

Seleccionar y diseñar evaluaciones formales del aprendizaje de los estudiantes

¿En qué consiste la práctica de seleccionar y diseñar evaluaciones formales del aprendizaje de los estudiantes?

Esta práctica consiste en la selección y diseño de evaluaciones sumativas efectivas que le permitan al educador recoger valiosa información sobre los aprendizajes de los y las estudiantes. Estas evaluaciones proveen a profesores y estudiantes de información que permite identificar lo que han aprendido los estudiantes, diseñar estrategias para planificar la enseñanza futura y tomar decisiones para la mejora de los aprendizajes.

Una evaluación tiene intencionalidad sumativa cuando el educador usa las evidencias de aprendizaje para dar cuenta del logro alcanzado por los estudiantes durante un período de tiempo (unidad, semestre, año). Se deben recoger evidencias suficientes para determinar el nivel de logro de cada uno de los aprendizajes definidos y certificarlos (Sanmartí, 2007).

Seleccionar y diseñar evaluaciones formales del aprendizaje de los estudiantes implica (Förster, 2018):

1. Identificar los objetivos de aprendizaje o aprendizajes esperados de los estudiantes. Supone, además, construir y/o seleccionar indicadores de evaluación que explicitan en términos evidenciables la o las acciones que deberá realizar un estudiante para considerar que ha alcanzado un objetivo de aprendizaje.
2. Diseñar la estrategia de evaluación y definir las evidencias necesarias. Supone decidir cuál es la mejor manera de evaluar los contenidos, habilidades y actitudes antes de enseñarlos. Una estrategia de evaluación considera:
 - Determinar quiénes serán los agentes de evaluación: el profesor o profesora (heteroevaluación), el mismo estudiante (autoevaluación) o compañeros a un mismo nivel jerárquico en la sala de clases (evaluación de pares). Una evaluación puede integrar más de un agente.
 - Seleccionar el procedimiento evaluativo. Entre los procedimientos o también llamadas situaciones de evaluación se encuentran: portafolio, estudio de casos, proyecto, organizadores gráficos, investigación, ensayos, informes, etc. Estos procedimientos permiten evaluar contenidos conceptuales, procedimentales y metacognitivos, además de habilidades de orden superior.
 - Seleccionar y diseñar los instrumentos. Entre los instrumentos de evaluación se encuentran:
 - Situaciones tipo prueba: cerrada, abierta (breve o extensa) y mixta. Mientras que la primera permite evaluar conocimientos factuales, conceptuales y, en algunos casos, procedimentales, las preguntas de respuesta abierta extensa permiten evaluar habilidades de orden superior (analizar, evaluar y crear).

- Registro de realización de tareas: escalas de valoración, listas de cotejo, rúbricas, registros de observación. Estos instrumentos están al servicio de las situaciones de desempeño, en tanto permite su corrección y calificación.
- Seleccionar un referente. El referente de una evaluación se refiere al elemento que se utilizará para emitir un juicio de valor sobre el nivel de aprendizaje obtenido por un estudiante. Entre los referentes están:
 - Ideográfico: las evidencias se comparan con el rendimiento obtenido anteriormente por el estudiante.
 - Normativo: las evidencias se comparan con el rendimiento obtenido por otros estudiantes que fueron evaluados (por ejemplo, por estudiantes del curso o la generación).
 - Criterial: las evidencias que se recogen durante la evaluación de los estudiantes se comparan con los objetivos e indicadores de evaluación.

Al construir evaluaciones sumativas el educador deberá resguardar los siguientes criterios de calidad:

Validez (Förster & Rojas-Barahona, 2008):

- Validez de contenido: el instrumento debe recoger información sobre lo que se ha propuesto evaluar (conocimientos, habilidades y/o actitudes).
- Validez instruccional: el instrumento debe incluir preguntas y actividades que los estudiantes han tenido oportunidad de aprender durante las clases. Está directamente relacionada a la validez de contenido.
- Validez consecuencial: las consecuencias derivadas de la aplicación de un instrumento de evaluación deben ser acordes con el propósito con que fue creado.

Confiabilidad: La evaluación debe ser precisa. Para ello, las evidencias que se recogen con un instrumento de evaluación deben ser suficientes para emitir un juicio sobre el nivel de logro de sus aprendizajes. Asimismo, se deben cuidar las condiciones de aplicación del instrumento, evitando aspectos que puedan desconcentrar a los estudiantes (como ruido, poca luz, mala ventilación, etc.), asegurar que tengan el tiempo suficiente y la calidad del instrumento (letra, calidad de las imágenes, etc.), de tal forma que no haya factores externos que incidan en su rendimiento (Förster & Rojas-Barahona, 2008).

Objetividad: Tanto los instrumentos de evaluación como el juicio que se emite a partir de la información recogida con ellos deben ser imparciales y libres de sesgos.

Lo que no es Seleccionar y diseñar evaluaciones formales:

- Hacer evaluaciones para regular el comportamiento de los estudiantes. La evaluación debe servir para la toma de decisiones en pos de la mejora de los aprendizajes y no para controlar el clima de aula.
- Diseñar evaluaciones que constituyen un hito aislado y que no se relacionan con lo que los estudiantes tuvieron oportunidad de aprender, practicar y ejercitar durante el transcurso de la unidad didáctica u otro periodo de enseñanza. Al no asegurar la coherencia entre el proceso de aprendizaje y la evaluación se corre el riesgo de que los estudiantes no tengan la oportunidad

de demostrar cuánto han aprendido. La evaluación debe ser parte integral del proceso de aprendizaje (Förster, 2018). Por ello, la planificación debe tener en cuenta de un modo integrado tanto el diseño de la enseñanza como el de las estrategias de evaluación.

- Diseñar solo una evaluación o pregunta para evaluar un aprendizaje. Se debe utilizar más de un método de evaluación y/o pregunta dentro de una prueba para garantizar la obtención de una imagen más completa del dominio que un estudiante tiene de un conjunto de aprendizajes (Förster, 2018). Esto permite aumentar la confiabilidad, teniendo tendencias consistentes sobre el desempeño de el o la estudiante y reducir las inconsistencias provocadas por errores de medición que se producen por factores externos, como una inestabilidad emocional momentánea, enfermedad, etc.
- Seleccionar la evaluación sin considerar las habilidades, contenidos y/o actitudes que deben ser dominados por los estudiantes en la unidad que se debe evaluar. Al no considerar los objetivos de aprendizaje, se corre el riesgo de no evaluar efectivamente lo que debe ser evaluado y que los resultados no sirvan para tomar decisiones.
- Utilizar siempre el mismo tipo de procedimiento y/o instrumento de evaluación. La evaluación condiciona la forma en que los y las estudiantes aprenden (Sanmartí, 2007), por ello, se deben presentar instrumentos y desafíos variados que sean pertinentes para la edad, la etapa del desarrollo de los estudiantes y el contenido (declarativo, procedimental y/o actitudinal) que se desea evaluar. De lo contrario, los estudiantes condicionarán su aprendizaje al tipo de evaluación que utiliza el o la docente.
- Diseñar evaluaciones en que predomine la heteroevaluación. Los y las estudiantes deben ser partícipes de su proceso de evaluación, para lo cual se debe utilizar la evaluación de pares y la autoevaluación; esto promoverá ejercicios de reflexión, metacognición y les permitirá conocerse a sí mismos (Anijovich, 2012). Asimismo, la participación de los estudiantes permitirá obtener mayores fuentes de información para hacer el juicio de valor relacionado a la evaluación.
- Utilizar evaluaciones diseñadas para otros estudiantes u otros contextos. Si se traducen o adaptan instrumentos de evaluación realizados en otro contexto se debe ajustar a la realidad y características de los estudiantes. De lo contrario, se pueden generar errores y distorsiones en los resultados producto de la incomprensión (Förster, 2018).

Estrategias docentes para seleccionar y diseñar evaluaciones formales:

- **Definir propósitos:** definir los propósitos de la evaluación identificando claramente qué decisiones se tomarán a partir de sus resultados. Por ejemplo: ajustar la planificación, realizar reforzamiento para algunos estudiantes, dar actividades de profundización, etc.
- **Determinar el referente:** definir cómo se interpretarán los resultados: referente criterial, normativo o ideográfico, de acuerdo a los objetivos de evaluación: evaluar el aprendizaje de los y las estudiantes con respecto a sí mismos (ideográfico), en función del rendimiento del curso (normativo) o en función de criterios previamente establecidos (criterial). Por ejemplo,

al usar un referente criterial para evaluar un trabajo de investigación, el profesor o profesora comparará el desempeño de cada estudiante con un conjunto de criterios (contenidos en una rúbrica o escala de apreciación) que surgen a partir de los indicadores de evaluación previamente definidos. Si se desea utilizar un referente ideográfico, se debería comparar el desempeño del estudiante usando dos entregas de la investigación, por ejemplo, un borrador que recibe retroalimentación y un producto final que integre las observaciones realizadas.

- **Determinar el agente:** determinar si en la evaluación participará el docente, los estudiantes o ambos. Se debe tomar la decisión pensando en el tipo de información que se quiere recoger y ponderar qué agente puede proporcionarla. Para ello, se deberá determinar claramente cuáles serán los criterios que evaluarán cada uno de ellos. Por ejemplo, en la elaboración en grupos de una maqueta del sistema digestivo, los estudiantes pueden evaluar la participación y responsabilidad de cada uno de los integrantes (coevaluación), mientras que el docente puede hacerlo a partir de aspectos como la incorporación y representación de las estructuras del sistema digestivo (heteroevaluación). Por último, cada estudiante puede evaluar si cumplió con el rol que le fue asignado dentro del grupo y si participó activamente dando ideas y/o resolviendo problemas (autoevaluación).
- **Determinar el procedimiento e instrumento de evaluación:** con el fin de seleccionar los más adecuados, se debe tomar en consideración la naturaleza de los contenidos, habilidades y actitudes que serán evaluados (Armstrong, 2010). Se sugiere utilizar una taxonomía de aprendizaje que permita contar con una definición clara de los tipos de conocimientos y habilidades que se evaluarán. Por ejemplo, para evaluar la habilidad de “argumentar”, se sugiere presentarle a los estudiantes una situación o problemática a partir de la cual deberán elegir una postura y sustentarla a partir de los contenidos revisados durante el transcurso de las clases.
- **Diseñar estrategias evaluativas que consideren contextos auténticos:** Implica evaluar el conocimiento, procedimiento y/o actitudes utilizando problemas o situaciones reales, de manera tal de darle valor al conocimiento para la reflexión y la acción (Diaz Barriga, 2005). Por ejemplo, para evaluar un objetivo que implique realizar cálculos que involucren las cuatro operaciones, proponer a los estudiantes que hagan un presupuesto para el paseo de final de curso, argumentando distintas opciones y la cuota que se debería pagar mensualmente para poder financiarlo.
- **Revisar ejemplares:** indagar ejemplos de evaluaciones de los objetivos e indicadores seleccionados. Para ello, se puede revisar los programas del Ministerio de Educación u otros ejemplares que se encuentran en línea. Se sugiere siempre utilizar instrumentos validados y de instituciones reconocidas.
- **Seleccionar instrumentos estandarizados:** Un instrumento de evaluación estandarizado posee normas para la creación de la prueba, su aplicación y la interpretación de sus resultados. Dado que han sido confeccionadas a través de un proceso riguroso que incluye la participación de expertos y su pilotaje, poseen altos índices de validez y confiabilidad. Asimismo, permite comparar los resultados entre distintas muestras de estudiantes (Backhoff, 2018). Para seleccionar la evaluación estandarizada más idónea, se debe revisar el constructo (contenido, habilidad y/o actitudes) que mide sus propósitos y tener completa claridad de cómo interpretar sus resultados.

- **Diseñar evaluaciones que consideren los siguientes elementos:**

- **Asegurar la cobertura de la evaluación:** para las situaciones tipo prueba, elaborar una tabla de especificaciones del instrumento en donde se precisen los objetivos, habilidades y contenidos de cada reactivo así como la respuesta esperada (Förster & Rojas-Barahona, 2008). Revisar contrastando con los objetivos de la unidad para evitar dejar afuera un aspecto relevante.

- **Asegurar el rigor:** asegurar que la evaluación sea desafiante, incluyendo preguntas y actividades que permitan la aplicación, análisis y reflexión de los contenidos de acuerdo a las capacidades de los estudiantes. Asimismo, deben estar alineadas al currículum y a las evaluaciones estandarizadas nacionales e internacionales (Bambrick-Santoyo, 2018). Se sugiere reflexionar en conjunto con otros docentes para ajustar el rigor de las evaluaciones de todas las asignaturas.

- **Crear pautas y rúbricas de evaluación:** crear, previo a la aplicación del instrumento, pautas y/o rúbricas con un vocabulario comprensible por los estudiantes y socializarlas para que puedan desarrollar su trabajo de manera guiada. Se deben incluir todos los aspectos que serán evaluados y calificados. Por ejemplo: en un ensayo escrito se evaluará redacción, uso de evidencias, referencias, aspectos de formato, etc. En el caso de un desempeño práctico, como la realización de un laboratorio, además de evaluar el planteamiento de la hipótesis, el diseño experimental y discusión de resultados, se deben evaluar aspectos de orden procedimental como el uso de los equipos y de medidas de seguridad dentro del laboratorio.

- **Crear pautas y rúbricas para la autoevaluación o evaluación de pares:** deben ser claras, de fácil aplicación y con criterios concretos y observables. Asimismo, deben ser claras para la edad de los y las estudiantes. Por ejemplo, en cursos más pequeños, se pueden utilizar escalas gráficas o que contengan símbolos en vez de escalas numéricas. Asimismo, se sugiere utilizar economía del lenguaje, planteando pautas con criterios concisos.

- **Ajustar la evaluación:** adecuar los instrumentos a los antecedentes y experiencias previas de los estudiantes. Evitar sesgos provocados por factores externos a la evaluación tales como la cultura, la etapa del desarrollo, la etnia, el género, la situación socioeconómica, el idioma, los intereses especiales y las necesidades educativas especiales (Förster, 2018). Para ello, se debe revisar, entre otras cosas, las situaciones o contextos, ejemplos, imágenes y el vocabulario utilizado.

- **Asegurar validez instruccional:** incluir preguntas o actividades que sean similares a las actividades realizadas por los estudiantes en clase para evitar ponerlos en una situación de desventaja (Förster & Rojas-Barahona, 2008). Por ejemplo, en la asignatura de matemáticas, se pueden adaptar problemas que se hayan trabajado los estudiantes durante las clases cambiando los datos o el contexto en que se presentaron.

- **Asegurar validez semántica:** utilizar un lenguaje conocido y trabajado durante la unidad o las clases por los estudiantes para evitar errores por incomprensión de instrucciones o preguntas. Se sugiere tener un registro de los conceptos que fueron utilizados durante las

clases, a modo de glosario, para nuevas palabras que puedan confundir a los estudiantes. Asimismo, en casos o problemas que los y las estudiantes deben resolver, el profesor deberá seleccionar conceptos de uso cotidiano (por ejemplo, usar la palabra lápiz en vez de bolígrafo) o, bien, agregar un pie de página con una definición en caso de que sea necesario.

- **Asegurar validez de contenido:** someter la evaluación a criterios de “jueces”. Pedir al jefe técnico u otro profesor que la revise el instrumento de evaluación y, de ser necesario, hacer cambios en base a sus sugerencias.
- **Asegurar la confiabilidad:** asegurar que el ambiente de aplicación de una evaluación esté libre de distracciones y presiones que interfieran en el desempeño de los estudiantes.
- **Aumentar confiabilidad:** evaluar a los estudiantes en reiteradas oportunidades respecto de un mismo aprendizaje, ya sea en diferentes instancias o en varios ítems dentro de una situación evaluativa.
- **Asegurar precisión de corrección:** asignar, antes de la aplicación del instrumento, los puntajes en función de la relevancia y nivel de complejidad de los aprendizajes.
- **Asegurar objetividad:** elaborar pautas de respuesta o corrección antes de la aplicación del instrumento para asegurar un referente objetivo al momento de corregir.
- **Disminuir sesgos:** en caso de respuestas escritas, es recomendable seguir un orden en la corrección que permita evaluar con el mismo rigor el desempeño de todos los y las estudiantes. Por ejemplo, revisar todas las respuestas a la pregunta 1, luego continuar con la 2 y así sucesivamente.

Referencias

Amstrong, P (2010). Bloom's Taxonomy. Vanderbilt University Center for Teaching. Recuperado el 28 de octubre de 2022 de <https://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/>

Anijovich, R. (2012). Evaluar para aprender. Conceptos e instrumentos. Aique Educación.

Backhoff Escudero, Eduardo (2018). Evaluación estandarizada del logro educativo: contribuciones y retos. Revista Digital Universitaria (RDU). Vol. 19, núm. 6 noviembre-diciembre. DOI: <http://doi.org/10.22201/codeic.16076079e.2018.v19n6.a3>

Bambrick – Santoyo, P. (2018). *Las Palancas del Liderazgo Escolar 2.0. Una guía práctica para construir colegios exitosos*. Corporación de la Educación Aptus Chile.

Bowen, R. S. (2017). Understanding by Design. Vanderbilt University Center for Teaching. Recuperado el 26 de octubre de 2022 de <https://cft.vanderbilt.edu/understanding-by-design/>.

Díaz Barriga, Frida. (2005). Enseñanza situada: Vínculo entre la escuela y la vida. México: McGraw Hill.

Fisher, M. R., Jr., & Bandy, J. (2019). Assessing Student Learning. Vanderbilt University Center for Teaching. Recuperado el 26 de octubre de 2022 de <https://cft.vanderbilt.edu/assessing-student-learning/>.

Förster, C. (2018). *El poder de la evaluación en el aula. mejores decisiones para promover aprendizajes*. Ediciones Universidad Católica de Chile.

Förster, C. & Rojas-Barahona, C. (2008). Evaluación al interior del aula: Una mirada desde la validez, confiabilidad y objetividad. *Rev. Pensamiento Educativo*, Vol. 43, 2008. pp. 285-305.

Ministerio de Educación de Chile (2018). *Orientaciones para la implementación del Decreto 67/2018 de Evaluación, Calificación y Promoción Escolar*. <https://bibliotecadigital.mineduc.cl/bitstream/handle/20.500.12365/14279/orientaciones%20decreto%2067.pdf?sequence=1&isAllowed=y>

Sanmartí, N. (2007). 10 ideas clave. Evaluar para aprender. Graó.

Yale Poorvu Center for Teaching and Learning (s.f). *Formative and Summative Assessment*. Recuperado el 26 de octubre de 2022 de <https://poorvucenter.yale.edu/Formative-Summative-Assessments>

Sugerencia para citar este documento

Observatorio de Buenas Prácticas Pedagógicas, Facultad de Educación, Universidad del Desarrollo. (2023) "Seleccionar y diseñar evaluaciones formales".

Este documento fue elaborado por Antonia Ortega y revisado por Kiomi Matsumoto, María Luisa Salazar y Catalina Correa el marco del Observatorio de Buenas Prácticas Pedagógicas de la Facultad de Educación de la Universidad del Desarrollo, Santiago, Chile.

Última actualización febrero 2023.

Facultad de Educación UDD

Facultad de Educación UDD