. En este caso, nosotros como docentes debemos generar una instancia y actuar como mediadores en la situación, ellos deben poder exponer sus argumentos y acordar la clasificación de su desempeño; si no se llega a un acuerdo, seremos nosotros quienes demos la resolución final.
9. Dejar tiempo suficiente para realizar la evaluación. Este punto es realmente importante porque en general nosotros conocemos los criterios de evaluación y las categorías de graduación del

desempeño (si estamos usando una rúbrica) y, por tanto, podemos emitir el juicio de forma rápida. Los estudiantes necesitan leer, analizar y luego marcar en la pauta. Si queremos que la evaluación a los pares sea seria y rigurosa, debemos dar el tiempo necesario para que la realicen en esta forma. Por ejemplo, si les pedimos que se evalúen entre grupos para generar una evaluación única, deberán discutir entre ellos los criterios para llegar a un juicio de consenso y eso toma más tiempo.

Lo que hemos planteado sobre las distintas formas de involucrar a los estudiantes supone que nuestras actividades de clases integren estas acciones, y que la evaluación es un proceso fluido y separado de la calificación.

Diversos autores han propuesto y socializado estrategias concretas para trabajar la evaluación de pares, algunas de las cuales se presentan a continuación (Dunn, 2013; Race et al., 2005):

Estrategia 1: Dos estrellas y un deseo. Los estudiantes deben identificar dos aspectos positivos del trabajo de un compañero y luego expresar una sugerencia (deseo) sobre lo que el compañero podría hacer la próxima vez para mejorar algún aspecto específico del trabajo (figura 2a). Se recomienda especialmente para abordar tareas de escritura y producción de textos. Por ejemplo “Quiero darte una estrella para el comienzo de tu historia porque es entretenida y dan ganas de seguir leyendo y otra estrella por la forma en que describiste la casa porque cumple con los criterios ficción. Sugiero que hables más sobre Raquel porque no se entiende tan bien por qué se come el pastel”.

Antes de utilizarla de forma masiva debemos modelar esta estrategia varias veces, con ejemplos del trabajo de estudiantes de otros años, y luego pedirles que usen la estrategia en duplas con sus trabajos. Cuando comenzamos a usarla debemos comprobar que lo están haciendo bien y podemos pedir a los pares que lo han hecho bien que se lo muestren a todo el curso.

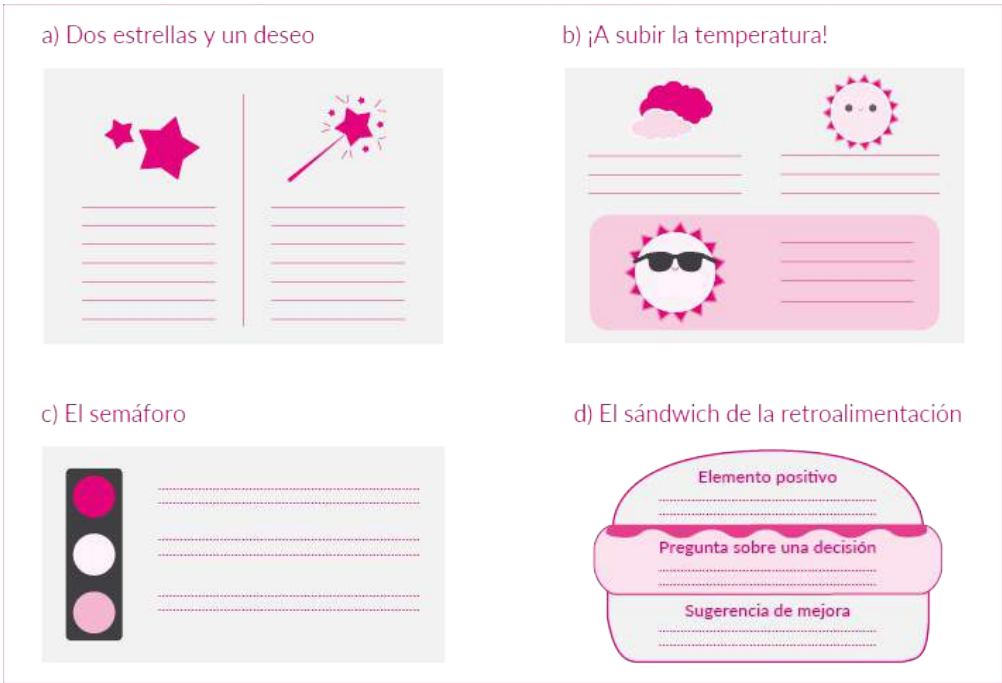


FIGURA 2: Ejemplos de estrategias para trabajar la evaluación de pares

Estrategia 2: Más, menos y ¿qué sigue? Esta estrategia consiste en que los estudiantes comenten a sus pares lo que hicieron bien en relación con los criterios de evaluación establecidos, y lo que podrían hacer mejor. También deben sugerir cómo podrían abordar estos aspectos más débiles. Requiere mayor dominio de los procesos de evaluación, ya que los estudiantes analizan los trabajos de sus pares en función de criterios específicos asociados a la disciplina y la tarea, por lo que se sugiere utilizarla en reemplazo de la estrategia anterior (Dos estrellas y un deseo), que cumple la misma función, pero es menos compleja. También se puede usar como parte de una autoevaluación, en que los estudiantes utilizan la pregunta "¿Qué sigue?" para establecer una meta de aprendizaje personal.

Estrategia 3: Del frío al calor, ¡a subir la temperatura! Esta estrategia se basa en la retroalimentación de tres aspectos: 1) positivos o bien logrados, que serían los puntos cálidos, 2) deficientes o menos logrados, que serían los puntos fríos y 3)

consejos o sugerencias para mejorar su trabajo o subir la temperatura de estos puntos fríos. En esta estrategia se sugiere usar un formato de cartilla donde estén explícitos los criterios de evaluación y las zonas de escritura para el estudiante (figura 2b).

Estrategia 4: El semáforo. Esta estrategia se basa en poner comentarios o sugerencias usando los colores del semáforo, los estudiantes deben marcar con verde los aspectos que están logrados, con amarillo los que requieren mejoras pero están bien encaminados y con rojo lo que está incorrecto y necesita cambios significativos. Siguiendo la idea de incorporar la evaluación formativa, se recomienda utilizar esta estrategia en trabajos que están en etapa de elaboración, de forma tal que los estudiantes puedan abordar estas mejoras en su producto final. Se puede implementar con distintos formatos, por ejemplo, en una hoja se dibuja un semáforo y el estudiante escribe al lado de cada color de luz los aspectos asociados a cada nivel de logro (figura 2c). También se puede hacer con papelitos adhesivos de estos tres colores y el estudiante pega la marca; esta segunda opción requiere que el par evaluador entregue además una retroalimentación oral para explicar por qué puso esa marca.

Estrategia 5: Marcas de pares. Esta estrategia se considera más avanzada en cuanto a complejidad porque consiste en que el estudiante revise la tarea como si fuera el profesor y, por tanto, realice comentarios y marcas en todo el trabajo, asociadas tanto a aspectos disciplinares como formales. Requiere que el grupo tenga un dominio intermedio de los criterios de evaluación para que pueda aplicar estos criterios y la retroalimentación que reciba el estudiante evaluado sea efectiva. Se recomienda implementar esta estrategia como una instancia intermedia, en complemento con la evaluación del docente. Es muy útil en la revisión de informes o textos.

Estrategia 6: El sándwich de la retroalimentación. Esta estrategia es similar a la de “Dos estrellas y un deseo” y se basa en hacer tres comentarios: dos positivos, asociados a los aspectos mejor logrados (pan) y uno que identifica un aspecto que debe ser

mejorado (relleno). También se puede realizar usando las capas del sándwich para estructurar distintos elementos de la retroalimentación, poniendo primero un comentario positivo general, luego una pregunta que apunte a algún aspecto deficitario (¿por qué usaste…?, ¿por qué decidiste hacer… en vez de…?), y después poner una sugerencia de mejora (Dunn, 2013) (figura 2d). Cuál de los dos modelos se utilice dependerá de la edad, experiencia en evaluación de pares y capacidad de análisis de los estudiantes; la primera forma se recomienda más para niños pequeños (primer ciclo) y la segunda, para estudiantes de cursos mayores (segundo ciclo hacia arriba). Hay que tener en cuenta en esta estrategia el equilibrio y peso de los comentarios, ya que se podrían hacer observaciones positivas irrelevantes y que el comentario negativo sea estructural o de tanto peso que destruya el trabajo del estudiante o, viceversa, que sea un comentario negativo de algún aspecto secundario del trabajo que no aporte significativamente a su mejora global.

Estrategia 7: Uso de una rúbrica. Para las actividades de evaluación más amplias llevadas a cabo durante un período de tiempo, la rúbrica, que se ha conversado con los estudiantes al inicio de la tarea, puede utilizarse como base para discusiones en duplas sobre el progreso de la tarea. Si los alumnos tienen claros los objetivos y criterios de evaluación descritos en la rúbrica pueden dar a sus compañeros retroalimentación útil que les permita mejorar.

Estrategia 8: Plenario de los amigos. Esta técnica se trabaja como una actividad en duplas o grupos pequeños, en la que se elige a un representante que cuenta al resto del curso el juicio y comentarios de la evaluación realizada al trabajo de otros al final de la clase. Por ejemplo, los estudiantes realizan un trabajo individual, este trabajo es intercambiado y evaluado entre pares, y luego se reúnen, conversan sobre su evaluación y en el plenario presentan al resto del curso algunos elementos que les parecieron interesantes de compartir. Esta estrategia requiere un espacio en la clase que debe estar planificado y se deben proveer los recursos para que los

estudiantes que presentan en el plenario muestren ejemplos de sus trabajos y que sus compañeros vean cómo aplicaron el criterio. Una forma interesante y rápida es utilizando la tecnología disponible, por ejemplo, que saquen una foto de su trabajo (con su celular) y lo presenten con un proyector al curso, o lo manden al grupo y lo vean directo en sus celulares. También se podrían sacar fotocopias de los trabajos para repartir a los alumnos, lo importante es que puedan acceder de alguna manera a la evidencia. Esta estrategia tiene como beneficio que, al compartir los comentarios, aumentan los ejemplos de desempeños que los estudiantes ven y, por tanto, tienen mayores oportunidades de aprendizaje en una misma instancia.

Estrategia 9: ¡A conversar! Esta estrategia consiste en solicitar a los alumnos que marquen el trabajo de un par poniendo solo una cruz junto a cualquier error, pero sin identificar lo que está mal. Al terminar la revisión se reúnen con su par y deben dialogar sobre estas marcas explicando por qué consideran que es un error y el par evaluado puede argumentar si no está de acuerdo con el juicio. Estas marcas deben utilizar los criterios de evaluación previamente discutidos para la tarea, no son comentarios improvisados desde lo que cada estudiante cree sobre el desempeño del otro, tienen un estándar a la base.

Estrategia 10: Desarrollando lenguaje para evaluar a los pares. Es clave desarrollar en los estudiantes habilidades para realizar la evaluación de un compañero. Entre estas habilidades, el lenguaje con el que hacen los comentarios es clave para generar un ambiente de armonía y confianza que potencie y valide esta experiencia y el *feedback* recibido. En ese sentido, la premisa: “No es solo lo que dices, sino la forma en que lo dices” se transforma en el elemento crucial de cualquier evaluación. Una opción, cuando los niños son más pequeños, es poner carteles en la sala con ejemplos de frases que ellos deben usar (cuadro 2).

CUADRO 2: Ejemplos de frases para desarrollar lenguaje adecuado en la evaluación de pares

“Me gusta esta parte, pero ¿has pensado en...?”.

“¿Qué te hizo usar esta palabra / frase / conector / metáfora y no otra?”.
“La mejor parte es cuando...”.
“Creo que la próxima vez deberías pensar...”.
“Creo que has logrado estos dos criterios de éxito, pero no estoy seguro acerca del tercero. ¿Qué piensas?”.
Adaptado de Dunn (2013)

¿QUÉ SE PUEDE EVALUAR A TRAVÉS DE LA EVALUACIÓN DE PARES?

Igual que con la autoevaluación, se pueden evaluar prácticamente todas las tareas de desempeño que un estudiante realiza, ya que la evaluación de pares supone la incorporación del estudiante como agente evaluador y no un instrumento o técnica en sí misma. Lo importante es que el docente seleccione criterios válidos para lo que se está evaluando y que los estudiantes tengan claros estos criterios y la expectativa de desempeño en la tarea.

Algunas tareas en las que la evaluación de pares se ha visto que presenta beneficios para el aprendizaje del estudiante o consideraciones para que su implementación sea efectiva son:

1. Presentaciones orales temáticas. Los estudiantes pueden evaluar aspectos comunicativos de una presentación y también contenidos de la asignatura. Para incorporar este segundo elemento, los que actúan como evaluadores deben tener suficiente dominio del contenido para poder hacer juicios al respecto.
2. Informes. La evaluación de pares en este tipo de trabajos les ayuda a reconocer buenas y malas prácticas en la redacción de informes, su estructura, coherencia interna y diseño. Como vimos anteriormente, les permite tomar conciencia de aspectos bien y mal abordados por sus compañeros y cuando reciben su retroalimentación incorporan los comentarios que les hacen y también lo que vieron como evaluadores.

3. Diseño y planificación de ensayos. Antes de realizar un ensayo se puede solicitar a los estudiantes que diseñen la estructura y contenidos que tendrá y esta planificación se somete a evaluación de pares. Esta instancia permite al estudiante evaluado recibir retroalimentación de lo que está pensando hacer y así los comentarios le ayudarán a refinar y ajustar su diseño. Esta actividad tiene un impacto importante en un tiempo relativamente breve.
4. Trabajos prácticos. Cuando el estudiante es evaluado en tareas que implican habilidades o destrezas psicomotoras, la presencia del docente mientras está practicando la habilidad puede resultar intimidante, sin embargo, necesita recibir retroalimentación para saber cómo lo está haciendo. En este contexto, la evaluación que realiza un compañero resulta un aporte, ya que es de baja presión para quien está siendo evaluado.
5. Actuaciones y presentaciones orales. Los estudiantes pueden aprender mucho acerca de sus propias habilidades de ejecución evaluando las actuaciones o presentaciones de otros. La evaluación de pares les ayuda a comprender mejor los criterios que se utilizarán y les permite mirar cómo otros los aplican en la práctica, aprendiendo por observación y también al ser más conscientes de la interpretación y aplicación de dichos criterios.

VENTAJAS Y LIMITACIONES DE LA EVALUACIÓN DE PARES

Ventajas

1. Cuando se trata de evaluar el desarrollo de habilidades prácticas que implican desempeños que requieren varios ensayos para lograrlo, involucrar a los estudiantes a través de la evaluación de pares puede ser menos amenazante que cuando es el docente quien evalúa (Race et al., 2005). En general, nosotros como profesores no podemos estar al lado de un estudiante durante un período de tiempo prolongado para entregarle retroalimentación de su desempeño a medida que avanza, por tanto, la evaluación

formativa que podemos hacer “uno a uno” es mucho más limitada que si tenemos la “ayuda” de todos los estudiantes para que se observen unos a otros. Además, que un compañero sea quien observa genera menos ansiedad y resulta menos intimidante para el estudiante que está siendo evaluado.

2. Hay mayor apropiación de los estudiantes en relación con lo que se espera de una tarea y los criterios con los que serán evaluados.
3. Se valoriza la participación y responsabilidad de los estudiantes en el aprendizaje de sus compañeros.
4. Los estudiantes desarrollan habilidades cognitivas de orden superior (análisis y evaluación), ya que deben aplicar los criterios de evaluación a la tarea o desempeño de un compañero. Esto implica analizar los elementos bien desarrollados y deficitarios en la tarea, emitir un juicio y proponer alternativas de mejora, todas habilidades de pensamiento complejas.
5. Reduce la carga de trabajo del docente, ya que permite que el estudiante reciba *feedback* oportuno haciendo que incorporar la lógica de la evaluación formativa en el aula no implique más trabajo para el profesor.

Limitaciones

1. La implementación requiere tiempo, el cual debe estar planificado dentro de la sesión de la clase. Esto puede considerarse por algunos docentes como una pérdida de tiempo o una recarga de trabajo, ya que se requiere elaborar la actividad evaluativa y los recursos que sean necesarios.
2. Asignar una calificación asociada a la evaluación de pares puede tener problemas de confiabilidad si los criterios no son explícitos y no se solicita al par evaluador que justifique su juicio, puesto que la nota puede estar influida por el grado de amistad que tenga con el estudiante evaluado.

3. Los alumnos no siempre tienen el conocimiento para aplicar los criterios de evaluación, por lo que es responsabilidad del docente asegurar la idoneidad de los criterios y de los estudiantes para evaluar a sus compañeros.
4. Para que una evaluación de pares sea exitosa se requiere enseñar a los estudiantes cómo hacerlo, enseñar el lenguaje que se debe utilizar y generar en la sala de clases un ambiente de respeto y confianza. Estas condiciones de base en algunos contextos o cursos específicos pueden ser difíciles de alcanzar, con lo que se pone en riesgo el desarrollo de la actividad.

EN SÍNTESIS, como podemos ver, hay muchas maneras de intencionar que nuestros estudiantes se conviertan en evaluadores efectivos de sus pares y de ellos mismos. Vale la pena perseverar con la evaluación de los compañeros como agentes evaluadores porque los beneficios para el aprendizaje en su clase pueden ser enormes, lo que permite además poder realizar una evaluación formativa sustentable en el tiempo sin aumentar nuestra carga de trabajo.

En el capítulo siguiente analizaremos los criterios técnicos para elaborar y seleccionar instrumentos de calidad que permitan recoger evidencias de los aprendizajes de los estudiantes.

BIBLIOGRAFÍA

Andrade, H. & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education, 32*(2), 159-181.

Black, P. J. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles Policy and Practice, 5*(1), 7-73.

Boud, D. & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment and Evaluation in Higher Education, 31*(4), 399-413.

Boud, D. (1995). *Enhancing learning through self-assessment.* London: Kogan Page.

Brew, A. (1995). Self assessment in different domains. In D. Boud (Ed.), *Enhancing Learning Through Self Assessment* (pp. 129-154). London: Kogan Page.

Candy, P. C., Crebert, G. & O'Leary, J. (1994). Developing lifelong learners through undergraduate education. *National Board of Employment, Education and Training Commissioned Report, 28.* Canberra: Australian Government Publishing Service.

Dochy, F. J. R. C., Segers, M. & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: a review. *Studies in Higher Education, 24*(3), 331-350.

Dunn, D. (2013). Using peer assessment in the primary classroom. *Teach Primary*, 45-46. Recuperado de http://www.teachprimary.com/resource_uploads/a-problem-shared.pdf

Earl, L. & Katz, S. (2006). Rethinking classroom assessment with purpose in mind. Recuperado de https://www.wncp.ca/media/40539/rethink.pdf

Kirby, N. F. & Downs, C. T. (2007). Self-assessment and the disadvantaged student: potential for encouraging self-regulated learning? *Assessment & Evaluation in Higher Education, 32*(4), 475-494.

Kitsantis, A., Reisner, R. A. & Doster, J. (2004). Developing self-regulated learners: goal setting, selfevaluation and organizational signals during acquisition of procedural skills. *The Journal of Experimental Education, 72*(4), 269-288.

McMillan, J. H. & Hearn, J. (2008). Student self-assessment: the key to stronger student motivation and higher achievement. *Educational Horizons, 87*(1), 40-49.

Munby, S., Phillips, P. & Collins, R. (1989). *Assessing and recording achievement.* Oxford: Blackwell.

Noonan, B. & Duncan, R. (2005). Peer and self-assessment in high schools. *Practical Assessment, Research & Evaluation, 10*(17), 1-7.

Panadero, E. & Brown, G. T. (2017). Teachers' reasons for using peer assessment: positive experience predicts use. *European Journal of Psychology of Education, 32*(1), 133-156.

Pintrich, P. R. & Schrauben, B. (1992). Students' motivational beliefs and their cognitive engagement in classroom academic tasks. *Student Perceptions in the Classroom, 7*, 149-183.

Pintrich, P. R. & De Groot, E. V. (1990). Motivational and self-regulated learning components

of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33-40.

Pintrich P. R. & Schunk, D. H. (1996). *Motivation in education: theory, research and applications.* Englewood Cliffs, NJ: Merrill.

Race, P., Brown, S. & Smith, B. (2005). *500 tips on assessment.* New York: Routledge Falmer.

Ross, J. (2006). The reliability, validity and utility of self-assessment. *Practical Assessment, Research & Evaluation, 11*(10), 1-13.

Sebba, J., Crick, R. D., Yu, G., Lawson, H., Harlen, W. & Durant, K. (2008). Systematic review of research evidence of the impact on students in secondary schools of self and peer assessment, technical report. In *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Spiller, D. (2012). *Assessment matters: self-assessment and peer assessment*. New Zeeland: Teaching Development Unit, the University of Waikato.

Topping, K. (2009). Peer-assessment. *Theory into Practice, 48*(1), 20-27.

Topping, K. J. (2013). Peers as a source of formative and summative assessment. In J. H. McMillan (Ed.), *The SAGE handbook of research on classroom assessment* (pp. 395-412). Thousand Oaks CA: SAGE.

Towler, L. & Broadfoot, P. (1992). Self-assessment in primary school. *Educational Review, 44*(2), 137-151.

Wiliam, D. (2010). The role of formative assessment in effective learning environments. In H. Dumont, D. Istance & F. Benavides (Eds.), *The Nature of Learning: using Research to inspire Practice* (pp. 135-155). Paris: OECD.

Young, D. & Tamir, P. (1977). Finding out what students know. *The Science Teacher, 44*, 27-28.

* Este capítulo está basado en los resultados del proyecto Fondecyt de Iniciación Nº 11140713, financiado por Conicyt, Chile.

[1] Ejemplo adaptado de https://www.exemplars.com/assets/files/How_Do_I_Know_Met_Standard.pdf

CAPÍTULO 7

Instrumentos para la evaluación de aprendizajes, ¿con qué evaluar?

Carla E. Förster
Sandra Zepeda Aguirre
Claudio Núñez Vega

INTRODUCCIÓN

Al planificar la estrategia evaluativa y luego al diseñar o seleccionar los instrumentos de evaluación estamos respondiendo a la pregunta ¿con qué o cómo evaluar? Así, la calidad y pertinencia de estos instrumentos tendrá directa relación con la calidad de la información que recogemos respecto del aprendizaje de nuestros alumnos. Si el instrumento tiene problemas de construcción o de alineamiento con los objetivos de aprendizaje, entonces tomaremos decisiones a partir de estos resultados que no darán cuenta de los aprendizajes reales, con lo que nuestras estrategias pedagógicas no serán tan efectivas como esperábamos.

En este capítulo abordaremos aspectos generales asociados a la construcción de situaciones de evaluación en formato de pruebas e instrumentos para registrar y evaluar tareas de desempeño, como son las listas de cotejo, escalas de valoración y rúbricas, analizando sus usos, ventajas y limitaciones. Esperamos que los tópicos aquí tratados den lineamientos que aporten a la elaboración de instrumentos que cuenten con validez, confiabilidad y objetividad en el contexto de la evaluación en aula.

SITUACIONES TIPO PRUEBA

La forma más tradicional de evaluar sumativamente en el contexto escolar es a través de pruebas, la mayoría de las veces escritas, aunque podemos encontrar también interrogaciones orales. Cualquiera sea el formato, lo que esperamos es que los ítemes que constituyen este tipo de evaluaciones estén bien construidos en cuanto a estructura y que tengan relación directa con los objetivos de aprendizaje. Daremos en cada apartado precisiones o particularidades a considerar cuando los ítemes son utilizados de forma individual (fuera de una prueba) con intencionalidad formativa.

TABLA 1: Formato de tabla de especificaciones para una prueba o guía de ejercicios

Objetivo de aprendizaje (OA)	Contenidos	Indicador de evaluación o logro	Habilidad cognitiva	Tipo de ítem	Número de ítem	Peso o porcentaje	Número del ítem en el instrumento
En esta columna está el o los OA que serán evaluados en la prueba o guía. Estos OA son los mismos incluidos en la planificación de la unidad, no corresponden a una nueva elaboración.	Corresponden a los temas abordados en el OA que serán evaluados. En general son los contenidos conceptuales o procedimentales en su dimensión cognitiva los que se incluyen acá. También se podrían incluir contenidos actitudinales, pero hay que tener en cuenta que los estaremos evaluando solo a nivel declarativo y no conductual.	Son los indicadores definidos en la planificación de la unidad, es posible incluir otros o reducirlos si se considera que la forma en que se abordó el OA en clases no es coherente con el indicador de evaluación.	Corresponde al nombre de la categoría de la dimensión de procesos cognitivos (ej. comprender, analizar, crear). Esta columna tiene como propósito chequear que en la prueba o guía se esté abordando diversidad de habilidades cognitivas, por lo que se necesita la categoría global y no el verbo del indicador, ya que pueden haber distintos indicadores que apunten al mismo proceso cognitivo.	En esta columna se pone el o los tipos de ítemes con los que se evaluará cada indicador. Recuerde que este debe ser coherente con la categoría taxonómica, p. ej., si mi indicador de evaluación apunta a diseñar algo, no puedo elegir un ítem de selección múltiple. Para esto es importante tener claras las características y limitaciones de cada tipo de ítem.	En esta columna se indica cuántos ítemes tendré de cada tipo y asociado a cada indicador. De esta forma puedo resguardar el énfasis evaluativo que estoy dando en esta prueba o guía. Es importante acá recordar que según el tipo de ítemes, la puntuación es diferente, por lo que el número en sí mismo no representa el peso, para eso está la siguiente columna.	Porcentaje aproximado que representa el grupo de ítemes en ese indicador respecto del total. Esto permite asegurar la relevancia que estoy otorgando al indicador en la evaluación y chequear que sea acorde con el énfasis y tiempo destinado durante las clases.	Esta columna se llena cuando ya tengo el instrumento diseñado y ensamblado, es decir, cuando ya tengo la versión final lista para aplicarlo. Corresponde al número y/o letra de cada ítem en el instrumento (ej. I. 1, 1. a, 2. c). Recuerde que los términos pareados constituyen un solo ítem.

Para elaborar una prueba, lo primero que debemos hacer es diseñarla, esto implica construir la tabla de especificaciones que guiará la construcción de los ítemes de nuestra prueba. Este primer paso es relevante por cuatro razones: 1) permite identificar qué objetivos de aprendizaje, de los que enseñamos, vamos a evaluar y qué relevancia o peso les vamos a dar, 2) hace posible visualizar la complejidad cognitiva que queremos evaluar en esta prueba, 3) ayuda a dar coherencia al tipo de ítem que elaboraremos para evaluar cada objetivo de aprendizaje y 4) permite organizar el análisis de resultados luego de aplicar la prueba.

Esta tabla de especificaciones, si bien está pensada para diseñar una prueba, también es útil cuando construimos una guía de ejercicios o de práctica para que desarrollen los estudiantes en clases o en su casa, ya que así cuidamos que en la ejercitación estemos incluyendo todos los objetivos de aprendizaje que esperamos lograr, con lo cual resguardamos la validez instruccional. La tabla 1 incluye un ejemplo de formato de tabla de especificaciones con la descripción de lo que debiera ir en cada columna.

Con la tabla de especificaciones definida tendremos una idea más precisa de la prueba o guía de ejercicios que queremos construir y así no nos desviaremos de nuestro propósito, que es recoger evidencias del logro de los aprendizajes de nuestros estudiantes.

Veamos ahora las características y normas de construcción de los distintos tipos de ítemes que se pueden incorporar en una situación tipo prueba. En la figura 1 se presenta una clasificación de ellos.

I. Ítemes de respuesta cerrada

I.1 Selección múltiple

Los ítemes de selección múltiple corresponden a aquellos en los que el estudiante debe seleccionar una opción correcta entre varias alternativas que se le proporcionan, la cual completa el enunciado o responde a la pregunta que se planteó en el encabezado del ítem. Pueden tener una estructura simple o una compuesta. Como se observa en el ejemplo (figura 2), la estructura simple supone una

pregunta y sus opciones de respuesta, mientras que la compuesta tiene la pregunta, una lista de posibles respuestas y luego las opciones que combinan las respuestas presentadas.

FIGURA 1: Clasificación de los distintos tipos de instrumentos y técnicas de evaluación

Estructura simple	Estructura compuesta
1. ¿Qué tipo de representación de la Tierra representa la imagen? a) Planisferio b) Globo terráqueo c) Mapamundi d) Plano	2. ¿Cuál(es) de las siguientes características compartían las civilizaciones maya, azteca e inca? I. Juegos ceremoniales II. Sacrificios humanos III. Religión monoteísta IV. Grandes construcciones a) Solo IV b) II y III c) I, II y IV d) I, II, III y IV

FIGURA 2: Formatos de estructura de un ítem de selección múltiple

Algunos elementos a considerar en su construcción son:

Concordancia gramatical. Se debe mantener la concordancia gramatical entre el enunciado y las opciones, y evitar dar pistas que puedan ayudar al estudiante a descartar opciones sin tener conocimiento sobre la temática evaluada.

Calidad de la pregunta. Cuando se formula un enunciado, este debe ser planteado como pregunta; muchas veces se presentan como completación de la oración, pero esta forma complejiza la comprensión de lo que se solicita al estudiante. Por ejemplo si formulamos el siguiente enunciado “El teorema de Pitágoras es: ..... ”, la concordancia gramatical con las opciones se dificulta, además se pierde el foco de qué es lo que nos interesa que el estudiante demuestre como aprendizaje. Al formularlo como pregunta se precisa lo que queremos como evidencia de aprendizaje: “¿Qué plantea el teorema de Pitágoras?

Calidad del contexto o estímulo. También podemos poner antes de la pregunta algún estímulo o contexto, que puede ser un texto o fuente, un gráfico, una imagen. Lo que tenemos que tener en

cuenta aquí es cuánto aporta este estímulo al ítem, a veces escribimos una introducción que solo agrega carga lectora al estudiante o añade una distracción, si es una imagen, porque la pregunta se podría responder sin esa información. El ejercicio es simple: tapa con tu mano esa parte del ítem y deja solo la pregunta, si la puedes responder, entonces es innecesaria y debes borrarla.

Calidad de las opciones. Lo primero es asegurar que el ítem tenga una única respuesta correcta entre las opciones y si posee más de una opción correcta, señalarlo al estudiante para que elija la mejor. En este caso, debería haber puntuaciones intermedias en las que la más correcta tiene el puntaje máximo y las parcialmente correctas tienen puntajes intermedios distintos a las opciones incorrectas.

En cuanto a la redacción de las opciones, a continuación presentamos algunas sugerencias:

- Las opciones deben ser lo más cortas posible, si una misma frase se repite en todas, esta debería incluirse en la pregunta, para evitar redundancia en las opciones de respuesta.
- Evitar el uso de la opción “todas las anteriores” si es la correcta. La razón es que basta que el alumno identifique dos opciones como posibles para que la marque. Del mismo modo, basta con que identifique una que no es correcta para que la descarte como opción.
- Evitar utilizar la opción “ninguna de las anteriores” si es la correcta, puesto que no sabremos si realmente la respuesta del estudiante refleja dominio del contenido. Por ejemplo, en situaciones que impliquen un cálculo, tanto quien llega al resultado correcto como quien comete un error, pueden marcar esa opción porque no está su resultado en las opciones.
- Las opciones deben ser homogéneas en cuanto al contenido que se está evaluando. Por ejemplo, si se pregunta sobre una obra musical, que todas las opciones hagan alusión a nombres de obras musicales y no mezclar con compositores.

- Si las opciones son numéricas, como fechas, cálculos matemáticos, medidas, se sugiere ordenarlas en orden ascendente o descendente para facilitar al estudiante encontrar la opción visualmente.
- Se sugiere mantener aproximadamente la misma extensión en todas las opciones, evitando que la respuesta correcta sea la más larga.
- Utilizar solo distractores u opciones incorrectas plausibles, que den cuenta de errores conceptuales o preconceptos del contenido a evaluar. El estudiante debe pensar que es una opción real y no descartarla por obvia o porque es demasiado distante de lo que se está preguntando.
- Evitar redactar los enunciados y las opciones en doble negación, por ejemplo ¿Cuál de las siguientes opciones no es falsa?

El número de opciones de respuesta. Existe la creencia de que lo correcto es tener 4 opciones en Enseñanza Básica y 5 en Enseñanza Media, con el propósito de disminuir la probabilidad de responder correctamente por azar. Esta regla viene de una lógica psicométrica que no es válida para evaluar aprendizajes a nivel de aula porque nosotros no tenemos un equipo de constructores de ítemes ni hacemos estudios de pilotaje antes de aplicar una prueba, construimos en poco tiempo y, por tanto, hacer una quinta opción que sea de calidad puede resultar muy difícil o a veces imposible. Por esto, nuestra propuesta es que las opciones sean 3 o 4 según la edad de los estudiantes y que estas sean de calidad en términos de su construcción.

Incorporación de imágenes. Siempre puede resultar atractivo para los niños pequeños poner imágenes en una prueba, especialmente cuando no tienen desarrollada completamente la lectura de textos. También una imagen puede permitir aumentar la complejidad cognitiva de una pregunta cerrada, pero debemos recordar que la impresión en la mayoría de los colegios es en blanco y negro y que al fotocopiarla puede terminar siendo una gran mancha ilegible. Del

mismo modo, no preguntes por colores en las imágenes que no lo tendrán, por ejemplo, ¿Qué tipo de circulación sanguínea representa la parte roja del esquema?, ¿Qué quiere decir el autor con la palabra que está marcada en verde en el texto?, ¿Por qué el afiche incorporó el morado y no el rojo?, ¿Cuál de las siguientes frutas es roja?

Asignación de puntajes. En este tipo de ítemes la norma es que haya una opción correcta y se le asigne 1 punto si está correcta y 0 punto si está incorrecta. Sin embargo, hay ocasiones en las que se busca medir desarrollos intermedios de un aprendizaje y entonces se puede solicitar al estudiante que marque la "más adecuada" para una situación específica y, en ese caso, se podrían asignar puntajes diferenciados entre las opciones parcialmente correctas. Ese tipo de asignación supone considerar que estas preguntas tienen más peso que otras que no tienen dicha lógica. Lo que no debería ocurrir es que un ítem tenga 2 puntos si es correcto y 0 si es incorrecto, ya que no se respetaría la continuidad de los puntajes puesto que ningún estudiante tendría la posibilidad de tener el puntaje intermedio (1 punto).

Instrucciones. Si se presentan varios ítemes de selección múltiple que conforman una sección, se debe dar instrucciones comunes respecto de la forma de responder. Por ejemplo: "Lea atentamente cada uno de los siguientes ítemes y encierre con un círculo la letra que corresponda a la opción correcta" o "A continuación se presentan 15 ítemes de selección múltiple. Lea cada uno de ellos y seleccione la opción correcta marcando la letra que corresponda en el casillero de la hoja de respuestas".

Ventajas	• Permiten cubrir un amplio número de contenidos. • La revisión es fácil, rápida y objetiva (con baja probabilidad de sesgo) por parte del profesor y/o estudiante. • El grado de dificultad del ítem se puede regular a partir de la similitud o complejidad de las opciones que se utilicen. • Favorece el análisis de los resultados de la prueba de forma rápida y simple. • Se pueden analizar patrones de respuestas incorrectas y tomar decisiones pedagógicas si los distractores son errores comunes.

	• Permite generar un banco de preguntas y luego combinarlas para ensamblar nuevas pruebas desechando las que tienen problemas en su aplicación.
Limitaciones	• Hacer buenas preguntas implica tiempo y habilidad. • Comúnmente evalúan habilidades cognitivas de bajo nivel taxonómico, como reconocer y comprender. Si hay contexto, podrían incluir la habilidad de analizar. • No permite evaluar la capacidad de organizar y expresar ideas ni la creatividad de los alumnos. • Es fácil desviarse y preguntar aspectos específicos e irrelevantes del contenido. • Los resultados pueden estar sesgados por la habilidad de comprensión lectora y por la experticia en este tipo de ítem. • Entrega una retroalimentación limitada de la comprensión del estudiante.

I.2 Verdadero o falso

Consiste en una afirmación o frase sobre la que el estudiante debe emitir un juicio en forma dicotómica, es decir, responder Sí o No, Verdadero o Falso, Correcto o Incorrecto.

Redacción de las afirmaciones. Dado que el ítem es una frase, la forma en que se presenta constituye la clave de un buen ítem. Algunas sugerencias para su elaboración son:

- Evitar que la afirmación sea extensa y compleja, pues se agrega una demanda de comprensión lectora innecesariamente. Por ejemplo, “Un cuento es una narración breve de un suceso imaginario con fines morales o recreativos, de contenido expectante, cuya acción se intensifica y aclara en su mismo desenlace”.
- Evitar que la afirmación sea trivial u obvia. Por ejemplo, “Einstein era un científico”; “los peces viven en un ambiente acuático”.
- Resguardar que la afirmación sea directa, sin dobles negaciones (ej. “Los hongos no obtienen nunca las sustancias nutritivas de otros seres vivos”).
- Evitar que la afirmación contenga dos o más aseveraciones (ej. “Las actividades de turismo pronto superarán a las de hotelería, que se han incrementado en los últimos años”).

- Resguardar que las afirmaciones no contengan palabras “trampa” que sean falsas. Por ejemplo, en la oración “En la escuadra chilena estaban los blindados Blanco y Cochrane, las fragatas Esmeralda, O’Higgins, Chacabuco y Abtao”; la palabra “fragatas” no corresponde porque eran corbetas.
- La afirmación debe ser taxativa, es decir, que no haya posibilidad de interpretación.
- Evitar que la afirmación sea parcialmente correcta o que la respuesta implique un “depende”. Por ejemplo, en la oración “El agua hierve a
  100 °C”, si los estudiantes saben que la altura influye en la temperatura de ebullición de los líquidos, es probable que respondan que es Falso, mientras que un estudiante que tiene menos conocimiento contestará que es Verdadero.
- Palabras como “más”, “poco”, “grande” o “bueno” se deberían evitar porque no son suficientemente precisas.

Cabe señalar que estas sugerencias son adecuadas cuando los ítemes serán incorporados en una prueba y, por tanto, se debe resguardar que tengan una única respuesta correcta. Si quisiéramos utilizar este formato para estimular la discusión y la capacidad de argumentación de los estudiantes, no es necesario cumplir estas reglas, por el contrario, se esperaría frases ambiguas y que ambas opciones de respuesta fueran correctas dependiendo de la forma en que se justifiquen.

Instrucciones. Cuando se usa este tipo de ítem es necesario incluir instrucciones precisas para responder. Hay que resguardar que exista coherencia entre la instrucción y el formato del ítem, por ejemplo, si decimos “Encierra en un círculo la V o la F, según corresponda” deben estar estas letras al inicio de la afirmación, aunque parezca obvio. En algunas ocasiones ponemos una línea para que el estudiante rellene con una letra y ahí se genera la inconsistencia.

Cambiar el ítem de cerrado a mixto. Muchas veces se utiliza la frase “justifique su respuesta” o “justifique las falsas” como una forma de

reducir el azar; cuando hacemos esto, el ítem deja de ser cerrado y la habilidad cognitiva a la que se asocia cambia, por lo que hay que asegurarse de que esto quede plasmado en la tabla de especificaciones.

Asignación de puntajes. Este tipo de ítemes es dicotómico y se debería puntuar con 0 si la respuesta es incorrecta y con 1 si es correcta. Cuando se transforma en mixto hay que definir una pauta para revisar esta segunda parte y se debe asignar un puntaje adicional a esta justificación. El problema se produce cuando solo se justifican las falsas porque los estudiantes ven el puntaje total de la sección y pueden calcular cuántas hay y luego responder en función de eso.

Ventajas	• Permiten evaluar una gran cantidad de contenidos. • La corrección es muy fácil y hace posible obtener y procesar la información en un tiempo corto. • Es una buena opción de uso en situaciones de evaluación de carácter formativo, ya que permite identificar los aprendizajes de los estudiantes en aspectos específicos de un determinado contenido. • La carga lectora es baja y permite responder en tiempo reducido.
Limitaciones	• Los ítemes evalúan habilidades cognitivas de bajo nivel de complejidad. • Es difícil elaborar buenos ítemes, muchas veces evalúan detalles o información específica de escasa importancia global. • Tiene una alta probabilidad de respuesta al azar (50%), por lo que el estudiante se siente con confianza para adivinar. Para compensar y lograr mayor confiabilidad se debe tener mayor cantidad de ítemes. Se sugiere más de 20 si la prueba no incluye otro tipo de ítem, resguardando que mida uno o dos objetivos de aprendizaje como máximo para asegurar la confiabilidad.

I.3 Términos pareados

Este tipo de ítemes tiene una estructura en dos columnas y consiste en “una serie de estamentos llamados premisas, los cuales se listan hacia abajo en la columna de la derecha, y una serie de respuestas que se listan en la columna de la izquierda” (Cunningham, 1998, p. 90). Son útiles para evaluar el recuerdo y la comprensión de los estudiantes en relación con un tema a través de la asociación entre los términos y sus significados; según el tipo de premisas y

respuestas que se utilicen, será la complejidad del proceso cognitivo que estos realicen. Por ejemplo, autores y sus obras literarias, fechas y hechos, científicos y sus descubrimientos, conceptos y sus definiciones, órganos y sus funciones, nombres de compuestos y sus fórmulas químicas. Cunningham (1998) señala que los ítemes de términos pareados funcionan como un ítem de selección múltiple con más opciones en que las respuestas correctas de un par se transforman en distractores para los otros pares. La estructura del ítem incluye una instrucción, la columna de premisas y la columna de respuestas. Se puede adaptar este formato si los niños son pequeños para que en vez de llenar el espacio de respuesta en la premisa, unan con una línea los elementos de ambas columnas.

Entre las sugerencias para su construcción están:

Instrucciones de respuesta. Las instrucciones constituyen un aspecto clave para orientar la respuesta del estudiante, debemos indicar lo que esperamos que asocie y la forma en que esperamos que responda. Por ejemplo, “Encuentra la pareja que corresponde uniendo con una línea los personajes con su cuento”, “Une cada parte del sistema digestivo con su función. Para esto escribe la letra que corresponda (columna B) en el espacio disponible de la columna A”, “En el espacio que está entre paréntesis al lado de cada batalla, escribe la fecha que corresponda”.

Aunque no es deseable, en caso de poner opciones de respuesta que se pueden repetir, debemos colocarlo explícitamente en la instrucción. Por ejemplo, “Pon en el espacio correspondiente el número del movimiento artístico al que pertenece cada obra. Puedes repetir alguno de los movimientos”.

Información puesta en cada columna

- Para identificar los pares, las premisas se numeran y las respuestas se identifican con letras.
- Los contenidos deben ser homogéneos. En cada columna tienen que incluirse contenidos de un mismo tipo (ej. autores y sus obras, científicos y sus descubrimientos, hechos históricos

y sus fechas, definiciones y sus conceptos, órganos y sus funciones). Cuando los contenidos son diferentes, se cambia el foco de la evaluación, y lo que medimos no es la comprensión del estudiante respecto de un aprendizaje específico, sino su habilidad para discriminar y descartar posibilidades de pareo.

- La columna de respuesta debe tener un número mayor de distractores para elegir que el número de premisas para responder; con esto evitamos que el estudiante responda al azar o realice pareo por descarte.
- Poner como máximo 10 premisas en un ítem. Cuando se incluyen más, se complejiza la tarea innecesariamente por la carga en la memoria de trabajo y, además, se hace más difícil construirlas para que se mantenga homogéneo el contenido.
- Revisar que cada respuesta tenga solo una correspondencia, o explicitar en la instrucción que se puede repetir.
- Las respuestas deben ser frases cortas, y en caso de ser definiciones, se recomienda que no sobrepasen las dos líneas.
- Presentar el contenido de la columna de respuestas en un orden lógico o cronológico. Se espera que el esfuerzo que nuestros estudiantes inviertan en responder el ítem se focalice en el contenido y no en aspectos secundarios como la búsqueda de la respuesta. Para esto se pueden poner los nombres en orden alfabético y las fechas en orden cronológico.
- Resguardar que el ítem esté completo en una misma página y que no quede cortado en su versión final para imprimir.

Veamos unos ejemplos de ítemes que tienen problemas en su construcción:

1. Une con una línea la **ropa** con la **estación del año** que corresponde.

Algunas reglas de construcción que se cumplen en este ítem son:

- Las instrucciones que indican qué es lo que se debe hacer.
- La presencia de los elementos en dos columnas, en que una agrupa a las estaciones y otra a la ropa.

El problema surge en los dibujos, se recurre a una representación prototípica y no al conocimiento ni experiencias de los estudiantes:

- Las flores y el sol no discriminan suficiente entre primavera y verano.
- En el lado de la ropa, el chaleco puede ser usado en todas las estaciones según el lugar geográfico, en verano hace frío en la playa en la tarde, por ejemplo.
- La ropa de playa incluye elementos que nos son vestimenta (baldes y anteojos).
- Las columnas no tienen título que indique lo que contienen.
- Las columnas tienen igual cantidad de elementos, por lo que se pueden parear por descarte.

Instrucción: Coloca en los espacios de la columna A las letras de la columna B que corresponden a la definición de cada concepto.

Columna A	**Columna B**
1. ..... Vía láctea	A. Grandes cuerpos esféricos que no tienen luz propia y que giran en torno a una estrella.
2. ..... Planeta	B. Conjunto de planetas que giran en torno a una estrella.
3. ..... Estrella	C. Esferas de gases a altas temperaturas que liberan enormes cantidades de energía en forma de luz y calor.
4. ..... Cometa	D. Lugar donde está ubicado nuestro sistema solar y que tiene forma de espiral.
5. ..... Satélite natural	E. Cuerpo pequeño sin luz propia que gira alrededor de los planetas.
	F. Cuerpo que tiene un núcleo formado por rocas y hielo, y una cola que se hace visible a medida que se acerca al sol.

En este ejemplo, los elementos que se cumplen son:

- La instrucción es clara.
- Cada columna tiene contenidos homogéneos.
- Las definiciones son de una extensión adecuada.
- En la columna B hay una respuesta más que las premisas de la columna A.

Lo que no se cumple es:
• Las columnas no tienen un título que indique qué contienen.

Ventajas	• Permite evaluar el grado de asociación, clasificación y relación que hacen los estudiantes en un determinado contenido. • Son ítemes rápidos de revisar. • Se puede abarcar una amplia cobertura de contenidos en poco tiempo. • La revisión no tiene sesgo de quien corrige. • Hay menor probabilidad de adivinar la respuesta correcta en comparación con otros tipos de ítemes.
Limitaciones	• Se evalúan habilidades cognitivas de bajo nivel de complejidad (recordar y comprender). • La construcción está limitada a contenidos que permitan la variabilidad suficiente para construir ambas columnas. • Si se incorporan demasiadas opciones de respuesta, el ítem puede dejar de evaluar el conocimiento relevante de la asignatura y estar midiendo la capacidad de recordar lo leído.

II. Ítemes de respuesta abierta o de desarrollo

Son ítemes en los que los estudiantes deben elaborar una respuesta, generalmente en forma escrita. Se denominan ítemes de respuesta abierta o construida porque, en la mayoría de los casos, para responder hay más de una forma correcta, y requieren que los estudiantes “construyan” o desarrollen sus propias respuestas sin el apoyo que constituyen las opciones de respuesta (Tankersley, 2007). Los ítemes de este tipo pueden variar, desde los que solicitan una respuesta simple de una palabra o un par de frases hasta los que requieren la construcción de un ensayo.

Dependiendo del tipo de ítem, podemos tener una variedad de alternativas de respuesta (puede haber tantas respuestas como estudiantes respondan), por lo que se consideran una buena opción para evaluar tanto el dominio de la disciplina como la creatividad. Pueden ser más difíciles y de mayor complejidad cognitiva que las de respuesta cerrada porque el estudiante debe recurrir a sus conocimientos para elaborar una respuesta, a diferencia de un ítem

de respuesta cerrada, en que puede evocar la información o reconocerla entre las opciones (Popham, 2003).

Hay dos grandes categorías en este tipo de ítemes, las de respuestas breves, cuando los estudiantes responden con una palabra o frase, y las de respuesta extensa, cuando los estudiantes elaboran un texto, esquema, dibujo, etc. Estas últimas implican mayor complejidad cognitiva en su respuesta.

II.1 Ítemes de respuesta breve

En general son ítemes de estructura sencilla para los que se espera una respuesta de unas pocas palabras. Este tipo de ítemes, aunque es de respuesta construida, mide habilidades cognitivas de baja complejidad y tiene características similares a las de respuesta cerrada.

Las principales ventajas que presentan estos ítemes son que se evita el exceso de palabras por parte del estudiante y esto facilita la corrección y ayuda a que sea más objetiva; evitan el factor “azar” porque el estudiante debe dar una respuesta en vez de elegirla, y permiten la evaluación de un amplio número de contenidos o de detalles específicos de estos. Entre las limitaciones que presenta este tipo de ítem está la facilidad de caer en la memorización y repetición de información y, por tanto, evalúan baja complejidad cognitiva.

Cunningham (1998) señala que lo primero al elaborar este tipo de ítemes es seleccionar los objetivos de aprendizaje que pueden ser evaluados con este formato; evitar que las preguntas se transformen en trivias, para lo cual la respuesta correcta debe tener un correlato en algún grado de competencia respecto del objetivo de aprendizaje, y asegurar que el ítem solo tendrá una interpretación posible por parte de los alumnos.

Se distinguen distintos tipos: las simples, de identificación, de completación, de asociación, de ordenamiento y jerarquización.

Simples: corresponden a preguntas simples que implican una respuesta puntual, como por ejemplo, un nombre, una fecha, un número. Una recomendación dada por Cunningham (1998) en relación con las preguntas que involucran una fecha como respuesta

es precisar la unidad en la que se espera dicha respuesta, por ejemplo, si yo espero que me digan el año de un evento, debería preguntar “en qué año…” en vez de “cuándo…”, pues la respuesta podría ser ambigua (ej. podrían responder “el siglo pasado” en vez de “en 1945”).

Algunos ejemplos de este tipo de ítem son:

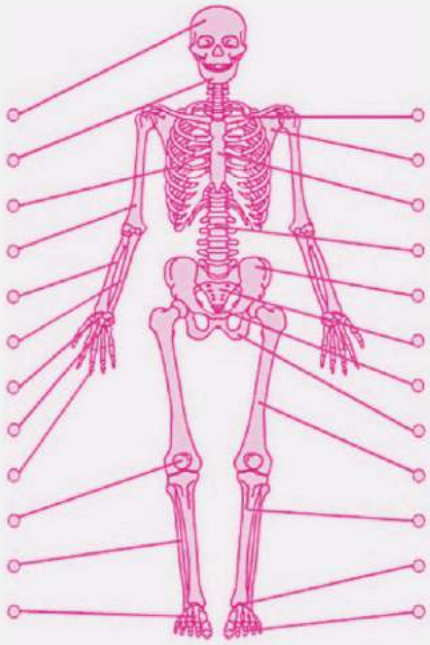
- ¿Cuáles son los elementos químicos que forman el agua?
- Nombra cuatro medios expresivos artísticos.
- ¿Cuáles son los componentes de la oración?
- ¿Quién fue el primer emperador del imperio romano?
- ¿Cuánto es la suma de los ángulos interiores y exteriores de un triángulo?
- Mencione tres propiedades fisicoquímicas del agua.

De identificación: este tipo de ítem consiste en presentar al estudiante una información, en general gráfica o auditiva, y solicitarle que identifique a qué corresponde, respondiendo en el espacio asignado para ello.

En estos ítemes, la respuesta del alumno no debe superar dos o tres palabras en cada elemento a identificar y es necesario asegurar que el espacio sea suficiente. Hay que recordar que este formato evalúa habilidades cognitivas simples, por lo que se deben seleccionar conceptos que sean relevantes en el aprendizaje de la temática abordada.

Un ejemplo de Ciencias Naturales y otro de Lenguaje se analizan a continuación:

En el siguiente esquema del esqueleto humano, señale el nombre de las estructuras que aparecen rotuladas.



En este ítem se evidencian algunos errores:

- No hay suficiente espacio para responder.
- El esquema no permite visualizar con claridad qué hueso es el que está marcado.
- La instrucción no es comprensible para la edad de los estudiantes.

En este ítem se cumple con los criterios de construcción y permite evaluar un conocimiento de la estructura de un tipo de texto y con esa evidencia podemos luego seguir avanzando en que elaboren un texto con este formato. Si vemos que no tienen claros los conceptos, deberíamos reforzarlos para que entiendan las instrucciones de tareas más complejas asociadas a futuros aprendizajes.

De completación: consiste en la presentación de oraciones a las cuales se les han dejado espacios para que el estudiante complete con una palabra que dé sentido a dicha afirmación. Los espacios deben corresponder a conceptos o palabras que resulten claves para el aprendizaje de los estudiantes y que permitan evaluar si lo han logrado. Si estas oraciones están bien elaboradas, pueden ser una buena herramienta para evaluar comprensión de los conceptos y no solo recordar alguna información. Una consideración es que en las instrucciones debe explicitarse si el espacio corresponde a una palabra o puede ser más de una (debería evitarse esto último). Cunningham (1998) plantea en relación a este tipo de ítemes que una práctica común pero inaceptable es extraer frases del libro de texto y borrar una o dos palabras, ya que al sacarla de contexto se tiende a perder su significado y, además, supone que el estudiante debería

memorizar el texto, situación que no tiene ningún sentido. En los ejemplos que están a continuación se grafica esta diferencia (tabla 2):

TABLA 2: Ejemplos de ítemes de respuesta abierta breve de completación

Un buen ejemplo: en este caso, las palabras que faltan están asociadas a aspectos claves del arte bizantino, se entrega suficiente información en el texto para no dar espacio a otra interpretación.	Complete las siguientes afirmaciones, escribiendo en cada espacio la o las palabras que faltan. El arte bizantino se desarrolló en la ciudad de *(Bizancio)* y alcanzó gran apogeo en la técnica del *(mosaico)* , destacándose el de Justiniano y Teodora, que están ubicados en el mausoleo de *(San Vitale / Rávena)* .
Un mal ejemplo: en este ítem, la oración elegida no releva lo importante de este hecho histórico, ya que se ubica en el contexto de la historia de Chile, por tanto, debería estar planteada dejando sin completar la fecha o el hecho (esclavitud indígena) o agregar algo asociado a la institución que lo decretó (Real Audiencia).	Complete en el espacio la palabra que falta. En 1608 se establece la esclavitud de los indígenas de *(Chile)* ; esto refleja los intereses económicos que se tejen en torno a la *(colonia)* .

De asociación: son ítemes que están estructurados de forma tal que el estudiante deba asociar conceptos entre sí, dando evidencias de aprendizaje que permiten evaluar sus habilidades para relacionar elementos de un contenido específico. Para su construcción se deben considerar elementos del mismo tipo en cada columna. Un ejemplo se presenta a continuación (figura 3):

Escriba frente a cada obra el tipo de texto que corresponda:

Obra	Tipo de texto
1. La liebre y la tortuga	1. .....
2. Los duendes y el zapatero	2. .....
3. La plaza tiene una torre	3. .....
4. De oruga a mariposa	4. .....

FIGURA 3: Ejemplo de ítem de asociación para Lenguaje

Ordenamiento y jerarquización: consiste en una lista de elementos (conceptos, hechos, nombres, características, etc.) y se le solicita al estudiante que las ordene utilizando algún criterio que se explicita en la instrucción. Los requisitos y cuidados que debemos tener en su elaboración son similares a los ya mencionados, algunas especificidades se presentan a continuación:

Instrucción: las instrucciones orientan la respuesta del estudiante, debemos indicar la forma en que esperamos que responda. Por ejemplo, “Escriba en el espacio correspondiente el orden en que se producen las distintas fases de la mitosis. 1 es la primera etapa y 4 es la última”.

Elaboración del ítem

- Asegurar la coherencia entre el aprendizaje esperado o el indicador de logro y el ítem que se está elaborando.
- La secuencia debe tener solo una posibilidad de orden correcta.
- La lista no debe exceder los 10 elementos, ya que la carga cognitiva para retener la información complejiza el ítem cambiando el foco desde el dominio disciplinar hacia un dominio cognitivo general.
- La lista debe incluir al menos 4 elementos para disminuir la posibilidad de responder correctamente al azar.
- Si se incluyen figuras para ordenar, por ejemplo, de más grande a más pequeña, se deben situar en una misma línea para que el estudiante no piense que una figura es más pequeña porque está más atrás.

Veamos un ejemplo:

1. Ordene las distintas administraciones político-geográficas de Chile según su nivel de complejidad. Escriba el número en el espacio correspondiente. 1 es el nivel más complejo y 4, el nivel más simple.
   a. ..... Región
   b. ..... Provincia

c. ..... Comuna

d. ..... País

Ventajas	• Permite evaluar el grado de comprensión que tienen los estudiantes respecto de un determinado contenido o procedimiento. • Los ítemes son rápidos de revisar. • La revisión tiene poco sesgo de quien corrige. • El estudiante tiene muy baja carga de escritura. • Es útil para evaluar secuencialidad en Historia y Lenguaje.
Limitaciones	• Se evalúan habilidades cognitivas de bajo nivel de complejidad. • La construcción está limitada a contenidos que permitan hacer jerarquías o seguir una secuencia. • Si el estudiante se equivoca en una respuesta, entonces tendrá todo el ítem erróneo aunque sepa la secuencia en su mayor parte.

II.2 Preguntas de respuesta extensa

Las preguntas de respuesta extensa, también llamadas “de ensayo”, se usan para medir el conocimiento y las habilidades de razonamiento complejo de los estudiantes (Livingston, 2009). Para responder, el estudiante no tiene límite de palabras ni restricciones de estructura. Hay que tomar en cuenta la amplitud de la pregunta, para que la respuesta sea acorde con lo que se desea.

Revisaremos a continuación algunos criterios a considerar para la construcción de ítemes de respuesta abierta o construida:

Instrucción: las instrucciones dan lineamientos generales que orientan al estudiante en la forma que se espera que responda. En caso de ser ítemes cuya instrucción está incluida en la pregunta, no es necesario redactar una instrucción aparte que resultará redundante. Lo importante es entregar la mayor claridad posible al estudiante respecto de qué esperamos que realice para demostrarnos lo que aprendió.

Calidad de la pregunta

- Elaborar los ítemes considerando la planificación realizada en la tabla de especificaciones permite tener claridad del plan de evaluación y la habilidad cognitiva que se desea evaluar. Es importante que tengamos a la vista que cada pregunta evalúa uno o más indicadores. Con esto nos será más fácil saber qué debe contener el ítem que construyamos.
- Escribir la pregunta o tarea en la forma más clara y específica posible respecto del nivel de detalle esperado en las respuestas. Debe especificar claramente la tarea que se le solicita al estudiante, es decir, debe detallar el número de argumentos/fundamentación o aspectos requeridos. Por ejemplo, oraciones como “Explique 2 efectos causados por…” son más precisas que “Explique algunos efectos de…”; del mismo modo, cuando decimos “Describa 3 características de…” intencionamos el número de elementos y la profundidad con la que esperamos que lo haga (describir), mientras que decir “Cuáles son las características de…” supone que debe señalarlas todas, pero no describirlas.
- La redacción y vocabulario de los ítemes deben ser claros y acordes con la edad y características de los estudiantes a los que serán aplicados.
- Se debe preferir el empleo de preguntas en lugar de frases a completar. Por ejemplo, “¿Cuál es la principal causa del calentamiento global?” es mejor que “El calentamiento global es producido por: ..... ” porque acota la respuesta del estudiante a lo que esperamos.
- La pregunta debe implicar habilidades cognitivas superiores de pensamiento. Es necesario constatar que las respuestas no sean información aprendida de memoria, ya que se pierde el potencial de este tipo de ítemes.
- La pregunta no debe incluir información sesgada o que guíe a una respuesta incorrecta, pues con eso estamos engañando al estudiante y no habrá ganancia para nadie, el docente recoge información de mala calidad y el estudiante no demuestra lo

que sabe respecto del tema. En estas “trampas”, los estudiantes que tienen mayor dominio tienden a salir perjudicados porque conocen más sobre el tema y, por tanto, piensan en las excepciones que pueda haber.

- Los ítemes deben apuntar a respuestas elaboradas por los estudiantes a partir de sus experiencias de aprendizaje, información y textos disponibles. No es ético ni se resguarda la validez instruccional si las preguntas abordan temáticas que los estudiantes no tuvieron la posibilidad de aprender.
- Evite comenzar las preguntas con “quién” “qué” “cuándo” “dónde” “mencione” o “enumere”, pues estos términos limitan las respuestas a aquellas de bajo nivel cognitivo; prefiera inicios de preguntas o palabras como “por qué” “describa” “explique” “compare” “analice” “critique” “evalúe”. Estos inicios deben tener además relación directa con el indicador de evaluación. Ejemplos de buenas preguntas que pueden ser adaptadas son las del método socrático.
- El ítem debe estar expresado de tal manera que la pregunta sea abierta y tenga múltiples respuestas posibles que representen la elaboración personal de los estudiantes.

Tiempo de respuesta

- Los ítemes de respuesta abierta tienen como propósito evaluar habilidades cognitivas de orden superior, por lo cual debemos darle al estudiante suficiente tiempo para pensar y escribir.
- Un estudiante demora tres veces más tiempo en responder una evaluación que el profesor que la elaboró. El *Eberly Center for Teaching Excellence and Educational Innovation* (2015) recomienda responder antes la evaluación y considerar este tiempo haciendo los ajustes en extensión o en dificultad de los ítemes si el tiempo calculado fuese mayor que el disponible.
- Informar a los estudiantes el tiempo de que disponen para responder les ayudará a utilizarlo más eficientemente.

Podemos darles algunas recomendaciones de cuánto demorarse por ítem para que alcancen a responder todo.

Pauta de corrección y asignación de puntajes

- Al elaborar los ítemes hay que pensar en cuáles serían las posibles respuestas que esperamos como correctas. En función de esto se puede ajustar lo que estamos pidiendo en el ítem y también permite hacer un esbozo de la rúbrica de revisión cuando se redacta la pregunta. Tanto las rúbricas como otras pautas para revisar deben considerar la variedad de posibles respuestas de los estudiantes y ser coherentes con el contexto y el enunciado.
- Para la corrección se sugiere listar los aspectos que se espera que estén contenidos en la respuesta y luego elaborar una rúbrica que describa los niveles de desempeño. Al hacer este ejercicio en conjunto con la elaboración de los ítemes podemos ver si los puntajes que estamos pensando asignarle a cada criterio son coherentes con la profundidad y número de elementos que esperamos en la respuesta. Por ejemplo, puede ser que yo solicite que el alumno describa 3 características de algo, y a eso le asigno 4 puntos, pero al hacer la graduación de desempeño, me doy cuenta de que la división del puntaje será compleja, ya que si responde correctamente 1 o 2 de las características, tendría que asignar puntajes con decimales (1,33 puntos cada una). Así, puedo decidir asignar 3 o 6 puntos al ítem según el peso que quiera en el total de la prueba y con eso genero una mejor distribución de los puntajes parciales, asignando puntajes enteros a cada descripción. Otra opción de distribución es que yo decida poner 2 puntos a las características nombradas (1 punto si no están todas correctas) y 3 puntos a la calidad de la descripción (1 punto a cada una), con lo cual el puntaje máximo del ítem debería ser 5 puntos. Como ven, no hay una única forma de poner el

puntaje, lo importante es que esta decisión sea coherente con lo que estoy midiendo en el ítem.

- No se recomienda que los estudiantes puedan elegir las preguntas que quieren responder, a menos que así lo requiera lo que queremos evaluar. Cuando los alumnos pueden elegir solo algunas de las preguntas responderán solamente aquellas en las que se sienten más preparados y eso no nos permite evaluar comparativamente sus respuestas para tomar decisiones pedagógicas posteriores. La única opción válida sería generar duplas de ítemes que se consideren paralelos en cuanto a lo que están evaluando y que el estudiante pueda elegir uno de esa dupla, pero esto significa un esfuerzo adicional para el docente, ya que debe hacer el doble de ítemes de buena calidad.

Livingston (2009) sugiere algunos ejemplos de situaciones evaluativas en distintas asignaturas que, en los ítemes de respuesta construida, son una buena opción para evaluar habilidades cognitivas complejas, los que se presentan en el cuadro 1:

CUADRO 1: Ejemplos de tareas útiles como ítemes de respuesta abierta extensa
Lenguaje: Escribir un ensayo comparando y contrastando dos poemas, historias o roles de personajes de obras literarias.
Matemáticas: Escribir una ecuación matemática para resolver un problema presentado con palabras o esquemas.
Ciencias Naturales: Describir cómo ocurre algún proceso biológico en una planta o explicar la habilidad de las plantas para sobrevivir en distintas condiciones ambientales.
Música: Escuchar una melodía o progresión de acordes y escribirla correctamente en notación musical.
Ciencias Sociales: Escribir un ensayo comparando dos instancias de procesos sociales o políticos que ocurrieron en distintos tiempos y regiones del mundo.
Adaptado de Livingston (2009)

	• Crear una respuesta es una tarea de mayor complejidad cognitiva que seleccionarla de una lista de opciones. • El resultado de aprendizaje de los estudiantes constituye una aproximación más cercana al dominio real de conocimiento de la disciplina.

Ventajas	• Admiten la organización libre de la respuesta por parte del estudiante y eso le da mayor flexibilidad para responder. • Evitan el factor azar en la respuesta de los estudiantes. • Son más fáciles de construir que una pregunta cerrada.
Limitaciones	• La corrección de las respuestas tiene un mayor grado de subjetividad. Para evitar esto se sugiere la utilización de pautas de corrección y especialmente rúbricas. • La formulación de preguntas puede ser ambigua. Al requerir una respuesta amplia se puede caer en la imprecisión respecto de lo que se desea que el estudiante responda. • La revisión requiere mayor tiempo que los ítemes de respuesta cerrada o respuesta abierta breve.

III. Ítemes mixtos

Consisten en ítemes que incorporan aspectos tanto de respuestas cerradas como abiertas. Es común encontrar este ítem como una adaptación a los ítemes de Verdadero o Falso, y en los ítemes de Selección múltiple, en los que se solicita al estudiante que justifique su decisión o, en el caso de Matemáticas, se le pide que desarrolle el ejercicio al lado para ver el procedimiento que realizó para llegar a ese resultado.

En rigor, estos ítemes se transforman en respuesta abierta (breve o extensa) según será lo que se solicite, y deben seguirse los mismos criterios y sugerencias antes mencionados.

LISTAS DE COTEJO

También se denominan listas de control, listas o pautas de chequeo o *check list.* Es un instrumento de evaluación que consiste en la enumeración de una serie de elementos cualitativos o cuantitativos cuya presencia o ausencia se desea constatar. Corresponde a una lista de características o conductas que se espera que el estudiante cumpla en la ejecución o aplicación de un proceso, destreza o producto, a través de la cual se verifica el logro de un aprendizaje. La lista de cotejo tiene como objetivo recoger información sobre la conducta del estudiante mediante la observación.

El propósito de las listas de cotejo como instrumento de evaluación al servicio de la mejora de los aprendizajes de nuestros estudiantes se

centra en:

- Proporcionar una herramienta para el registro sistemático de observaciones.
- Entregar al estudiante una herramienta para su autoevaluación.
- Proporcionar al alumno los criterios de evaluación que se considerarán en la revisión de su desempeño o trabajo. Estos deberían ser entregados con anticipación para que él los considere en la preparación de su trabajo y no solo como instancia de retroalimentación de sus logros.
- Clarificar los ajustes necesarios en las oportunidades de aprendizaje que damos a nuestros estudiantes a partir de evidencia concreta sobre sus logros actuales.

Elaborar listas de cotejo no es fácil, ya que debemos considerar varios criterios para que el instrumento sea de calidad y nos proporcione evidencia confiable de los aprendizajes de nuestros estudiantes. Algunos de estos criterios se comparten con las escalas de valoración, por lo que revisaremos los específicos a continuación, los generales solo los nombraremos y los desarrollaremos con más detalle luego de ver las escalas.

Planificación de la lista: esto es fundamental; para hacer una buena lista debemos tener claro el objetivo de aprendizaje que queremos evaluar y la tarea que realizará el estudiante para demostrarlo.

- Los descriptores de desempeño que se incluyan en la lista de cotejo deben ser coherentes y posibles de observar con la tarea. Para asegurar esta coherencia es relevante elaborar una tabla de especificaciones.
- Identificar qué tipo de información o características serán útiles para los propósitos que tenemos ayuda a lograr mayor coherencia y que el instrumento sea válido. Algunos contenidos típicos que se incorporan en las listas de cotejo

son: uso del lenguaje del estudiante, hábitos de trabajo, estrategias de aprendizaje, interacción en el aula.

- Es clave en la planificación decidir a quién vamos a observar (estudiantes individuales o grupos), la frecuencia con que aplicaremos la lista (una o más de una vez), y cuándo lo haremos (en qué clases u ocasiones específicas).

Características de los descriptores: los descriptores son la lista de características que serán chequeadas en el desempeño o producto del estudiante, por tanto:

- Deben ser directamente observables.
- La lista debe estar completa, esto es, recoger el conjunto de características observables del objeto a evaluar, evitando la insuficiencia de aspectos (lo que disminuye la confiabilidad) y la excesiva minuciosidad. Este último punto es de particular importancia porque cuando los aspectos a evaluar son muy específicos, es difícil para el evaluador observarlos, por tanto tiende a marcar lo mismo en todos los aspectos que se parecen, lo que se denomina “efecto halo” (Corbetta, 2003; Himmel, Olivares & Zabalza, 1999).
- Se sugiere confeccionar la lista en conjunto con otras personas y generar consenso respecto de los aspectos que se evaluarán.
- Si se utilizan listas sacadas de internet, es importante verificar la fuente y chequear los criterios básicos de construcción, pues hay una gran cantidad de “listas de cotejo” que no lo son y que corresponden a escalas de apreciación o no tienen un sustento teórico que las apoye, por tanto, no constituyen un elemento confiable ni válido. Sin embargo, pueden ser de gran utilidad para extraer ideas y adaptarlas a los propios requerimientos y necesidades.
- Los descriptores deben tener un lenguaje claro y preciso, es decir, evitar la interpretación de quien las utiliza. Cuando los

descriptores son escuetos (p. ej. “trabajo en equipo”) o tienen un lenguaje inadecuado para quien las utilizará, puede llevar a interpretaciones que disminuyen la confiabilidad y validez del instrumento. Se sugiere considerar la experticia o familiaridad de quien aplicará el instrumento.

- Los descriptores deben evaluar solo un aspecto, ya que se marca la presencia o ausencia, por tanto, al incluir más de un elemento, no es posible diferenciar si la marca corresponde a la ausencia de la característica o a un desempeño parcial del descriptor.
- Los descriptores deben tener una redacción gramaticalmente coherente con las opciones de respuesta. Si los descriptores son de distinta naturaleza se los puede separar en grupos y hacer minilistas utilizando distintas opciones de respuesta.
- La redacción debe ser lo más corta posible sin sacrificar la comprensión del contenido.
- La presentación de los descriptores debe tener una secuencia lógica que puede estar asociada con la observación temporal de la característica en el desempeño, es decir, si estoy revisando un trabajo escrito, los descriptores deberían tener un orden similar al de su aparición en el texto, facilitando así la aplicación.

Sobre la aplicación: una lista de cotejo puede ser de calidad en cuanto a su construcción, pero perder su validez y confiabilidad si se aplica mal. Algunas sugerencias de aplicación se dan a continuación:

- La información recogida es individual, ya que se espera recabar información del desempeño de cada estudiante. Si se utiliza para evaluar productos grupales, el grupo funciona como una unidad individual y los criterios deben ser solo focalizados en el grupo y no en aspectos de los integrantes de manera individual.

- Se pueden usar durante una sesión o en varias, lo que permite una observación sistemática, planificada y orientada en una dirección determinada.
- En una pauta de cotejo, para cuidar la validez de la información, se sugiere aplicarla en más de una oportunidad antes de emitir un juicio valorativo, ya que lo que se observa está en una lógica de presencia/ausencia.

Formatos de presentación y opciones de respuesta: hay diferentes formatos de listas de cotejo, que van más allá de la estructura básica de las tres columnas: 1) descriptores, 2) presencia y 3) ausencia. Algunas opciones de respuesta y variaciones al formato tradicional son:

- Las opciones de respuesta siempre deben ser dicotómicas, por ejemplo, Presente - Ausente, Sí - No, Logrado - No logrado, Observado - No observado.
- Columna de No observado: se marca en caso de no ser posible observar el aspecto en esa oportunidad.
- Columna de Observaciones o comentarios: se agrega un espacio frente a cada descriptor o grupo de descriptores que permite al evaluador escribir comentarios asociados a lo que está observando. Esta práctica permite dar una retroalimentación más precisa al estudiante y también le posibilita al docente reconstruir un hecho pasado, de ser necesario.

En la figura 4 tenemos un ejemplo de lista de cotejo elaborada para evaluar la aplicación de la técnica de pintura con témpera. Como se puede observar, el descriptor 1 considera para su cumplimiento una serie de elementos que el estudiante debe llevar. Si no trae los materiales, los otros descriptores no pueden ser evaluados, pero ¿qué pasa si lleva solo un pincel en vez de dos? En rigor debería ser marcado “No”, ya que no cumple el criterio en su totalidad, pero es muy probable que si esto lo aplicamos en la realidad, le pondríamos

“Sí” pensando en que lo que le faltó no es tan relevante para realizar la tarea. Esta flexibilización es la que hace que el criterio se vuelva ambiguo, ya que dependerá de quien lo aplique, si realiza esta excepción u otra. Esa es la razón de que los descriptores solo respondan a un elemento. Lo mismo ocurre con el descriptor 2.

El descriptor 3, en cambio, es un buen ejemplo, ya que si bien se emplea una redacción ambigua en apariencia “Utiliza en buena forma…”, se define qué significa eso (mezclar de manera homogénea la tempera y el agua); lo mismo se evidencia en el descriptor 4.

Los descriptores 8 y 9 agregan la frase “sin dificultad”, con lo cual hacen ambiguo el contenido, ya que el grado de dificultad es una percepción subjetiva de una persona frente a una tarea, por tanto, que un observador externo, en este caso el docente, emita un juicio respecto de esto, no corresponde; además, implícitamente se alude al “talento” del estudiante. Si en cambio dijera “Aplica aguadas correctamente: utiliza pintura diluida en agua para dibujar un boceto de lo que pintará” hace referencia a la técnica y eso es posible de evaluar. Del mismo modo, en el descriptor 9, lo que nos interesa evaluar es que el estudiante logra mezclar y obtener los colores que desea. Lo que está implícito en esta “facilidad” es la comprensión de los colores primarios, secundarios y terciarios, quien los domina llegará más rápidamente al color que desea. Este descriptor podría ser “Obtiene los colores que se propone” o “Utiliza los conceptos teóricos para obtener colores secundarios y terciarios”.

Otro elemento a considerar en el análisis de este instrumento es que los primeros dos descriptores obedecen a aspectos formales y no directamente al indicador de logro, ya que no son parte del aprendizaje sino elementos de responsabilidad. Si bien pueden ser incluidos en la lista, para dar un mensaje al estudiante de su importancia, en el análisis de los resultados, al emitir un juicio sobre el logro del indicador deberían ser excluidos para no distorsionar la evidencia.

Indicador de logro: Aplica la técnica de la pintura con témpera		
CONDUCTAS A OBSERVAR	**SÍ**	**NO**
1. Trae los materiales requeridos para la técnica: témpera de colores primarios, dos pinceles de distinto tamaño, bloc, vaso para el agua, mezclador.		
2. Trae los utensilios de aseo requeridos: carpeta cubremesa, trapo para limpiar los pinceles.		
3. Utiliza en buena forma el agua: mezcla homogéneamente la témpera con el agua.		
4. Utiliza de buena manera el pincel: lo satura por completo de color.		
5. Limpia adecuadamente el pincel cuando cambia de color.		
6. Utiliza distintos grosores de pincel según corresponda.		
7. Aplica la pintura respetando márgenes y contornos.		
8. Aplica aguadas sin dificultad.		
9. Obtiene los colores que se propone sin dificultad.		

FIGURA 4: Lista de chequeo para evaluar la aplicación de la técnica de pintura con témpera

ESCALAS DE VALORACIÓN

Las escalas de valoración o de apreciación son instrumentos de registro compuestos por un conjunto coherente de descriptores que se consideran indicadores de un concepto más general (Corbetta, 2003). Las escalas permiten a los docentes emitir un juicio respecto del grado o la frecuencia en que el estudiante muestra ciertos comportamientos, habilidades y estrategias. Además, establecen los criterios que serán evaluados y generalmente presentan entre tres y cinco opciones de respuesta para describir la calidad o frecuencia del trabajo del alumno.

Estos elementos o criterios corresponden a una “operacionalización” de un concepto abstracto que no puede ser observado directamente, ante los cuales el docente y/o el estudiante indican el grado o frecuencia en que se evidencia cada uno de ellos en una tarea o desempeño, utilizando un código de valoración preestablecido. Un ejemplo de esta operacionalización es la forma en que se evalúa el

“respeto hacia el otro”. El respeto es un concepto abstracto que no se puede medir directamente, sin embargo, hay conductas observables y medibles que se pueden considerar elementos de respeto, como estar en silencio mientras otro habla o seguir turnos para opinar en una conversación. Es necesario tener claro que estos elementos no abarcan el concepto en su totalidad, pero nos permiten valorar dicho concepto, en mayor o menor medida, dependiendo de la calidad de nuestros descriptores.

Las escalas surgen como respuesta a la necesidad de medir conceptos complejos que no siempre son observables, como el amor, la autoestima, la felicidad, la religiosidad, el respeto o la ansiedad, permitiendo tener una aproximación a estos constructos a través de la observación de conductas, comportamientos y opiniones, y así inferir que estos actos dependen de las características de las personas (Corbetta, 2003). Aunque surgen desde la Psicología, fueron adoptados en el ámbito educativo y hoy constituyen una herramienta muy útil para evaluar actitudes, conductas y desempeños en tareas concretas.

Los propósitos de las escalas de valoración como instrumento de evaluación al servicio de la mejora de los aprendizajes son similares a los de las listas de cotejo antes mencionadas:

- Proporcionan una herramienta para el registro sistemático de observaciones, pero con un juicio graduado y no dicotómico.
- Entregan al estudiante una herramienta para su autoevaluación, especialmente en la calidad de sus trabajos.
- Proporcionan al alumno los criterios de evaluación que se considerarán en la revisión de su desempeño. Estos deberían ser entregados con anticipación para que él los considere en la preparación de su trabajo y no solo como instancia de retroalimentación de sus logros.
- Permiten identificar cuáles son los ajustes que hay que hacer en la enseñanza para generar oportunidades de aprendizaje a

nuestros estudiantes a partir de evidencia concreta sobre sus logros actuales.

TIPOS DE ESCALAS

Las escalas pueden ser numéricas, gráficas o descriptivas según la forma en que se expresen las categorías de respuesta.

Escalas numéricas: en este tipo de escalas, la persona indica el grado o la persistencia con que se expresa el descriptor, marcando un número en un rango que por lo general va entre 3 y 5 puntos. En las instrucciones se debe indicar que la característica evaluada está en grado creciente, es decir, mientras mayor sea el número en la escala, la característica del descriptor está más presente o es más relevante. Por ejemplo, si quisiéramos evaluar el concepto general de “convivencia” podríamos utilizar una escala y observar cómo es la conducta de los estudiantes durante el recreo y generar un perfil para cada uno. En la figura 5 podemos ver que todos los descriptores, con excepción del 1, tienen lo deseable en el máximo puntaje, ya que describen características o conductas positivas, mientras que el descriptor 1 indica lo deseable en el puntaje mínimo, pues lo que se describe es una conducta negativa. Si bien no hay una norma respecto de esto, es importante que en el análisis se considere que el “perfil” de comportamiento ideal no está en 5 para toda la escala.

Indicador de logro: Sigue las normas de conducta social para una sana convivencia escolar

Nombre: *Javier Rojas*

DESCRIPTORES	Categorías				
1. Durante los juegos es agresivo/a con sus compañeros.	1 ✓	2	3	4	5
2. Acepta perder sin pelear con sus compañeros.	1	2	3	4 ✓	5
3. Conversa con todos sus compañeros.	1	2	3	4 ✓	5
4. Comparte su colación.	1	2	3 ✓	4	5
5. Presta sus juguetes a otros.	1	2	3	4	5 ✓
6. Respeta turnos en los juegos.	1	2	3	4 ✓	5
7. ...	1	2	3	4	5

FIGURA 5: Escala numérica para evaluar convivencia escolar

Es importante señalar que en estas escalas, los números no tienen un significado descriptivo, es decir, que el alumno esté en 4 no significa que casi siempre realice esta conducta, sino que en una escala de 1 a 5, ese niño está cercano a la máxima intensidad de la acción evaluada.

Escalas gráficas: las escalas gráficas tienen dos formas de representación: 1) en un perfil gráfico de una escala numérica o descriptiva (figura 6) o 2) para un descriptor hay una escala cuyas categorías son imágenes que representan dicha graduación y el estudiante o el docente valoran el grado de ejecución de una tarea marcando en el ícono o imagen que representa su apreciación (figura 7).

Indicador de logro: Sigue las normas de conducta social para una sana convivencia escolar

Nombre: *Javier Rojas*

DESCRIPTORES	Categorías				
1. Durante los juegos es agresivo/a con sus compañeros.	1 ✓	2	3	4	5
2. Acepta perder sin pelear con sus compañeros.	1	2	3	4 ✓	5
3. Conversa con todos sus compañeros.	1	2	3	4 ✓	5
4. Comparte su colación.	1	2	3 ✓	4	5
5. Presta sus juguetes a otros.	1	2	3	4	5 ✓
6. Respeta turnos en los juegos.	1	2	3	4 ✓	5
7. ...	1	2	3	4	5

FIGURA 6: Perfil gráfico del estudiante en la evaluación de su conducta en convivencia

FIGURA 7: Ejemplo de escala gráfica

Escalas descriptivas: estas valoran la ejecución de una tarea o la opinión frente a un tema utilizando como categorías de respuesta expresiones como “siempre”, “usualmente”, “a veces” y “nunca”, pertinentes al aspecto que se esté evaluando. Entendiendo que en el aprendizaje siempre existe la posibilidad de mejorar, las escalas descriptivas constituyen un buen instrumento para que los estudiantes identifiquen sus fortalezas y necesidades específicas usando estas categorías. Cuanto más precisas y descriptivas sean las palabras para cada punto de la escala, más confiable será la herramienta.

Las escalas de valoración son más eficaces para retroalimentar al estudiante cuando se utilizan descriptores con medidas claramente comprensibles, como la frecuencia. Cuando se habla de “calidad”, las escalas son menos efectivas porque las categorías no contienen suficiente información sobre qué criterios se indican en cada uno de estos puntos y, por lo tanto, se recomienda utilizar rúbricas que describan en detalle estos desempeños parciales.

La gradación puede indicar frecuencia, intensidad o calidad de una tarea determinada; algunos ejemplos de categorías de respuesta se presentan en la tabla 3:

TABLA 3: Ejemplos de categorías de respuesta para escalas de valoración

Tipo de categoría	Ejemplos
Frecuencia	• Siempre, casi siempre, a menudo, casi nunca, nunca. • Habitualmente, generalmente, esporádicamente. • Muchas veces, a veces, algunas veces, rara vez.
Intensidad	• Mucho, bastante, poco, nada. • Muy de acuerdo, de acuerdo, en descuerdo, muy en desacuerdo. • Muy satisfecho, satisfecho, neutral, insatisfecho, muy insatisfecho.
Calidad de la tarea	• Deficiente, suficiente, bueno, excelente. • No logrado, parcialmente logrado, logrado, muy bien logrado.

Analicemos la escala que se presenta en la figura 8; las instrucciones van dirigidas a estudiantes de Enseñanza Básica, pero el lenguaje que se utiliza pareciera orientado a un adulto. Como se señala en las instrucciones, el propósito de esta escala es identificar las estrategias de aprendizaje que utilizan los estudiantes cuando preparan una evaluación, pero los descriptores 2 y 5 se centran en cómo los estudiantes se esfuerzan cuando se enfrentan a una evaluación y no en sus estrategias para estudiar previo a ella. Ahora bien, en cuanto a la elaboración de los descriptores, el 2 se podría decir con menos palabras (ej. “Durante una evaluación me esfuerzo por tener una buena nota”), y en el 3 se hace alusión a un paso previo, decir que utiliza las palabras subrayadas supone que con anterioridad leyó el texto o su cuaderno y subrayó palabras claves. Las categorías de respuesta son coherentes gramaticalmente con los descriptores, pero las abreviaciones que se ponen en la tabla podrían confundir a los estudiantes o hacerlos perder tiempo al tener que volver a la instrucción, especialmente, si están en los primeros años de Enseñanza Básica.

Instrucciones: Las afirmaciones que se presentan a continuación tienen por objetivo identificar las estrategias de aprendizaje que utilizas cuando estudias para una evaluación. Es posible que recurras a cada una de las estrategias de estudio que se indican a continuación, con mayor o menor frecuencia. Te pedimos que señales con una "X" la respuesta que muestre la frecuencia con que normalmente usas las estrategias de aprendizaje mientras estudias. Para ello establecimos cuatro categorías posibles:

N = Nunca,

O = Ocasionalmente,

MV = Muchas veces,

S = Siempre.

Para responder, lee atentamente la frase que describe la estrategia.

DESCRIPTORES	**Categorías**			
1. Elaboro esquemas para comprender la materia.	N	O	MV	S
2. Durante la realización de una evaluación me esfuerzo por obtener una calificación positiva.	N	O	MV	S
3. Al estudiar construyo esquemas utilizando las palabras subrayadas o los resúmenes hechos.	N	O	MV	S
4. Comprendo mejor la materia cuando presto atención en clases.	N	O	MV	S
5. Me esfuerzo por analizar las figuras, gráficos y cuadros presentes en los exámenes.	N	O	MV	S
6. ...	N	O	MV	S

FIGURA 8: Escala de valoración descriptiva para evaluar estrategias de estudio en estudiantes de Enseñanza Básica

RECOMENDACIONES PARA LA CONSTRUCCIÓN DE LISTAS DE COTEJO Y ESCALAS DE VALORACIÓN

Elabore una tabla de especificaciones: igual que para la elaboración de una prueba, estos instrumentos requieren una planificación de su estructura, dimensiones y conceptualización de ellas. A diferencia de las de prueba, acá no se incluyen las habilidades cognitivas asociadas a la tarea porque el instrumento no tiene ese propósito, pero es relevante definir qué se entiende por cada dimensión o subdimensión. En la tabla 4 se presenta la explicación de cada columna y en la tabla 5 se ejemplifica esta estructura:

TABLA 4: Tabla de especificaciones para la elaboración de una lista de cotejo o de una escala de valoración

Dimensión	**Subdimensión**	**Conceptualización**	**Número del descriptor en el instrumento**

En esta columna se coloca el nombre de la dimensión, debe ser manejable, por lo que se recomienda que sea corto pero representativo del contenido.	Esta columna es optativa, solo si la dimensión es muy amplia y se requiere precisar algunas distinciones que se consideran relevantes, entonces se pueden agregar categorías específicas de la dimensión.	En esta columna se espera que se defina breve pero comprensiblemente qué se entenderá bajo el nombre de la dimensión, si bien el nombre es representativo, esta columna da el sustento teórico o marco de referencia en el cual se basa la dimensión.	Esta columna se elabora al finalizar el instrumento y queda listo para su aplicación. Se puede poner acá el número de los descriptores que pertenecen a cada dimensión y subdimensión o se pueden colocar los descriptores completos. Es importante considerar que con esta columna se tiene identificada la pertenencia original de los descriptores a una dimensión, facilitando el análisis de la información recogida.

TABLA 5: Tabla de especificaciones para elaborar una escala de percepción de la violencia en adolescentes

Dimensión	Descripción	Número del descriptor en el instrumento
Violencia entre pares	Justificación de la violencia entre iguales como reacción y valentía.	1-5-8-11-13-14-16
Violencia doméstica	Creencias sexistas y justificación de la violencia doméstica.	2-4-7-9-15-18-21-22
Violencia a minorías	Intolerancia y justificación de la violencia hacia minorías como castigo por ser así.	3-6-10-12-17-19-20

Construya el párrafo introductorio e instrucciones de respuesta considerando la pertinencia. Para esto tenga en cuenta:

- Las indicaciones para la aplicación del instrumento (si es para observar desempeños, requiere un protocolo de aplicación).
- El objetivo de la aplicación.
- Las instrucciones de respuesta.
- La forma en que se analizará la información.
- La explicación de los conceptos que constituyen las opciones de respuesta (cuando son descriptivas).

Elabore los descriptores o afirmaciones considerando los requerimientos de lenguaje y redacción de cada tipo de instrumento. Hay que considerar aquí también cómo será la aplicación (autorreporte o aplicado por el docente o un par), el propósito del instrumento y las definiciones establecidas en la tabla de especificaciones.

Seleccione la categoría de respuesta más apropiada al propósito y la situación evaluativa. Primero, definir si es intensidad, frecuencia o calidad y luego cuáles categorías son más adecuadas. En general se colocan 4 categorías cuando el punto central o neutro no aporta con información relevante y se pone esta categoría intermedia cuando nos interesa saber si el estudiante tiene alguna posición respecto de un tema o no. En el caso de las listas de cotejo, solo son dos categorías de respuesta. En ambos tipos de instrumento hay que resguardar la coherencia gramatical de los descriptores con las categorías de respuesta.

Comparta su trabajo con otros. Siempre es importante que otro par mire nuestro trabajo y actúe como juez experto. Cuando hemos trabajado mucho en un instrumento tendemos a pasar por alto palabras o detalles que se podrían mejorar. Cuando lo ve alguien que no ha participado en la elaboración, puede aportarnos con dudas o comentarios que enriquecerán nuestro instrumento.

Pertinencia de los análisis de la información recogida. Según el tipo de escala, hay diferentes análisis que se pueden realizar. Por ejemplo, si esta evalúa más de un aprendizaje esperado, el análisis debería hacerse para cada aprendizaje o dimensión y no juntar todas las respuestas y hacer un análisis de la escala en su totalidad. Para evitar “caer en esta tentación” se recomienda construir miniescalas unidimensionales, es decir, que todo lo que se evalúe en ellas corresponda a un solo aprendizaje; lo mismo se puede realizar con las listas de cotejo.

Las escalas de valoración son de gran utilidad para evaluar la ejecución de tareas por parte de los estudiantes y sus actitudes frente a un tema en particular, son relativamente fáciles de construir, su aplicación requiere poco tiempo, y la corrección y análisis es de baja

complejidad. Sin embargo, tienen limitaciones dadas por su carácter apreciativo, que implica necesariamente subjetividad de la información recogida. Por esta razón, se debe tener en cuenta este contratiempo al momento de construir el instrumento, tratando de lograr objetividad y fiabilidad de los resultados, y también calificar y tomar decisiones a partir de los resultados obtenidos.

LA RÚBRICA

Cuando tenemos que evaluar desempeños prácticos como una exposición oral, una obra de artes visuales o una ejecución musical, nos enfrentamos al problema de cómo resguardar la objetividad. Generalmente ocupamos una escala de valoración o una lista de cotejo y asignamos puntuaciones a un conjunto de criterios o aspectos que consideraremos. Pero ¿qué pasa cuando la pauta no se ajusta a lo que estamos observando?, ¿qué pasa cuando el producto final es más que las partes que estamos evaluando con la pauta? Otro elemento importante y que cada vez se intenciona más es la evaluación de pares, pero ¿cómo asegurar que todos los alumnos estén usando el mismo criterio de evaluación y que gradúen en igual medida la complejidad del desempeño? Nuevamente, la respuesta es usar una buena pauta de evaluación, no obstante, si un indicador fuera “El volumen de la voz es adecuado” en una presentación oral, debo asegurar que todos los auditores (o la mayoría, al menos) asignen la misma puntuación y eso implica que para todos esté claro qué significa “adecuado” en ese contexto, para estudiantes de cierta edad, etc. Si debemos realizar esta aclaración para cada indicador, la tarea puede ser interminable y, por tanto, inviable para ser aplicada en una actividad de clases.

En este contexto, la rúbrica parece ser un instrumento de evaluación más robusto que las listas de cotejo y las escalas de valoración, pues permite tener explícitamente descritos los desempeños esperados y la graduación de dichos desempeños asociándolos a una categoría que puede ser cuantitativa (un puntaje) o cualitativa (un nivel). Aunque la tendencia actual es asumir la rúbrica como el instrumento de evaluación por excelencia, dada su versatilidad y potencialidad

pedagógica (Blanco, 2008), es necesario tener en consideración sus alcances, pues puede condicionar al profesor en cuanto a los elementos evaluados. Así, una rúbrica no es un instrumento universal, requiere adaptaciones de acuerdo con el contexto de la tarea que será evaluada y las características de los estudiantes a los que se aplicará (Spence, 2010).

QUÉ ES UNA RÚBRICA

El término “rúbrica” proviene del latín *rubricae* o rojo ocre. Este color indica la relevancia de un elemento o aspecto que debe ser explicado, pues se considera importante (Balch, Blanck & Balch, 2016). De esta manera, una rúbrica indicaría o valoraría aquello que es muy relevante.

Las rúbricas son matrices de evaluación o tablas de doble entrada que tienen en uno de sus lados los criterios que se evaluarán y en el otro, la graduación de la calidad o desarrollo del desempeño esperado para cada criterio (Mertler, 2001; Montgomery, 2000). Dicha graduación corresponde a una descripción específica del desempeño del estudiante en un proyecto, presentación, investigación, portafolio o cualquier tarea que implique la observación de distintos niveles de ejecución por parte de los alumnos. Estas descripciones son definidas como una guía para asignar un puntaje o categoría al estudiante respecto de su actuación o rendimiento en una tarea (Blanco, 2008; Mertler, 2001).

La rúbrica es una descripción narrativa concreta de los criterios considerados para emitir un juicio sobre el desempeño de un estudiante en una tarea evaluativa. Dichos criterios se establecen en varios niveles de competencia, los cuales van desde no cumplir con el requisito o criterio hasta demostrar su dominio de forma competente (Balch et al., 2016).

Las rúbricas son esquemas matrices explícitos que se usan para clasificar productos o comportamientos (evidencias) de los estudiantes, en categorías que varían a lo largo de un continuo (Allen, 2014).

Las rúbricas proporcionan los criterios para evaluar el trabajo de los estudiantes. Allen (2014) y Andrade (2000) sostienen que pueden utilizarse para evaluar una multiplicidad de productos o desempeños, y entre ellos se distinguen los ensayos, informes de investigación, portafolios, obras de arte, recitales, presentaciones orales, actuaciones y actividades de grupo. También podemos utilizarlas para que los propios estudiantes o sus pares las apliquen, además del docente (Andrade, 2000).

Las rúbricas son instrumentos versátiles, pues pueden usarse para diversos propósitos. Por una parte, se pueden utilizar para clarificar las expectativas de los estudiantes acerca de las metas de aprendizaje y los estándares de una tarea evaluativa, y por otra, para proporcionar retroalimentación formativa a los estudiantes y también para calificar sus tareas evaluativas y/o para evaluar cursos y programas (Allen & Tanner, 2006; Balch et al., 2016; Brookhart, 2013).

Churches (s. f.) sugiere que las rúbricas son herramientas que se pueden utilizar para proporcionar a los estudiantes retroalimentación efectiva bajo un enfoque de evaluación formativa.

Se plantea que las rúbricas están referidas a criterios y no a normas. Quienes evalúan con rúbricas se preguntan en qué nivel de logro se encuentra el desempeño o producto del estudiante, respecto del estándar definido, y no cómo está en comparación con los demás estudiantes (Allen, 2014).

Esta herramienta de evaluación cumple con los requisitos de fiabilidad (indicando criterios explícitos e indicadores de desempeño), de validez (coherencia con los resultados de aprendizaje declarados) y efectividad (transparencia, consistencia y retroalimentación detallada) (Balch et al., 2016).

APORTE DE LAS RÚBRICAS

A lo largo de este libro hemos examinado que el aprendizaje se despliega a partir del desarrollo de una evaluación formativa, con la que se le proporcione al estudiante una retroalimentación significativa

y efectiva. En este contexto, la rúbrica representa una herramienta que aporta a desarrollar ese ambiente de aprendizaje. A continuación examinaremos algunas de sus ventajas (Allen, 2014; Andrade, 2000; Brookhart, 2013):

- Pueden ser utilizadas para una evaluación formativa y sumativa. En la primera entrega retroalimentación al estudiante para enriquecer su aprendizaje, y en la segunda, clasifica las evidencias y asigna puntajes de manera válida y fiable.
- Aclaran las expectativas de los estudiantes. Ellos aprenden más cuando tienen claro lo que se espera y entienden lo que deben hacer o aprender para lograr el nivel satisfactorio.
- Motivan e involucran a los estudiantes a automonitorear su propio aprendizaje a partir del desarrollo de la tarea evaluativa.
- Permiten que el proceso de clasificación de las evidencias y asignación de puntuaciones (y su posterior calificación) sea más claro tanto para el docente como para el estudiante, por lo que:
    - Agilizan la calificación de las tareas evaluativas por parte de los profesores.
    - Reducen las quejas de los estudiantes acerca de la corrección de las tareas evaluativas.
- Ayudan al equipo pedagógico a crear mejores tareas o situaciones evaluativas que aseguren que los estudiantes muestren lo que se requiere que ellos desplieguen.
- Ayudan a los profesores a adaptar la enseñanza para satisfacer las necesidades y desafíos detectados en el desarrollo de las tareas evaluativas.
- Entregan a cada estudiante una retroalimentación explícita sobre las dimensiones principales del desempeño o producto evaluado. Así él conoce lo que está haciendo bien y lo que necesita para mejorar.
- A modo de conclusión, podríamos sostener que las rúbricas representan una poderosa herramienta para la enseñanza y la

evaluación. Ellas aportan a mejorar el aprendizaje del estudiante, pues ayudan a definir la “calidad” de la meta de aprendizaje y los conducen hacia su logro y, a su vez, aportan al profesor información relevante para orientar la enseñanza y el diseño de tareas evaluativas de sus estudiantes.

TIPOS DE RÚBRICAS

La literatura describe diversos tipos de rúbricas (Allen & Tanner, 2006; Andrade, 2011, Brookhart, 2013), algunas son las rúbricas holísticas y analíticas y otra tipología distingue entre genéricas y de tareas específicas. El uso de uno u otro tipo deberá ser coherente con el aspecto a evaluar y su intencionalidad, pues depende de qué y para qué se evalúa.

Rúbricas genéricas y específicas

Una rúbrica genérica contiene criterios que son generales a una tarea evaluativa concreta y pueden ser utilizados para tareas o actuaciones similares. Los criterios se evalúan por separado y por ello se asocia con una rúbrica analítica.

Una rúbrica de tarea específica evalúa criterios únicos de esa tarea en particular. En esta rúbrica, sus criterios y descriptores reflejan las características específicas del desempeño o producto en evaluación.

La división entre rúbricas genéricas y específicas es difusa, pues la rúbrica de tarea específica suele basarse en rúbricas genéricas. También es posible diseñar rúbricas híbridas que combinen características o criterios de ambos tipos.

Las rúbricas genéricas se pueden aplicar a una serie de tareas diferentes. Por ejemplo, están las que son utilizadas para evaluar una habilidad global de cualquier asignatura, como la oralidad o las relaciones interpersonales. Por lo anterior, una rúbrica genérica se podría aplicar a cualquier tarea evaluativa que esté diseñada para una misma habilidad o aprendizaje.

Las rúbricas híbridas –que combinan características de las genéricas y de tareas específicas– son muy útiles en la evaluación en aula

porque proporcionan retroalimentación a los alumnos tanto sobre las dimensiones amplias del aprendizaje como de su desempeño en las competencias y conocimientos específicos del aprendizaje en evaluación, alineadas ambas en un mismo instrumento.

Un ejemplo de rúbrica genérica para evaluar formativa y sumativamente el trabajo en equipo se presenta a continuación. Para desarrollarla se consideró el objetivo de aprendizaje transversal planteado en las bases curriculares de 7° Básico a II° Medio: “Trabajar en equipo de manera responsable, construyendo relaciones de cooperación basadas en la confianza mutua, y resolviendo adecuadamente los conflictos”.

La rúbrica que se presenta a continuación contiene cuatro criterios o dimensiones a evaluar (que son analíticas) y cuyo desempeño se ha graduado en cuatro niveles con sus respectivos descriptores.

RÚBRICA PARA EVALUAR Y RETROALIMENTAR EL TRABAJO EN EQUIPO

Criterio o dimensión	Categorías de desempeño			
	Excelente	**Satisfactorio**	**En proceso de logro**	**Insatisfactorio**
Organización del trabajo grupal	Los miembros del grupo organizan y priorizan tareas para el cumplimiento de la meta.	Los miembros del grupo organizan las tareas para el cumplimiento de la meta.	Los miembros del grupo dividen las tareas a realizar.	Los miembros del grupo no dividen tareas, todos hacen todo o piden ayuda al docente para que realice él la división de tareas.
	El grupo asigna roles para el logro de la meta, garantizando que la carga de trabajo sea definida equitativamente entre sus miembros.	El grupo asigna roles igualitariamente entre todos sus miembros, en función de la meta a lograr.	El grupo asigna roles a partir de las tareas que se dividieron sin considerar la meta a lograr.	El grupo asigna roles sin considerar las tareas definidas ni la meta a lograr.
	El grupo fija y compromete los plazos y fechas para la realización de cada una de las tareas, y las organiza en etapas y metas parciales.	El grupo precisa los plazos y fechas para la realización de cada una de las tareas en función de la meta a lograr.	El grupo fija los plazos y fechas para la realización de cada una de las tareas.	El grupo no define los plazos y fechas para la realización de cada una de las tareas, o bien, lo hace de modo incompleto, o las fechas o plazos no

				son viables.
Responsabilidad colectiva e individual (cumplimiento de lo propuesto y de los compromisos)	Responsabilidad colectiva: El equipo de trabajo desarrolla todas las tareas planificadas. El equipo cumple con los plazos definidos.	Responsabilidad colectiva: El equipo de trabajo desarrolla la mayoría (más del 75%) de las tareas planificadas. El equipo cumple con los plazos definidos.	Responsabilidad colectiva: El equipo de trabajo desarrolla menos del 75% de las tareas planificadas. El equipo cumple la mayoría de las veces con los plazos definidos.	Responsabilidad colectiva: El equipo de trabajo desarrolla la mitad o menos de las tareas planificadas.
	Responsabilidad individual: Cada estudiante asiste a las sesiones de trabajo definidas para el equipo (de 90 a 100%).	Responsabilidad individual: Cada estudiante asiste a la mayoría (más de 75%) de las sesiones.	Responsabilidad individual: Cada estudiante asiste a menos del 75% de las sesiones.	Responsabilidad individual: Cada estudiante asiste a la mitad de las sesiones.
	Cada estudiante se presenta puntualmente a todas las sesiones.			Cada estudiante no se presenta puntualmente a todas las sesiones.
	Cada estudiante entrega las tareas asignadas en forma completa.	Cada estudiante entrega la mayoría de las veces (75%) las tareas asignadas en forma completa.	Cada estudiante entrega pocas veces (menos de la mitad) las tareas asignadas en forma completa.	Cada estudiante entrega la mayoría de las tareas asignadas en forma incompleta o no las entrega.
	Cada estudiante entrega con anticipación las tareas asignadas.	Cada estudiante entrega a tiempo las tareas asignadas.	Cada estudiante entrega la mayoría de las veces (más de 75%) a tiempo las tareas asignadas.	Cada estudiante entrega fuera de plazo las tareas asignadas, o bien, depende de otros para desarrollar su trabajo.
Resolución de conflictos (cómo resuelven los problemas y dificultades durante el desarrollo del trabajo)	Los miembros del equipo reconocen un conflicto, dificultad o diferencia de opinión y cada uno busca y sugiere activamente soluciones.	Los miembros del equipo reconocen un conflicto, dificultad o diferencia de opinión y varios de ellos buscan y sugieren soluciones.	Los miembros del equipo no reconocen un conflicto, dificultad o diferencia de opinión y muy pocos buscan y sugieren soluciones.	Los miembros del equipo no reconocen un conflicto, dificultad o diferencia de opinión y no buscan ni sugieren soluciones para enfrentarlos.

	Los miembros del equipo toman la decisión que genere mayor adhesión y a la vez permite resolver el conflicto.	Los miembros del equipo toman la decisión que permite resolver el conflicto.	Los miembros del equipo resuelven parcialmente el conflicto utilizando estrategias que generan escasa o nula adhesión de sus integrantes.	Los miembros del equipo no resuelven el conflicto sino que recurren a otros externos al grupo para enfrentarlo.
Calidad de la interacción: relaciones interpersonales	El alumno siempre respeta los turnos de habla, se relaciona respetuosamente en las interacciones del grupo y es capaz de valorar los distintos puntos de vista aportados, reconoce errores y acepta sugerencias.	El alumno respeta los turnos de habla y mantiene una actitud respetuosa dentro del grupo la mayoría de las veces. Sin embargo, presenta dificultad para reconocer errores y dar sugerencias.	El alumno presenta dificultades para respetar los turnos de habla y mantener una actitud respetuosa frente a las interacciones del grupo. Asimismo, no logra valorar distintos puntos de vista, tampoco reconoce errores ni acepta sugerencias.	El alumno no respeta los turnos de habla, no se relaciona respetuosamente en las interacciones del grupo y no es capaz de valorar los distintos puntos de vista aportados, no reconoce errores ni acepta sugerencias.

Rúbricas holísticas

Las rúbricas holísticas son matrices en las que la descripción de las graduaciones incluye un grupo de criterios específicos que busca representar la globalidad de la acción (Mertler, 2001). Son adecuadas para la evaluación de tareas en que resulta difícil desagregarlas en partes para asignarles una puntuación, y en las que la calificación debería reflejar un logro general de los objetivos (Truemper, 2004). Las rúbricas holísticas usualmente son utilizadas cuando se pueden tolerar errores en alguna parte del proceso si estos no afectan la calidad general (Mertler, 2001). Un ejemplo de este tipo de rúbrica se presenta en la figura 9:

Analiza una afirmación argumentando desde la disciplina y su contexto histórico:	
Nivel	**Descripción**
4 Destacado	La tarea aborda la temática, la revisión bibliográfica es relevante y exhaustiva, la discusión incorpora elementos del contenido y del contexto histórico evidenciando una comprensión profunda de ellos; puede contener errores intrascendentes al tema.
3 Bien logrado	La tarea aborda la temática, la revisión bibliográfica es relevante pero incompleta, la discusión incorpora elementos del contenido y del contexto histórico; puede contener

	algunos errores menores.
2 Mínimamente logrado	La tarea aborda parcialmente la temática, la revisión bibliográfica es limitada pero pertinente, la discusión incorpora algunos elementos del contenido; puede contener algunos errores.
1 Deficiente	La tarea tiende a abordar la temática tangencialmente, la revisión bibliográfica es confusa y poco relevante, la discusión es general y carece de fundamentos; puede contener errores mayores.
0	La respuesta es vaga, la evidencia, inapropiada, y no se presenta argumentación.

FIGURA 9: Rúbrica holística para evaluar las habilidades de análisis y aplicación de un contenido disciplinar específico

La ventaja que tiene una rúbrica holística es que se enfoca en la calidad general del producto, la destreza o el entendimiento de un contenido o una habilidad específica (evaluación unidimensional), lo que redunda en una mayor rapidez del proceso de evaluación, ya que el profesor tiene que examinar el producto del alumno solo una vez para evaluarlo en sentido general (Mertler, 2001). Esto genera como consecuencia que este tipo de rúbricas sea usado principalmente en evaluaciones de naturaleza sumativa.

Su principal desventaja es que una descripción global del desempeño tiene menos precisión y puede dejar en una misma categoría a estudiantes que están en un desempeño casi completo de dicha categoría y a estudiantes que la alcanzan de manera incipiente. Además, los alumnos reciben una retroalimentación menos detallada y, por tanto, tienen menos orientaciones respecto de lo que deben abordar para mejorar su rendimiento en tareas similares.

Rúbricas analíticas

La rúbrica analítica consiste en una matriz de doble entrada en la que el producto o desempeño que se evaluará puede ser desagregado en partes y lo que interesa es levantar evidencias del rendimiento del estudiante en estas dimensiones específicas, para las cuales existe una graduación de descripciones propia. De esta forma, no hay una puntuación global, sino un conjunto de puntuaciones parciales que pueden ser analizadas de manera desagregada o sumarse y generar una puntuación total (Blanco, 2008; Truemper, 2004). Las rúbricas

analíticas se utilizan principalmente cuando se espera una evaluación multidimensional, lo que ocurre en tareas en las que las partes del proceso son relevantes y se espera un rango acotado de respuestas posibles por parte de los estudiantes (Mertler, 2001). La rúbrica de trabajo en equipo es un ejemplo de este tipo de rúbrica, otro ejemplo se presenta en la figura 10:

Criterio	0	1	2	3
Uso de conectores	No presenta	Algunos conectores no enlazan conceptos o enlazan conceptos no relacionados	La mayoría de los conectores enlaza y relaciona conceptos	Todos los conectores enlazan y relacionan conceptos
Relaciones entre conceptos	No presenta o son incorrectas	Solo presenta relaciones verticales	Se presentan relaciones horizontales	Se presentan relaciones cruzadas y horizontales
Otros criterios…				

FIGURA 10: Rúbrica analítica para evaluar aprendizajes asociados a la elaboración de un mapa conceptual

Entre las ventajas de este tipo de rúbrica se encuentran la calidad de la retroalimentación a los estudiantes, quienes pueden saber con mayor precisión en qué están teniendo dificultades y qué elementos ya tienen logrados en su trabajo, por lo que son un buen instrumento para realizar evaluaciones formativas o de proceso (Arter & McTighe, 2001). Además, es posible crear un “perfil” de las fortalezas y debilidades específicas de los alumnos y trabajarlas en clases.

Las principales desventajas que presenta este tipo de rúbrica tienen relación con el tiempo, ya que el proceso de evaluación puede ser substancialmente más lento tanto en la construcción como en su aplicación, porque requiere que el profesor examine el producto varias veces para evaluar diferentes habilidades o características individualmente.

Pasos para diseñar y elaborar una rúbrica

Este instrumento de evaluación requiere un tiempo de planificación y de elaboración que puede ser bastante largo, sin embargo, una vez construido, facilita notablemente el tiempo de corrección o de evaluación de los estudiantes. Los pasos que se describen a continuación son una compilación de la propuesta de Mertler (2001) y Mineduc (2006):

Paso 1: Definir de forma preliminar qué dimensiones del aprendizaje se evaluarán o valorarán.

Implica examinar nuevamente los objetivos de aprendizaje que se esperan lograr con la tarea. Esto le permite al profesor equiparar la estructura de puntajes con los objetivos y las instrucciones. Se recomienda que las dimensiones o aspectos a evaluar sean organizados en un orden que puede ir desde el más importante hasta el menos importante.

Paso 2: Identificar características observables específicas que usted quiere (al igual que las que no quiere) que los estudiantes demuestren en sus productos o procesos. Se tienen que especificar las características, habilidades o comportamientos que deben evidenciar los desempeños o productos de los estudiantes, al igual que los errores comunes que no debieran estar presentes.

Paso 3: Hacer una lluvia de ideas de características que describan cada atributo. Identificar las maneras de describir desempeños sobre la media, en la media y bajo la media, para cada característica observable identificada en el paso 2.

Paso 4a: Para rúbricas holísticas se sugiere escribir detalladas descripciones narrativas de un trabajo excelente y de un trabajo pobre incorporando cada característica o atributo en ellas. Se debe explicitar el nivel más alto y más bajo de desempeño combinando las descripciones de todas las características.

Paso 4b: Para rúbricas analíticas, escribir detalladas descripciones narrativas de un trabajo excelente y uno pobre para cada característica o atributo individual. Describir el nivel más alto y más bajo de desempeño usando los descriptores para cada característica o atributo por separado.

Paso 5a: Para rúbricas holísticas se debe completar la rúbrica con la descripción de otros niveles del continuo que va desde excelente hasta pobre para las características colectivas. Se sugiere escribir descripciones para todos los niveles intermedios de desempeño.

Paso 5b: Para rúbricas analíticas se debe completar la rúbrica describiendo otros niveles del continuo que va desde excelente hasta pobre para cada característica. Se recomienda escribir descripciones para todos los niveles intermedios de desempeño y para cada característica o atributo por separado.

Paso 6: Escoger ejemplos de trabajos de alumnos que dan cuenta de cada nivel. Esto ayuda al profesor a asignar los puntajes y en el futuro dichos trabajos pueden representar hitos o ejemplificaciones de los niveles.

Paso 7: Revisar la rúbrica cuanto sea necesario. El profesor debe estar preparado para reflexionar sobre la efectividad de la rúbrica y revisarla antes de implementarla nuevamente.

PROCESO DE APLICACIÓN DE LA RÚBRICA

A continuación se detallan los aspectos clave que deben considerarse antes de utilizar la rúbrica como instrumento de evaluación:

a. Defina y describa brevemente una tarea que desee evaluar con una rúbrica.
b. Señale los objetivos de aprendizaje involucrados en la tarea, y considere todos los aprendizajes que la tarea evaluará.
c. Utilizando los criterios de construcción ya descritos, construya una rúbrica alineada con los objetivos de aprendizaje.
d. Evalúe la rúbrica antes de utilizarla. Brookhart (2013) recomienda hacerlo considerando los siguientes criterios de calidad:
    - Al examinar las dimensiones se debieran evaluar los siguientes aspectos:
        - ¿Cada dimensión cubre los elementos importantes del desempeño del estudiante?

    - ¿La dimensión captura algún tema clave en su enseñanza?
    - ¿Las dimensiones están claras?
    - ¿Las dimensiones son distintivamente diferentes unas de otras?
    - ¿Las dimensiones representan las habilidades que el estudiante ya posee (como organización, análisis, usar convenciones)?
- Al examinar los descriptores o descripciones de cada dimensión se debieran evaluar los siguientes aspectos:
    - ¿Las descripciones igualan las dimensiones?
    - ¿Las descripciones son claras y diferentes entre ellas?
    - Si usa puntos, ¿hay un fundamento claro para asignarlos a cada dimensión?
    - Si usa una rúbrica con niveles de desempeño, ¿las descripciones son apropiadas e igualmente ponderadas a lo largo de esos niveles?
- Al examinar la escala completa de la rúbrica se debieran evaluar los siguientes aspectos:
    - ¿Los descriptores bajo cada nivel verdaderamente representan ese nivel de desempeño?
    - ¿Son las escalas (como ejemplar, competente, principiante) alentadoras y aun así informativas sin ser negativas ni desmotivadoras?

SUGERENCIAS PARA EL USO DE RÚBRICAS EN EL AULA

Rescatamos los planteamientos de Brookhart (2013), Andrade (2000) y Allen (2014), quienes señalan una serie de recomendaciones para implementar el uso de rúbricas en el aula, aumentando su potencial para colaborar con el aprendizaje de los estudiantes.

Calificar usando una rúbrica. Entréguela al inicio de la situación evaluativa para que los estudiantes conozcan las expectativas que usted tiene y cómo serán evaluados. Destine tiempo para discutir la rúbrica con sus estudiantes antes de que comiencen a desarrollar su trabajo.

Desarrolle una rúbrica con sus estudiantes para una tarea evaluativa como, por ejemplo, un proyecto de grupo. Los estudiantes pueden entonces monitorearse a sí mismos y a sus compañeros usando criterios acordados y que ellos ayudaron a desarrollar. Muchos profesores han constatado que los estudiantes elaboran criterios de evaluación y sus estándares son más elevados para ellos mismos de los que el propio docente definiría.

Solicite a los estudiantes que apliquen su rúbrica a algunos productos de muestra antes de crearlos. Los profesores reportan que los estudiantes son muy precisos al hacer esto, y este proceso tiende a ayudarlos a evaluar sus propios productos en la medida que los están desarrollando. Su habilidad para evaluar, editar y mejorar los borradores es muy importante.

Solicite a los estudiantes que intercambien borradores de su trabajo y que entreguen los comentarios a sus compañeros usando la rúbrica. Luego dé a los estudiantes unos días para que reelaboren sus borradores a partir de la retroalimentación, antes de entregar al profesor el producto final.

Solicite a los estudiantes que autoevalúen sus productos usando la rúbrica y que entreguen la autoevaluación con la tarea. Luego, el profesor evalúa el trabajo final y posteriormente pueden comparar las brechas entre las evaluaciones generadas por ambos.

EN SÍNTESIS, el tipo de instrumento que se utilice debe ser pertinente a los objetivos de aprendizaje que se pretenden evaluar. Como docentes debemos recordar que las técnicas e instrumentos de evaluación tienen limitaciones y ventajas, por lo que es importante

que tengamos conocimiento profundo de ellos para poder seleccionar o construir el que mejor se ajuste a nuestra necesidad.

A continuación profundizaremos en los organizadores gráficos como instrumento de evaluación para realizar evaluaciones formativas iniciales y continuas, integrando a los estudiantes como agentes evaluativos en el proceso de aprendizaje.

BIBLIOGRAFÍA

Allen, M. (2014). Using rubrics to grade, assess and improve student learning. *Strengthening Our Roots: Quality, Opportunity & Success Professional Development Day*. Miami-Dade College. Recuperado de http://www.mdc.edu/sailearn/documents/4.1%20Rubric%20Workshop%20Handout-Mary%20Allen.pdf

Allen, D. & Tanner, K. (2006). Rubrics: tools for making learning goals and evaluation criteria explicit for both teachers and learners. *CBE-Life Sciences Education*, *5*(3), 197-203.

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, *57*(5), 13-19.

Arter, J. & McTighe, J. (2001). *Scoring rubrics in the classroom: using performance criteria for assessing and improving student performance*. Thousand Oaks CA: Corwin Press.

Balch, D., Blanck, R. & Balch, D. H. (2016). Rubrics-sharing the rules of the game. *Journal of Instructional Research*, *5*, 19-49.

Blanco, Á. (2008). Las rúbricas: un instrumento útil para la evaluación de competencias. En L. Prieto (Coord.). *La enseñanza universitaria centrada en el aprendizaje: estrategias útiles para el profesorado* (pp. 171-188). Barcelona: Octaedro-ICE.

Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. Alexandria VA: ASCD.

Churches, A. (s. f.). *A guide to formative and summative assessment and rubric development*. Recuperado de https://edorigami.wikispaces.com/file/view/Assessment.pdf/555452441/Assessment.pdf

Corbetta, P. (2003). *Metodología y técnicas de la investigación social*. México DF: McGraw Hill.

Cunningham, G. (1998). *Assessment in the classroom: constructing and interpreting texts*. London: Falmer Press.

Eberly Center for Teaching Excellence and Educational Innovation (2015). *Creating exams*. Recuperado de https://www.cmu.edu/teaching/assessment/assesslearning/creatingexams.html

Himmel, E., Olivares, M. A. & Zabalza, J. (1999). Hacia una evaluación educativa. *Conceptos actuales sobre la evaluación de aprendizajes*, vol. I. Santiago: Pontificia Universidad Católica de Chile y Mineduc.

Livingston, S. (2009). Constructed response test questions: why we use them; how we score them. *R&D Connections*, *11*. Princeton NJ: Educational Testing Service.

Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, *7*(25), 1-10.

Mineduc (2006). *Evaluación para el aprendizaje: enfoque y materiales prácticos para lograr que sus estudiantes aprendan más y mejor*. Santiago de Chile: Mineduc.

Montgomery, K. (2000). Classroom rubrics: systematizing what teachers do naturally. *The Clearing House*, *73*, 324-328.

Popham, W. J. (2003). *Test better, teach better: the instructional role of assessment*. Alexandria VA: ASCD.

Spence, L. K. (2010). Discerning writing assessment: insights into an analytical rubric. *Language Arts*, *87*(5), 337-347.

Tankersley, K. (2007). *Tests that teach: using standardized tests to improve instruction*. Alexandria VA: ASCD.

Truemper, C. M. (2004). Using scoring rubrics to facilitate assessment and evaluation of graduate-level nursing students. *Journal of Nursing Education*, *43*(12), 562-564.

CAPÍTULO 8

La evaluación con organizadores gráficos: cómo y cuándo utilizarlos

Paola Marchant-Araya

INTRODUCCIÓN

Este capítulo tiene como propósito presentar los organizadores gráficos más utilizados por los profesores en el aula, y discutir sus usos tanto en la enseñanza como en la evaluación. En él se describen las características, ventajas y limitaciones de distintos tipos de organizadores y sus formas de representar el conocimiento. Asimismo, se analizan los aspectos centrales a considerar en el contexto de la evaluación con organizadores gráficos, dado que, por lo general, se desconoce qué evaluar, con qué intencionalidad, en qué contextos, quiénes participan en esta evaluación, así como las implicancias generales que tiene evaluar los aprendizajes de los estudiantes con organizadores gráficos. El capítulo se organiza en dos partes; la primera describe los aspectos conceptuales que permiten comprender desde dónde provienen los organizadores gráficos a nivel teórico y las características principales de cada uno de ellos. La segunda parte entrega distinciones respecto de cómo evaluar aprendizajes utilizando organizadores gráficos y plantea recomendaciones en el contexto del aula.

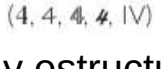
ANTECEDENTES TEÓRICOS

Cuando los niños y niñas empiezan a asistir a la escuela ya han adquirido ciertas reglas de lenguaje y algunos conceptos centrales para el aprendizaje escolar posterior. Así, el proceso de aprendizaje se va construyendo de manera que ellos hacen propios algunos métodos para ordenar acontecimientos y objetos de forma que les permiten ver nuevas regularidades y a su vez reconocer los términos que estas representan. Ese proceso continúa durante toda la vida.

Desde la perspectiva socioconstructivista, el aprendizaje es un proceso en el cual el estudiante elabora una representación interna del conocimiento, que está constantemente abierta al cambio en cuanto se construye a partir de las experiencias. Esta perspectiva plantea que los alumnos dan sentido al mundo cuando establecen

relaciones y coordinaciones entre lo que están aprendiendo y lo que ellos saben o han experimentado (Condemarín & Medina, 2000; Bronfenbrenner, 1987).

Lo significativo para los estudiantes, entonces, es lo que estos han vivido, combinando pensamientos, sentimientos y actividades. Debido a que la secuencia de experiencias de una persona es única, cada uno construye significados propios e idiosincrásicos[1], por lo que mientras más disímiles sean las experiencias de dos personas, más difícil será para estas compartir significados (Novak, 1998). Así, el aprendizaje significativo tiene, por definición, muchos factores personales involucrados; Ausubel, Novak y Hanesian (1983) distinguen tres tipos:

- El aprendizaje conceptual: corresponde a la percepción del aprendiz respecto de la regularidad en que se fundamenta el concepto (objetos, eventos, situaciones o propiedades), que tiene criterios comunes y que se designa mediante algún símbolo. Por ejemplo, el concepto de “número cuatro” irá aumentando en complejidad desde una única forma hasta múltiples formas de escribirlo  en la medida que el aprendiz amplíe su vocabulario y estructuras cognitivas.
- El aprendizaje representacional: el aprendiz reconoce una palabra, signo o símbolo como una etiqueta para un objeto específico, evento o categoría de eventos, por ejemplo, los sustantivos son aprendidos a través del aprendizaje representacional.
- El aprendizaje proposicional: este aprendizaje va más allá de la comprensión de las palabras aisladas o combinadas y requiere la capacidad del aprendiz para captar el significado de ideas expresadas en proposiciones. Las proposiciones son dos o más palabras combinadas para formar una afirmación acerca de un evento, objeto o idea, que pueden ser válidas, no válidas o sin sentido. La riqueza de significado que tenemos para un concepto crece exponencialmente con el número de proposiciones válidas que relacionan un concepto con otros.

Tanto el aprendizaje conceptual como el representacional son procesos que se dan en forma paralela, debido a que son idiosincrásicos y dependen del contexto. Así, los significados conceptuales se combinan para formar significado proposicional y ambos se pueden representar a través de su etiqueta o nombre.

En el aula es común la utilización de representaciones gráficas para identificar los conocimientos de los estudiantes. Estas representaciones comunican la estructura conceptual de un dominio disciplinar específico, el que contiene tanto las ideas fundamentales como las interrelaciones que el estudiante hace (Campos, 2005). Las representaciones gráficas pueden dar cuenta del razonamiento espacial de los estudiantes, que es la base del conocimiento abstracto y de la inferencia (Tversky, 2011). También permiten transmitir significados más directamente que un conjunto de palabras, y facilitan la retención, la memorización, el establecimiento de conexiones entre conceptos, la resolución de problemas y la comunicación de la información, haciéndola más comprensible para los estudiantes (Moreland, Dansereau & Chmielewski, 1997; Tversky, 2011; Uba, Oteiku & Abiodun-Eniayekan, 2016; Uba, Oteiku, Onwuka & Abiodun-Eniayekan, 2017).

Las representaciones gráficas que elaboran los alumnos constituyen una fuente de evidencias que podemos utilizar para monitorear el proceso de aprendizaje y para usarlo como recurso de enseñanza que favorezca distintos procesos cognitivos (Alshatti, Watters & Kidman, 2010; Campos, 2005; Preciado, s. f.; Tversky, 2011). Algunos usos en el aula se describen a continuación:

- Identificar ideas fundamentales y la complejidad de las relaciones entre conceptos que hacen los estudiantes.
- Conducir la comprensión de conceptos y relaciones.
- Facilitar la integración y la síntesis de información.
- Favorecer la representación de situaciones-problema para poder abordarlas y buscar una solución.

- Generar una organización de los contenidos de manera personal o colectiva.
- Favorecer la discusión grupal y la reflexión.
- Estimular la metacognición y habilidades cognitivas superiores en general.
- Producir imágenes mentales que los estudiantes puedan usar posteriormente.
- Integrar el conocimiento previo de los estudiantes con lo nuevo que se desea que aprendan.
- Motivar el desarrollo conceptual de los alumnos.

A continuación revisaremos distintos tipos de representaciones gráficas del conocimiento, relevando lo que cada uno de estos organizadores puede aportar en los procesos de aprendizaje de los estudiantes.

LOS ORGANIZADORES GRÁFICOS COMO TIPO GENÉRICO

Esta técnica de representación del conocimiento es la más antigua que encontramos en la literatura. Fue Barrón, en 1969, quien la introdujo a partir de sus estudios sobre las ideas propuestas por Ausubel sobre aprendizaje significativo y por Bruner acerca de las formas de representación del conocimiento. Ausubel (1978, 2002; Ausubel, Novak & Hanesian, 1983) desarrollaron la idea de “organizador de avanzada”, de carácter textual y en prosa, que servía de andamiaje entre el nuevo conocimiento y el adquirido con anterioridad por los estudiantes. Con este tipo de organizador se buscaba activar conocimientos previos para enfrentar una tarea de aprendizaje y, de esta manera, poder discriminar entre el nuevo conocimiento y las ideas o conceptos similares o contradictorios que existían en la estructura cognitiva del estudiante.

Barrón (1969) planteó el cambio de formato textual a gráfico y propuso una metodología para la enseñanza del vocabulario,

desarrollando así la idea de un organizador visual y textual o “visión estructurada del contenido” (p. 32).

En la actualidad, los organizadores gráficos han dado lugar a una serie de otras formas de representación, las cuales se definen como “representaciones visuales del conocimiento que presentan información, rescatando aspectos importantes de un concepto o materia dentro de un esquema usando etiquetas” (Preciado, s. f., p. 2).

Existe acuerdo en que los organizadores gráficos son herramientas que permiten a los estudiantes comprender, clasificar, organizar y conectar información, siendo tales procesos enseñados y evaluados en el aula (Alshatti et al., 2010; Bartels, 1995; Campos, 2005; Reyes, 2011; Uba et al., 2017; Villalustre-Martínez & Del Moral-Pérez, 2010). Para Campos (2005), “el organizador gráfico es una representación esquemática que presenta las relaciones jerárquicas y paralelas entre los conceptos amplios e inclusivos, y los detalles específicos” (p. 30) y se construye en el mismo nivel de lectura del nuevo material y no a nivel más alto y genérico de abstracción, constituyéndose en una herramienta de enseñanza para favorecer el aprendizaje significativo (Villalustre-Martínez & Del Moral-Pérez, 2010) (tabla 1).

TABLA 1: Clasificación de los organizadores gráficos según estructura y temporalidad en el proceso de enseñanza

Criterio	Tipo	Descripción
Temporalidad de uso en el proceso de enseñanza	Organizador gráfico previo	El estudiante lo elabora antes de presentar el nuevo conocimiento.
	Organizador gráfico posterior	El estudiante lo elabora después de recibir la nueva información.
Estructura	Organizador gráfico de líneas y palabras	Independientemente del tiempo, los estudiantes definen su organizador combinando líneas con palabras.
	Organizador gráfico de tipo pictórico	Independientemente del tiempo, los estudiantes definen su organizador a partir de dibujos o esquemas.
	Organizador gráfico mixto	Los estudiantes combinan líneas, palabras y dibujos.

Si bien estos organizadores gráficos presentan una serie de bondades que favorecen su utilización en el aula, también tienen algunas limitaciones:

Fortalezas	• Su estructura y visualización permiten la asimilación y retención de nuevo contenido facilitando la comprensión de los estudiantes. • Permiten esquematizar y organizar la información o contenido específico. • Presentan información concisa y sintética. • Permiten demostrar las relaciones entre diversos conceptos. • Permiten reconocer conocimiento previo y experiencias anteriores de los estudiantes. • Permiten integrar conocimientos previos con los que se están aprendiendo para construir nuevo conocimiento. • Dan evidencia de una visión integral del nuevo aprendizaje. • Pueden ser trabajados colaborativamente, favoreciendo la interacción social. • Permiten dar evidencia de manera holística de la representación del conocimiento de los estudiantes.
Limitaciones	• Pueden forzar a los estudiantes a seguir una estructura de manera lineal repitiendo modelos previos, limitando su creatividad, imaginación y pensamiento independiente. • Pueden disminuir la motivación intrínseca de los estudiantes para escribir en prosa si se utilizan de manera única o preferente. • Pueden usarse para obligar a los estudiantes a incorporar un modelo de pensamiento que no les es propio, por ejemplo, la forma de pensar del profesor, especialmente en la evaluación. • Los formatos jerárquicos[2], aunque simples, son rígidos y no permiten a los estudiantes incluir sus ideas o puntos de vista.

Adaptado de Daniels, Zemelman y Steineke (2007)

A partir de este tipo de representación del conocimiento fueron surgiendo otras alternativas en el contexto del aula escolar, algunas de las cuales se presentan a continuación.

MAPA MENTAL

Desde la Psicología, Buzan (2004) propone su idea de mapa mental a inicios de los ’70. Esta consiste en una expresión del pensamiento en la que “a partir de una imagen central se ramifican los principales elementos de un determinado tema mediante una estructura nodal conectada” (Villalustre-Martínez & Del Moral-Pérez, 2010, p. 18), es

decir, una función natural de la mente humana. Según Buzan y Buzan (1996; Buzan, 2004), los mapas mentales se caracterizan por:

- El tema se cristaliza en una imagen central.
- La imagen central irradia los principales temas o asuntos de forma ramificada.
- Las ramas comprenden una imagen o una palabra impresa sobre una línea asociada: los puntos de menor importancia también están representados como ramas más simples adheridas a las ramas de nivel superior.
- Las ramas forman una estructura nodal conectada.

En su estructura, los mapas mentales suelen combinar imágenes (ya elaboradas o construidas); figuras geométricas; líneas rectas, quebradas y curvas; palabras, códigos, números y colores. Todo lo anterior, diseñado por el estudiante, es decir, “hecho a mano” según la imaginación, creatividad y racionalidad de a quién representa. La figura 1 muestra un ejemplo cuyo contenido se refiere a un mapa mental. Como todos los organizadores, este también tiene fortalezas y limitaciones, las cuales se presentan a continuación:

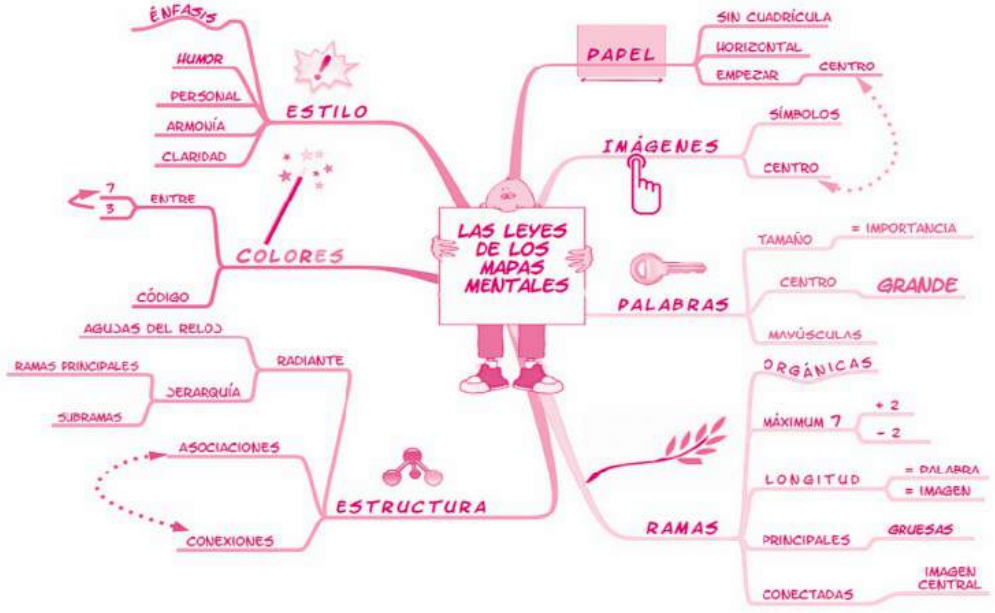


FIGURA 1: Ejemplo de mapa mental[3]

Fortalezas	• Permite organizar y aclarar lo que se sabe de un tema o concepto. • Permite ordenar el conocimiento previo sobre un determinado tema. • Permite reorganizar la estructura cognitiva. • Planear, comunicar y resolver problemas. • Estudiar de marea rápida y eficiente. • Mejora el recuerdo y la memoria y permite tener una visión global de la información. • Potencia la capacidad mental de las personas. • Útil para tomar notas personales en clase.
Limitaciones	• Puede ser difícil de usar por personas que son extremadamente lógicas, porque les cuesta confiar en su creatividad e intuición. • La creación de un mapa mental puede tomar mucho tiempo, el cual se considera un bien escaso. • Puede ser difícil de comprender para otros. Un mapa creado y personalizado supone un patrón de pensamiento propio, y al momento de considerarlo un instrumento de evaluación, hay que tener precaución con los criterios que se utilizan en su revisión. • Los patrones generales no necesariamente ayudan a la memorización. • Representa principalmente relaciones jerárquicas. • Puede llegar a ser demasiado complejo y perder de vista el panorama general.

Adaptado de Eppler (2006) y Agustina, Ngadiso y Rochsantiningsih (2013)

MAPA SEMÁNTICO

Fue propuesto por Pearson y Johnson a fines de los ’70 y difundido por Heimlich y Pittelman (1990). Su propósito es relacionar conceptos de una manera libre. Los mapas semánticos, también llamados “grafos léxicos”, son una estrategia que expresa en forma gráfica la “estructura categórica de una información o contenido a través de relaciones e ideas, conceptos o palabras fundamentales que integran un concepto mayor y que lo definen y explican” (Campos, 2005, p. 37).

El mapa semántico se estructura a partir de conceptos verbales y no verbales. Los verbales se presentan en figuras geométricas para simbolizar los nodos dentro de los cuales van las palabras e ideas, y las asociaciones entre conceptos mayores y menores se representan a través de líneas de interrelación (sólidas o flechas) que sirven para unir o relacionar los conceptos (Campos, 2005; Sinatra, Stahl-Gemake & Morgan, 1986).

Existen dos tipos de mapas semánticos (libres y fijos). Los mapas libres son desarrollados por cada estudiante sin una estructura predefinida y dependen de su creatividad. El formato suele usar la escritura de una idea o concepto clave al centro, dentro de alguna figura geométrica y, alrededor, se ubican los conceptos o características que se asocian con el concepto central. En la primera línea de desagregación, los conceptos se escriben con mayúscula, y los niveles siguientes en minúscula. Por su parte, los mapas fijos tienen una estructura predeterminada que da cuenta de un propósito en la organización de la información y es posible distinguir cuatro categorías (Sinatra et al., 1986) (figura 2):

1. Organización narrativa o secuencial: los estudiantes describen cada episodio en las figuras respectivas, siguiendo una estructura secuencial para evidenciar los eventos ocurridos.
2. Temático o descriptivo: se describen en detalle las personas, lugares o cosas alrededor de un tema central y las conexiones de sus componentes. Los tópicos o características más importantes se ubican directamente conectados al tema central.
3. Comparación o contraste: se presenta el discurso a partir de sus elementos diferenciadores y semejantes en cada lado del mapa en los nodos respectivos.
4. Clasificación: se muestran las relaciones entre conceptos, propiedades y atributos presentando conceptos macro y los que se desprenden de ellos.

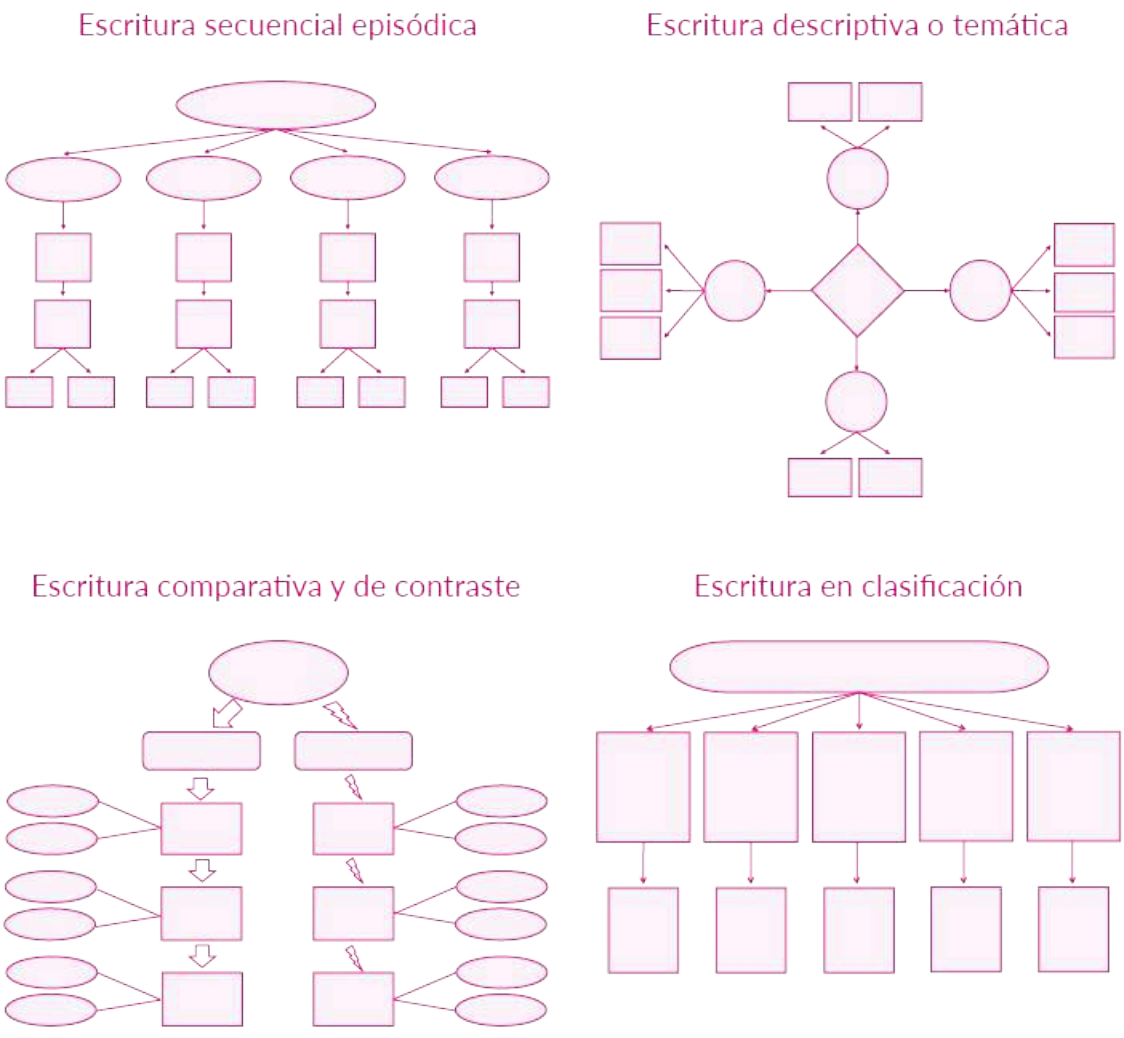


FIGURA 2: Ejemplos de estructuras de mapas semánticos fijos

A continuación se presentan las fortalezas y limitaciones de los mapas semánticos:

Fortalezas	• Mejoran el vocabulario de los estudiantes y su conocimiento sobre el significado de las palabras. • Favorecen la comprensión de nueva información. • Permiten organizar ideas antes de la redacción de un texto. • Permiten integrar diferentes partes de un contenido. • Sirven como material de estudio. • Son útiles para hacer visible la relación entre palabras.
	• No son fáciles de aplicar para quienes están aprendiendo a usarlos, requieren entrenamiento y hay que enseñar explícitamente cómo elaborarlos.

Limitaciones	• Puede ser difícil descubrir la relación entre ideas. • Las estructuras que van de arriba hacia abajo pueden dificultar la comprensión a quienes no tienen un pensamiento lineal.

Adaptado de Eppler (2006) y Agustina et al. (2013)

DIAGRAMA UVE DE GOWIN

Esta técnica heurística fue ideada por Gowin en 1988 y aparece como una alternativa adecuada para identificar los elementos conceptuales y metodológicos que se relacionan recíprocamente en la construcción del conocimiento, y es útil para resolver un problema o entender un procedimiento (Novak & Gowin, 2002; Campos, 2005; Guerra & Naranjo, 2016). El nombre “Uve” responde a su forma de representación en V, tal como se aprecia en la figura 3. En este diagrama es posible distinguir cuatro elementos centrales de su estructura:

1. Acontecimientos/objetos que son registrados y los conceptos que se asocian a estos hechos se ubican en el vértice de la Uve y son el punto de partida de la construcción del conocimiento.
2. Preguntas centrales o determinantes permiten iniciar la actividad entre los dos campos de la Uve; se ubican al centro.
3. Aspectos conceptuales que son conocidos, aplicables, útiles, necesarios para entender los acontecimientos y dar respuesta a las preguntas; se ubican en el lado izquierdo del diagrama.
4. Aspectos metodológicos o métodos utilizados que completan la descripción. Aquí se describen los juicios de valor comprometidos; se ubican en el lado derecho del diagrama.

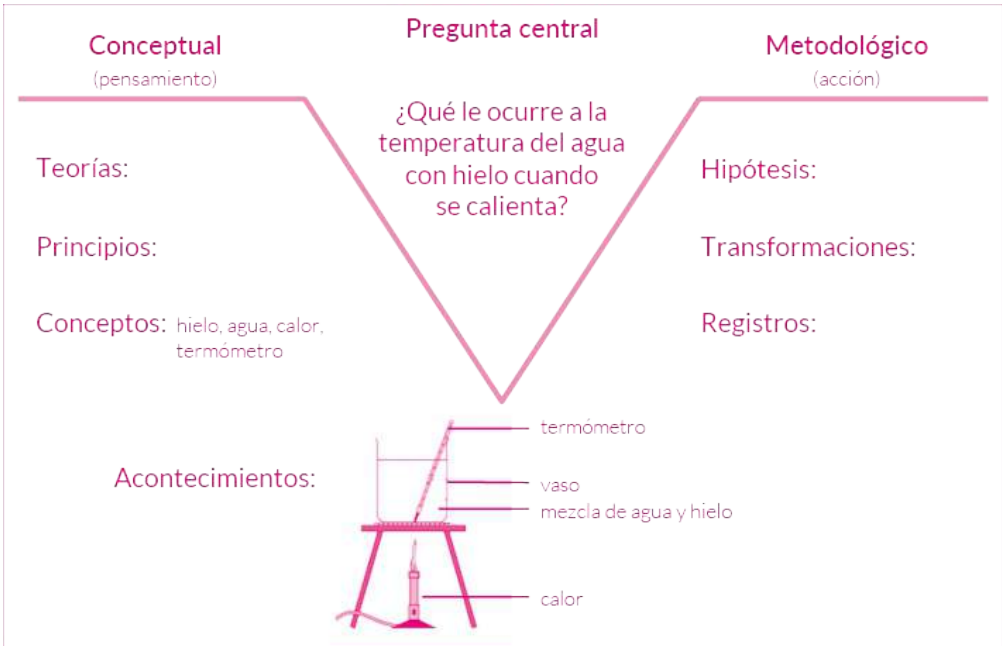


FIGURA 3: Ejemplo de un diagrama Uve de Gowin

La Uve de Gowin tiene como propósito generar aprendizajes significativos a partir de la relación entre lo que los estudiantes ya saben y los nuevos conocimientos que se están desarrollando, comprendiendo e integrando a la estructura cognitiva. En este sentido, la Uve permite a los estudiantes reconocer:

1. Qué acontecimientos o qué objetos se están observando.
2. Qué conceptos de los que ya conocen pueden relacionar con estos objetos y acontecimientos
3. Qué clase de registros vale la pena realizar.

En el sistema escolar, este tipo de representación del conocimiento es utilizado en distintas disciplinas, sin embargo, es en Ciencias donde se usa mayormente, estimulando a los estudiantes a plantearse una pregunta o preguntas que guíen el desarrollo de una indagación científica. En el contexto de la evaluación, Novak y Gowin (2002) observaron que estudiantes que habían construido un diagrama Uve antes de entrar al laboratorio de Ciencias tuvieron mejor aprovechamiento del tiempo en actividades relacionadas con la

tarea de laboratorio, que quienes no lo utilizaron. Las principales ventajas y limitaciones de este diagrama se presentan a continuación:

Fortalezas	• Facilita la construcción de nuevo conocimiento. • Permite la integración del conocimiento previo con los nuevos aprendizajes. • Favorece la comprensión en los procesos de resolución de problemas. • Permite interrelacionar aspectos conceptuales y metodológicos con las interrogantes que se puedan plantear sobre un tema. • El estudiante requiere poner en acción habilidades cognitivas complejas (interpretación, análisis, síntesis y evaluación del conocimiento por sobre la memorización).
Limitaciones	• Para evaluar con la Uve, los estudiantes deben estar entrenados, de lo contrario resulta ser una técnica muy compleja para ellos dada su costumbre a formatos tradicionales. • Se suelen confundir los conceptos con las teorías y los principios. • El requerir a los estudiantes poner en juego habilidades cognitivas complejas puede llevarlos a una sensación de dificultad inabordable y a que abandonen la tarea, por lo que se requiere monitoreo de este aspecto por parte del docente.

MAPA CONCEPTUAL O MAPA DE CONCEPTOS

Los orígenes de los mapas conceptuales se encuentra en los planteamientos de Novak y Gowin (2002), quienes, inspirados por Ausubel (1978) y la teoría del aprendizaje significativo, buscaban proponer aproximaciones que permitieran dar sentido al aprendizaje de los estudiantes, estimulando la reflexión sobre la estructura y el proceso de producción del conocimiento o del metaconocimiento. Es así como se concibe en los conceptos y el aprendizaje proposicional una base sobre la que se construyen significados propios e idiosincrásicos (Novak, 1998). Los mapas conceptuales aparecen, en este sentido, como una herramienta de mediación entre el que enseña, el que aprende y el que evalúa, de modo que los aprendizajes sean efectivamente significativos y que sea reconocida su importancia y aporte a la enseñanza (Alshatti et al., 2010; Campos, 2005; Eppler, 2006; Rovira, 2016). Los mapas conceptuales han sido definidos como:

- “Un recurso esquemático para representar un conjunto de significados conceptuales incluidos en una estructura de

proposiciones” (Novak & Gowin, 2002, p. 33).

- “Representaciones gráficas de conceptos de un dominio específico de conocimiento, construidas de tal forma que las interrelaciones entre los conceptos son evidentes” (Cañas et al., 1997, p. 2).
- “Diagramas jerárquicos que reflejan la organización conceptual de una disciplina o parte de ella (...) pueden ser entendidos como una estrategia para ayudar a los estudiantes a aprender y a los profesores a organizar el material de trabajo” (Campos, 2005, p. 23).

Escaño y Gil de la Serna (1999) agregan que esta representación es, además de gráfica, esquemática y fluida, y que promueve y traduce la organización de las ideas sobre un determinado contenido. “Promueve” porque su elaboración (incluso su lectura) requiere reparar en lo esencial, organizar las ideas y establecer relaciones entre ellas. “Traduce” porque refleja la manera como el conocimiento está almacenado en nuestra mente. Los mapas conceptuales son también efectivos en favorecer la retención y la memorización a través de las conexiones que se establecen, ya que permiten aplicar el conocimiento de manera organizada (Moreland et al., 1997).

Para Novak (1998), la elaboración de mapas conceptuales es un proceso personal basado en conocimientos y experiencias individuales o al menos con un significado propio. Por esta razón, el resultado final de los mapas de dos personas podría ser muy diferente. En ese mismo sentido, el mapa conceptual que alguien elabora no representa de manera idéntica (como si fuera una fotografía) la estructura de los aprendizajes y conocimientos en su mente (Novak & Gowin, 2002), por lo que su evaluación no puede ser con fines sumativos o de certificación y menos calificada. Su uso debe ser procesual, de modo que cada estudiante pueda dar evidencia variada y en distintos momentos de su aprendizaje.

La técnica para elaborar mapas conceptuales se basa en los conceptos o “términos”, los cuales deben ser extraídos de una fuente (texto, discurso del profesor, hecho, etc.) y organizados visualmente a

través de líneas que significan relaciones, indicando para cada una de ellas las palabras de enlace correspondientes al tipo de relación. De acuerdo con Novak y Gowin (2002), hay tres características que diferencian a los mapas conceptuales de otros recursos gráficos:

1. Jerarquización: los conceptos están dispuestos por orden de importancia o de "inclusividad". Las ideas más generales o inclusivas ocupan la parte superior de la estructura y las más específicas, junto con los ejemplos, la parte inferior. Los conceptos que tienen un mismo nivel jerárquico deben estar en la misma fila o altura.
2. Selección: son una síntesis o resumen que contiene lo más significativo de un tema. Se pueden elaborar submapas que vayan ampliando diferentes partes o subtemas del tema principal.
3. Impacto visual: "Un buen mapa conceptual es conciso y muestra las relaciones entre las ideas principales de un modo simple y vistoso, aprovechando la notable capacidad humana para la representación visual" (Novak, 1998, p. 106). Mientras más visual se haga el mapa, la cantidad de contenido que se logra memorizar aumenta y se acrecienta la duración de esa memorización, ya que se desarrolla la percepción, lo que beneficia, con la actividad de visualización, a estudiantes con problemas de atención.

Los elementos básicos de un mapa conceptual son:

- Los conceptos: regularidades en los acontecimientos o en los objetos que se designan mediante un término. Es decir, ideas o palabras que refieren a acontecimientos.
- Las proposiciones: constituyen la unidad semántica más pequeña que tiene valor de verdad. Estas se forman cuando se unen los conceptos con las palabras-enlace.
- Las palabras-enlace: son palabras que unen los conceptos y señalan los tipos de relación existentes entre ellos. También

son llamados “conectores”, por su función de conectar los conceptos.

En el mapa se organizan dichos elementos relacionándolos gráficamente y formando cadenas semánticas, es decir, con significado. En su forma más simple, un mapa conceptual está formado por dos conceptos conectados por una palabra-enlace (Novak & Gowin, 2002).

Los conceptos nunca se repiten, van dentro de óvalos o rectángulos y las palabras-enlace se ubican cerca de las líneas de relación. Es conveniente escribir los conceptos con letra mayúscula y las palabras-enlace en minúscula; pueden ser distintos a los utilizados en el texto siempre y cuando se mantenga el significado de la proposición. Para las palabras-enlace se pueden emplear verbos, preposiciones, conjunciones u otro tipo de nexo conceptual. Estas palabras son las que le dan sentido al mapa, incluso para personas que no conozcan mucho el tema. La figura 4 muestra la estructura de un mapa conceptual:

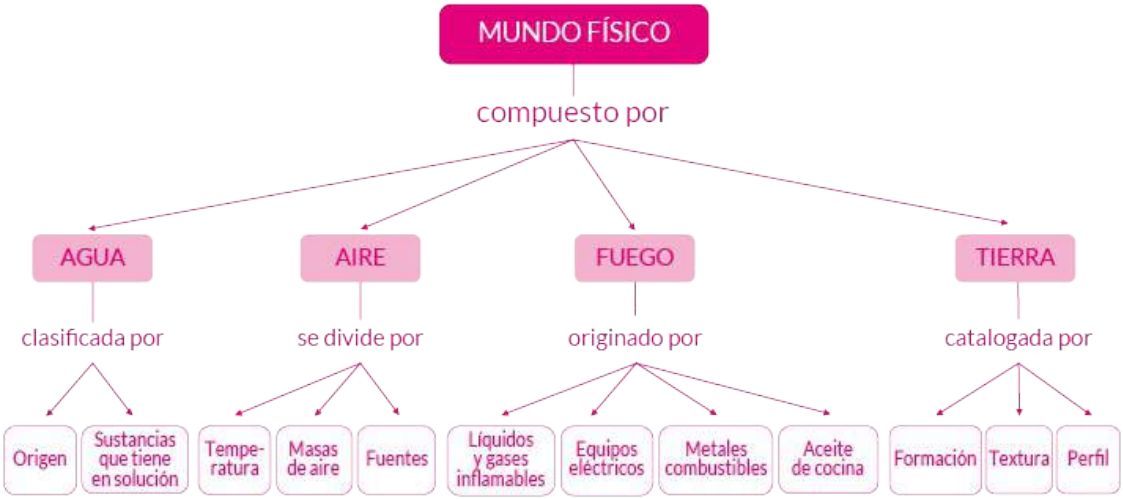


FIGURA 4: Ejemplo de mapa conceptual (adaptado de Guerra Naranjo, 2016)

Es fundamental considerar que en la construcción del mapa conceptual, lo importante son las relaciones que se establezcan entre

los conceptos a través de las palabras-enlace, que permiten configuran un significado verdadero (“valor de verdad”) sobre el tema estudiado, es decir, si estamos construyendo un mapa conceptual sobre “la célula”, la estructura y relaciones de este deben llevar a representar ese concepto y no otro.

Es importante resguardar que los alumnos hayan tenido experiencia en la construcción de mapas conceptuales antes de ser evaluados mediante esta técnica (Eppler, 2006) y que el texto o contenido utilizado para la construcción del mapa sea significativo para el estudiante. En Matemáticas, los mapas conceptuales son un aporte para dar cuenta de manera visual de las conexiones que se producen, y así evaluar cómo están entendiendo los estudiantes un contenido (Bartels, 1995). Las fortalezas y limitaciones de los mapas conceptuales se presentan a continuación:

Fortalezas	• Permiten planificar la enseñanza identificando qué conceptos son los relevantes a abordar y ayudan a los estudiantes a comprender la visión global de lo que están aprendiendo. • Constituyen una herramienta que sirve para ilustrar la estructura cognoscitiva o de significados que tienen los estudiantes. • Permiten trabajar y retroalimentar los errores conceptuales del estudiante, así como facilitar la conexión de la información con otros conceptos relevantes. • Permiten la diferenciación progresiva entre conceptos, sobre todo si se elaboran en diferentes momentos del desarrollo del tema. • Favorecen la integración o asimilación de nuevas relaciones entre conceptos. • Proporcionan una organización lógica y estructurada de los contenidos. • Favorecen la creatividad, la resolución de problemas y la autonomía. • Permiten lograr un aprendizaje interrelacionado al no aislar los conceptos, las ideas de los alumnos y la estructura de la asignatura. • Fomentan la negociación al compartir y discutir significados cuando se confeccionan en forma grupal. • Son un referente, un buen elemento gráfico cuando se desea recordar un concepto o un tema. • Permiten relacionar las partes (el todo) unas con otras.
	• La riqueza de los conceptos depende en parte del capital sociocultural del estudiante. • Para elaborar mapas conceptuales se requiere dominar la información y los conocimientos (conceptos) con los que se va a trabajar. • Es mayor el tiempo invertido en la evaluación de los mapas conceptuales, en parte porque se debe enseñar de forma explícita a hacerlos y porque no hay un referente común esperado (pauta única) cuando son revisados; hay que comprenderlos en forma individual.

Limitaciones	• Cuando son formatos no directivos (como entregarles una hoja en blanco), a los estudiantes les cuesta situarse sin una estructura dada. • La evaluación con mapas conceptuales es un proceso eminentemente subjetivo y arbitrario en la puntuación y no debería ser calificado. • Las estructuras que van de arriba hacia abajo pueden no ser adecuadas para representar o estructurar contenidos secuenciales tales como procesos o líneas de tiempo. • Tienen un nivel de dificultad medio a alto para los estudiantes, debido a la complejidad de la red de relaciones que implican. Por esto requieren mucho entrenamiento. • Tienden a ser idiosincrásicos, por lo que se dificulta su comprensión y evaluación en el aula. • Los patrones generales no necesariamente ayudan a la memorización.

LÍNEAS DE TIEMPO O TÉCNICA DE REPRESENTACIÓN Y DESARROLLO EN EL TIEMPO

Este tipo de representación visual es una técnica utilizada para mostrar secuencias cronológicas, eventos, hitos o pasos ordenados temporalmente. Fue creada con el propósito de facilitar la descripción y la comprensión, dado que se van mostrando hechos y momentos históricos tal como ocurrieron en el tiempo (Preciado, s. f.; Villalustre-Martínez & Del Moral-Pérez, 2010). Su estructura comprende una línea recta horizontal graduada en unidades de tiempo. El tiempo puede ser representado de manera clásica, con la fecha exacta, o también se pueden utilizar dibujos, figuras, fotos, íconos u otra forma de representación gráfica que su autor elija. Un ejemplo de línea de tiempo se presenta en la figura 5. Su estructura comprende:

- El tema representado, el que siempre corresponde a un título claramente definido.
- Una línea de tiempo recta, graduada en ambos sentidos.
- Eventos importantes asociados al tema que se han desarrollado a lo largo del tiempo.
- Fecha en que ocurrieron los eventos (puede ser exacta o aproximada).
- Intervalo de tiempo adecuado para comprender el tema.

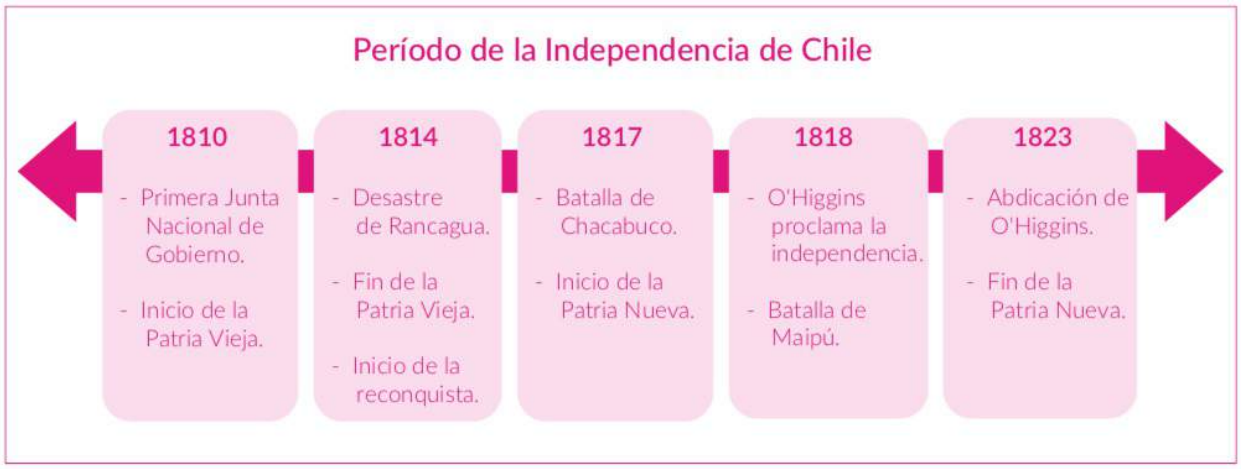


FIGURA 5: Ejemplo de línea de tiempo[4]

A continuación se presentan las fortalezas y limitaciones de las líneas de tiempo:

Fortalezas	• Permiten trabajar con convenciones temporales. • Permiten reforzar patrones de información. • Favorecen la retención y comprensión de información. • Favorecen la idea de simultaneidad en el aprendizaje de la Historia. • Permiten organizar recuerdos y conexiones entre ellos.
Limitaciones	• Los estudiantes deben conocer unidades de medida del tiempo, cómo se establecen las divisiones temporales, convenciones temporales (antiguo, nuevo, etc.) para que se pueda aprovechar la herramienta. • Se debe tener conocimiento de cómo visibilizar la duración de procesos y el exceso de acontecimientos. • De ser evaluado el diseño o la creatividad en la construcción, esto debe ser informado a los estudiantes previamente y entregarles los criterios respectivos.

DIAGRAMA DE VENN

Fue desarrollado en 1880 por Venn (Campos, 2005), es un esquema utilizado para representar relaciones entre conjuntos de información (intersección, inclusión y disyunción). Es muy utilizado en las asignaturas de Matemáticas, Ciencia y Lenguaje y permite a los estudiantes contrastar ideas representadas en distintos textos, buscar patrones para conectar lo que ya saben con la nueva información que

están aprendiendo y profundizar en su proceso de comprensión y análisis (Dreher & Gray, 2009). Por medio de la unión de círculos, permite al alumno establecer similitudes y diferencias entre conceptos, ideas o textos. En el sistema escolar, el diagrama de Venn es muy usado para representar grupos de ideas que comparten o no propiedades comunes. En la tabla 2 se analizan algunos ejemplos asociados a tareas específicas en las que se utiliza este organizador gráfico (figuras 6, 7 y 8):

TABLA 2: Ejemplos de uso del diagrama de Venn en distintas asignaturas

Tarea de clasificación de seres vivos:

Jorge y Anita tienen un conjunto de imágenes de animales que deben clasificar y separar en tres grupos usando los criterios que ellos definan. El resultado de su trabajo se presenta en la figura 6.

Como se puede apreciar, en la parte de los círculos que no tienen unión entre ellos se escribe aquello que hace único al conjunto; en este caso los niños definieron como conjuntos los animales que son aves, los que vuelan y los que nadan pero que no son aves, y en la zona en que se unen dos o más círculos se escribe aquello que comparten.

FIGURA 6: Diagrama de Venn sobre clasificación de los seres vivos

Tarea de resolución de problemas:

De un total de 99 personas, 5 hablan inglés y español, 7 español y alemán y 8 inglés y alemán. Si el número de personas que hablan alemán, español e inglés es 1, 2 y 3 veces mayor que el número de personas que hablan los 3 idiomas respectivamente, ¿cuántas personas hablan español?

Este tipo de problemas que demandan una comprensión lectora importante puede ser traducido a un diagrama que ayuda a entender de manera gráfica lo que se está planteando y qué es lo que se pide. La figura 7 muestra el desarrollo de la tarea. Según cómo los estudiantes pongan la información en el diagrama, se pueden detectar errores de ejecución de forma fácil y rápida para el docente.

FIGURA 7: Ejemplo de resolución de problema matemático

Planteamiento del problema:

SA + 8 + 7 + X = 2X

SE + 5 + 7 + X = 3X

SI + 5 + 8 + X = 4X

SA + SE + SI + 40 + 3X = 9X

SA: solo alemán

SE: solo español

SI: solo inglés

Tarea: Compare las características de un texto discursivo y argumentativo.

En general, para este tipo de pregunta se solicita una tabla, sin embargo, para un estudiante, el diagrama puede ser visualmente más significativo y para el docente, más rápido de revisar (figura 8).

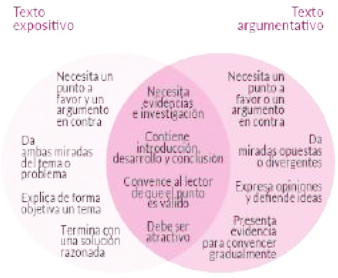


FIGURA 8: Ejemplo de diagrama para comparar textos discursivos y argumentativos

Para utilizar un diagrama de Venn como instrumento de evaluación es clave determinar cuáles son los conceptos que serán comparados o relacionados. Una opción de tarea es dar los conceptos a los estudiantes y que ellos sean quienes investiguen sobre el tema y ordenen la información en un diagrama; otra opción es que el docente entregue los conceptos y una lista de elementos sin separar y que los alumnos los clasifiquen en la parte de los círculos que corresponda.

Fortalezas	• Permite al estudiante sistematizar y sintetizar su conocimiento sobre un tema. • Es útil para comparar y contrastar dos o tres conjuntos de elementos. • Permite ilustrar situaciones de intersección, inclusión y disyunción de elementos. • Se pueden traducir textos a un formato de síntesis que facilita la comprensión, especialmente en Lenguaje y Matemáticas.
Limitaciones	• Con más de tres grupos de elementos es complicado manejar la información, especialmente en las intersecciones. • Su efectividad depende de la capacidad de los profesores para crear tareas desafiantes con este formato. • Su efectividad dependerá de cómo los docentes comprendan su uso y potencialidades para guiar el proceso de aprendizaje de los estudiantes. • No permite evidenciar relaciones entre conceptos.

ESTRUCTURA DE PROCESO O DIAGRAMA DE FLUJO

Los diagramas de flujo son representaciones gráficas de las operaciones y actividades que se desarrollan en secuencia de pasos en un procedimiento, siguiendo un camino lógico, y fueron utilizados

por primera vez por Gilberth en Estados Unidos (Sáenz & Martínez, ٢٠١٤). Para Preciado (s. f.), los esquemas son diagramas de flujo cuando “los símbolos utilizados se conectan en una secuencia de instrucciones o pasos indicada por medio de flechas”. Estudios confirman la efectividad de este tipo de organizador para desarrollar habilidades en la traducción de textos y representación esquemática, y permiten a los estudiantes mostrar evidencia de sus habilidades analíticas y de procesamiento (Alshatti et al., 2010). Es muy utilizado para representar procedimientos y hacer que los alumnos desarrollen pensamiento lógico a través de preguntas y respuestas, afirmaciones Sí o No u otros elementos que conectan las relaciones y jerarquías. En todo diagrama de flujo podemos encontrar los siguientes elementos (Sáenz & Martínez, 2014):

- Inicio de proceso.
- Especificación de las etapas del proceso.
- Acciones o actividades de cada etapa.
- Resultados.
- Fin del proceso.

Para cada una de estas actividades se utilizan símbolos específicos que denotan los elementos y acciones que se considerarán en el proceso. Los símbolos más característicos se muestran en la tabla 3:

TABLA 3: Símbolos utilizados en un diagrama de flujo

Característica	Símbolo
Este símbolo representa lo general, se utiliza para iniciar y finalizar un diagrama de flujo.	
Se utiliza para representar un resultado o insumo. También se usa para indicar una operación o acción realizada.	
Representa toma de decisiones y ramificaciones.	
Es usado para conectar o unir el flujo a otra parte del diagrama.	
Los símbolos se unen con líneas que en la punta tienen una flecha que indica la dirección en que va la información del proceso.	

La figura 9 muestra un ejemplo de diagrama de flujo. Si bien este organizador es muy útil para evaluar la comprensión de procedimientos o generar protocolos de trabajo con los estudiantes, como todo organizador gráfico, tiene sus fortalezas y limitaciones:

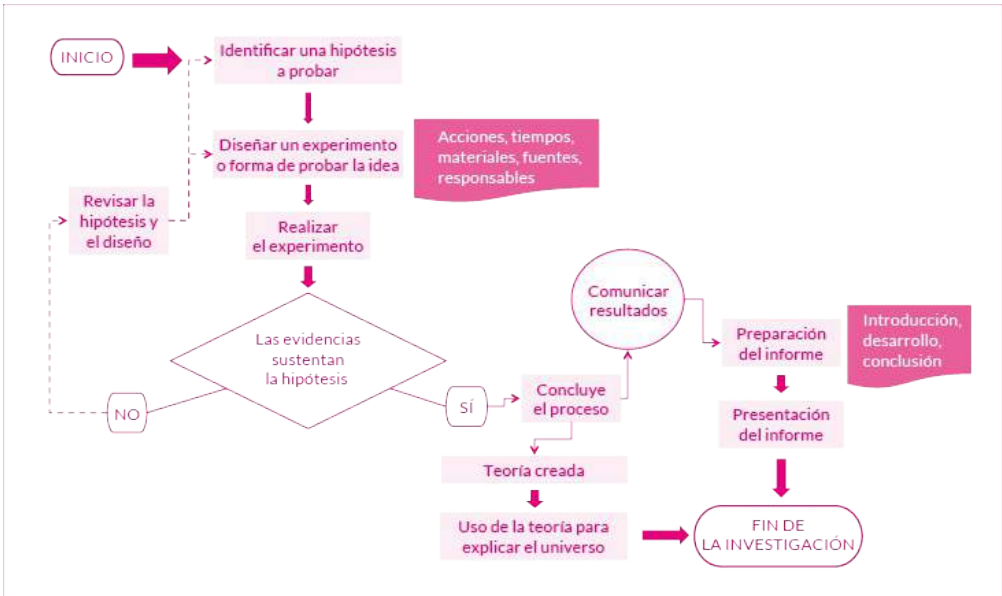


FIGURA 9: Diagrama de flujo[5]

Fortalezas	• Favorecen el razonamiento lógico y de proceso en los estudiantes. • Ayudan en la comprensión de un proceso, los componentes de este, las relaciones e interacciones. • Permiten diagnosticar fallas en los procesos o desarrollo de procedimientos. • Favorecen la comprensión de un fenómeno a partir de sus causas y efectos. • Favorecen la comprensión y la abstracción de la información. • Pueden ser trabajados en cualquier ámbito de la didáctica, en especial para explicar procesos. • Permiten identificar los problemas y las oportunidades de mejora de un proceso.
Limitaciones	• Su uso es masivo en Educación Superior, a nivel técnico profesional y universitario, pero su aplicación en el aula en contextos escolares debe ser mediada por el profesor en forma reiterada. • Utilizar muchos niveles entre los conceptos puede complejizar el flujo de la información y la comprensión de los procesos involucrados por parte de los estudiantes. • Pueden ser difíciles de seguir si existen diferentes caminos.

	• Ilustran el flujo de un proceso, pero no siempre se evidencia su estructura.

ESTRUCTURAS DEL TIPO CAUSA-EFECTO: LA ESPINA DE PESCADO

Llamado también “diagrama de Ishikawa” por el apellido de su creador, su nombre se debe a la forma visual que toma esta representación del conocimiento, organizando de ese modo las causas que puede tener un problema específico. Tiene como propósito identificar los factores, variables o aspectos que afectan un resultado para comprender un problema y luego proponer una solución. En un estudio con este tipo de organizador en la asignatura de Ciencias se evidenció su utilidad tanto en profesores como en estudiantes. Ambos mostraron mayor comprensión del contenido después de aprender a utilizar este diagrama y mayor capacidad de implementarlo en tareas escolares (Lu, Tsai & Hong, 2008). Los elementos clave de la estructura de este organizador gráfico son (Campos, 2005):

- Causas: son los factores que influyen, determinan o explican un resultado. Las causas tienen distintos niveles jerárquicos según su cercanía al centro. Las principales aparecen en recuadros y están unidas por líneas horizontales con flechas sólidas que se conectan con los efectos. Cada causa o factor puede ser explicado por otra causa de segundo orden, las que se unen a las primeras y así sucesivamente. Como se puede apreciar en la figura 12, las flechas varían según el nivel de desagregación en que se encuentren.
- Efectos: son consecuencias y resultados atribuidos a las causas y se describen en un recuadro con un término o unidad semántica mínima.
- Problema: en la cabeza se identifica el efecto principal o problema, mientras que en las espinas se ubican las posibles causas, tal como se aprecia en la figura 10.

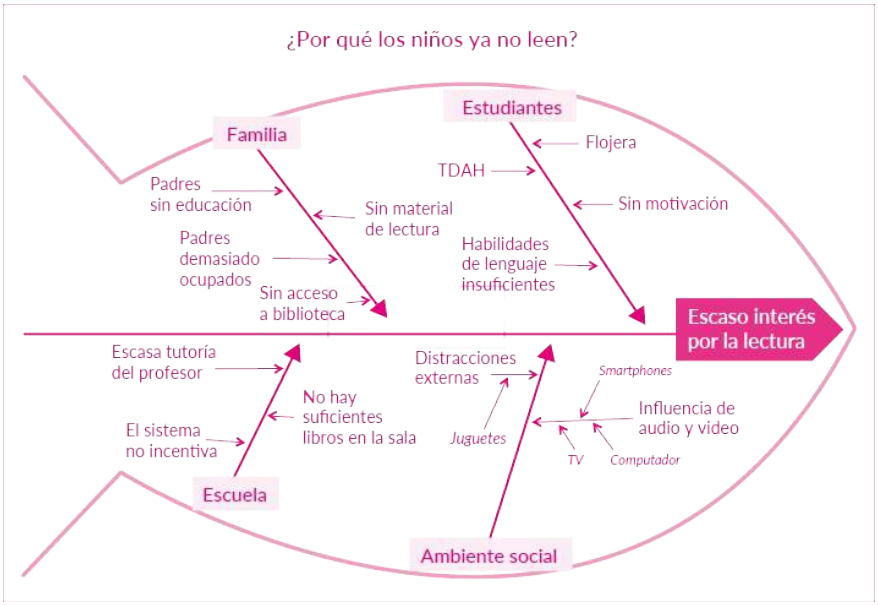


FIGURA 10: Ejemplo de la espina de pescado como representación gráfica

Fortalezas	• Incentiva discusiones en base a problemáticas concretas. • Permite identificar causas verdaderas que afectan los resultados. • Es una buena estrategia para la resolución general de problemas. • Favorece el pensamiento reflexivo en los estudiantes. • Favorece la organización del pensamiento para identificar causas y efectos principales y secundarios. • Desarrolla habilidades evaluativas, la síntesis de información y la toma de decisiones.
Limitaciones	• Considera solo relaciones lineales. • Supone un conocimiento previo y entrenamiento de los estudiantes para ser utilizada en la enseñanza y evaluación.

Los organizadores gráficos descritos hasta este momento son considerados los más comunes para usar en el ámbito escolar, no obstante, hay muchos otros que se pueden utilizar, algunos de los cuales se presentan en la figura 11:

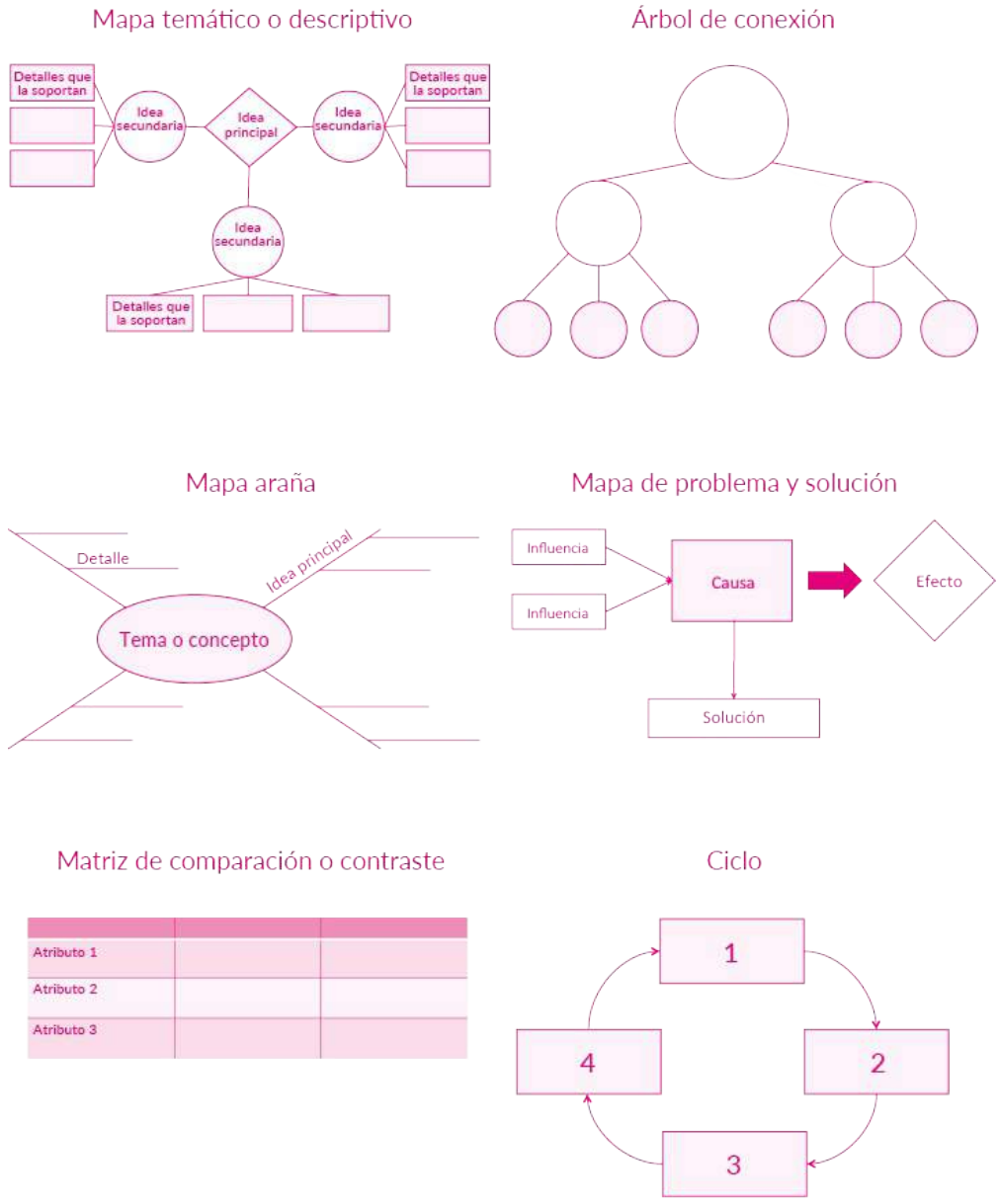


FIGURA 11: Ejemplos de otros tipos de organizadores gráficos[6]

Si bien es reconocido el aporte de los organizadores gráficos para el aprendizaje, se debe tener presente que su efectividad dependerá del nivel académico de los estudiantes, de la complejidad del organizador y del momento en que este se aplica (Alshatti et al., 2010). Una combinación de ellos en el aula puede ser muy útil y aunque la carga cognitiva sea alta para los estudiantes, constituyen herramientas de bajo costo para identificar conocimientos e ideas previas, monitorear

la comprensión de conceptos durante el desarrollo de una unidad y desarrollar habilidades transversales de aplicación, síntesis e integración de información en una o más asignaturas (Eppler, 2006).

CÓMO EVALUAR CON ORGANIZADORES GRÁFICOS Y OTRAS FORMAS DE REPRESENTACIÓN CONCEPTUAL

Como ya hemos visto en capítulos anteriores, la evaluación para el aprendizaje ha puesto el foco en la importancia de recoger evidencia que permita conocer el proceso y avance de los aprendizajes, con el propósito de retroalimentar y motivar a los estudiantes a alcanzar la meta final (Wiliam, 2011). Es así que se deben dar a los alumnos múltiples oportunidades para que puedan evidenciar su aprendizaje, reflexionar y aprender de estas evidencias. Los organizadores gráficos, como forma de representación del conocimiento, son un tipo de oportunidad de aprendizaje que puede cumplir el propósito de desafiar, motivar, entretener y favorecer aprendizajes nuevos en ellos en el mismo momento de su ejecución. Es por esto que se consideran una forma efectiva no solo de enseñanza sino también de evaluación (Preciado, s. f.).

Al elaborar una representación gráfica, los estudiantes ponen en acción la intuición y el análisis, conectándose con su producto desde un sentido de realidad, ya que las relaciones que se establecen en estos organizadores siguen una lógica significativa para ellos (Mayer, 1989). Dado lo anterior, es fundamental enseñar a los alumnos este tipo de representación y modelarla a través de ejemplos ya realizados, como también construirla en conjunto, de forma tal que podamos tener certeza de que la representación evidencia aprendizajes, y que la elaboración de la tarea evaluativa no es una limitación para el estudiante. Asimismo, es importante considerar que la representación del conocimiento es algo personal, propio de cada estudiante, por lo que buscar estandarizar el conocimiento a través de los organizadores gráficos u otras formas de representación es reducir las potencialidades de estos y de los propios estudiantes,

limitando sus capacidades para desarrollar habilidades cognitivas superiores. Esta es la principal razón por la que no se recomienda calificar sumativamente cualquier tipo de organizador gráfico.

¿QUÉ EVALÚO EN UN ORGANIZADOR GRÁFICO?

Si consideramos que la representación del conocimiento es personal e idiosincrásica, lo evaluable, dependiendo del tipo de organizador gráfico que se utilice, debería considerar aspectos como:

- Uso correcto de las palabras o conceptos.
- Asimilación de ideas claves.
- Pertinencia de las definiciones si existiesen.
- Claridad de la estructura en términos globales.
- Consistencia, pertinencia y veracidad de las relaciones.
- Organización, clasificación y síntesis de los conceptos.
- Integración de la nueva información.
- Habilidad para comunicar información (conceptos y relaciones).

Una experiencia de evaluación en la asignatura de inglés evidenció que a través de organizadores gráficos no solo se puede evaluar la organización y clasificación de información, sino también, cómo se conecta el conocimiento previo con conocimiento nuevo y la reconstrucción de información como base para la escritura de textos (Reyes, 2011). Otras investigaciones señalan que se puede evaluar conocimiento declarativo y procedimental así como habilidades cognitivas de análisis (Rice, Ryan & Samson, 1998).

¿CON QUÉ INTENCIONALIDAD SE EVALÚAN LOS ORGANIZADORES GRÁFICOS?

En la literatura especializada se encuentran diversas experiencias de evaluación con mapas conceptuales y otros organizadores gráficos, utilizando variadas intencionalidades evaluativas, en las que se reconocen elementos distintivos (Rovira, 2016):

- Intencionalidad formativa: se reconoce que la mejor y más recomendable opción de evaluación con organizadores gráficos es de carácter formativa (Villalustre-Martínez y Del Moral-Pérez, 2010), ya sea con propósitos formativos iniciales, de modo que el profesor pueda reconocer y explorar conocimientos previos, así como también, para monitorear y retroalimentar el proceso de enseñanza y aprendizaje, favoreciendo la autonomía y metacognición de los estudiantes (Preciado, s. f.).
- Intencionalidad sumativa: en el sistema escolar, los mapas conceptuales y otros organizadores gráficos suelen utilizarse con fines de certificación y, por ende, se califica a los estudiantes a partir de sus producciones. Sin embargo, dado lo descrito en este capítulo, eso no es lo recomendado. Utilizarlos con intencionalidad sumativa puede servir como cierre de un proceso de elaboraciones y revisiones previas, aunque lo aconsejable no es calificar lo que hacen los estudiantes, ya que la nota lleva a que estos busquen responder a las expectativas del profesor, más que a reflejar sus propias ideas, conceptos y estructura de pensamiento, lo que afecta la evidencia reportada por estos.

¿QUÉ AGENTES PARTICIPAN EN LA EVALUACIÓN CON ORGANIZADORES GRÁFICOS?

Es claro que en cualquier evaluación, además del profesor, se debe incluir a los estudiantes como agentes evaluadores activos, entendiendo que tienen roles y funciones distintas en el proceso de aprendizaje.

El profesor: tiene la responsabilidad de conducir y retroalimentar los aprendizajes de los estudiantes y debe recoger evidencia válida y confiable de su aprendizaje. Una acción concreta en la que se puede implementar los organizadores es, por ejemplo, que el docente proponga a los estudiantes elaborar un organizador sobre un tema específico que va a comenzar a enseñar, para identificar los conocimientos previos acerca de un contenido en alguna asignatura.

Después de trabajar los conceptos, el profesor les pide a los alumnos que analicen su diagrama y hagan los cambios que consideren pertinentes para representar su actual conocimiento del tema. Por último, el profesor evalúa la comprensión y retención de los estudiantes acerca del tópico trabajado.

El estudiante a sí mismo: este es el actor clave del proceso y, por tanto, tiene la responsabilidad y el derecho de monitorear y autorregular su aprendizaje. Por ejemplo, el profesor pide a los estudiantes que elijan un concepto clave en Ciencias y desarrollen un mapa para mostrar todos los conceptos asociados y sus relaciones. Luego, se discuten en clase estos conceptos y relaciones y se les entregan preguntas metacognitivas acerca de las decisiones tomadas, del tipo, ¿por qué decidiste utilizar este concepto clave?, ¿qué habilidades y conocimientos tenías que te permitieron construir tu mapa?, ¿harías cambios en tu mapa ahora que vimos con más detalle los conceptos? Los estudiantes escriben sus reflexiones y ajustan su mapa.

Los pares o grupos de estudiantes: el aprendizaje colaborativo es una oportunidad para generar nuevo conocimiento, que se ha descrito como ventaja en la mayoría de los organizadores gráficos, por lo que se debe hacer uso de la evaluación de pares para potenciar los aprendizajes. Por ejemplo, el profesor entrega a los estudiantes una lista de conceptos y se les pide que construyan un organizador gráfico. Posteriormente, el docente recoge los organizadores y los redistribuye entre los alumnos de la sala junto con una rúbrica o pauta de cotejo en la que se presentan las conexiones y relaciones correctas entre los conceptos. La idea es que cada estudiante tenga un trabajo de otro compañero. Cada uno analiza y evalúa con rúbrica el organizador respectivo, y luego en duplas los estudiantes conversan acerca de sus evaluaciones y dan retroalimentación al compañero autor del mapa, con el propósito de corregir en conjunto sus relaciones y que así mejoren la comprensión de las mismas.

Otro ejemplo puede ser, el profesor les pide a grupos de estudiantes que en forma colaborativa y libre elaboren mapas semánticos a partir de un texto leído. Luego cada grupo de estudiantes presenta al curso

su mapa y explica las relaciones con conceptos y relaciones establecidas. Los grupos de alumnos deben escribir un comentario evaluativo acerca de los trabajos de sus compañeros, con criterios dados por el docente previamente y entregarlo una vez terminada la presentación. Al finalizar, los grupos analizan los comentarios recibidos y deciden si hacer o no cambios a su representación, justificando las modificaciones.

Es muy relevante que una tarea en la que se utiliza cualquier organizador gráfico sea dialogada y retroalimentada, ya que, de lo contrario, será vista como una mera actividad sin conexión con el proceso de enseñanza y aprendizaje, desaprovechando las potencialidades de este.

¿NECESITO UN INSTRUMENTO ADICIONAL PARA EVALUAR ORGANIZADORES GRÁFICOS?

Si la respuesta es sí, ¿cuál sería el más adecuado?

1. Lista de cotejo: instrumento de carácter dicotómico, suele ser utilizado para poder registrar presencia o ausencia de las características descritas en los indicadores, que en este caso serían los elementos solicitados en un organizador gráfico. Con este tipo de instrumentos hay que resguardar que los indicadores no sean un conjunto de aspectos formales y que se pierda la evaluación del contenido de la asignatura.
2. Escala de valoración: si se desea retroalimentar a los estudiantes en cuanto al avance o calidad de su trabajo, una escala es un buen recurso. Se registra el nivel cualitativo (muy satisfactorio, satisfactorio, insatisfactorio o muy insatisfactorio) o cuantitativo (1, 2, 3, 4, 5), en que el organizador gráfico cumple con los aspectos solicitados. Igual que en el punto anterior, es necesario asegurar que los criterios sean relevantes y no solo aspectos formales o de estructura.
3. Rúbrica: este instrumento es el más utilizado en la literatura para evaluar los organizadores gráficos, y el que recomendamos,

porque presenta mayor detalle de los aspectos a evaluar, niveles de desempeño y, por tanto, mejor retraolimentación para los estudiantes. A continuación se muestra un ejemplo de rúbrica para evaluar mapas conceptuales en Matemáticas en la tabla 4.

TABLA 4: Rúbrica analítica para la evaluación con mapas conceptuales (Bartels, 1995)

Dimensiones	3 puntos	2 puntos	1 punto	0 punto	Observaciones
Conceptos y terminología	Muestra un entendimiento del concepto o principio matemático y usa notación y terminología adecuada.	Comete algunos errores en la terminología empleada y muestra algunos vacíos en el entendimiento del concepto o principio.	Comete muchos errores en la terminología y muestra vacíos conceptuales profundos.	No muestra ningún conocimiento en torno al concepto tratado.	
Conocimiento de las relaciones entre conceptos	Identifica todos los conceptos importantes y demuestra un conocimiento de las relaciones entre estos.	Identifica importantes conceptos pero realiza algunas conexiones erradas.	Realiza muchas conexiones erradas.	Falla al establecer cualquier concepto o conexión apropiada.	
Habilidad para comunicar conceptos a través del mapa conceptual	Construye un mapa conceptual apropiado y completo, incluyendo ejemplos, colocando los conceptos en jerarquías y conexiones adecuadas e indicando relaciones en todas las conexiones. El resultado final es un mapa fácil de interpretar.	Coloca la mayoría de los conceptos en una jerarquía adecuada estableciendo relaciones apropiadas la mayoría de las veces, dando como resultado un mapa fácil de interpretar.	Coloca solo unos pocos conceptos en una jerarquía apropiada e indica solamente unas pocas relaciones entre los conceptos, dando como resultado un mapa difícil de interpretar.	Produce un resultado final que no es un mapa conceptual.	

Bartels (1995) sugiere que para evaluar mapas conceptuales es recomendable aportar con comentarios del profesor en cada dimensión evaluada, en especial porque el mapa cuenta con varios

niveles jerárquicos y el estudiante necesita saber dónde podrían estar los errores en las relaciones o en la comprensión de los conceptos.

Siguiendo la propuesta de Novak y Gowin (2002), para la evaluación con mapas conceptuales se recomienda tener presente los siguientes aspectos cuando evalúe con cualquiera de los organizadores gráficos propuestos:

1. Definir una tarea

   Por medio de esta se invita a los estudiantes a representar la organización de su conocimiento en un tópico específico. La tarea tiene una demanda determinada, unas restricciones y una estructura de contenidos de lo que se espera que realicen los alumnos. En general, corresponde a la guía que acompaña la tarea o la instrucción dada para realizarla.

2. Un formato de respuesta

   En la tarea evaluativa se debe incluir el formato de respuesta, el cual dará las orientaciones al estudiante de cómo realizarla. Es común que el docente entregue un formato más directivo, que consiste en la presentación de un esquema, con nodos y líneas de enlace, el cual debe ser completado por el alumno, incluyendo en algunos casos una lista de palabras o conceptos que debe ordenar. Si queremos seguir la lógica original de los autores, se debería considerar un formato menos directivo que, a diferencia del anterior, no contiene ningún esquema ni se proporcionan los conceptos a incluir en el mapa, sino que se entrega una hoja en blanco en la cual los alumnos deben construir un mapa a partir de un contenido (Ruiz-Primo, 2000).

3. Un sistema de puntuación para ser evaluado

Dependiendo del tipo de organizador gráfico que se esté utilizando, se pueden considerar aspectos tales como:

a. Organización jerárquica de su estructura cognitiva: no existen organizadores correctos, sino jerarquías de relaciones realizadas por los individuos en cuanto a los significados que se establezcan

y que motivan estas relaciones. El profesor puede identificar si las estructuras son demasiado generales o demasiado concretas, cualquiera de estas alternativas indica error de comprensión o de integración en los conceptos subordinados (Novak, 1998).

b. Diferenciación progresiva: este elemento hace referencia a lo que Ausubel (1978) ha establecido como aprendizaje significativo, en que los nuevos conceptos alcanzan mayor signficado a medida que se adquieren nuevas relaciones o vínculos proposicionales. En este sentido, los conceptos nunca se aprenden totalmente sino que se están re-aprendiendo constantemente, modificando o haciéndose más explícitos.

c. Reconciliación integradora: este concepto corresponde a la mejora en el aprendizaje significativo de quien aprende, el cual reconoce nuevas relaciones entre conjuntos de conceptos o proposiciones. La nueva relación se capta cuando se produce una alteración circunstancial en el significado de un concepto del tipo *insight*. Los errores conceptuales son desplazados por un nuevo conocimiento.

Los mismos autores, partiendo de los conceptos anteriores como base fundamental, proponen los siguientes criterios de puntuación de los mapas conceptuales, los que aplican también para otros tipos de organizadores gráficos a nivel formativo:

a. Las proposiciones, es decir, los conceptos con las palabras-enlace apropiadas que nos indicarán las relaciones válidas o erróneas. Se asigna 1 punto por cada proposición válida y significativa.

b. La jerarquización, siempre en el sentido de que los conceptos más generales incluyan a los más específicos. Se asignan 5 puntos por cada nivel válido de la jerarquía.

c. Las relaciones cruzadas, que muestran conexiones entre conceptos pertenecientes a partes diferentes del mapa conceptual. Asignar 10 puntos por cada ligazón cruzada que sea tanto válida como significativa. Asignar 2 puntos por cada ligazón

cruzada que sea válida pero no ilustre una síntesis entre grupos relacionados de proposiciones. Estas conexiones cruzadas muestran creatividad por parte de los estudiantes.

d. Los ejemplos, en ciertos casos, para asegurarse de que los estudiantes han sabido comprender la expectativa de lo que es concepto y lo que no es. Asignar 1 punto a cada uno.

Ontoria (2000) sugiere que cada docente sea el que construya su propia escala numérica y sus propios criterios de puntuación, ya que las existentes poseen criterios complicados y se encuentran en una puntuación en base 100. Dicho autor propone que siendo los mapas conceptuales principalmente un instrumento de aprendizaje, se pueden negociar los criterios con los estudiantes, lo que será de gran utilidad para ellos.

VALIDEZ Y CONFIABILIDAD EN EL PROCESO DE EVALUACIÓN CON ORGANIZADORES GRÁFICOS

Evaluar con mapas conceptuales u otras maneras de representación del conocimiento no es tarea fácil y son múltiples las formas de uso de asignación de puntajes. En algunos estudios, el sistema de puntaje contempló únicamente contar el número de determinados componentes de un mapa. En otros, los evaluadores solo calificaron una parte de los mapas conceptuales. Por lo tanto, la confiabilidad de este tipo de instrumentos dependerá de factores variados (Klein, Chung, Osmundson & Herl, 2002; Rice et al., 1998; Ruiz-Primo, 2000; Ruiz-Primo, Shultz & Shavelson, 1997).

Los principales factores que pueden ser fuente de error en un proceso de evaluación con organizadores gráficos son, por ejemplo, variaciones en la competencia del estudiante en su representación del conocimiento con organizadores gráficos; variaciones en el dominio del contenido de quienes evalúan con esta técnica; la consistencia en la asignación de puntajes, lo que dependerá de la rúbrica u otro sistema que se utilice para evaluar. Asegurar la consistencia es uno de los aspectos más dificiles, ya que el profesor tiene la tarea de evaluar muchos mapas y corre el riesgo de que se descalibre su juicio

evaluativo, afectando la confiabilidad del proceso (McClure, Sonak & Suen, 1999).

Es frecuente encontrar en la literatura (Bartels, 1995; McClure et al., 1999; Ruiz-Primo et al., 1997) el uso de mapas maestros para evaluar los organizadores gráficos: esto tiene la ventaja de hacer más confiable el proceso, pero la desventaja de limitar las potencialidades de los estudiantes para desarrollar habilidades superiores, la creatividad y la intuición, como ya hemos señalado en este capítulo. Lo importante es que los profesores tengan presente esta información cuando tomen decisiones de evaluación en el aula.

En algunos casos, los docentes podrían utilizar mapas modelo que han sido validados por otros profesores de su misma disciplina o por colegas que tienen experiencia en la construcción de organizadores gráficos, asemejando el proceso de "validación de expertos". Si bien estos modelos son profusamente usados en la literatura, cabe preguntarse: ¿todos los expertos en un área comparten la misma estructura de conocimiento? Acton, Johnson y Goldsmith (1994) mostraron que las estructuras de los expertos son altamente variables. Y las diferencias en sus mapas se deben a que la estructura de conocimiento refleja no solamente el dominio (contenido) sino también un esquema personal de pensamiento y de cognición (estrategias para resolver problemas), por lo que pedirles a los estudiantes que generen un mapa único o semejante al mapa modelo pareciera no ser lo más adecuado.

En el caso de los mapas conceptuales, concordamos con Novak y Gowin (2002) en que ellos tienen validez de constructo en términos de la teoría de evaluación. Los autores plantean que "existe una correspondencia entre la evaluación del funcionamiento cognitivo y lo que predice nuestra teoría sobre lo que debería ser la organización cognitiva resultante de los aprendizajes significativos" (p. 132).

Un criterio importante para evaluar la validez de contenido es el juicio del experto respecto de la representatividad de los conceptos utilizados en la evaluación. Otro criterio es el juicio del evaluador relacionado con la precisión de los organizadores de los estudiantes dentro de un dominio. En el contexto del aula se recomienda que los

profesores en forma colectiva y colaborativa diseñen la tarea de enseñanza y evaluación, así como que también revisen y acuerden en conjunto qué y cómo retroalimentar a los estudiantes.

Finalmente, está demostrada la importancia de utilizar instrumentos de evaluación variados, que permitan recoger evidencia, explorando formas nuevas y más significativas de demostrar los aprendizajes para los estudiantes, como lo son los organizadores gráficos en general. Incluso, algunos estudios han mostrado relaciones consistentes entre los puntajes de los mapas conceptuales y el logro académico de los estudiantes (p. ej. Agustina et al., 2013; Ruiz-Primo et al., 1997), es decir, los alumnos que presentan mejor rendimiento académico en un área específica lograrían dar evidencia pertinente y suficiente de su comprensión acerca de lo que se está evaluando, por sobre quienes presentan un desempeño disminuido. Otros estudios refuerzan la relación positiva entre desempeños de los estudiantes y uso de organizadores gráficos en la enseñanza y aprendizaje en Ciencia, Matemáticas y especialmente Lenguaje (Agustina et al., 2013; Lu et al., 2008; Uba et al., 2017), lo que estimula su utilización en el aula escolar.

EN SÍNTESIS, para evaluar con organizadores gráficos, el profesor debe tener experiencia en su uso, especialmente si estos tendrán calificación, ya que hay una incidencia directa en las oportunidades de aprendizaje que tengan los estudiantes. En el caso de los mapas conceptuales, un formato directivo demora más tiempo en ser evaluado que un formato no directivo, en general los docentes se demoran más en revisar los mapas de los estudiantes cuando los comparan con un mapa modelo, que cuando lo hacen a partir de la comprensión directa del mapa elaborado por los alumnos, pero en este caso se corre el riesgo de no usar el mismo criterio de revisión para todos. Dado lo anterior, resulta relevante que la guía de instrucciones de la tarea con organizadores gráficos sea clara y pauteada, y se deje un formato libre para no restringir a los estudiantes y favorecer los aprendizajes.

En el siguiente capítulo profundizaremos en los criterios de validez y confiabilidad que deben ser resguardados en el diseño de una estrategia evaluativa y en la elaboración de instrumentos de evaluación para asegurar que la evidencia de aprendizaje que recogemos dé cuenta del progreso y logro de las metas que nos hemos propuesto que alcancen nuestros alumnos.

BIBLIOGRAFÍA

Acton, W. H., Johnson, P. J. & Goldsmith, T. E. (1994). Structural knowledge assessment: comparison of referent structures. *Journal of Educational Psychology*, *86*(2), 303-311.

Agustina, Y., Ngadiso & Rochsantiningsih, D. (2013). The effectiveness of semantic mapping to teach reading viewed from students' intelligence. *English Teaching*, *1*(1), 23-36.

Alshatti, S., Watters, J. & Kidman, G. (2010). Enhancing the teaching of family and consumer sciences: the role of graphic organizers. *Journal of Family and Consumer Sciences Education*, *28*(2), 14-35.

Ausubel, D. (1978). *Psicología educativa*. Ciudad de México: Trillas.

Ausubel, D. (2002). *Adquisición y retención del conocimiento. Una perspectiva cognitiva.* Barcelona: Paidós.

Ausubel, D., Novak, J. D. & Hanesian, H. (1983). *Psicología educativa: un punto de vista cognoscitivo*. Ciudad de México: Trillas.

Barrón, A. (1969). The use of vocabulary as an advance organizer. In H. Herber & P. Sanders (Eds.), *Research in reading in the content areas: first year report.* Siracuse NY: Syracuse University, Reading and Language Arts Center. Recuperado de http://files.eric.ed.gov/fulltext/ED037305.pdf#page=34

Bartels, B. (1995). Promoting mathematics connections with concept mapping. *Mathematics Teaching in the Middle School*, *1*(7), 542-549.

Bronfenbrenner, U. (1987). *Ecología del desarrollo humano.* Barcelona: Paidós.

Buzan, T. (2004). *Cómo crear mapas mentales: el instrumento clave para desarrollar tus capacidades mentales que cambiará tu vida*. España: Urano.

Buzan, T. & Buzan, B. (1996). *El libro de los mapas mentales: cómo utilizar al máximo las capacidades de la mente.* Barcelona: Urano.

Campos, A. (2005). *Mapas conceptuales, mapas mentales y otras formas de representación del conocimiento.* Colección Aula Abierta. Colombia: Magisterio.

Cañas, A., Ford, K., Coffey, J., Reichherzer, T., Suri, N., Carff, R., Shamma, D., Hill, G., Hollinger, M. & Mitrovich, T. (1997). *Colaboración en la construcción de conocimiento mediante mapas conceptuales.* Pensacola: Institute for Human and Machine Cognition, University of West Florida.

Condemarín, M. & Medina, A. (2000). *Evaluación auténtica de los aprendizajes: un medio para mejorar las competencias en lenguaje y comunicación*. Santiago de Chile: Editorial Andrés Bello.

Daniels, H., Zemelman, S. & Steineke, N. (2007). *Content-area writing: every teachers guide.* USA: Heinemann Educational Books.

Dreher, M. & Gray, J. (2009). Compare, contrast, comprehend: using compare-contrast text structures with ELLs in K-3 classrooms. *The Reading Teacher*, *63*(2), 132-141.

Eppler, M. (2006). A comparison between concept maps, mind maps, conceptual diagrams and visual metaphors as complementary tools for knowledge construction and sharing. *Information Visualization*, *5*, 202-210.

Escaño, J. & Gil de la Serna, M. (1999). Los mapas conceptuales, un recurso para ser feliz. *Revista Aula de Innovación Educativa*, *78*, 48-57.

Guerra, F. & Naranjo, M. (2016). Estado del arte en el tema de los organizadores gráficos en la representación de esquemas y diagramas. *UNIMAR*, *34*(2), 43-60.

Heimlich, J. & Pittelman, S. (1990). *Los mapas semánticos: estrategias de aplicación en el aula.* Madrid: Editorial Antonio Machado.

Klein, D., Chung, G., Osmundson, E. & Herl, H. (2002). *The validity of knowledge mapping as a measure of elementary students scientific understanding*. California: National Center for Research on Evaluation, Standards, and Student Testing (CRESST), University of California.

Lu, C., Tsai, C. & Hong, J. (2008). Use root cause analysis teaching strategy to train primary pre-service science teachers. *US-China Education Review*, *5*(12), 47-53.

Mayer, R. (1989). Models for understanding. *Review of Educational Research*, *1*(59), 43-64.

McClure, J., Sonak, B. & Suen, H. (1999). Concept map assessment of classroom learning: reliability, validity and logistical practicality. *Journal of Research in Science Teaching*, *36*(4), 475-492.

Moreland, J., Dansereau, D. & Chmielewski, T. (1997). Recall of descriptive information: the roles of presentation format, annotation strategy and individual differences. *Contemporary Educational Psychology*, *22*(4), 521-533.

Novak, J. D. (1998). *Learning, creating and using knowledge: concept maps as facilitative tools in school and corporations*. Mahwah NJ: Lawrence Erlbaum Associates.

Novak, J. D. & Gowin, D. B. (2002). *Aprendiendo a aprender*. Barcelona: Ediciones Martínez Roca.

Ontoria, A. (2000). *Mapas conceptuales, una técnica para aprender*. Madrid: Narcea.

Preciado, G. (s. f.). Recopilación: organizadores gráficos. *Orientación Educativa*. Recuperado de http://prepajocotepec.sems.udg.mx/sites/default/files/organizadores_graficos_preciado.pdf

Reyes, C. (2011). Connecting knowledge for text construction through the use of graphic organizers. *Colombian Applied Linguistics Journal*, *13*(1), 7-19.

Rice, D., Ryan, J. & Samson, S. (1998). Using concept maps to assess student learning in the science classroom: must different methods compete? *Journal of Research in Science Teaching*, *35*(10), 1103-1127.

Rovira, C. (2016). Theoretical foundation and literature review of the study of concept maps using eye tracking methodology. *El Profesional de la Información*, *25*(1), 59-73.

Ruiz-Primo, M. A. (2000). On the use of concept maps as an assessment tool in science: GAT we have learned so far. *Revista Electrónica de Investigación Educativa*, *2*(1), 29-52.

Ruiz-Primo, M. A., Shultz, S. E. & Shavelson, R. (1997). Concept map-based assessment in science: two exploratory studies. CSE Technical Report 436. California: National Center for Research on Evaluation, Standards, and Student Testing (CRESST), University of California.

Sáenz, B. & Martínez, L. (2014). Diagramas de flujo. En L. Martínez, P. Ceceñas & V. Ontiveros, *Lo que sé de: mapas mentales, mapas conceptuales, diagramas de flujo y esquemas* (pp. 116-131). Durango, México: Red Durango de Investigadores Educativos.

Sinatra, R., Stahl-Gemake, J. & Wyche, N. (1986). Using semantic mapping after reading to organize and write original discourse. *Journal of Reading*, *30*(1), 4-13.

Tversky, B. (2011). Visuospatial reasoning. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge Handbook of Thinking and Reasoning* (pp. 209-240). New York: Cambridge University Press.

Uba, E., Oteiku, E. & Abiodun-Eniayekan, E. (2016). Towards embedding graphics in the teaching of reading and literature in Nigeria. *International Journal of Education Investigations*, *3*(7), 106-119.

Uba, E., Oteiku, E., Onwuka, E. & Abiodun-Eniayekan, E. (2017). A research-based evidence of the effect of graphic organizers on the understanding of prose fiction in ESL classroom. *Sage Open*, 1-9.

Villalustre-Martínez, L. & Del Moral-Pérez, E. (2010). Mapas conceptuales, mapas mentales y líneas temporales: objetos “de” aprendizaje y “para” el aprendizaje en Ruralnet. *Revista Latinoamericana de Tecnología Educativa RELATEC*, *9*(1), 1527.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, *37*, 3-14.

[1] Los significados idiosincrásicos son aquellos aspectos significativos para los estudiantes desde lo particular o característico de ellos. En el caso de una comunidad escolar, corresponden a las características comunes y compartidas por esta.

[2] Un tipo de formato jerárquico son los “mentefacto conceptuales”, que llevan a los estudiantes a representar una estructura de proposiciones (cuatro operaciones intelectuales), que dan prioridad a las ideas fundamentales y se desechan las secundarias (Reyes, 2011).

[3] Diseño de Boukobza (2010). Extraído de www.ibermapping.es

[4] Extraído de https://www.portaleducativo.net/quinto-basico/507/Que-es-una-linea-de-tiempo-como-se-organizan

[5] Ejemplo extraído de http://bernalunea.blogspot.cl/

[6] Adaptado de http://www.cast.org/ncac/GraphicOrganizers3015.cfm

CAPÍTULO 9

Criterios de calidad para la construcción de instrumentos de evaluación*

Carla E. Förster
Cristian A. Rojas-Barahona

INTRODUCCIÓN

Como profesores, permanentemente debemos emitir juicios respecto del progreso y logro en los aprendizajes que están alcanzando nuestros estudiantes en las diferentes temáticas que enseñamos. Desde la mirada de la evaluación que se presenta en este capítulo, este juicio supone que con anterioridad recogimos información de manera sistemática y confiable, que la analizamos correctamente y que la contrastamos con un referente previamente establecido (como los objetivos de aprendizaje e indicadores de evaluación que definimos), por tanto, la conclusión a la que lleguemos y las decisiones posteriores que tomemos serán válidas y confiables. Sin embargo, en más de una ocasión, seguramente nos quedamos con la sensación de que los estudiantes sabían más de lo que reflejan los resultados de una prueba, o que estos resultados no son consistentes con lo que observamos en clases.

Hay que tener presente que las evidencias de aprendizaje que recogemos siempre están permeadas por las características psicológicas, socioculturales, económicas y físicas de cada estudiante, y pueden estar influidas por problemas pasajeros, circunstancias particulares del contexto o situaciones propias de la aplicación de los instrumentos con los que se recoge dicha información (Luckett & Sutherland, 2000; Salinas, 2002; Sanmartí, 2007). En este sentido, la calidad de una evaluación, es decir, la consistencia, adecuación y pertinencia de las situaciones, procedimientos e instrumentos que se utilizan para evaluar el desempeño de los estudiantes debiese considerar la validez, confiabilidad y objetividad como elementos esenciales al momento de interpretar los resultados y así tomar decisiones adecuadas respecto de qué y cómo enseñar y retroalimentar al estudiante (Brookhart, 2003; Hattie, 2015; Himmel, Olivares & Zabalza, 1999; McMillan, 2003).

Este capítulo tiene por objetivo analizar desde una perspectiva teórico-práctica los conceptos de validez y confiabilidad de una evaluación, y proponer sugerencias que permitan a los docentes resguardar estos criterios de calidad en futuras evaluaciones al interior de sus aulas (figura 1).

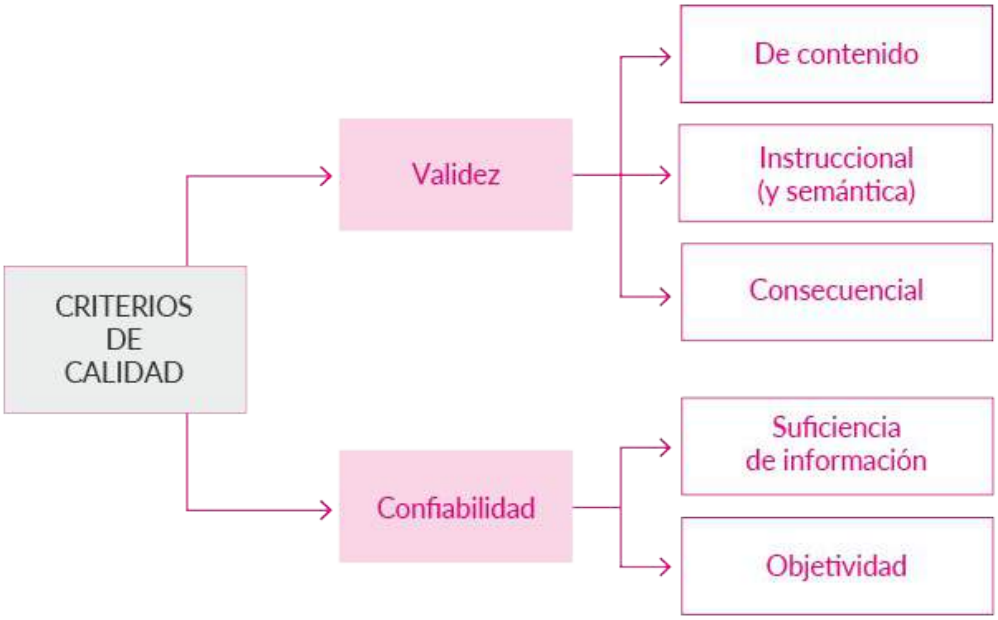


FIGURA 1: Organización de los criterios de rigor en la construcción de instrumentos y situaciones de evaluación

VALIDEZ

El primer elemento a considerar en la calidad de una evaluación es la validez de la información recogida con un instrumento o en una situación evaluativa. Este concepto tiene su origen en la medición y se ha desarrollado ampliamente desde el enfoque psicométrico, en el cual se plantea que una prueba es válida “si mide lo que pretendía medir”, no obstante, esta definición genérica debe complementarse con la finalidad para la que fue construida, es decir, si la interpretación de una determinada puntuación permite señalar correctamente una conclusión respecto del fin o constructo que mide dicha prueba (Hogan, 2004). Concordante con esta idea, Messick (1989) define la validez como “un juicio evaluativo integrado del grado

en el cual la evidencia empírica y teórica soporta la adecuación y pertinencia de las inferencias y acciones basadas en los puntajes de un test o en otros modos de evaluación" (p. 13). Así, por ejemplo, la pregunta ¿la Prueba de Selección Universitaria (PSU) es válida? no es una buena pregunta, más bien, correspondería decir: ¿la PSU es válida para predecir los resultados de los estudiantes durante su primer año de universidad? o mejor aún ¿en qué grado es válido el puntaje de la PSU para predecir el resultado de los estudiantes en su primer año de universidad? En otras palabras, la validez no es una propiedad intrínseca de un instrumento o situación evaluativa, sino una propiedad de la interpretación y los usos que se realiza de la información obtenida a partir del instrumento (Valverde, 2000).

El concepto de validez utilizado en el enfoque psicométrico tiene sentido solo para pruebas a gran escala, aplicadas a un número considerable de estudiantes, en que se espera que los resultados generen una distribución normal, es decir, que la mayoría de las personas obtenga puntuaciones medias y una proporción menor se ubique en los extremos. Pero en la sala de clases no se tiene como objetivo que los resultados de los alumnos se distribuyan de esta forma, por el contrario, se espera una curva asimétrica en que la mayoría (o todos) esté sobre el puntaje de aprobación de un determinado aprendizaje, ya que esto indicaría que lograron aprender lo que se esperaba para ese momento (edad, curso, período del año, etc.). Por esta razón, Stiggins (2001) evita hablar de "validez" en el contexto de la sala de clases y reemplaza este término por "evaluación en aula de alta calidad" (*high-quality classroom assessment*).

Diversos autores plantean que en la sala de clases el significado de validez no es el tradicional, ya que la evidencia estadística es muy difícil de obtener (Luckett & Sutherland, 2000; McMillan, 2003; Moss, 2003; Stiggins, 2001). Más aún cuando la mayoría de los profesores no tiene conocimientos en teoría de la medición, producto de la escasa presencia de esta disciplina en su formación inicial y, por tanto, muy pocas herramientas para hacer adaptaciones y reconocer las limitaciones de la "validez psicométrica" en la sala de clases

(Brookhart, 2003). Asimismo, Moss (2003) se pregunta hasta qué punto la utilización de la validez propia de la medición psicométrica asegura una orientación hacia las temáticas relevantes en el aula y cómo aportarían otras perspectivas teóricas a generar un marco sólido para obtener evaluaciones válidas en ese contexto.

Brookhart (2003), por medio de un cuadro comparativo, realiza una clara diferenciación de la validez de una evaluación según el contexto para el cual fue diseñada (tabla 1). Destaca principalmente en la validez el propósito y la importancia del contexto de aplicación de la evaluación. A gran escala, el propósito es sacar conclusiones del desempeño de un gran número de estudiantes para los fines específicos que tenga la evaluación diseñada (en general, toma de decisiones a nivel institucional o nacional), mientras que a nivel de aula, la evaluación tiene como fin monitorear y certificar el desempeño de los estudiantes entregando evidencias respecto de sus progresos en los objetivos de aprendizaje y de la efectividad de la enseñanza. En cuanto al contexto, las evaluaciones a gran escala no lo consideran como una variable relevante, mientras que a nivel aula sí es un factor a considerar, debido a que importan, por ejemplo, los conocimientos previos del alumno, su contexto sociocultural y las oportunidades de aprendizaje que los estudiantes han tenido antes de la evaluación.

TABLA 1: Comparación de la validez en evaluaciones a gran escala y en la sala de clases

Criterio	Evaluación a gran escala	Evaluación a nivel de aula
Interpretación de datos y acciones tomadas	Son externas al proceso de medición.	Son internas al proceso de medición.
Participación de los estudiantes	Los estudiantes son el “objeto” sobre el que se realizan las observaciones, no tienen injerencia en la evaluación.	Los estudiantes son observadores conjuntamente con los profesores; las mediciones corresponden a procesos formativos de evaluación.
Beneficios para los estudiantes	No reciben un beneficio directo de la evaluación, los análisis no son individuales.	Reciben beneficio directo, ya que la información de la evaluación es individual.

Propósito de la evaluación	El objetivo es realizar una inferencia estadísticamente significativa sobre el rendimiento de los estudiantes y/o la utilización eficaz de esa información para una finalidad específica.	El objetivo es comprender cómo es el desempeño de los estudiantes comparado con el "ideal" (definido en los objetivos de aprendizaje) y/o el uso efectivo de esa información para el aprendizaje de cada alumno.
Contexto	El contexto de la medición es un constructo irrelevante.	El contexto de la medición es un constructo relevante.
Contenido evaluado	Las especificaciones del contenido describen un dominio disciplinar.	Las especificaciones de contenido reflejan el dominio (objetivo de aprendizaje) y la enseñanza (modos, actividades).
Elaboración y aplicación del test	La elaboración y administración del test es estandarizada y se controla el efecto de las prácticas y diferencias culturales.	La elaboración y aplicación del test está mediada por las creencias de los profesores, sus prácticas y conocimiento del contenido y de los estudiantes (p. ej., diferencias culturales y lingüísticas).
Relación con la enseñanza	La evaluación es independiente de la enseñanza, se puede relacionar con diferentes contextos a través de cruces de datos.	La evaluación es parte de la enseñanza, una buena evaluación es un "episodio genuino de aprendizaje".

Adaptado de Brookhart (2003)

Aunque es claro el hecho de que la validez de una evaluación en la sala de clases no obedece a una lógica psicométrica (Brookhart, 2003), es necesario relevar algunos elementos de los distintos tipos, categorías o aspectos relacionados con la validez, que se han descrito en diversas clasificaciones; las más clásicas son la validez de constructo, de contenido y la relacionada al criterio (Allen & Yen, 2002; Gorin, 2007; Hogan, 2004; Moss, 2007; Sireci, 1998). La validez de constructo es el grado en que un instrumento permite medir una construcción teórica (p. ej., la inteligencia, la creatividad, la comprensión de lectura o la competencia matemática); la validez del contenido refiere al grado en el que una medición refleja realmente la variable para la que se ha diseñado y la validez del criterio corresponde a la predictibilidad que tiene una prueba para inferir el resultado de los estudiantes en una situación futura (Allen & Yen, 2002).

No existe acuerdo en la comunidad científica respecto de si hay una validez (la de constructo) que incluye a las otras o si son de

naturaleza diferente y excluyente y, por tanto, deben estar separadas (Borsboom, Mellenberg & Van Heerden, 2004; Gorin, 2007; Moss, 2007; Sireci, 1998). En este sentido, se rescatan las palabras de Valverde (2000), quien señala que “Existe una gran cantidad de opciones en cuanto al tipo de evidencia que se puede acumular y reportar. Cada tipo de evidencia ilumina y da apoyo a distintas facetas de la validez” (p. 26). Así, en este capítulo no se pretende profundizar en la discusión ni adherir a una clasificación en particular, sino considerar cuáles de estas categorías o aspectos sería relevante tener en cuenta para resguardar la calidad de la evaluación en el contexto educacional, específicamente en la sala de clases. A continuación se presenta una revisión conceptual de las tipologías fundamentales en la evaluación al interior de la sala de clases y se indican algunos ejemplos de situaciones de evaluación descritas por los estudiantes, en las cuales se evidencia la falta de cada tipo de validez.

VALIDEZ DE CONTENIDO

La validez de contenido es un concepto que ha resultado controversial a través del tiempo (ver resumen de la evolución del concepto en Sireci, 1998). Teóricos como Messick (1989) señalan que es parte de la validez de constructo y que dada su naturaleza “cualitativa”, es técnicamente incorrecto llamarla “validez”, ya que no corresponde a una nomenclatura psicométrica, y proponen hablar de relevancia, representación o cobertura de los contenidos. Gorin (2007), por su parte, plantea que un constructo se mantiene independiente del contexto de evaluación, mientras que un contenido es dependiente de dicho contexto. Hogan (2004) indica que para determinar la validez de constructo de un instrumento se requiere analizar su estructura interna[1] y los procesos de respuesta de cada ítem; como ambos procedimientos son muy difíciles de llevar a cabo en el aula y el contexto es un elemento importante que condiciona las situaciones de evaluación, hemos decidido presentar la validez de contenido como un elemento clave a considerar para resguardar la calidad de una evaluación.

Esta validez se refiere a la correspondencia que existe entre los contenidos, habilidades u objetivos de aprendizaje que evalúa el instrumento y el campo de conocimiento al cual se atribuye dicho contenido (Brualdi, 1999; García, 2002; Hogan, 2004; Lukas & Santiago, 2004). En el ámbito educacional se ha trabajado ampliamente en generar índices de validez de contenido en pruebas a gran escala, considerando variables como: a) la opinión de expertos (profesores, investigadores, curriculistas, directivos) en cuanto a la cobertura curricular de las pruebas y la enseñanza de dichos contenidos, b) la muestra de estudiantes a quienes se les aplicó la prueba, c) el número de ítemes de cada contenido, entre otras (Crocker, Llabre & Miller, 1988). Sin embargo, estos modelos estadísticos diseñados para calcular índices de validez de contenido no son aplicables a las situaciones evaluativas que se realizan al interior del aula, dado que no cumplen con los requisitos de base que sostienen dichos modelos.

En ese sentido, para un docente resguardar que sus evaluaciones tengan validez de contenido implica garantizar que las situaciones evaluativas incluyan los contenidos y habilidades correspondientes con el Objetivo de aprendizaje del cual pretende recoger evidencias. Se deben considerar los contenidos y habilidades más relevantes para dar cuenta del logro de un aprendizaje, es decir, seleccionar una muestra pertinente y representativa del universo de contenidos cubiertos en un curso, una clase o un conjunto de ellas. Sireci (1998), luego de una exhaustiva revisión bibliográfica, plantea que hay consenso en al menos cuatro elementos que caracterizan el concepto de validez de contenido: 1) la definición de un dominio disciplinar, 2) la relevancia de dicho dominio, 3) la representatividad del dominio y 4) los procedimientos de construcción del instrumento o situación evaluativa. Esto es particularmente importante porque las conclusiones que se pueden sacar a partir de una evaluación solo son válidas para lo que se evaluó, esto es, los puntajes o categorías que obtenga cada estudiante deben provenir de tareas y estímulos de calidad, coherentes con los propósitos conceptuales para los que fueron elaborados. Por tanto, certificar lo que un estudiante sabe y es

capaz de hacer o no en un período escolar determinado tiene directa relación con los contenidos y habilidades que fueron evaluados.

A continuación presentamos algunos ejemplos de experiencias evaluativas consideradas negativas por los estudiantes y que tienen relación con la falta de validez de contenido:

> “Durante un curso, el profesor dio bastantes lecturas para una prueba y en ella preguntó datos demasiado específicos, palabras exactas citadas en los textos para identificar de qué autor podían ser y nombres y fechas que, si bien eran interesantes o importantes, para una evaluación de tantos contenidos eran muy poco relevantes”.

> “Para mí fueron negativas todas aquellas evaluaciones en las que se preguntaban citas textuales de la bibliografía, que no están relacionadas con el marco teórico general del curso. Por ejemplo, ¿a qué se refiere el autor con la frase…?”.

> “Recuerdo una prueba que involucraba muchos contenidos, pero que nos habían anticipado no sería memorística, sin embargo, a la hora de responder, era evidente que las respuestas debían ser literales, pues las preguntas eran muy específicas y no apuntaban a lo que en el curso se había considerado como relevante”.

> “Me ha tocado responder evaluaciones en las que las preguntas no tienen nada que ver con los contenidos que entraban, parecieran evaluar cualquier cosa menos lo que debiesen”.

> “Recuerdo en Matemáticas cuando nos ponían unos ejercicios que se podían resolver solo si se te ocurría el truco adecuado. Creo que es negativo porque el aprender no depende de si te sabes o no el truco, sino de resolver y plantear bien un problema”.

> “En una clase nos pidieron hacer unos gráficos de la economía de la región y luego el profesor evaluó más la estética que si estaban o no bien elaborados, y a los compañeros que hicieron los gráficos con colores les fue mejor que a mi grupo, que teníamos los gráficos en blanco y negro… ahí aprendí a preocuparme más por la estética que por los contenidos”.

Como se puede apreciar en estos relatos, los estudiantes recienten que se les evalúe contenidos demasiado específicos, carentes de significado o que no tienen relación con el dominio disciplinar esperado.

A continuación se presentan algunas sugerencias para resguardar la validez de contenido de una situación evaluativa, que corresponden a procedimientos y acciones que un profesor podría llevar a cabo en la elaboración de sus propias evaluaciones. Si bien hay sugerencias en el ámbito estadístico que también se pueden realizar, no han sido consideradas aquí:

1. Elaborar una tabla de especificaciones del instrumento: como vimos en capítulos anteriores, la tabla de especificaciones de la elaboración de una prueba, una lista de cotejo o una escala de valoración permite visualizar si se están cubriendo todos los contenidos o elementos que se espera abordar, si las habilidades asociadas a dichos aspectos son coherentes con las metas de aprendizaje trazadas y si los ítemes y descriptores corresponden a los contenidos y habilidades que se esperaban. A partir de este análisis también es posible establecer si existe una graduación de la complejidad en cuanto a las habilidades evaluadas y si los ítemes y descriptores son los más apropiados para medir una habilidad determinada.
2. Verificar a través del criterio de jueces: dado que en el aula somos nosotros quienes construimos o adaptamos los instrumentos y situaciones evaluativas, muchas veces se nos pasan por alto aspectos que para quien no ha visto antes dicha evaluación son más evidentes. Por esta razón, se recomienda mostrar a otros profesores el instrumento con su pauta de corrección y la tabla de especificaciones para que nos den su opinión respecto de si la situación de evaluación es adecuada para medir los aprendizajes e indicadores definidos. Puede ser que tengamos problemas en la redacción de un ítem o de un

descriptor y que los aprendizajes que queremos evaluar no se reflejen en nuestro instrumento.

VALIDEZ INSTRUCCIONAL

Según Hogan (2004), esta validez corresponde a una aplicación particular de la validez de contenido y es conocida también como "validez curricular". Dice relación con lo que los estudiantes han tenido oportunidad de aprender durante las clases para responder correctamente en una evaluación (Crocker et al., 1988). Cabe señalar que el nombre "instruccional" (*instruction* en inglés) hace referencia a la enseñanza y no, como se suele interpretar, a las instrucciones de una tarea evaluativa.

En el ámbito educativo, este tipo de validez es clave, dado que representa el grado de relación entre lo que se enseña y lo que se evalúa. Cuando esta relación es débil se presentan dos problemas, por una parte, los estudiantes no tienen posibilidad de demostrar lo que aprendieron durante las clases y, por otra, son evaluados en aspectos que no se les enseñaron (Himmel et al., 1999; McMillan, 2003). Esta última idea se ve reflejada especialmente cuando se cambia el énfasis de lo que se trabajó, por ejemplo, en clases se enseñan los conceptos, sus definiciones y luego en la evaluación se les pide a los alumnos que apliquen dichos conceptos en situaciones que nunca han sido trabajadas durante la clase aludiendo a que se espera que sean capaces de hacerlo como parte de la "construcción" de su propio aprendizaje. Así, diremos que una evaluación tiene validez instruccional cuando contiene situaciones evaluativas coherentes con las actividades de aprendizaje realizadas por los estudiantes.

Cuando este criterio no se cumple se evidencian también problemas éticos relacionados con la interpretación que se hace de los resultados, como responsabilizar a los estudiantes por no saber contenidos que no han tenido la oportunidad de aprender, y responsabilizar a los docentes por el bajo logro de sus estudiantes sin proporcionarles las condiciones, materiales o la capacitación necesaria para enseñar los contenidos (Valverde, 2000). Si bien este

problema se presenta principalmente en pruebas a gran escala, también se da en las interpretaciones que hacen los equipos directivos sobre los resultados de una prueba de nivel o una prueba externa aplicada a varios cursos.

En este apartado se incluye lo que se ha llamado “validez semántica”, que consiste en que las situaciones de evaluación contienen términos cuyo significado es conocido y compartido entre el constructor del instrumento (en el caso del aula, el profesor) y los estudiantes. Muchas veces se enseña un contenido usando un término y luego en la evaluación se pregunta utilizando un término “sinónimo” que no ha sido trabajado en clases. En estricto rigor el contenido fue enseñado, pero frente a una respuesta errónea de los estudiantes, cabe preguntarse si no dominan el contenido o si fue la palabra que no comprendieron lo que provocó dicha respuesta.

A continuación presentamos algunas situaciones de evaluación percibidas como negativas por los estudiantes y que aluden a la validez instruccional.

> “Muchas veces los profesores evalúan materia o contenidos que no correspondían a lo que habíamos visto en las clases. En las pruebas usaba palabras raras que no habíamos visto ni estaban en el libro y cuando alguien preguntaba qué significaba, te humillaba diciendo que uno debería saberlo porque era cultura general”.

> “Fue negativo para mí cuando las clases teóricas no iban acordes a las evaluaciones de laboratorio y nos evaluaban materia que aún no habíamos visto, hay muchas actividades de laboratorio que aún no logro comprender del todo”.

> “Las pruebas en un curso no correspondían a lo que el profesor enseñaba en clases, enseñaba mal, creaba un clima de terror en la sala y luego nos preguntaba cosas que no había enseñado o que había pasado en el curso del lado”.

> “… la primera prueba que nos hicieron no correspondía con lo que trabajábamos y aprendíamos en clases. Solo hacíamos teoría y la prueba fue de aplicación, con muchos detalles, y el profesor no hacía ejercicios ni destacaba esos elementos. En suma, lo que aprendí no me

sirvió para la prueba”.

“He tenido profesores que evalúan en sus pruebas cosas (contenidos) que han enseñado mal, a la rápida o sencillamente no han tratado ni en clases ni en los textos referidos, asumiendo que los vimos en otros cursos”.

Como se puede observar, las opiniones de los estudiantes aluden directamente al desempeño profesional del docente, por tanto, las acciones que se sugieren a continuación para resguardar la validez instruccional tienen relación con las prácticas de enseñanza que se utilizan habitualmente:

1. Velar por la coherencia entre lo enseñado y lo evaluado: esto supone resguardar que las situaciones de evaluación contengan los contenidos vistos en las actividades de aprendizaje realizadas. Es clave llevar un registro de lo que se enseñó en cada clase. Para elaborar las situaciones de evaluación no basta con utilizar la calendarización o planificación, a menos que esta sea actualizada de manera continua. Normalmente las dinámicas propias de un grupo de estudiantes generan diferencias entre los cursos de un mismo nivel, con algunos se puede profundizar o volver atrás en una temática y con otros no; también pueden haber actividades no planificadas que producen desfases temporales y si no tenemos claros estos eventos, podríamos evaluar contenidos que aún no han sido enseñados o que abordamos en un curso, pero en otro aún no.
2. Velar porque las situaciones de evaluación sean equivalentes a las actividades de aprendizaje realizadas: esto no significa repetir la guía que se realizó en clases y ahora ponerle una nota, por el contrario, esta sugerencia apunta a generar evaluaciones que consideren habilidades similares a las ya trabajadas. Si durante el proceso de enseñanza se intencionaron actividades para lograr los aprendizajes esperados o metas de aprendizaje determinadas para esos estudiantes, luego, en la evaluación, se busca monitorear o certificar si se lograron o no para poder

retroalimentar al estudiante e iniciar un nuevo ciclo. Se sugiere, por una parte, hacer uso de la planificación de clases y que esta sea una guía flexible de apoyo a la preparación de la enseñanza y, por otra, utilizar la tabla de especificaciones poniendo énfasis en que las situaciones de evaluación sean coherentes con lo que cada profesor trabajó con sus estudiantes. En este último punto, los "jueces expertos" no tienen competencia, ya que no estuvieron durante el proceso, por lo que la coherencia entre lo enseñado y lo evaluado solo depende de la autoevaluación que realice el profesor respecto de su propia práctica.

3. Cuidar que el lenguaje utilizado en las situaciones de evaluación sea conocido por los estudiantes: esto implica tener especial cuidado en utilizar términos familiares para los estudiantes, tanto en las instrucciones como en el lenguaje técnico propio de la disciplina que se desea evaluar. Es importante hacer notar que el propósito de una evaluación es obtener información del aprendizaje de los estudiantes y las "palabras trampa" solo generan información de mala calidad respecto de ese fin.

VALIDEZ CONSECUENCIAL

Este tipo de validez se relaciona con las consecuencias y secuelas intencionales y no intencionales que tendrá el uso e interpretaciones que se dará a la información obtenida en la evaluación (Hogan, 2004; McMillan, 2003; Moss, 1997). Aunque no hay acuerdo entre los autores respecto de si esta validez es pertinente al ámbito psicométrico (Borsboom et al., 2004; Hogan, 2004; McMillan, 2003), en opinión de Moss (2003) y McMillan (2003), debería ser la principal consideración al momento de tomar decisiones respecto de una evaluación en la sala de clases.

Los aspectos consecuenciales de la validez son especialmente importantes en una evaluación cuando las interpretaciones de la información pueden implicar consecuencias adversas para los participantes (Brualdi, 1999), por ende, reconocer esta validez podría ayudar a controlar esos aspectos. Como vimos en la tabla 1, Brookhart (2003) plantea que esta validez se asocia a los beneficios

que tiene la evaluación para los estudiantes. En las evaluaciones a gran escala, las consecuencias se enmarcan en la toma de decisiones y reformulación de políticas públicas (Schutz & Moss, 2004) y tienen implicancias éticas y sociales (Borsboom et al., 2004) que hacen relevante mirar la validez desde este punto de vista, pero no inciden en los estudiantes individualmente. En el aula, en cambio, las consecuencias tienen relación directa con las dinámicas de enseñanza/aprendizaje que se dan entre profesores y alumnos, ya que hay beneficios o consecuencias adversas individuales. En ese sentido, Brookhart (2003) señala que la integración de la enseñanza y la evaluación debe ser tomada seriamente, y entender que una situación evaluativa es una oportunidad más para lograr aprendizajes.

Por tanto, se podría decir que la validez consecuencial a nivel de aula alude a los efectos de la evaluación sobre la enseñanza y los aprendizajes de los estudiantes (Himmel et al., 1999), por consecuencia, se relaciona con los propósitos para los cuales se diseñó la evaluación. McMillan (2003) plantea una serie de preguntas que, de ser respondidas positivamente, indicarían que una evaluación tiene validez consecuencial:

- ¿Los estudiantes adquieren una comprensión profunda de lo que están aprendiendo mientras se preparan para la evaluación?
- ¿Los estudiantes creen que son capaces de aprender nuevos conocimientos después de autoevaluarse en ejercicios de práctica de una evaluación?
- ¿Los estudiantes son capaces de transferir sus conocimientos en nuevas situaciones?
- ¿La decisión de realizar revisiones adicionales formativas previas a la entrega de una tarea evaluativa lleva a un mayor aprendizaje del estudiante?

Estas preguntas apuntan a propósitos diferentes y no siempre se darán todas al mismo tiempo, lo importante es tener en cuenta que la evaluación en la sala de clases siempre tiene consecuencias y, por

tanto, los docentes deben pensar en ellas antes, durante y después de realizar una evaluación para sacar el mayor provecho de esta instancia y minimizar los efectos negativos que se puedan generar. Entre las consecuencias imprevistas de una evaluación, y que le restan validez, destacan la falta de motivación de los estudiantes para participar en una actividad de evaluación, generar un ambiente de competencia entre ellos, resentimiento de los alumnos por el formato de la evaluación y que interioricen un error conceptual por causa de la tarea realizada (Brookhart, 2003; McMillan, 2003).

Algunas situaciones percibidas como negativas por los estudiantes, en que se evidencia la validez consecuencial, se presentan a continuación:

> “En un trabajo en grupo, en el cual se evaluó a todos por igual, uno de mis compañeros no se preparó bien, se puso nervioso y la exposición no fue la mejor, pero los otros tres compañeros hicieron su presentación de buena forma. Por las razones anteriores, la nota del grupo fue un 5 y con esto se premió al compañero deficiente y se perjudicó a los otros que sí lo hicimos bien”.

> “En Ciencias Naturales teníamos que hacer una investigación sobre el Fenómeno del Niño, mi papá es biólogo marino y vio que el libro estaba malo, le dijimos a la profesora y yo en mi investigación lo expliqué correctamente, pero me saqué un 4,3 porque no lo puse como decía el libro. Lo que más me molesta es que nos obligaron a aprendernos algo que estaba malo porque la profesora no quiso asumir su error”.

> “En Matemáticas nos hicieron responder una guía de ejercicios para la casa que llevaría una nota en el curso, yo la respondí un poco a la rápida y tuve un 70% de logro. Después al colegio se le ocurrió que los que habían tenido más de 80% estarían en un grupo avanzado de preparación para la PSU, lo que me pareció súper injusto porque nunca dijeron que la usarían para eso, de saberlo, la habría respondido más a conciencia, yo necesitaba ese curso intensivo para entrar a Ingeniería y tuve que pagar uno aparte”.

> “Recuerdo pruebas demasiado extensas en que las preguntas eran de memorización de contenidos y no había oportunidad de pensar y analizar nada. Me aprendía las cosas y me fue bien, pero no me acuerdo de nada,

es como si no hubiese hecho el curso y ahora lo voy a necesitar para mi futuro laboral".

"En Historia las pruebas eran todas de selección múltiple y tenían unos párrafos gigantes para leer, daba lo mismo estudiar o no porque lo que se pedía era comprensión de lectura, yo soy lenta para leer así que como siempre me iba mal, dejé de estudiar para ese curso".

Puesto que las consecuencias de una evaluación pueden ser negativas y de largo plazo para nuestros estudiantes, las sugerencias que se dan para resguardar la validez consecuencial se basan en su mayoría en acciones relacionadas con la definición y planificación de una evaluación:

1. Definir claramente los propósitos y usos de la evaluación: como esta validez se basa en las implicancias que tiene un proceso evaluativo para los estudiantes, es importante definir previamente la finalidad que tiene realizarla, es decir, ¿para qué evaluamos? En este sentido se presentan tres grandes fines: 1) diagnóstico formativo (p. ej., identificar los conocimientos previos de los estudiantes, reconocer conductas de entrada, ajustar la planificación al contexto), 2) formativo de proceso (p. ej., mejorar los aprendizajes y la motivación de los estudiantes, monitorear su progreso en los aprendizajes definidos, mejorar el proceso de enseñanza) y 3) sumativo (p. ej., certificar los aprendizajes de los estudiantes, informar a los padres el desempeño de sus hijos a fin de año). Es importante tener presente que un mismo instrumento puede ser válido consecuencialmente para una instancia y no tener ninguna validez en otra.
2. Identificar claramente las evidencias que darán cuenta de los propósitos de la evaluación: como ya se ha señalado anteriormente, la validez consecuencial está determinada por la correspondencia entre la información que se recoge y los fines para los que se utilizará. En este sentido, la sugerencia es tener una buena planificación que permita resguardar dicha

correspondencia, lo que implica no hacer una evaluación y luego definir para qué la puedo usar. Un ejemplo de esto son las “pruebas sin aviso” que se realizan en una clase, cuya motivación obedece más a un control disciplinario o castigo que a la evaluación de los aprendizajes esperados. Estas evaluaciones, por lo general, son improvisadas y se agregan como una calificación, cuyas consecuencias son, en su mayoría, negativas para los estudiantes.

CONFIABILIDAD

Otro elemento que se considera al momento de analizar la calidad de una evaluación es la confiabilidad o precisión de la información que se obtiene con esta. A diferencia de la validez, la confiabilidad solo se relaciona con la consistencia de la medición, al margen de lo que se mida exactamente (Flórez, 1999; Hogan, 2004); así, puedo tener un instrumento muy confiable, pero que no tiene ninguna validez.

El concepto de confiabilidad hace referencia a la consistencia, exactitud y estabilidad de los resultados, a las inferencias que se pueden realizar, y tiene directa relación con las conclusiones y posterior toma de decisiones (Luckett & Sutherland, 2000; McMillan, 2003). La confiabilidad, en una lógica psicométrica, implica que el instrumento entrega resultados similares cuando se repite su aplicación en las mismas circunstancias y mismas personas. Se puede verificar de diversas maneras, pero las más comunes son el coeficiente alfa de Cronbach, que entrega la consistencia interna de un instrumento; la correlación test-retest, que evalúa la consistencia entre dos mediciones aplicadas al mismo sujeto en distintos momentos, y las formas paralelas, que evalúan el grado de correlación que hay entre dos versiones de una misma prueba.

Este criterio de calidad tiene su origen en la psicometría; se dice que una prueba será confiable si consistentemente genera la misma puntuación, o una similar, al ser aplicada a un individuo, y su puntuación es replicable con un margen de error pequeño. Moss (2003) señala que en las evaluaciones de aula (específicamente

pruebas), a diferencia de las evaluaciones a gran escala, la confiabilidad no es un tema relevante y que muy pocas veces se considera, ya que en la sala de clases los profesores tienen múltiples instancias para recoger información complementaria de un aprendizaje. Smith (2003) reafirma esta idea señalando que los profesores no calculan coeficientes alfa ni correlaciones test-retest ni formas paralelas, por una parte, porque no tienen un número de casos suficiente (más de 100 alumnos en una misma aplicación) y, por otra, porque las evaluaciones son realizadas a un estudiante solo una vez en un tiempo específico. No obstante, lo que es fundamental en este caso es que se espera que el alumno cambie su aprendizaje de una semana a otra, elemento que va contra lo esperado en psicometría respecto de la estabilidad de los rasgos que se miden.

Por tanto, en el contexto de la sala de clases, donde no interesa generar un orden o jerarquía de los estudiantes en las actividades que se evalúan, la confiabilidad desde una mirada psicométrica no es pertinente (Brookhart, 2003; Smith, 2003). Además, muchas evaluaciones se califican con categorías (p. ej. excelente, satisfactorio, regular, insatisfactorio), lo que generaría un problema para hacer algún cálculo de confiabilidad “tradicional”. Moss (2003) plantea que lo correcto a nivel de aula es analizar la “suficiencia de información”, es decir, la confiabilidad está representada por tener suficiente evidencia de un aprendizaje y que permita tomar decisiones con el menor margen de error. Esta evidencia debe no solo corresponder al mismo contenido, sino tener un mismo nivel de exigencia, lo que parece ser lo más difícil de cumplir (Smith, 2003). Brookhart (2003) es muy clara al momento de diferenciar la confiabilidad a gran escala y a nivel de aula (tabla 2). A gran escala, la confiabilidad se entiende como la consistencia entre las dimensiones que componen un instrumento de medición, mientras que a nivel de aula es vista como suficiencia o saturación de la información. En el primer caso se busca una clasificación estable a lo largo de un continuo y a nivel de aula se busca obtener información estable entre el ideal del aprendizaje (lo que espera el profesor de sus alumnos) y lo real alcanzado. Popham (2014) es claro en señalar que

“si un profesor construye sus propias pruebas en el aula con rigurosidad, estas serán lo suficientemente confiables para sustentar las decisiones que tome basadas en los resultados obtenidos (...) si bien necesita al menos estar informado sobre el significado fundamental de confiabilidad, no sugiero que haga que sus propias pruebas de aula pasen por algún tipo de revisión de confiabilidad” (p. 89).

TABLA 2: Comparación de la confiabilidad de evaluaciones a gran escala y en la sala de clases

Criterio	Evaluaciones a gran escala	Evaluación en el aula
Definición	Es la consistencia interna entre dimensiones relevantes de un instrumento.	Es la suficiencia de información sobre el aprendizaje de un estudiante en una estrategia evaluativa.
Propósito	Tener una clasificación estable de los estudiantes en una escala de puntuación o una categorización estable a lo largo de un continuo de progreso.	Contar con información consistente acerca de la diferencia entre el desempeño de un estudiante y el “ideal”, tal como se define en los Objetivos de aprendizaje.

Adaptado de Brookhart (2003)

Entre los factores que pueden influir en la confiabilidad de una evaluación destacan: a) el número de observaciones o evidencias de un aprendizaje, ya que mientras más instancias de evaluación del mismo aprendizaje (número de ítemes en una prueba, situaciones formativas, etc.), mayor será la “consistencia interna” del proceso de evaluación y más confiable será la conclusión respecto del logro del estudiante; b) las características de la aplicación, referidas a la claridad de las instrucciones, tiempo destinado para responder y las disposiciones del espacio físico, y c) la precisión de la revisión y puntuación de la tarea evaluativa (Himmel et al., 1999; Hogan, 2004). Este último punto será tratado más adelante en forma independiente bajo el título de “objetividad”, dada la relevancia que tiene en el ámbito educativo.

Los estudiantes, frente a este criterio de calidad, señalan que sus malas experiencias evaluativas se centran principalmente en la falta de oportunidades para demostrar su aprendizaje en distintas

instancias. Algunos ejemplos dados por ellos se presentan a continuación:

> “En un taller en la universidad solo había una prueba al final y con eso se jugaba la nota del curso. Encuentro que eso es una mala evaluación, porque hay muchas cosas que pueden afectar la respuesta que uno da y justo ese día puede pasar algo que haga que uno responda mal, aunque haya estudiado”.
>
> “Una situación negativa de evaluación fue cuando tenía que disertar y estaba enferma con fiebre y me sentía muy mal, la profesora me retó porque mi exposición fue muy mala. Me dio mucha rabia, porque yo me sabía la materia y no tuve otra oportunidad de demostrarlo”.
>
> “En un curso donde las pruebas eran una pregunta por cada tema y si justo no te sabías eso específico, perdías con esa pregunta, aunque te supieras el tema”.

Para resguardar que las evaluaciones sean confiables se debe tener en cuenta que el fin último de este criterio es que las conclusiones y posteriores decisiones pedagógicas estén basadas en evidencia suficiente que dé cuenta del progreso y logro de los aprendizajes para no cometer errores o al menos minimizarlos. En este sentido, las acciones sugeridas a continuación proporcionan una guía para lograr evaluaciones confiables:

1. Aplicar al mismo estudiante variadas situaciones evaluativas que midan el mismo aprendizaje: dar la posibilidad al estudiante de demostrar su desempeño en un mismo aprendizaje, a través de reiteradas oportunidades, nos permite tener la certeza de que el aprendizaje fue o no logrado (Riggan & Oláh, 2011). Esto se puede realizar en distintas tareas evaluativas o poniendo en un mismo instrumento diversas situaciones que evalúen lo mismo.
2. Velar por la claridad de los ítemes e instrucciones: esta sugerencia tiene relación con la calidad de la información que recogemos, ya que si es el instrumento no es adecuado, nuestras

evidencias no serán confiables y las decisiones que tomemos pueden tener sesgo o ser erróneas. En este sentido, hay ocasiones en que un ítem de una prueba o el descriptor de una escala son ambiguos, o hacemos una guía para un trabajo que nos queda confusa y los estudiantes no la comprenden o entienden otra cosa. Si esto ocurre, la información no será acorde a los aprendizajes que los estudiantes tienen y debemos considerarlo al momento de revisar, puntuar e interpretar la información para tomar decisiones que pueden ir desde eliminar el ítem hasta ampliar las posibles respuestas correctas o cambiar los criterios de evaluación en el trabajo. Lo importante es no perder de vista el objetivo de aprendizaje que se está evaluando.

3. Velar porque el ambiente de aplicación sea adecuado en cuanto a recursos, espacios y tiempo: es muy importante tener en cuenta esta sugerencia cuando se realizan evaluaciones en ambientes que no son los óptimos, por ejemplo, hay música en el patio, hace mucho calor en la sala, hay una actividad recreativa programada a continuación, el estudiante está enfermo, etc., pues los resultados de la evaluación estarán influidos por estos aspectos y puede que no representen los aprendizajes reales de los estudiantes.
4. Velar por la precisión de la revisión: esta sugerencia se refiere a la objetividad de la revisión y puntuación de la tarea evaluativa y será tratada con detalle a continuación.

OBJETIVIDAD

La objetividad o precisión de la revisión en un proceso evaluativo es un elemento clave asociado a la confiabilidad de una evaluación, sin embargo, decidimos tratarla en forma separada, dada la relevancia e impacto que tiene al interior de la sala de clases. Esto no supone que sea independiente de la confiabilidad.

Por lo general, la objetividad se entiende como la calidad de un objeto en sí, independiente de las consideraciones o juicios personales. Si llevamos la objetividad al ámbito evaluativo, esta supone que tanto los instrumentos como el juicio que se emite a partir de la información

recogida con ellos sean imparciales. En relación con este tema, Calatayud (1999) afirma que quien cree que la evaluación de los estudiantes es una acción objetiva se embarca en una tarea imposible. La autora plantea que la evaluación es una práctica compleja, que involucra no solo el dominio de una técnica sino también posee una carga moral y valórica importante y, por tanto, es una actividad ligada a las creencias personales de los docentes, lo quieran o no. Desde esta perspectiva, la imparcialidad no es posible y cada juicio que se emite tiene un componente de apreciación propia de quien lo expresa. La preocupación por la objetividad de la evaluación no es un tema nuevo, estudios asociados a las expectativas de los docentes, como los de Rosenthal y Jacobson (1968, efecto Pigmalión), de Spear (1984) y de Rubie-Davies (2006), muestran que los estudiantes considerados de “bajo rendimiento” tienen menos tiempo para contestar en una interrogación oral, reciben menos ayuda o “pistas” cuando preguntan alguna duda en una prueba y obtienen menos retroalimentación de su desempeño que los estudiantes considerados de “alto rendimiento”, para los cuales los docentes tienen mayores expectativas. Cabe señalar que estas expectativas, en general, están condicionadas por características personales de los estudiantes, como el género, el nivel socioeconómico, la raza o etnia, la apariencia física y los patrones de lenguaje oral.

Estudios como el de López, Llopis, Llorens, Salinas y Soler (1983) dan cuenta de cómo diferentes profesores de Matemáticas o Física asignan puntuaciones o calificaciones diferentes a un mismo ejercicio, lo cual se podría atribuir a diferentes niveles de exigencia en la corrección, que no debería afectar a los estudiantes si todos los ejercicios son revisados por la misma persona (Gil-Pérez & Vilches, 2008), mientras que otras evidencias indican que un mismo profesor puede puntuar o calificar de manera diferente la respuesta de un estudiante dependiendo del momento (Hoyat, 1962). Aunque se asume que la revisión nunca será ciento por ciento objetiva, es conocido por todos que los profesores tienen escaso tiempo para corregir trabajos y pruebas de sus estudiantes, y que lo hacen en los

pocos espacios físicos y temporales que tienen disponibles, algunos comienzan en el colegio durante la mañana, pero pueden terminar al día siguiente en su casa de madrugada. Si no se generan acciones concretas que eviten distorsionar el criterio de corrección, las posibilidades de cambiar la exigencia y la respuesta esperada pueden ser enormes.

Estos problemas de objetividad se ven reflejados en las opiniones de estudiantes respecto de experiencias de evaluación que resultaron ser negativas para su proceso de aprendizaje:

> “En un taller se hacía controles de lectura con 3 preguntas sin mayor explicación, al recibir mi control corregido observé que no tenía todo el puntaje en ninguna de las preguntas, al preguntar por qué, el profesor solo me dio la respuesta correcta y según yo era lo mismo pero con otras palabras”.

> “Fue una prueba en un ramo donde casi todo el curso se sacó un rojo. Nadie entendía por qué estaba todo mal. Las preguntas fueron muy específicas, y la pauta era muy estricta. No quedó claro qué estaba malo y qué era correcto”.

> “En la clase de Arte, en mi caso presenté una pintura que fue bien evaluada, la hice a la rápida y no me tomó mayor trabajo, en cambio, a uno de mis compañeros lo evaluó muy mal, siendo que su pintura era mucho más trabajada, novedosa y había seguido todas las instrucciones. Como no había pauta, no pudo alegar nada”.

> “En Enseñanza Media, todas las evaluaciones de Historia se realizaban con cuestionarios, el problema radicaba en que las pruebas quedaban en poder del profesor y no se tenía derecho a corrección”.

> “Un trabajo final, a pesar de que mi trabajo fue considerado uno de los mejores del curso en la presentación, fue mal evaluado porque me encontraba en malas relaciones con el profesor”.

> “Un profesor calificó como insatisfactorio mi trabajo, pero al compararlo con otros notamos que estaba mejor que los que tenían calificaciones satisfactorias. Esta experiencia fue negativa porque pesó más el criterio del profesor, que no es conocido ni entendido por los alumnos”.

Dado que la objetividad de una evaluación consiste en resguardar la ausencia de sesgos o apreciaciones subjetivas en la interpretación de las evidencias y/o procesos que las generaron, se sugieren algunas acciones que un docente puede realizar fácilmente en el aula:

1. Informar a los alumnos la intencionalidad de la evaluación y los aprendizajes que se evaluarán: la doble función de esta acción es que, por una parte, el docente debe hacer explícito el propósito de sus evaluaciones en la planificación y, por otra, para sus estudiantes no hay ambigüedades en qué se les va a evaluar y con qué fin, por tanto, la evaluación deja de ser una "caja negra" y se convierte en una herramienta de aprendizaje con normas claras y bien definidas.
2. Dar a conocer a los alumnos los criterios de evaluación: si entendemos que la evaluación es parte del proceso de enseñanza/aprendizaje y, por ende, se constituye como una herramienta para mejorar los aprendizajes de los estudiantes, estos deben saber claramente con qué criterios se les va a evaluar. Preparar una presentación oral que va a ser evaluada en cuanto al uso de las habilidades comunicativas es muy diferente que prepararla para ser evaluado en el dominio conceptual de dicha presentación, por tanto, los énfasis que un estudiante pone y el tipo de aprendizaje que desarrolla estarán en directa relación con la pauta de evaluación que se le haya entregado. Si la pauta y sus criterios son ambiguos o desconocidos, el estudiante realizará lo que cree que debe hacer y si eso no concuerda con lo que el profesor evalúa, la evidencia no será confiable y las conclusiones, erróneas.
3. Elaborar pautas de respuesta o de revisión: es común elaborar pruebas o situaciones de evaluación y luego, al momento de revisarlas, hacemos un punteo de las ideas que deberían estar presentes en la respuesta o desempeño del estudiante, pero es un hecho que el criterio y la exigencia varían producto de diferentes causas (p. ej., el momento del día, el contexto en el

cual se corrige, etc.), por tanto, tener una pauta clara y precisa ayuda a mantener la consistencia en la aplicación y evitar los sesgos propios de un proceso de corrección. Además, permite que los estudiantes conozcan esos criterios y puedan contrastarlos con sus resultados, haciendo transparente el proceso de evaluación y permitiéndoles autoevaluarse, reconocer sus progresos y saber qué se esperaba en cada caso.

4. Establecer previamente los criterios de asignación de puntajes en función de la relevancia y nivel de complejidad de los aprendizajes: este elemento es clave en dos sentidos: primero, permite al docente asegurar, previamente a la aplicación de la evaluación, que los ítemes reflejen, en términos de extensión y complejidad, la relevancia que se les asigna con un puntaje determinado en relación con el puntaje total de la evaluación y, en segunda instancia, los estudiantes, cuando la responden, toman decisiones respecto de la dedicación y orden con la que se enfrentan a los ítemes en función del peso que tiene cada uno. La ausencia de esta asignación genera ambigüedad para los alumnos respecto de los énfasis que da el docente a cada aspecto que se está evaluando. Sabemos que nuestro criterio y asignación de puntajes puede variar mientras vamos revisando las evaluaciones, y una respuesta que tiene, por ejemplo, 0,75 puntos cuando comenzamos se transformará en 0,5 o 0,8 puntos cuando estemos terminando si no tenemos una pauta clara y precisa que nos ayude a mantener la consistencia.

La evaluación es parte del proceso de enseñanza y aprendizaje al interior del aula y, como tal, puede marcar a un estudiante tanto positiva como negativamente. En esta revisión se ha presentado cómo la ausencia de validez y confiabilidad en una situación evaluativa puede generar desmotivación, aprendizajes erróneos y sentimientos de injusticia en los estudiantes, que en nada favorece un ambiente de aprendizaje que resulte significativo para ellos.

En general, los profesores no calculan la confiabilidad, el error estándar de una medición o los coeficientes de validez y

discriminación de una evaluación. Estas técnicas son propias del desarrollo de pruebas estandarizadas y de aplicación a gran escala, pero tienen una importancia limitada en la evaluación en el aula. Aquí rescatamos los elementos que pueden ser claves y necesarios de resguardar para hacer evaluaciones que aporten al aprendizaje de los estudiantes, como son la validez de contenido, la validez instruccional, la validez consecuencial, la confiabilidad y la objetividad. Como se discutió anteriormente, la validez no se plantea en términos absolutos, sino en grados, y depende del propósito para el que fue creada la evaluación. De esta forma, los resultados serán más o menos válidos para sacar conclusiones respecto de ese fin. Tanto la validez como la confiabilidad aumentan o disminuyen según sea la calidad de las evidencias de aprendizaje que las sustentan.

Aun cuando la evaluación de aprendizajes al interior del aula es una tarea compleja que requiere tiempo y esfuerzo, es fundamental entenderla como parte integral del proceso de enseñanza y no verla como una instancia separada que va al final del proceso. Si se asume esta posición, habrá coherencia entre la planificación y la práctica, en la selección de los contenidos y habilidades que se enseñan y que se evalúan, mejorando prácticas evaluativas en las que escasean la validez y la confiabilidad.

EN SÍNTESIS, la calidad de la evidencia que recogemos para monitorear y certificar el aprendizaje de nuestros estudiantes depende de la calidad de los instrumentos y situaciones evaluativas que usemos. Mientras más alineada esté una evaluación con los Objetivos de aprendizaje y las oportunidades que el estudiante tuvo para aprenderlos, mejor será la evidencia y nos permitirá tomar decisiones pedagógicas que sean más efectivas para mejorar los aprendizajes. No debemos perder de vista también que los instrumentos de revisión, como las rúbricas, le dan consistencia a la revisión y nos ayudan a mantener un mismo criterio, por lo que nuestras evaluaciones serán más objetivas.

En el capítulo final abordaremos una propuesta metodológica de

análisis de casos que permite sistematizar evidencias obtenidas a partir de distintas fuentes, que configuran una situación problemática o una necesidad a nivel escolar y contrastarla con referentes teóricos para hacer una propuesta de mejora.

BIBLIOGRAFÍA

Allen, M. J. & Yen, W. M. (2002). *Introduction to measurement theory.* Long Grove: Waveland Press.

Borsboom, D., Mellenbergh, G. J. & Van Heerden, J. (2004). The concept of validity. *Psychological Review, 111*(4), 1061-1071.

Brookhart, S. (2003). Developing measurement theory for classroom assessment purposes and uses. *Educational Measurement: Issues and Practice, 22*(4), 5-12.

Brualdi, A. (1999). *Traditional and modern concept of validity*. Washington DC: ERIC/AE Digest.

Calatayud, M. A. (1999). La creencia en la objetividad de la evaluación: una misión imposible. *Aula Abierta, 73*, 205-222.

Crocker, L., Llabre, M. & Miller, M. D. (1988). The generalizability of content validity ratings. *Journal of Educational Measurement, 25*(4), 287-299.

Flórez, R. (1999). *Evaluación pedagógica y cognición*. Bogotá: Mc Graw Hill.

García, S. (2002). La validez y la confiabilidad en la evaluación del aprendizaje desde la perspectiva hermenéutica. *Revista de Pedagogía, 23*(67), 297-318.

Gil-Pérez, D. & Vilches, A. (2008). ¿Qué deben saber y saber hacer los profesores universitarios? En M. I. Cebreiros & N. Casado (Eds.), *Novos enfoques no ensino universitario* (pp. 25-43). Vigo: Universidad de Vigo.

Gorin, J. S. (2007). Reconsidering issues in validity theory. *Educational Researcher, 36*(8), 456-463.

Hattie, J. (2015). We aren't using assessments correctly: there's a distinction between formative and summative assessments. *Education Week, 35*(10), 23-28.

Himmel, E., Olivares, M. A. & Zabalza, J. (1999). *Hacia una evaluación educativa. Aprender para evaluar y evaluar para aprender* (vol. I). Santiago: Ministerio de Educación - Pontificia Universidad Católica de Chile.

Hogan, T. (2004). *Pruebas psicológicas. Una introducción práctica.* México DF: El Manual Moderno.

Hoyat, F. (1962). *Les examens*. Institut de l'UNESCO pour l'Education. Paris: Bourrelier.

López, N., Llopis, R., Llorens, J. A., Salinas, B. & Soler, J. (1983). Análisis de dos modelos evaluativos referidos a la química de COU y selectividad. *Enseñanza de las Ciencias, 1*(1), 21-25.

Luckett, K. & Sutherland, L. (2000). Assessment practices that improve teaching and learning. In S. Makoni (Ed.), *Improving teaching and learning in higher education: a handbook for Southern Africa* (pp. 98-130). Johannesburg: Witwatersrand University Press.

Lukas, J. F. & Santiago, K. (2004). *Evaluación educativa*. Madrid: Alianza.

McMillan, J. H. (2003). Understanding and improving teachers' classroom assessment

decision making: implications for theory and practice. *Educational Measurement: Issues and Practice, 22*(4), 34-43.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (pp. 13-103). New York: Macmillan.

Moss, P. A. (1997). The role of consequences in validity theory. *Educational Measurement: Issues and Practices, 17*(2), 6-12.

Moss, P. A. (2003). Reconceptualizing validity for classroom assessment. *Educational Measurement, Issues and practice, 22*(4), 13-25.

Moss, P. A. (2007). Reconstructing validity. *Educational Researcher, 36*(8), 470-476.

Popham, W. J. (2014). *Classroom assessment: what teachers need to know*. Boston: Pearson Education.

Riggan, M. & Oláh, L. N. (2011). Locating interim assessments within teachers' assessment practice. *Educational Assessment, 16*(1), 1-14.

Rosenthal, R. & Jacobson, L. (1968). *Pygmalion in the classroom*. New Jersey: Rineheart & Winston.

Rubie-Davies, C. M. (2006). Teacher expectations and student self-perceptions: exploring relationships. *Psychology in the Schools, 43*(5), 537-552.

Salinas, D. (2002). *¡Mañana examen! La evaluación entre la teoría y la realidad*. Barcelona: Graó.

Sanmartí, N. (2007). *10 ideas clave. Evaluar para aprender*. Barcelona: Graó.

Schutz, A. & Moss, P. A. (2004). Reasonable decisions in portfolio assessment: evaluating complex evidence of teaching. *Education Policy Analysis Archives, 12*(33). Recuperado de http://epaa.asu.edu/ojs/article/view/188

Sireci, S. G. (1998). The construct of content validity. *Social Indicators Research, 45*, 83-117.

Smith, J. K. (2003). Reconsidering reliability in classroom assessment and grading. *Educational Measurement: Issues and Practice, 22*(4), 26-33.

Spear, M. G. (1984). Sex bias in science teachers' ratings of work and pupils characteristics. *European Journal of Science Education, 6*, 369-377.

Stiggins, R. J. (2001). *Student-involved classroom assessment*. New Jersey: Prentice-Hall.

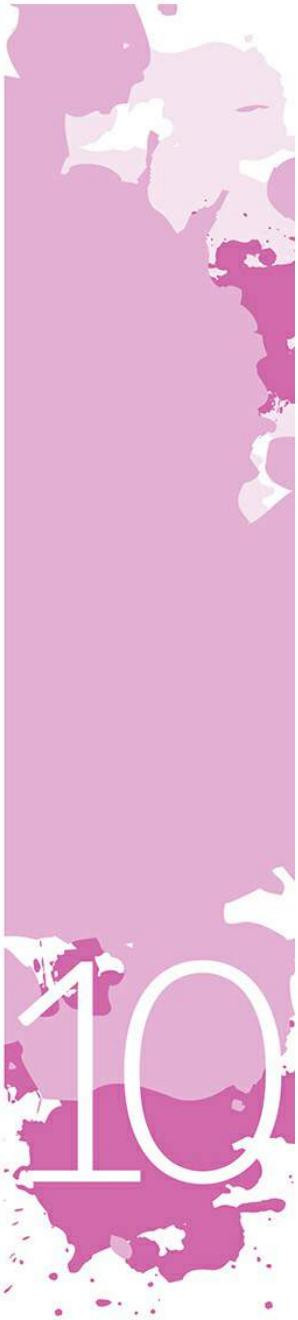
Valverde, G. (2000). La interpretación justificada y el uso apropiado de los resultados de las mediciones. En P. Ravela (Ed.), *Los próximos pasos: ¿hacia dónde y cómo avanzar en la evaluación de aprendizajes en América Latina?* (pp. 21-30). Lima: GRADE/PREAL.

* Este capítulo fue publicado originalmente en *Pensamiento Educativo. Revista de Investigación Educacional Latinoamericana*, vol. 43, N° 2, 2008. Reproducido con la autorización del editor © 2017.

[1] Consistencia interna o confiabilidad de dicho instrumento, la que se calcula estadísticamente con KR-20 o alfa de Cronbach y las dimensiones comunes que subyacen al modelo a través de análisis factorial.

CAPÍTULO 10

Una metodología de análisis de casos para orientar la toma de decisiones en el contexto escolar



Paola Marchant-Araya
Carla E. Förster

INTRODUCCIÓN

El propósito de este capítulo es presentar un método de análisis que, aplicado a contextos reales y siguiendo una serie de etapas, permita tanto a los equipos docentes y técnicos de las organizaciones escolares como a quienes realizan asesoría externa a las instituciones educativas, comprender una realidad particular de un modo holístico y desde múltiples perspectivas para proponer caminos de acción que orienten la mejora y/o favorezcan la sustentabilidad de sus decisiones. A través del análisis, las instituciones, programas o proyectos educativos podrán reconocer y valorar cómo sus acciones y decisiones han configurado la realidad que se vive cada día en el aula, sus causas, consecuencias, relaciones y tensiones. Se propone un ejercicio de articulación entre marcos conceptuales, evidencia y reflexión crítica, en que la participación de los actores educativos será relevante, ya que en conjunto se harán emerger las decisiones pedagógicas de cambio. Este método de análisis de casos es flexible y favorece en los equipos el desarrollo de habilidades cognitivas superiores, tales como el pensamiento reflexivo y analítico, la capacidad de síntesis, la evaluación y la creación. Así también, promueve la toma de decisiones basada en evidencias, tan necesaria en las instituciones escolares de hoy (Mendoza, 2006).

CONCEPTUALIZACIÓN DE LA METODOLOGÍA DE ANÁLISIS DE CASOS

¿QUÉ ES UN CASO?

En la literatura podemos encontrar distintas definiciones de lo que se entiende por un caso, pero hay consenso en que no es necesariamente equivalente a un individuo o a un estudio cualitativo *per se* (Ragin, 1992; Coller, 2000); “un buen caso presenta un tema que provoca interés y promueve la empatía con los personajes centrales” (Center for Teaching and Learning Stanford University, 1994, p. 1).

El caso corresponde a una unidad singular de análisis que puede ser una organización, programa o proyecto educativo, la experiencia de un colectivo o de una persona que ha vivido algo singular, acontecimientos presentes o pasados. Un caso puede ser una experiencia de cambio organizacional, la implementación de una nueva política, un incidente o problema vivido. También el currículum de un establecimiento educacional puede configurar un caso, así como los métodos de evaluación de aprendizajes usados por los profesores de un departamento o ciclo. Es decir, un caso es una situación real que se vive a nivel micro, meso o macro (Stake, 1998; Mendoza, 2006; Simons, 2011).

Para Coller (2000), un caso es “un objeto de estudio con unas fronteras más o menos claras que se analiza en su contexto y que se considera relevante, bien sea para comprobar, ilustrar o construir una teoría o una parte de ella, bien sea por su valor intrínseco” (p. 29). Stake (1998), por su parte, entiende el caso como un sistema integrado, algo específico, complejo y en funcionamiento, una singularidad o una particularidad. Es decir, se constituye como caso de estudio un fenómeno que requiere ser comprendido en profundidad. Para Rusque y Castillo (2009), un caso “es la narración de una situación real o simulada que se presenta en el contexto en que se produce y que se analiza sobre la base de unos criterios preestablecidos” (p. 53).

Por lo tanto, en este capítulo entenderemos un caso como:

> La exposición de hechos, datos, circunstancias, puntos de vista opuestos,
> el contexto de una situación exitosa o problemática que ha enfrentado o se encuentra enfrentando una organización, programa o proyecto educativo o social; la experiencia de un grupo de personas de una institución, sean estudiantes, profesores y/o el equipo directivo; la respuesta, tensiones y acciones que se han dado en la situación, y los efectos o implicancias que surgen a partir de tales respuestas, es decir, hechos con cierta complejidad y en movimiento, dadas las decisiones y acciones que se han y están tomando en el contexto específico. Un caso es una singularidad con límites claros por tres razones: por su propia naturaleza, esto es, dadas sus

particularidades en contexto; porque está referido a un momento en el tiempo y a un lugar, y porque involucra a un grupo de personas que viven una experiencia de relaciones en común.

¿QUÉ ES UN ESTUDIO DE CASO?

Los estudios de casos son aproximaciones metodológicas que se han desarrollado durante décadas en distintas modalidades (figura 1) y con diferentes fines para contribuir al conocimiento individual, grupal, organizacional, social y político, relacionado con un fenómeno particular (Stake, 1998; Mendoza, 2006; Mertens, 2015; Yin, 2009). Existen estudios de casos en el contexto de la investigación (Covarrubias-Papahiu, 2016; Guizardi & Garcés, 2014; Rambla, 2013); como metodología de enseñanza y aprendizaje en contextos de aulas escolares y universitarias (Brooke, 2006; Chion, 2013; Coll, Mauri & Onrubia, 2006; Davis & Wilcock, 2003; Golich, Boyer, Franko & Lamy, 2000; Instituto Tecnológico Monterrey, s. f.; Servicio de Innovación Educativa, 2008); en el campo del Derecho (Goyas & Monzón, 2016); la gestión de empresas y negocios (Ellet, 2007; Farhoomand, 2004; Richardson, 2013; Rivera, 2000); la Medicina (Freeman, 1997; Edelstein, Wallensten & Kühlmann-Berenzon, 2014); en el desarrollo de estudios organizacionales y en el análisis de políticas públicas (Mendoza, 2006).

Los orígenes de los estudios de casos están en la Antropología, la Sociología, el Trabajo Social, la Psicología, la Historia y en el campo del Derecho y la Medicina. También han tenido un gran desarrollo en la Educación y la evaluación educativa, en el contexto de las innovaciones curriculares, programas y proyectos como un medio para evaluar su efectividad (Coller, 2000; Simons, 2011; Yin, 2009). En la figura 1 se presentan los tres contextos que agrupan los estudios de casos en el ámbito educacional.

FIGURA 1: Esquema que caracteriza la variedad de estudios de casos

En el contexto de la investigación, los estudios de caso surgen a mediados del siglo XIX, en Europa, con el propósito de conocer las condiciones de vida de los grupos que vivían en contextos vulnerables. Se buscaba, a partir de este conocimiento, contar con evidencia que, complementando los datos de encuestas y mediciones, sirviera de fundamento para los cambios sociales (Coller, 2000; Ottenberger, 2000; Yin, 2009). Estos estudios utilizaban la entrevista en profundidad, la observación y el trabajo de campo como técnicas principales para recoger información. Posteriormente, a inicios del siglo XX en la Escuela de Chicago, los sociólogos de la época desarrollaron estudios de casos para abordar temáticas como la delincuencia y los grupos migrantes, ampliando la selección de distintas fuentes de información para la construcción de los casos de estudio, consolidándose como método de investigación hasta ahora (Coller, 2000; Stake, 1998; Yin, 2009; Mertens, 2015).

El estudio de caso es una aproximación holística y compleja hacia lo singular, lo particular, lo exclusivo en el contexto de la realidad (Stake, 1998). Tiene como propósito construir, describir, explicar y comunicar la historia de un caso objeto de investigación. Es un ejercicio de documentación e interpretación de una experiencia vivida por una institución, un programa, un proyecto, un sistema, grupos o personas particulares, y aporta información para la toma de decisiones

(Ottenberger, 2000; Simons, 2011; Yin, 2009). Se analizan las distintas perspectivas de cómo ocurren los fenómenos o situaciones de manera natural, sin intervención de un tercero, es decir, los sucesos se dan en acción y movimiento, en un contexto real (Simons, 2011). La interpretación de los hechos siempre es en contexto y dadas sus características metodológicas y finalidades; no busca la generalización estadística.

En investigación, los estudios de casos no son exclusivos de los enfoques cualitativos, también tenemos estudios de casos que implican enfoques cuantitativos o mixtos; en todos es fundamental la pregunta que se busca responder y si el foco es contemporáneo o no (Coller, 2000; Stake, 1998; Yin, 2009). Independientemente de lo anterior, la investigación con estudios de casos necesita cumplir con los criterios de rigurosidad y precisión en el manejo de los datos empíricos que se utilicen (Yin, 2009). Se debe buscar, por tanto, evitar falsas interpretaciones y representaciones, de modo de asegurar la validez de los resultados y conclusiones. Para ello, la técnica de triangulación[1] y la técnica de discusión pública son las más recomendadas (Stake, 1998; Coller, 2000). Ambas ayudan a comprobar y contrastar la información recabada a partir de diversas fuentes (informantes y documentos).

En el contexto de la enseñanza y el aprendizaje, los estudios de caso se remontan al año 1880 en la Escuela de Leyes de la Universidad de Harvard. Este tipo de estudios, también llamados “análisis de casos”, se ha utilizado en la enseñanza con el propósito de desarrollar en los estudiantes la capacidad de pensar y aprender por sí mismos; promover la metacognición y la autorreflexión; el análisis, la síntesis, la abstracción, el razonamiento y el trabajo en equipo, así como desarrollar una visión multidimensional de la realidad y promover el aprendizaje activo (Davis & Wilcock, 2003; Flyvbjerg, 2004; Instituto Tecnológico de Monterrey,
s. f.; Mendoza, 2006).

Con foco en el logro de determinados aprendizajes, el análisis de casos o *teaching case* se ha utilizado para analizar hechos, problemas o situaciones que se dan en la realidad o que han sido

construidas de manera ficticia, con el propósito de que los estudiantes discutan y debatan acerca de estas (Lane, 2007), reconozcan sus causas y efectos, aplicando a su propia vida las habilidades y conocimientos que surgen de la reflexión grupal y proponiendo soluciones pertinentes (Mendoza, 2006; Yin, 2009). Dada la dinámica grupal del análisis, no se espera una única solución, sino que se pone énfasis en la capacidad de los alumnos de descomponer una situación en sus múltiples partes para luego organizarlas y dar forma a una nueva comprensión de la situación dada, es decir, lograr aprendizajes y el desarrollo de habilidades cognitivas complejas (Morra & Friedlander, 2001; Flyvbjerg, 2004).

Cabe destacar que en una aproximación cualitativa del estudio de caso, su validez y carácter probativo no dependen de la capacidad de representación, sino de su propia realidad, de su autenticidad y no de su frecuencia, es decir, se espera desarrollar la generalización analítica y particularización del caso (Stake, 1998; Coller, 2000; Flyvbjerg, 2004; Yin, 2009). En este sentido, Stake (1998) plantea que la generalización que se hace en los estudios de casos, llamadas también “generalizaciones naturalistas”, son “conclusiones a las que se llega mediante la implicación personal en los asuntos de la vida o mediante una experiencia vicaria tan bien construida que las personas sienten como si ellas mismas la hubieran tenido”
(p. 78).

En el contexto de las organizaciones podemos encontrar distintos tipos de estudios de casos, en primer lugar y el más clásico, las evaluaciones de programas, las que, según Stake (1998), son en sí mismas estudios de casos, ya que abordan una particularidad en un contexto real con una mirada crítica y una variedad de fuentes. Las evaluaciones con estudios de casos, para Morra y Friedlander (2001), ponen el énfasis tanto en la implementación del programa como en los efectos del mismo, promoviendo la participación y colaboración, con el propósito de fortalecer la identificación con un proyecto de mejora.

Otro tipo de estudio de caso es el “Uso sabio de datos” o “Data wise” (Bambrick-Santoyo, 2010; Boudett, City & Murnane, 2005; Parker-

Boudett, City, Murnane & Steele, 2006; Ebbeler, Poortman, Schildkamp & Pieters, 2017). Este tipo de análisis busca favorecer el uso constructivo y regular de los datos de evaluaciones para el mejoramiento de la enseñanza y el aprendizaje en directivos y docentes de centros escolares para favorecer la toma decisiones y el mejoramiento continuo. Tanto la evaluación como la metodología del uso sabio de datos buscan involucrar a la escuela en los procesos de análisis y generación de los resultados, que permitan que la comunidad aprenda y tome las decisiones que sean necesarias para seguir su desarrollo y favorecer los aprendizajes de los estudiantes.

Como lo ha planteado Santos-Guerra (2001), una escuela que aprende no considera a sus miembros como una suma de individualidades con tareas específicas, sino que es una comunidad de aprendizaje que trabaja por un proyecto común. Es en este contexto que surge la metodología de análisis de casos para la toma de decisiones que presentamos en este capítulo, como una posibilidad real de que la comunidad en conjunto analice la historia de sus acciones, sus resultados, logros, fracasos, aspectos críticos y tensiones, para identificar una problemática y planificar líneas de acción que permitan abordarla y mejorar. Así, surge como una oportunidad para la comunidad educativa el trabajo en conjunto y la reflexión efectiva, superando la rutinización de las prácticas, la centralización excesiva, el miedo a la supervisión, la desmotivación de los profesores, la ausencia de autocrítica y el rechazo a la crítica (Santos-Guerra, 2001).

La metodología de análisis de casos para la toma decisiones educativas que presentamos se entenderá como:

> Un tipo de estudio de caso que busca, a través de una serie de etapas, de la reflexión colectiva de los equipos educativos y del uso de herramientas metodológicas mixtas, analizar las particularidades de un caso real, sus causas, consecuencias o efectos y singularidades, con base en conocimientos y evidencias, existentes y/o construidas, para llegar a comprender el caso en profundidad, aprender de él y proponer de manera sistemática estrategias que permitan la mejora y/o sustentabilidad de las acciones emprendidas dentro del contexto escolar.

Analizar un caso en su contexto supone un ejercicio de observación, de carácter holístico, para llegar a comprenderlo en su totalidad y no como una suma de partes (Ragin, 1992). Para Schramm (1971 en Yin, 2009), un estudio de casos intenta iluminar una decisión o un conjunto de decisiones: ¿por qué se tomaron?, ¿de qué forma se implementaron?, ¿con qué resultado?

El análisis de casos es una oportunidad para el desarrollo de habilidades de comunicación, trabajo en equipo, solución de problemas, ejercer la toma de decisiones y de evaluación (Davis & Wilcock, 2003; Rusque & Castillo, 2009). Estimula la reflexión crítica de los equipos y la autorreflexión de los participantes.

La metodología de análisis de casos permite detectar necesidades particulares de las organizaciones escolares y hacer vínculos entre la realidad y los marcos conceptuales o conocimientos que poseen quienes participan, avanzando hacia la generación de un nuevo conocimiento y aprendizajes colectivos que surgen en el propio proceso de análisis. También, compromete a los actores en la definición de acciones para la mejora y en la solución de los problemas con base en las evidencias recolectadas y analizadas.

El foco de este análisis está puesto en responder dos preguntas claves: ¿cómo? y ¿por qué?, más que ¿cuántos? y ¿con qué frecuencia? Esto porque el espíritu del análisis es abordar el caso desde una aproximación comprensiva e interpretativa, que pone el acento en la experiencia de los sujetos y en la exploración profunda de esta experiencia. Considera la historia vivida y los patrones de continuidad y cambio de la experiencia colectiva. Lo anterior no descarta, obviamente, el uso de datos cuantitativos o resultados de mediciones, por el contrario, se valora en cuanto evidencia que enriquece la comprensión de un fenómeno o realidad particular.

RESPONSABILIDADES DE LOS PARTICIPANTES EN EL ANÁLISIS DE CASOS

La participación de los distintos actores (analista externo, analista interno, comunidad educativa) en el análisis de casos debe ser promovida durante cada una de las etapas de dicho análisis. No nos referimos solamente a la participación en la entrega de información, que resulta ser lo habitual, sino a la participación efectiva en el sentido de ser parte del análisis, de la planificación y la implementación de las acciones posteriores. Estimular el autoconocimiento y la autorreflexión de las personas y equipos participantes favorecerá no solo la comprensión del caso sino también el compromiso por la sustentabilidad de las acciones que se desprendan del análisis. De este modo, el análisis de caso puede ser guiado o dirigido tanto por un agente interno como por uno externo.

El analista interno es un miembro de la comunidad que por asignación de esta responsabilidad desde un superior o por una atribución otorgada por un colectivo le corresponde dirigir el análisis y la propuesta de mejora. Este rol tiene grandes desafíos, ya que deberá administrar los distintos intereses y posiciones que se dan al interior de la organización, manteniendo un carácter íntegro y asegurando el cumplimiento de principios éticos tales como la dignidad de las personas, la justicia y la equidad, así como el valor de la confianza y el respeto. El analista interno deberá guiar cada etapa de análisis recabando toda la evidencia que sea necesaria desde distintas fuentes para construir el caso de estudio y, posteriormente, estimular la reflexión crítica del grupo o equipo
respectivo.

Asumir un rol de analista interno tiene la ventaja de estar cercano a las necesidades, la historia y la confianza de sus colegas, sin embargo también presenta la dificultad de estar personalmente involucrado con las experiencias vividas en la organización, así como también, los demás miembros de la organización conocen su posición e ideas y, por ende, pueden cuestionar la objetividad o imparcialidad de su análisis; por tal motivo, deberá encargarse de estar vigilante a sus propias creencias y preconcepciones, filiaciones y compromisos, para lograr tomar distancia y poder contribuir de un modo efectivo a la toma de decisiones. Se recomienda, en este sentido contar con un

pequeño equipo de apoyo de carácter heterogéneo, con el que se pueda discutir el desarrollo del análisis, o el apoyo de agentes claves, especialistas en el tema central del análisis a quien se puede solicitar una opinión externa una vez que el caso haya sido construido y en etapas anteriores al análisis.

El analista externo es un profesional especialista en la metodología de análisis de casos para la toma de decisiones, a quien se le ha solicitado abordar una problemática o tema de interés para la organización, equipos directivos, profesores, entre otros. Al analista y su equipo le corresponderá dirigir tanto el proceso de análisis como el diseño de la propuesta de mejora, por lo que se hace necesario recabar, en una primera entrevista exploratoria, los aspectos que iluminarán la búsqueda de evidencia desde distintas fuentes.

Asumir un rol de analista externo tiene la ventaja de aproximarse al caso con imparcialidad, colaborando con las personas de la organización para comprender la situación vivida y llegar a descubrir el foco central del caso más allá de los síntomas o aspectos secundarios, que suelen ser los que se aprecian con mayor claridad. Sin embargo, la dificultad de este rol es que no se cuenta con la confianza ni el conocimiento vivido respecto de las problemáticas o situaciones propias de la organización, por lo que resulta ser un desafío generar confianza y apertura respetando las dinámicas y códigos internos de relación sin imponer los propios. Junto con lo anterior, es rol del analista externo favorecer el aprendizaje y la creación de capacidades internas en los equipos y actores de los establecimientos, programas, proyectos, entre otros, para evitar la amenaza de la dependencia y el retroceso a etapas anteriores una vez que el analista abandona la organización.

ORGANIZACIÓN Y ESTRUCTURA DEL MÉTODO DE ANÁLISIS DE CASO PARA LA TOMA DE DECISIONES

RELEVANCIA DEL CASO

Antes de iniciar un análisis de caso debe hacerse explícito en la discusión con los miembros de la institución y en forma escrita por

qué estamos haciendo este análisis, qué queremos saber o qué queremos cambiar. Cuál es la necesidad o inquietud a la que responde el análisis, es decir, se debe determinar el horizonte en que se realiza. Luego es necesario tener certeza de estar frente a un caso, en los términos planteados con anterioridad, así como también, que se cumpla con los siguientes aspectos que denotan su relevancia para ser analizado:

1. Se trata de un caso singular, una particularidad que debe ser comprendida en su totalidad y complejidad, de un modo sistemático, a través de un muestreo motivado (Coller, 2000; Stake, 1998).
2. El interés del análisis no recae en un sujeto particular, sea este un investigador o un miembro de la organización, programa o proyecto, sino que en sí mismo constituye un fenómeno de interés por el alcance que tiene, por los actores involucrados, por sus efectos o consecuencias, por su naturaleza ejemplificadora de otros casos o su tipicidad.
3. A partir del análisis se pueden tomar decisiones de mejora o cambio, posibles de ser implementadas a corto o mediano plazo.

Si ya se cuenta con el caso, entonces se deben tener presentes ciertas condiciones de base para implementar la metodología de análisis de casos, tales como:

- Equipo directivo con disposición y compromiso.
- Acceso a información desde distintas fuentes y disponible para ser analizada.
- Actores activos participantes de los procesos de análisis y acciones
  de mejora.
- Analistas que conocen la metodología y los temas críticos que se trabajan en el análisis, para que la discusión vaya más allá

del sentido común o de las experiencias de vida cotidiana de los sujetos.

Contando con estas condiciones se puede avanzar hacia la construcción del caso propiamente tal, cumpliendo con las siguientes etapas:

Primera etapa: construcción del caso

Contextualización

El primer aspecto por describir es la contextualización. Esta debe dar cuenta del ambiente en el que se inserta el caso que se analizará con posterioridad. En el contexto se describe información tal como: tipo de institución, ubicación geográfica, recursos, cantidad de estudiantes y profesores, infraestructura, niveles de logro de aprendizajes en mediciones estandarizadas, evaluaciones internas y/o externas, clima del establecimiento, estructura organizacional, historia, entre otros aspectos. Se refiere a aspectos que permiten situar al lector en el caso, pero que son externos a este. Es parte también el contexto social, económico y político del momento, tanto en la organización como en el país, si corresponde, además de las características del barrio o localidad en que se desarrolle el caso.

La información se presenta de lo general a lo particular, intentando dejar al final del contexto lo que pueda tener directa relación con el caso para conectar ambas partes: contexto y caso.

Un ejemplo de información de contexto se presenta a continuación:

El caso a estudiar se desarrolla en un colegio de Educación Básica con jornada completa, que atiende a ٣٥٩ estudiantes desde prekínder hasta 8° Básico. El establecimiento está ubicado en la comuna de Valparaíso y forma parte de una fundación de educación. Esta fundación tiene, además, 2 colegios con enseñanza completa, ubicados en otras comunas de la región.

La composición socioeconómica de los estudiantes del colegio es primordialmente clase media baja, teniendo un índice de vulnerabilidad

mayor al 83,8%. Además, la institución se emplaza en una comuna con un alto índice de delincuencia.

El colegio cuenta con aportes de la ley de Subvención Escolar Preferencial (SEP), siendo clasificada por esta como un establecimiento emergente.

No ha mostrado “sistemáticamente buenos resultados educativos de sus alumnos de acuerdo a los instrumentos diseñados por el Ministerio de Educación” (ley SEP, 2008, p. 9). Asimismo, las evaluaciones estandarizadas indican que los alumnos del colegio tienen muy bajos resultados en comparación con estudiantes con un GSE similar.

La información del contexto puede obtenerse a partir de una entrevista preliminar con actores de la organización escolar, en la que se da cuenta de la necesidad, interés y motivación por realizar el caso. También se pueden solicitar antecedentes generales, informes o documentos que permitan involucrarse poco a poco en las experiencias vividas. El contexto puede ir complementándose y enriqueciéndose una vez que se tenga toda la información del caso.

Proceso de levantamiento de información

Para construir un caso y llegar a describirlo en su totalidad se debe planificar cómo se recogerá la información. Es así que, como punto de partida, se tiene que reconocer la necesidad que llevó a plantearse el análisis de caso y el tema o tópico central implicado en esta necesidad. Posteriormente se definen las dimensiones, subdimensiones o criterios que caracterizan la dimensión, lo que permitirá de manera exploratoria organizar la recogida y evitar que información que pueda ser relevante quede fuera. La desagregación de la dimensión en subdimensiones o criterios dependerá de la complejidad de los aspectos a abordar, puede haber dimensiones que se desagreguen y otras que no.

Posteriormente se definen las fuentes de información y los instrumentos y/o procedimientos que serán utilizados para recoger los datos. En la construcción de un caso se utilizan instrumentos y técnicas variadas, de carácter cuantitativo y cualitativo, esto es, entrevistas individuales y grupales, cuestionarios, pautas de registro

de observación, entre otros. Mientras más variada sea la evidencia recogida, mayor será la riqueza de la información disponible.

En los análisis de caso, las técnicas más apropiadas a utilizar serán la observación directa con sus pautas o procedimientos de registro, la entrevista, análisis de documentos y cuestionarios o escalas de apreciación.

Si se desea realizar un análisis de caso en una escuela donde existe una preocupación por la evaluación de los aprendizajes de los estudiantes, y sus resultados en mediciones, entonces se deberá planificar la recogida de información en una matriz como la siguiente (tabla 1):

TABLA 1: Matriz de recogida de información

OBJETO O NECESIDAD: ANALIZAR LAS PRÁCTICAS EVALUATIVAS DEL COLEGIO

DIMENSIONES: 1. Planificación de la estrategia evaluativa

Definición operacional (cómo se entiende la dimensión)	Subdimensiones	Criterio que caracteriza la dimensión	Fuentes de información	Técnicas o procedimientos de recogida de información	Técnicas de análisis de la información
Refiere a la definición de lo que se evalúa, la finalidad del proceso evaluativo, los momentos, quiénes participan en la evaluación, las situaciones e instrumentos seleccionados	1.1 Objetivos de Aprendizaje	1.1.1 Especificación del aprendizaje esperado: conocimientos conceptuales, procedimentales y/o actitudinales	• Docentes • Planificaciones de unidad • Reglamento de evaluación o similar	• Entrevistas individuales y/o grupales • Cuestionarios • Matriz de información de evidencia documental	• Análisis de contenido • Análisis estadístico descriptivo • Análisis documental
		1.1.2 Declaración de los indicadores de evaluación			
		1.1.3 Coherencia entre Objetivos de Aprendizaje e indicadores			
	1.2 Intencionalidad de la evaluación	1.2.1 Presencia en la unidad didáctica de la intencionalidad formativa inicial, formativa de proceso y sumativa			

TABLA 1: Matriz de recogida de información (cont.)

Definición operacional (cómo se entiende la dimensión)	Subdimensiones	Criterio que caracteriza la dimensión	Fuentes de información	Técnicas o procedimientos de recogida de información	Técnicas de análisis de la información
	1.3 Agentes de la evaluación	1.3.1 Tipo de agentes que participan en la evaluación según la intencionalidad			
		1.3.2 Razones para la selección de los agentes evaluadores			
	1.4 Situaciones e instrumentos de evaluación	1.4.1 Tipos de situaciones e instrumentos evaluativos utilizados			
		1.4.2 Frecuencia en el uso de las situaciones e instrumentos			
		1.4.3 Pertinencia de las situaciones e instrumentos con los Objetivos de Aprendizaje e indicadores a evaluar			

TABLA 1: Matriz de recogida de información (cont.)

OBJETO O NECESIDAD: ANALIZAR LAS PRÁCTICAS EVALUATIVAS DEL COLEGIO

DIMENSIONES: 2. Implementación de la evaluación

Definición operacional (cómo se entiende la dimensión)	Subdimensiones	Criterio que caracteriza la dimensión	Fuentes de información	Técnicas o procedimientos de recogida de información	Técnicas de análisis de la información
Refiere al conjunto de acciones que se realizan a lo largo del proceso de enseñanza/aprendizaje para recoger evidencia del nivel de avance y logro de los aprendizajes de los estudiantes	2.1 Diseño de las situaciones e instrumentos de evaluación	2.1.1 Presencia de planificación de la situación evaluativa y/o instrumento	• Docentes • Jefe UTP o similar • Tablas de especificaciones • Situaciones e instrumentos • Protocolo o procedimientos de aplicación • Reglamento de evaluación	• Entrevistas individuales y/o grupales • Cuestionarios • Matriz de información de evidencia documental	• Análisis de contenido • Análisis estadístico descriptivo • Análisis documental
		2.1.2 Coherencia interna en la planificación de la situación y/o instrumento			
		2.1.3 Presencia de procedimientos asociados al diseño			
	2.2 Aplicación de las situaciones e instrumentos	2.2.1 Existencia de protocolos de aplicación			
		2.2.2 Claridad de los procedimientos y/o protocolos de aplicación			

TABLA 1: Matriz de recogida de información (cont.)

OBJETO O NECESIDAD: ANALIZAR LAS PRÁCTICAS EVALUATIVAS DEL COLEGIO					
DIMENSIONES: 3. Decisiones evaluativas para el logro y mejora de los aprendizajes					
Definición operacional (cómo se entiende la dimensión)	**Subdimensiones**	**Criterio que caracteriza la dimensión**	**Fuentes de información**	**Técnicas o procedimientos de recogida de información**	**Técnicas de análisis de la información**
Refiere a la interpretación, juicio y toma de decisiones sobre los aprendizajes de los estudiantes	3.1 Interpretación y juicio acerca de los resultados de las evaluaciones	3.1.1 Tipos de análisis de resultados	• Docentes • Pautas de corrección • Trabajos revisados • Ejemplos de retroalimentación	• Entrevistas individuales y/o grupales • Cuestionarios • Matriz de información de evidencia documental	• Análisis de contenido • Análisis estadístico descriptivo • Análisis documental
		3.1.2 Tipos de juicios (cualitativos-cuantitativos)			
		3.1.3 Razones para la elección de los análisis y tipo de juicios realizados			
	3.2 Comunicación de los resultados de la evaluación	3.2.1 Forma de comunicación de los resultados	• Reglamento de evaluación		
		3.2.2 Determinación de los actores educativos informados			
	3.3 Toma de decisiones pedagógicas	3.3.1 Decisiones centradas en lo pedagógico			
		3.3.2 Decisiones centradas en lo administrativo			

Adaptado de Espinoza (2013)

Una vez que se ha planificado la recogida de información, se está en condiciones de diseñar los instrumentos para poder ser aplicados. En este proceso se deben respetar y cumplir los criterios de calidad

propios de la construcción, que se describen en la literatura especializada[2].

Tan importante como lo anterior es someter a juicio del experto (profesional especialista en el área temática y en la medición o evaluación de la misma) tanto la matriz de recogida de datos como los instrumentos construidos, antes de ser aplicados, para poder ajustar los aspectos que sean necesarios con el objeto de asegurar que la recogida de información sea desarrollada en un óptimo nivel de calidad.

En el momento de aplicar tales instrumentos o procedimientos es importante considerar aspectos éticos que resguarden la confidencialidad de la información y cuidar “no hacer daño” a quien entrega información tanto en el propio proceso de entrevista como en el manejo posterior de la información.

Al finalizar el proceso de recogida de información se realizarán los análisis de la misma a través de procedimientos estadísticos o técnicas cualitativas que permitan sistematizarla y organizarla adecuadamente. Para ello se sugiere considerar a especialistas en este tipo de análisis, en caso de que los equipos no cuenten con estas capacidades instaladas[3]. Ahora se está en condiciones de iniciar la configuración del caso.

Configuración del caso o encuadre

La configuración del caso propiamente tal consiste en un proceso que busca determinar la estructura y significado subyacente de los elementos constituyentes del caso, siguiendo la lógica de un relato histórico no en el sentido cronológico sino de orden y secuencia de sucesos que van contando una historia coherente (Simons, 2011). Como se señaló con anterioridad, tiene un sentido holístico que incorpora múltiples fuentes en un contexto real.

En la explicitación del caso se hacen visibles los puntos de vista opuestos, las dinámicas de cambio, éxitos y fracasos redactados en lenguaje accesible para que todos puedan entender de lo que se está hablando y se facilite el análisis posterior. Incluye observaciones,

incidentes y escenarios que faciliten al lector involucrarse en la situación (Farhoomand, 2004).

Es recomendable incorporar en el relato del caso citas textuales de las voces de los participantes, esto es, las propias palabras de las personas, señalando una característica genérica que las represente, como por ejemplo: profesor, estudiante, directivo, sin señalar el nombre de quien habla. También puede considerar información documental, de carácter oficial de los establecimientos o programas, debidamente citada.

La evidencia que se incluya en este apartado debe ser pertinente y focalizada, cumpliendo criterios tales como:

- Carácter descriptivo y explicativo, sin interpretación de los hechos: se describen los hechos objetivos y cómo fueron sucediendo, explicando desde la evidencia obtenida las decisiones y acciones tomadas, las razones que las personas señalan, sus percepciones, confianza y desconfianza, tal como se dan.
- Evidencia válida y confiable: la evidencia obtenida se encuentra libre de sesgo o error, se refiere a lo acontecido de manera verídica, no se debe al azar. Puede ser comprobada de distintas formas.
- Sin juicios u opiniones de sentido común: la información se apoya en evidencia, no es la mera opinión de una persona, sin fundamento.
- Sigue un hilo conductor lógico que permita tomar una línea argumentativa.
- La información es suficiente para el análisis.

Un ejemplo de narración de un caso se presenta a continuación:

Este año en la Prueba Internacional de Inglés a la que se somete el colegio, por primera vez en 8º año se obtienen buenos resultados, que dan cuenta de que la enseñanza efectivamente forma estudiantes

bilingües. Los padres y apoderados habían hecho muchos reclamos y la matrícula había ido en descenso hasta ahora. Luego de publicitar los resultados, la demanda para el siguiente año ya ha subido y al entrevistar a los padres, una de las razones que surge es esta mejora. Los directivos, asumiendo que el éxito en los resultados se debe a la labor que una profesora ha realizado, la eligieron como la mejor profesora del colegio (premio que se otorga anualmente)y le proponen que sea un modelo para sus colegas, especialmente en el área de evaluación.

Algunos de los aspectos destacados de la profesora es que alinea sus evaluaciones con las actividades que realizan en clase. Ella señala que las evaluaciones usadas son similares a las actividades que aparecen en los planes de clases.

Aunque todavía usa preguntas de opción múltiple, también incorpora muchas actividades auténticas para evaluar a los estudiantes, como matrices o guías de calificación (listas de chequeo, escalas de valoración, rúbricas holísticas y analíticas) para calificar los desempeños de los estudiantes y para retroalimentarlos, pero señala que esto le significa invertir mucho tiempo y que debe hacerlo en su casa. Lo bueno es que estas mismas guías las usa para hacer autoevaluación y coevaluación.

La profesora evalúa continuamente a sus estudiantes a través de preguntas informales en el aula y observándolos cuando trabajan en clase, diseña sus propios instrumentos de evaluación y adoptó unas matrices de calificación (rúbricas) que encontró en internet para calificar e interpretar los resultados de las evaluaciones y para retroalimentar a los estudiantes. Como estos instrumentos son de la web “educamejor”, ella confía a ojos cerrados en su validez. Se destaca que las rúbricas son en inglés, porque ella trabaja en un colegio bilingüe y aunque el nivel de los estudiantes en inglés no es muy alto, deben aprender a leerlas, esa es la normativa del colegio.

Un aspecto que la profesora menciona que ha incorporado cada vez con más fuerza es el uso de criterios de evaluación o indicadores de logro; en este sentido señala: “mis estudiantes saben claramente los objetivos de aprendizaje y las expectativas que se tienen. Esto los ayuda a guiar su proceso de aprendizaje”.

Pese a que los directivos del establecimiento se muestran entusiasmados con la oportunidad que significa los logros y la visión de la profesora, el reconocimiento le ha traído algunos problemas con sus colegas, que encuentran que es demasiado estructurada e individualista con sus

conocimientos. Pero como sus estudiantes están felices, a ella no le importan esos comentarios.

Después de unas semanas, la coordinadora académica del colegio inicia un proceso de observación de clases y va a la sala de la profesora modelo. Como resultado le informa algunos elementos que destacan y otros que se podrían mejorar. Entre los que sobresalen está el uso del material didáctico para fomentar la participación de los estudiantes y también el cierre que hace de su clase. Se presenta como una debilidad la falta de monitoreo permanente de los estudiantes, le señala que no es tan evidente y que debe intencionarlo a través de actividades de evaluación formativa más estructuradas y sistematizadas.

En una reunión con todos los profesores del primer y segundo ciclo básico, se presenta un documento síntesis elaborado con la experiencia sistematizada de la docencia de la profesora modelo. Las reacciones de los profesores son diversas, algunos lo consideran positivo, aunque les habría gustado que la forma de retroalimentar su propio trabajo hubiese sido diferente; señalan: “seguro ahora la evaluación será a la pinta de lo que ellos dicen”. Otros creen que no corresponde que les entreguen “modelos” cuando ellos han tenido buen Simce durante años y los papás no se quejan de lo que sus hijos aprenden. Existen dudas acerca de cómo y quién los evaluará y deciden hacer una reunión fuera del colegio para decidir con qué estrategia van a enfrentar las innovaciones y el procedimiento que se está instalando, en especial cuando no existen criterios claros para la observación de aula ni tampoco las consecuencias que tendrá una evaluación desfavorable en ese sentido.

Al finalizar la configuración del caso se sugiere responder a la siguiente pregunta:

¿La información presentada permite una comprensión acabada de la situación, sin dejar aspectos ambiguos o con falta de evidencias?

Si la respuesta a la pregunta anterior es afirmativa, se está en condiciones de avanzar a la siguiente etapa.

Segunda etapa: análisis del caso

Para desarrollar el análisis se debe poner en juego una serie de habilidades personales y colectivas de los equipos, tales como capacidad de reflexión, abstracción, síntesis, autocrítica, automonitoreo y pensamiento crítico.

Esta etapa del análisis considera cuatro desarrollos que contribuyen al proceso de descomposición propio del análisis:

1. Análisis FODA.
2. Preguntas que orientan la discusión.
3. Discusión de la evidencia con los aspectos conceptuales o referentes teóricos pertinentes al tema del análisis.
4. Definición del núcleo problemático o conflicto latente.

A continuación se describirá cada uno de estos aspectos con ejemplos:

a. Análisis FODA: este tipo de recurso analítico permite describir la situación interna y externa en la que se encuentra un caso. Es reconocido como positivo por quienes han utilizado esta metodología, ya que obliga a los participantes de los equipos no solo a poner el foco en las debilidades sino que también en las fortalezas y oportunidades, que siempre existen en toda realidad educativa.
    - Las fortalezas (F) son aquellos aspectos o hechos positivos, ventajas que se constituyen en recursos, habilidades e incluso actitudes de los participantes de los equipos. La formulación de las fortalezas debe cumplir con los siguientes criterios de calidad:
        1. Se levantan a partir de la información del caso, no se infieren.
        2. Se explicita por qué un hecho constituye una fortaleza, solo nombrarlo no es una fortaleza en sí misma, porque lo que puede ser positivo para alguien puede ser una debilidad para otros. Al

construirla es necesario plantear el hecho que se extrae del caso y justificar por qué ese hecho puntual constituye una fortaleza.

3. Debe existir coherencia entre el hecho que se plantea y su justificación, es decir, no se puede justificar planteando otra fortaleza o hacer una reflexión sobre la misma.

Ejemplo de fortaleza:

Existen variados instrumentos de evaluación que permiten recoger distintos tipos de evidencia de los aprendizajes de los estudiantes, favoreciendo así la objetividad de los juicios que emiten los docentes acerca de los niveles de logro alcanzados.

Es clave en este punto que los elementos seleccionados deben estar ocurriendo en el contexto real del caso.

- Las oportunidades (O) son aquellos aspectos favorables tanto a nivel interno del caso como del entorno ambiental o contexto que podrían ser aprovechados para mantener o mejorar las condiciones del mismo. Constituyen hechos o escenarios que, de seguir presentes en el futuro, resultarían positivos para el caso en cuestión.

La formulación de las oportunidades debe cumplir con los siguientes criterios de calidad:

1. Se levantan desde la información del caso, pero realizando una inferencia hacia el futuro.
2. Se explicita por qué constituyen una oportunidad, el solo nombrar el hecho no es un beneficio en sí mismo, se debe desarrollar la idea y justificar por qué ese hecho puntual constituye una oportunidad en ese contexto y en ese momento. Debe ser evidente que no ha sido considerado como un aspecto positivo en la problemática, porque sería una fortaleza.

3. Debe existir coherencia entre el hecho que se plantea y su justificación, esto es, no se puede justificar planteando otra oportunidad o hacer una reflexión de la misma.

Ejemplo de oportunidad:

La existencia en la actualidad de nuevas iniciativas realizadas desde el municipio (sostenedor) en torno a la evaluación de los aprendizajes puede constituirse en una oportunidad de recursos vía proyectos orientados a apoyar a los docentes en evaluación.

La experiencia que el establecimiento ha adquirido respecto del uso de datos de pruebas estandarizadas es una oportunidad para desarrollar una gestión de resultados efectiva, permitiendo diseñar e implementar acciones coordinadas que apunten a una mejora cualitativa en los aprendizajes que los estudiantes son capaces de alcanzar.

- Las debilidades (D) son aquellos aspectos “negativos”, áreas o puntos débiles que interfieren en el logro de los objetivos de la organización o de las personas involucradas. Constituyen barreras que impiden avanzar o mejorar.

La formulación de las debilidades debe cumplir con los siguientes criterios de calidad:

1. Se levantan de la información del caso, no se infieren.
2. Se explicita por qué un hecho constituye una debilidad, solo nombrarlo no es una debilidad en sí misma, porque lo que puede ser una debilidad para alguien puede ser una amenaza o fortaleza para otros. Al construir una debilidad es necesario plantear el hecho que se extrae del caso y justificar por qué ese hecho puntual constituye un aspecto negativo o perjudicial.
3. Debe existir coherencia entre el hecho que se plantea y su justificación, es decir, no se puede justificar planteando otra debilidad o hacer una reflexión de la misma.

Ejemplo de debilidad:

Los docentes cuentan con una visión castigadora de la evaluación, que genera un clima atemorizante en los estudiantes, afectando sus ganas de asistir a clases y de aprender.

Es clave en este punto que los elementos seleccionados deben estar ocurriendo en el contexto real del caso.

- Las amenazas (A) corresponden a aspectos desfavorables que se generan en torno al caso y que podrían atentar en contra de la situación, poner en riesgo o agravar los problemas. Son amenazas que pueden afectar adversamente al caso. Dado esto, es necesario disminuirlas o hacerlas desaparecer, al igual que las debilidades.

La formulación de las amenazas debe cumplir con los siguientes criterios de calidad:

1. Se levantan de la información del caso, infiriendo razonablemente sus consecuencias. Es decir, no se inventan, sino que surgen a partir de la evidencia conocida.
2. Se explicita por qué constituyen una amenaza, solo nombrar el hecho no constituye un perjuicio en sí mismo, es necesario desarrollar la idea y justificar por qué ese hecho puntual es una amenaza en el contexto del caso. Debe ser evidente que no ha sido considerado como un aspecto negativo en la problemática, porque sería una debilidad.
3. Debe existir coherencia entre el hecho que se plantea y su justificación, esto es, no se puede justificar planteando otra amenaza o hacer una reflexión de la misma.

Ejemplo de amenaza:

De continuar el establecimiento utilizando la información de los resultados de la evaluación como una medida de comparación y competencia entre los

estudiantes, se pondrá en riesgo la calidad de los aprendizajes de los mismos, deteriorando el proceso formativo.

a. Preguntas que orientan la discusión: una vez finalizado el FODA se plantea una serie de preguntas que ayudarán a realizar la discusión posterior. Se describen como un medio para generar un cuestionamiento o reflexión crítica de la evidencia previamente recogida. Para ello, las preguntas deben permitir ir desde lo concreto y evidente hacia lo latente y abstracto, que no siempre se encuentra descrito, sino que es descubierto por el analista y su equipo, dada la experiencia y dominio teórico del que disponen.

   Tales preguntas deben cumplir con los siguientes criterios:

   - La respuesta a la pregunta supone una elaboración compleja que no se responde con un Sí o No.
   - Deben responderse con la información presentada en el caso.
   - En su redacción, deben estimular una reflexión crítica y no ser solo una descripción del caso, tienen que ser tensionantes para el análisis.

   Sin el propósito de intentar mecanizar su construcción sino que solamente orientar la misma, se recomienda construir las preguntas centrándose en ¿qué?, ¿cómo? y ¿por qué?, poniendo el énfasis en un foco más explicativo que descriptivo de los fenómenos (Yin, 2009), tal como se presenta en los siguientes ejemplos:

   1. ¿Cómo se justifican las decisiones de los profesores en cuanto a los instrumentos de evaluación seleccionados y aplicados en el aula?
   2. ¿Cómo es la relación entre planificación e implementación de la estrategia evaluativa utilizada por el profesor para el contexto de los niños del curso?
   3. ¿Por qué las decisiones evaluativas tomadas por una directora, a pesar de estar centradas en resultados y no

considerar procesos, pueden mejorar los rendimientos de los estudiantes generando la satisfacción de profesores y alumnos?

4. ¿Qué aspectos de la cultura de la escuela pudieron favorecer el cambio evaluativo instalado por la dirección y frenar las resistencias propias de un proceso de cambio?
5. ¿En qué medida las decisiones tomadas por parte de la profesora dan cuenta de una estrategia planificada de evaluación de aprendizajes, o es solo una secuencia de actividades por cumplir para los niños del nivel?
6. ¿Cómo se explican los efectos que la evaluación estandarizada implementada en el colegio han provocado en la cultura escolar y en la dinámica de las decisiones estratégicas en torno al desarrollo de los aprendizajes de los estudiantes?
7. ¿Por qué el reconocimiento por parte de un establecimiento educacional a una profesora como "modelo en evaluación" genera tensiones en sus colegas y resistencia a ser evaluados?

b. Análisis propiamente tal o discusión: analizar es una habilidad cognitiva de jerarquía superior, que consiste en desmenuzar un objeto de estudio en busca de una comprensión o explicación del mismo en su contexto particular.
   En este proceso de análisis se ponen en juego, con mayor énfasis, dos habilidades cognitivas superiores: analizar y evaluar.

   Recordando a Anderson y Krathwohl (2001), la habilidad de analizar comprende descomponer un fenómeno o situación entre sus partes y determinar de qué manera se relacionan unas con otras. Este proceso a su vez implica diferenciar bajo distintas categorías los elementos implicados en el fenómeno, por ejemplo, lo relevante de lo menos relevante, lo característico de lo no característico, lo frecuente de lo menos frecuente, lo crítico de lo no crítico. Luego, organizar, es decir, volver a estructurar los hechos descritos en una nueva explicación de las relaciones

que se dan en el fenómeno o situación, dando cuenta de las tensiones o puntos de vista opuestos. Y atribuir, esto es, determinar lo esencial o el foco de la problemática o situación analizada. En este mismo sentido, en un proceso de análisis de caso, se evalúa la información obtenida, es decir, se emite juicio basado en evidencia, determinando un problema a intervenir o los reales efectos, consecuencias o resultados de las decisiones tomadas.

Este es el momento crítico de la metodología, todas las actividades y etapas anteriores fueron desarrolladas para este punto, en que se ponen en relación la evidencia recogida, las preguntas orientadoras, el FODA y los aspectos teóricos conceptuales para que, en una reflexión crítica, se discutan los elementos del caso. En el relato debiesen aparecer las tensiones, conflictos, decisiones, resultados de éxito y fracaso, entre otros ya mencionados. Pero no solo estos aspectos, la teoría ocupa un rol importante en esta fase de análisis y en las anteriores para que cada fenómeno sea entendido en su complejidad. Se debe comprender cómo ha sido abordado el tema de análisis en la literatura, sus características, sus componentes, cómo se ha explicado el fenómeno en el campo educativo y sus relaciones conceptuales.

El análisis deberá ser redactado como un relato descriptivo, integrado, fundamentado y apoyado en aspectos teóricos que asegurarán que la discusión vaya más allá de la propia experiencia, de la situación puntual o de la mera opinión o juicio sin base. Se sugiere describir el caso respondiendo las preguntas planteadas con anterioridad, e integrando los aspectos del FODA en la respuesta de manera parsimoniosa. Para fundamentar el análisis se debe apoyar en la literatura las explicaciones, las relaciones y los juicios.

Un ejemplo de un texto de discusión se presenta a continuación:

Retomemos la pregunta 7 anteriormente presentada en los ejemplos de preguntas:

¿Por qué el reconocimiento por parte de un establecimiento educacional a una profesora como “modelo en evaluación” genera tensiones en sus colegas y resistencia a ser evaluados?

Análisis propiamente tal o discusión:

A partir de la información o evidencia del caso expuesto se puede observar el énfasis que la organización escolar a través de sus directivos ha puesto en los resultados, más que en el proceso de aprendizaje de los estudiantes (Santos-Guerra, 1995) y cómo este énfasis se conecta también con la demanda de los padres, quienes exigen resultados para validar la calidad del establecimiento. Los resultados de la prueba Simce y de la Prueba Internacional de Inglés a la que se sometió el colegio son muestra de ello.

Se plantea que existe una preocupación por la evaluación, el seguimiento y el monitoreo de los aprendizaje, a partir, por ejemplo, de la observación de la coordinadora académica, sin embargo, la presión por parte de los directivos a los profesores es por los resultados de los niños y niñas. Se carece de un análisis profundo de los aprendizajes logrados por los estudiantes para orientar así la toma de decisiones, poniendo la responsabilidad del modelaje en una profesora del equipo, corriendo el riesgo de desaprovechar miradas más amplias sobre los procesos de aprendizaje de los estudiantes (Brown, 2004; Dylan, 2009).

Presentar a una profesora como modelo de enseñanza y evaluación en el contexto de la organización es visto de manera confusa y preocupante por parte de los docentes, quienes desconocen los criterios y los antecedentes bajo los cuales se le atribuyó este reconocimiento a la profesora. La única evidencia que parece visible son los resultados de una evaluación.

De acuerdo con Villalón (2006), se reconoce la evaluación como uno de los elementos clave del proceso de enseñanza/aprendizaje, por el volumen de información que facilita al profesor y por las consecuencias que tiene para el docente, el alumnado, el sistema educativo en que está integrado y la sociedad. Sin embargo, no es lo único, la gestión curricular, el liderazgo educativo y el clima son factores relevantes para asegurar el éxito de los propósitos de los colegios, por lo que se espera que los establecimientos desarrollen estrategias sistémicas para enfrentar los desafíos de mejora continua y no el uso de modelajes que si en efecto no resultan ser paradigmáticos, pueden provocar sentimientos de enojo, frustración, competencia, envidia, entre otros, todo lo cual amenaza la disposición de los profesores para involucrarse en procesos de mejora.

c. Definición del núcleo problemático o conflicto latente

Al finalizar el análisis se debe describir el foco problemático del caso, aquello que es necesario enfrentar, mejorar, reparar o resolver. O, por el contrario, la explicación principal del éxito. El foco problemático no refiere a aspectos evidentes que se han descrito en el caso, no es una debilidad relevada, o una secuencia de hechos, sino que corresponde a aquello que se encuentra latente en el caso y que constituye su explicación principal. Dicha explicación refleja la mirada del experto, quien en su definición integra la experiencia, la evidencia y su manejo conceptual. Así también, el foco deberá estar acompañado de la descripción de los actores que participan y las consecuencias que provoca el problema detectado en forma precisa. La redacción del problema se presenta de manera separada del análisis y a continuación de este.

El problema latente o foco central que corresponde al ejemplo presentado es:

El foco central de la discusión y el problema a abordar en el establecimiento dice relación con una visión restringida de la evaluación y sus efectos en el colegio por parte de directivos y algunos profesores, centrada en la certificación y la dimensión sumativa por sobre una visión de proceso y de carácter formativo. Esta concepción restringida atraviesa las decisiones de gestión y propuestas de acción, trayendo efectos negativos en el clima y en la disposición de los equipos de profesores.

Cuarta etapa: propuesta de mejora

La propuesta de mejora es una alternativa que presenta el analista para enfrentar la problemática del caso o asegurar la sustentabilidad de sus efectos positivos. Se traduce en una matriz con diferentes componentes que orientan al lector acerca del curso de acción de dicha propuesta.

Las organizaciones valoran que un proceso de análisis no solo quede en plantear una problemática sino que, por el contrario, se oriente la toma decisiones futuras con acciones concretas y organizadas para asegurar su implementación.

En este proceso de construcción se ponen en juego dos habilidades cognitivas de importancia para el trabajo de los equipos: evaluar y crear.

Al evaluar, los equipos, junto al especialista que asesora (interno o externo), determinan los mejores caminos para abordar la solución de un problema o la mejora sostenible en el tiempo a partir de evidencia obtenida en el análisis de caso. El proceso de crear se pone en juego al diseñar y planificar un camino elegido, asegurando que este responda al propósito para el cual fue pensado (Anderson, 2001).

Esta propuesta debe responder a los siguientes criterios de calidad:

1. Factibilidad: que se pueda llevar a cabo en el establecimiento o aula determinados.
2. Validez: que corresponda al problema central del análisis y no a uno colateral o inexistente.

3. Novedad: que no haya sido implementada en forma parcial o total con anterioridad. Por lo tanto, se requiere indagar en la historia del establecimiento y que sea diseñada para este caso, en su contexto particular.

Los componentes de una propuesta de mejora de calidad y las preguntas que se busca responder con estos son:

a. Propósito: ¿Qué busca solucionar la propuesta?, ¿cuáles son los objetivos de la propuesta?
    - Este apartado tiene por finalidad establecer lo que se busca como resultado de la propuesta. Es decir, aquí se plantean los objetivos de la misma.
    - Los objetivos orientan las acciones a seguir para lograr el propósito.
b. Actividades: ¿Qué actividades permiten el logro de los objetivos trazados para la etapa?, ¿su nivel de complejidad hace posible alcanzar la conducta deseada?
    - Deben ser acordes con los objetivos planteados.
    - Puede haber más de una acción por objetivo.
    - Deben ser factibles de realizar en el contexto específico del caso.
c. Medios de verificación: ¿Cuál es la evidencia que puede testimoniar la realización de las acciones planificadas?
    - Tiene que ver con la fidelidad de la implementación de la propuesta.
    - Son evidencias concretas que muestren que la actividad se realizará y la calidad de esa realización.
d. Recursos: ¿Qué medios materiales se requieren para el desarrollo de las actividades?
    - Corresponde a una lista de recursos que se requieren para llevar a cabo la actividad propuesta.
e. Responsables: ¿Qué actores involucrados en la situación problemática lideran la realización de la actividad?

- Es fundamental para la realización de cada acción que se defina una persona que sea la encargada de gestionarla, y que sea responsable de su implementación.

f. Indicadores: ¿Cómo podemos darnos cuenta de que se logra el propósito definido?
   - Aspectos medibles u observables que se tomarán en cuenta para evaluar el nivel de logro de las acciones propuestas. Por lo tanto, se describen indicadores de resultado.
   - Cada indicador (puede haber más de uno para cada acción) establece en forma concreta algún comportamiento o aspecto que debe ser observado.
   - Tienen relación directa con el propósito establecido para el componente estudiado.

Es importante que todos estos aspectos sean coherentes entre sí y respondan al problema o problemática central del caso.

La propuesta de mejora se presenta en formato de tabla con el propósito de asegurar la coherencia entre cada uno de sus aspectos (tabla 2).

Es muy importante la participación de los miembros de la organización en el diseño de la propuesta, para fidelizar y lograr la adhesión de la comunidad toda o la mayoría de esta en su implementación. Lo peor que podría pasar es que después de todo el esfuerzo del análisis y el diseño de la propuesta, esta se quede en el papel o en el escritorio de algún directivo del establecimiento.

TABLA 2: Formato de presentación de propuesta de mejora

Propósitos/objetivos	Actividades	Medios de verificación	Recursos	Responsables	Indicadores[4]
Se recomiendan no más de tres objetivos, ya que la propuesta debe ser acotada	Una o varias actividades para cada	Evidencias concretas de que las actividades	Corresponde a los recursos necesarios para lograr	Son personas específicas que tienen a cargo el logro de los objetivos, por lo	Corresponde a un desglose en metas o propósitos

para que sea factible	objetivo	fueron realizadas	cada objetivo	tanto no se refiere a la máxima autoridad del establecimiento	específicos asociados a cada objetivo

CONSIDERACIONES RESPECTO DE LOS ASPECTOS ÉTICOS DEL ANÁLISIS DE CASO

El análisis de caso asume como valor principal la dignidad humana y su integridad. A los participantes se les debe tratar con equidad, promoviendo el diálogo, la comprensión mutua, el respeto, la confianza y la confidencialidad. El analista (interno o externo) debe asegurar que las personas no se sientan vulneradas en sus derechos o susceptibles de sufrir algún daño y menos aún, provocarlo (House & Howe, 2001; Simons, 2011).

Si surgen dilemas éticos, se sugiere recurrir a los principios y prácticas éticas acordadas como comunidad o establecidas en el establecimiento; si estas no existieran, acordar en conjunto la manera de proceder apelando a valores y principios universales o a una ética relacional y situada. Estos dilemas suelen surgir a la hora de poner en juego intereses y posiciones de poder, por lo que desde un comienzo se deben explorar las tensiones posibles en el proceso, reconociendo cuál es la postura ética del analista y de la institución, programa, proyectos y/o personas involucradas en el caso (House & Howe, 2001; Simons, 2011).

Una opción que puede ayudar es situarse en una ética democrática (House & Howe, 2001) que reconoce y valora los principios de la imparcialidad, la justicia y la equidad, junto con la confidencialidad, la negociación y la accesibilidad como procesos éticos de acción. Acordar quiénes son los propietarios de los datos extraídos es una decisión que puede evitar conflictos posteriores.

Reconocer y tomar acciones a tiempo de cómo proceder éticamente a nivel metodológico también es importante, en especial, en el proceso de recolección de datos para la configuración del caso, en aspectos tales como: utilizar consentimientos informados, respeto en el proceso de entrevistas, confidencialidad y anonimato de información delicada,

ya sea extraída de documentación o relativa a personas (Danzin & Lincoln, 2011).

El informe final y las decisiones que de él emanen puede ser también una fuente de dilemas éticos, por lo que es necesario generar instancias de diálogo y reflexión para no perder de vista a las personas que puedan resultar involucradas (Simons, 2011).

INFORME FINAL DEL CASO

La escritura del caso involucra esfuerzos complejos y reflexivos. El relato realista del caso se describe siguiendo la cronología y/o biografía de este, los temas y las soluciones, más que problemas o hipótesis (Stake, 1998). Se consideran citas textuales de las voces de los actores, información gráfica, tablas o esquemas que organizan la información estadística o documental.

Escribir cualquier informe comprende cuatro macroactividades: planificación, organización, redacción y revisión (Lane, 2007). En este trabajo es importante describir el caso en función de sus etapas de desarrollo, cuidar que la redacción sea clara y, para ello, una buena estrategia es pedirle a otro que lo lea y comente las posibles mejoras en la precisión y redacción.

La formalidades consideran escritura clara, en tercera persona, destacando la información relevante, manteniendo un sentido lógico, el cual se puede estructurar como relatos de sucesos acontecidos, áreas temáticas o procesos abordados en la organización, que ayude al lector a comprender cómo fueron dándose los hechos y el análisis.

LIMITACIONES

La metodología de análisis de caso para la toma de decisiones, si bien es una herramienta muy efectiva para lograr orientar decisiones basadas en evidencia y reflexión crítica, no deja de tener algunas limitaciones que es importante destacar:

- Circunscrito al tiempo específico de la observación y al relato de los participantes, por lo que no se puede extrapolar a otras

etapas no vividas, sean futuras o pasadas.

- Imposibilidad de generalizar a nivel estadístico, dada la singularidad y particularidad del caso, y el contexto en que se da, no puede ser utilizado para generalizar e inferir otros contextos desde una lógica cuantitativa.

EJEMPLOS DE APLICACIÓN DEL MÉTODO DE ANÁLISIS DE CASO PARA LA TOMA DE DECISIONES:

Espinoza, L. (2013). *Análisis de caso de las prácticas evaluativas de las docentes de lenguaje y comunicación de enseñanza básica en un colegio bilingüe*. Tesis de magíster en Educación mención Evaluación de Aprendizajes. Facultad de Educación, Pontificia Universidad Católica de Chile.

Gaete, J. (2016). *Las prácticas de uso de la información que realiza el equipo directivo y docentes de lenguaje y matemática de segundo ciclo básico de un colegio particular pagado, a partir del Sistema de Evaluación de Progreso del Aprendizaje (SEPA)*. Tesis de magíster en Educación mención Evaluación de Aprendizajes. Facultad de Educación, Pontificia Universidad Católica de Chile.

EXPERIENCIA DE PROFESORES QUE HAN UTILIZADO ESTA METODOLOGÍA Y COMENTARIOS DE CADA UNA DE SUS PARTES

De la metodología en general

“Es una buena metodología, permite desarrollar las etapas siguiendo un orden lógico y creciendo en complejidad” (profesora de Lenguaje).

“La metodología me permitió hacer un análisis de manera ordenada y sistemática en la medida que me permitió ir paso a paso construyendo el caso desde el contexto y su configuración, preguntas y FODA hasta llegar al análisis propiamente tal y proponiendo una mejora o solución al problema” (educadora de párvulos).

“La metodología del análisis de casos es una estrategia y una herramienta muy útil en lo profesional; nos permite contar con un modelo para identificar problemáticas asociadas a la evaluación de los aprendizajes, analizar un caso uniendo la teoría y las prácticas evaluativas y, conjuntamente con eso, nos facilita la creación de propuestas de mejora” (educadora de párvulos).

“La estructura es muy favorable, debido a que es una herramienta totalmente viable de utilizar para identificar los nudos críticos de algún proceso en cualquier organización y, así, tomar decisiones para la mejora Para mí esta metodología es una forma de trabajo como más práctica y viable de implementar, ya que los cambios son viables, pequeños pero importantes y nacen de sus propios análisis. Además que las respuestas están en ellos mismos o dentro de la organización. Pocas veces hay que recurrir a buscarlas afuera, las propuestas están en los propios participantes, por lo cual este método les permite, en conjunto, analizar las situaciones problemáticas y darles soluciones” (director de colegio).

Del contexto de caso

“Es importante para poder entender en qué se enmarcaba el problema y poder pensar en una solución viable” (profesora de Lenguaje).

“Ayuda en la medida que permite situarnos en la realidad, tal cual se da el problema en la comunidad educativa a analizar. Permite visualizar otras variables que están interviniendo en el problema, que quizás no son determinantes pero que sí están presentes y son dignas de considerar a la hora del análisis” (educadora de párvulos).

“Creo que sirve para ir dando una orientación al resto de los pasos de la metodología. No obstante, creo que hay que dedicar más al FODA, contrastación evidencia/teoría y, desde luego, a la propuesta de mejora” (director de colegio).

De la configuración del caso

“Considero que el aspecto más fácil en la configuración de un caso es la organización de la información: decidir qué va primero o qué va después. Uno de los aspectos más complejos de la configuración del caso es el referido a los supuestos o inferencias; quitar los juicios al redactar la configuración del caso es un tanto complejo” (educadora de párvulos).

“En esta etapa es difícil dejar de lado los juicios personales cuando se es miembro del establecimiento, pero cuando se guía el análisis, como consultor, es más fácil” (director de colegio).

De la realización del FODA y la definición de preguntas orientadoras

“El FODA es una buena estrategia para iniciar el análisis, ayuda a pensar no solo en las debilidades, sino también en los aspectos positivos que siempre cuesta ver porque uno siempre se va a lo malo de las cosas. Las preguntas permiten levantar los problemas realmente relevantes y no caer en los detalles o en problemas derivados del problema real. De todos modos no es fácil construir el FODA ni las preguntas y se requiere mucho apoyo de las especialistas” (profesora de Lenguaje).

“El aporte del FODA viene dado porque permite identificar las debilidades o amenazas, aspectos que ayudan en el análisis del caso y en la propuesta de mejora (esto es lo que no está tan bien o lo que se debe mejorar/cambiar). Por otra parte, las fortalezas y oportunidades permiten resaltar los aspectos positivos con que cuenta la comunidad para hacer frente al problema, y también porque les da un mensaje esperanzador a quienes leen el informe: ‘no todo está tan mal en nuestra realidad’ u ‘¡oh!, esto no lo había pensado como una oportunidad’” (educadora de párvulos).

“Las preguntas orientadoras me permitieron conectar el FODA con el análisis, en la medida que las respondiera en virtud de los aspectos centrales del FODA mientras iba analizando el caso” (educadora de párvulos).

Es una herramienta fundamental, pues desde ahí se saca el material para la discusión/tensión" (director de colegio).

"

Del análisis del caso y la discusión

"El análisis propiamente tal me obligó a poner en acción mi capacidad reflexiva, analítica, de síntesis, de evaluación, descripción y redacción" (profesora de Lenguaje).

"El proceso me permitió desplegar habilidades relacionadas a la inferencia: al tener que detectar el problema; capacidad de análisis: al 'desglosar' el problema en el análisis; capacidad de síntesis: al unir todos los aspectos para realizar una propuesta de mejora; capacidad de redacción: al tener que escribir de manera clara y precisa la configuración, el FODA, la definición del problema y la propuesta de mejora. Capacidad de relacionar aspectos: al tener que vincular todas las secciones. En cuanto a conocimientos: todos los conocimientos adquiridos hasta ahora en diferentes instancias de formación" (educadora de párvulos).

"En esta etapa debo saber bien la teoría antes de guiar los análisis, tener actualizados mis conocimientos del tema de fondo, ya que muchas veces los participantes no tienen tiempo de leer" (director de colegio).

De la estrategia de mejora

"Creo que es la parte más importante del caso, ya que es lo que se busca con todo el trabajo previo" (profesora de Lenguaje).

"Es importante, puesto que el análisis de casos no solo se queda con 'este es el problema: esto está bien o esto no está tan bien por estas razones', sino que se da un paso más allá y se crea una estrategia que permite dar solución a lo anteriormente planteado. No solo se discute una situación, sino que además hay un real aporte al sugerir qué debería

hacerse y cómo” (educadora de párvulos).

“Es lo fundamental, pues la propuesta de mejora es el objetivo del análisis, ayuda a mejorar los procesos y a obtener buenos resultados. Sobre todo porque en las instancias de reflexión, ellos están recibiendo *feedback* inmediato de sus equipos de trabajo” (director de colegio).

EN SÍNTESIS, como hemos visto en el transcurso de este libro, la evaluación de y para el aprendizaje constituye una herramienta muy poderosa que los docentes podemos usar para desarrollar todo el potencial de nuestros estudiantes y motivarlos a conocer e integrar los conocimientos de las distintas asignaturas. Sin embargo, también puede ser un arma devastadora para ellos si no tenemos presentes los principios éticos y los criterios de calidad en la construcción de los instrumentos, debido a las consecuencias que estas faltas pueden tener.

En nuestro recorrido relevamos la importancia que tiene la alfabetización evaluativa para nuestro desarrollo profesional y la calidad de nuestra docencia. Si hacemos evaluaciones deficientes recogeremos evidencia de mala calidad y nuestras decisiones pedagógicas serán erradas. La invitación es a planificar y definir qué es lo que esperamos que nuestros estudiantes aprendan y realizar tareas de enseñanza y evaluación que conduzcan a esas metas, a involucrar a los estudiantes en su propio proceso de aprendizaje a través de la autoevaluación y evaluación de pares, a tener espacios planificados en nuestras clases para dar retroalimentación efectiva y oportuna, a diversificar las formas en que recogemos las evidencias de aprendizaje considerando la riqueza de la diversidad en el aula, y a trabajar con otros miembros de la comunidad escolar para perfeccionar nuestras prácticas evaluativas y para trabajar necesidades institucionales generando propuestas de mejora que sean viables y efectivas.

BIBLIOGRAFÍA

Anderson, L. W. & Krathwohl, D. R. (Eds.) (2001). *A taxonomy for learning, teaching and assessing: a revision of Bloom's taxonomy*. New York: Longman Publishing.

Bambrick-Santoyo, P. (2010). *Driven by data. A practical guide to improve instruction*.USA: Jossey-Bass.

Boudett, K. P., City, E. A. & Murnane, R. J. (Eds.). (2005). *Data wise: a step-by-step guide to using assessment results to improve teaching and learning*. Cambridge: Harvard Education Press.

Brooke, S. (2006). Using the case method to teach online classes: promoting socratic dialogue and critical thinking skills. *International Journal of Teaching and Learning in Higher Education*, *18*(2), 142-149.

Center for Teaching and Learning Stanford University. (1994). Teaching with case studies, *Stanford University Newsletter on Teaching*, *5*(2), 1-4.

Chion, A. (2013). Estudios de caso en la enseñanza de la biología y en la educación para la salud en la escuela media. *Bio-grafía*, *6*(10), 42-49.

Coll, C., Mauri, T. & Onrubia, J. (2006). Análisis y resolución de casos-problema mediante el aprendizaje colaborativo. *Revista de Universidad y Sociedad del Conocimiento*, *3*(2), 29-41.

Coller, X. (2000). *Estudio de casos*. Madrid: Centro de Investigaciones Sociológicas.

Corbetta, P. (2003). *Metodología y técnicas de investigación social*. Madrid: McGraw Hill.

Covarrubias-Papahiu, P. (2016). Representaciones docentes de la educación basada en competencias. Un estudio de caso. *Propósitos y Representaciones: Revista de Psicología Educativa*, *4*(2), 73-132.

Davis, C. & Wilcock, E. (2003). Developing, implementing and evaluating case studies in material science. *European Journal of Engineering Education*, *30*(1), 59-69.

Denzin, N. & Lincoln, Y. (2011). *The Sage handbook of qualitative research*. California: Sage Publications.

Ebbeler, J., Poortman, C., Schildkamp, K. & Pieters, J. (2017). The effects of a data use intervention on educators' satisfaction and data literacy. *Educational Assessment, Evaluation and Accountability*, *29*(1), 83-105.

Edelstein, M., Wallensten, A. Kühlmann-Berenzon, S. (2014). Is case-chaos methodology an appropriate alternative to conventional case-control studies for investigating outbreaks? *Oxford Academic: American Journal Epidemiology*, *180*(4), 406-411.

Ellet, W. (2007). *The case study handbook: how to read, discuss and write persuasively about cases*. Boston: Harvard Business School Press.

Espinoza, L. (2013). *Análisis de caso de las prácticas evaluativas de las docentes de lenguaje y comunicación de enseñanza básica en un colegio bilingüe*. Tesis de magíster

en Educación mención Evaluación de Aprendizajes. Facultad de Educación, Pontificia Universidad Católica de Chile.

Farhoomand, A. (2004). Writing teaching cases: a quick reference guide. *Comunications of the Association for Information Systems*, *13*, 103-107.

Flyvbjerg, B. (2004). Cinco malentendidos acerca de la investigación mediante estudios de caso. *Revista Española de Investigaciones Sociológicas*, *106*(1), 33-62.

Freeman, C. (1997). What is a case? Bringing to science education the established teaching tool of law and medicine. *Journal of College Science Teaching*, *27*(2), 92-94.

Gaete, J. (2016). *Las prácticas de uso de la información que realiza el equipo directivo y docentes de lenguaje y matemática de segundo ciclo básico de un colegio particular pagado, a partir del Sistema de Evaluación de Progreso del Aprendizaje (SEPA).* Tesis de magíster en Educación mención Evaluación de Aprendizajes. Facultad de Educación, Pontificia Universidad Católica de Chile.

Golic, V., Boyer, M., Franko, P. & Lamy, S. (2000). *The ABCs of case teaching: pew case studies in international affairs*. Georgetown University: Institute for the Study of Diplomacy.

Goyas, L. & Monzón, Y. (2016). Análisis de casos como modalidad de titulación en la carrera de Derecho de la Universidad Metropolitana del Ecuador. *Revista Conrado*, *12*(53), 125-130.

Guizardi, M. & Garcés, A. (2014). Estudios de caso de la migración peruana “en Chile”: un análisis crítico de las distorsiones de representación y representatividad en los recortes espaciales. *Revista de Geografía Norte Grande*, *58*, 223-240.

Hernández, R., Fernández, C. & Baptista, P. (2006). *Metodología de la investigación*. México: Mc Graw Hill.

House, E. & Howe, K. (2001). *Valores en evaluación e investigación social*. Madrid: Morata.

Instituto Tecnológico de Monterrey (s. f.). El estudio de casos como técnica didáctica. *Las Estrategias y Técnicas Didácticas en el Rediseño*. Dirección de Investigación y Desarrollo Educativo, Vicerrectoría Académica. Recuperado de http://sitios.itesm.mx/va/dide2/tecnicas_didacticas/casos/casos.pdf

Lane, J. (2007). *Case writing guide*. Schreyer Institute for Teaching Excellence. Pensilvania State. Recuperado de https://www.schreyerinstitute.psu.edu/pdf/CaseWritingGuide.pdf

Martinic, S. (1997). *Diseño y evaluación de proyectos sociales*. México DF: Comexani.

Mella, O. (2003). *Metodología cualitativa en Ciencias Sociales y Educación*. Chile: Primus.

Mendoza, A. (2006). *El estudio de casos: un enfoque cognitivo*. México: Editorial Trillas.

Mertens, D. (2015). *Research and evaluation in education and psychology, integrating diversity with quantitative, qualitative and mixed methods*. California: Sage.

Morra, L. & Friedlander, A. (2001). *Evaluaciones mediante estudios de caso*. Washington DC: Departamento de Evaluación de Operaciones del Banco Mundial.

Ottenberger, A. (2000). *El estudio de casos en la investigación social*. Chile: Universidad Tecnológica Metropolitana.

Parker-Boudett, K., City, E. A., Murnane, R. J. & Steele, J. L. (2006). *El proceso de mejoramiento y el liderazgo escolar haciendo "Uso Sabio de Datos"*, PREAL. Recuperado de https://es.scribd.com/document/351762748/Proceso-Mejoramiento-Liderazgo-Escolar

Ragin, C. (1992). Introduction: cases of "what is a case?". In C. Ragin & H. S. Becker, *What is a case? Exploring the Foundations of Social Inquiry* (pp. 1-15). Cambridge: Cambridge University Press.

Rambla, X. (2013). Las complejas geografías de la política educativa: tres estudios de caso. *Educação & Sociedade*, *34*(125), 1229-1249.

Richardson, A. (2013). *Effective case analysis: techniques for success in case-based learning an examinations*. Concord: Captus Press.

Ritchey, F. (2006). *Estadística para las Ciencias Sociales*. México: Mc Graw Hill.

Rivera, J. (Ed.) (2000). *Casos de empresas.* Ediciones Universidad Católica de Chile (material de clases). Santiago de Chile.

Rusque, A. & Castillo, G. (2009). *Método de caso, su construcción y animación*. Santiago: Universidad de Santiago de Chile.

Santos-Guerra, M. (2001). *La escuela que aprende*. Madrid: Morata.

Servicio de Innovación Educativa (2008). *El método del caso*. Universidad Politécnica de Madrid. Recuperado de http://innovacioneducativa.upm.es/guias/MdC-guia.pdf

Simons, H. (2011). *El estudio de caso: teoría y práctica*. Madrid: Morata.

Stake, R. (1998). *Investigación con estudios de casos*. Madrid: Morata.

Yin, R. K. (2009). *Case study research*. London: Sage.

[1] La técnica de triangulación se explica en Stake, 1998, pp. 94-101.

[2] Existe una variedad de manuales de investigación educativa en los que se describen los elementos que se deben considerar en la construcción de instrumentos o técnicas de recogida de información. También es posible encontrar esta información en textos de metodología de investigación social, tales como: Hernández, Fernández y Baptista (2006), Mella (2003), Corbetta (2003).

[3] Para los análisis se recomienda utilizar el texto de Hernández, Fernández y Baptista (2006) y/o el texto de Ritchey (2006).

[4] Un buen texto para aprender a construir indicadores es Martinic (1997).