

Evaluación de desempeño

Recursos de evaluación del Consejo Australiano para la Investigación Educativa 5



SERIE: DOCUMENTOS TÉCNICOS

***EVALUACIÓN DE DESEMPEÑO
RECURSOS DE EVALUACIÓN DEL CONSEJO
AUSTRALIANO PARA LA INVESTIGACIÓN
EDUCATIVA 5***



Serie Documentos Técnicos, 23

Consejo Directivo Ad Hoc

Carolina Barrios Valdivia, Presidenta
Fabiola León-Velarde Servetto
Daniel Alfaro Paredes

Secretaría Técnica

Haydee Chacón Cabanillas (e)

Cuidado de la edición

Dirección de Evaluación y Gestión del Conocimiento
Verónica Alvarado Bonhote, Directora
Diana Zapata Pratto, Especialista en Gestión de Publicaciones

Traducción

Cecilia Torres Llosa

Maquetación

Ángel García Tapia

Se terminó de imprimir en noviembre de 2017 en:

EDITORIAL SUPER GRÁFICA E.I.R.L.
Calle Luisa Beausejour 2049
Cercado de Lima

Hecho el Depósito Legal en la Biblioteca Nacional del Perú N.º 2017-14521
ISBN N.º 978-612-4322-25-9

Tiraje: 500 ejemplares
Primera edición
Lima, noviembre de 2017

© Sistema Nacional de Evaluación, Acreditación y Certificación de la Calidad Educativa

Calle Manuel Miota N.º 235 - San Antonio, Miraflores, Lima 18, Perú
Teléfonos: (+51 1) 637-1122, (51-1) 221-4826, (51-1) 221-4807 anexo 108
E-mail: sir@sineace.gob.pe / www.sineace.gob.pe

De la edición inglesa:

© **The Australian Council for Educational Research Ltd (ABN 19 004 398 1450) 2016**
© **El Consejo Australiano para la Investigación Educativa Ltd. (ABN 19 004 398 1450) 2016**

Título original: *Performances. Assessment Resource Kit*

Autor: Margaret Forster y Geoff Masters

Translated by permission of the Australian Council for Educational Research Ltd (ABN 19 004 398 1450). All rights reserved.
Traducido con autorización del Consejo Australiano para la Investigación Educativa Ltd. (ABN 19 004 398 1450). Todos los derechos reservados.

Se autoriza la reproducción total o parcial solo para propósitos educativos dentro del territorio nacional.
Distribución gratuita. Prohibida su venta.

ÍNDICE

Presentación.....	6
1. ¿Qué es la evaluación del desempeño?.....	10
1.1. Resultados de "realización".....	10
1.2. Contextos y propósitos.....	15
1.3. Planear evaluaciones de desempeño.....	16
1.4. Calificación del desempeño.....	18
1.5. Estimar y reportar las ubicaciones en un mapa de progreso.....	19
2. Observaciones informales en el salón de clases.....	23
2.1. Enfocar las observaciones.....	24
2.2. Registrar observaciones.....	27
2.3. Estimación y reporte del logro.....	34
3. "Eventos" de evaluación del desempeño en el salón de clases.....	37
3.1. Planeación de eventos de evaluación de desempeño.....	38
3.2. Asegurar una profunda evidencia.....	40
3.3. Calificación y registro.....	44
3.4. Estimar y reportar la ubicación en un mapa de progreso.....	55
4. Cuando la comparación importa.....	60
4.1. Planeamiento de la evaluación del desempeño.....	62
4.2. Calificación de evaluaciones del desempeño.....	68
4.3. Resumir y reportar el rendimiento.....	71
5. Diseñando la evaluación del desempeño.....	76
5.1. Planeamiento de la evaluación del desempeño.....	76
5.2. Enfocar y registrar observaciones / Calificar el desempeño.....	78
5.3. Ubicar a los estudiantes en un mapa de progreso.....	79
6. Un resumen del proceso de diseño de la evaluación del desempeño.....	80

PRESENTACIÓN

Entre los años 2009 y 2015, el SINEACE estuvo abocado a la elaboración de estándares de aprendizaje, con la finalidad de contribuir con el Ministerio de Educación y diversos actores a lograr mejores aprendizajes en los estudiantes de la educación básica regular; durante dicho proceso, se conoció la experiencia australiana de evaluación de los aprendizajes realizada por el Consejo Australiano para la Investigación Educativa (ACER). Su trabajo inspiró la elaboración de los estándares de aprendizaje en forma de mapas de progreso, lo que significó un cambio de paradigma en la manera de enseñar, toda vez que no indican tanto qué debe enseñarse, sino qué debe aprender un estudiante.

La pertinencia para el Perú de la experiencia de ACER, motivó que el SINEACE considerara la conveniencia de poner al alcance de los docentes los folletos que comprendían el Kit de Recursos de Evaluación, que había sido un valioso aporte para la elaboración de estándares de aprendizaje. Para ello suscribimos un convenio con ACER, que permite poner al alcance de los lectores de habla castellana el quinto número de la serie, el cual rescata la importancia de la evaluación de desempeño, es decir, la que se realiza cuando los estudiantes se involucran en una actividad, la evaluación inmediata de una presentación, comportamiento e interacción.

En la evaluación de desempeño, los profesores monitorean el progreso del estudiante sobre un mapa de progreso preconstruido de habilidades,

conocimiento y comprensión. Los profesores hacen observaciones sobre el desempeño del estudiante, las que constituyen la “evidencia” utilizada para estimar el nivel de los estudiantes. Mientras mayor sea el número de observaciones, más confiable será la conclusión sobre el nivel de desempeño de cada uno. Todo ello sin perder de vista la forma en que el género, el historial cultural y la personalidad pueden influir en el desempeño del estudiante, para asegurarse de que los indicadores de logro sean justos.

Hay muchos contextos para —y tipos de— evaluaciones de desempeño, cada una de las cuales brinda distintos tipos de evidencia que los profesores pueden usar para estimar los niveles de rendimiento del estudiante. Se listan aquí los temas que los profesores deben considerar cuando diseñan una evaluación de desempeño; que incluyen preguntas de propósito evaluador, métodos para calificar y registrar el desempeño del estudiante, y formas para estimar la ubicación de un estudiante en un mapa de progreso.

La publicación explora diferentes tipos de evaluación de desempeño: la que se realiza día a día en el salón de clase; estrategias para planear, calificar y registrar presentaciones; o la que se da en contextos donde la comparación de evaluaciones es particularmente importante. En el desempeño de los estudiantes de alto rendimiento es importante que las evaluaciones de desempeño sean comparables de estudiante a estudiante, de evaluador a evaluador y de escuela a escuela. En estos casos, se desea *coincidencia curricular, legitimidad, estandarización y generalización*.

Cuando se califica el desempeño para inferir el nivel de logro del estudiante, se deben considerar dos características de evaluación: el método de calificación y la comparación de calificaciones (fiabilidad entre evaluadores). Las calificaciones de desempeño pueden compararse a partir de:

- Definir cuidadosamente el criterio para evaluación del desempeño.
- Asegurar consistencia de criterio en la interpretación.

- Utilizar un número de calificaciones.
- Identificar y tomar en cuenta las diferencias sistemáticas sobre los evaluadores.

Para obtener la mayor información posible de la observación cotidiana, los profesores necesitan obtener un sistema de registros que sea organizado, fácil de usar y eficiente. En la publicación se describen ampliamente los sistemas que se pueden utilizar para registrar las observaciones.

El SINEACE difunde este libro, que se espera aporte conocimientos acerca de las diversas maneras de evaluar el desempeño de los estudiantes, permitiendo profundizar tanto en los diferentes contextos como en los aspectos que pueden influir en los resultados de la evaluación, y en la importancia de la comparabilidad de las evaluaciones; todo ello de relevancia para los actores educativos en su conjunto.

Consejo Directivo Ad Hoc
SINEACE

**EVALUACIÓN DE DESEMPEÑO
RECURSOS DE EVALUACIÓN DEL CONSEJO
AUSTRALIANO PARA LA INVESTIGACIÓN
EDUCATIVA 5**

1. ¿QUÉ ES LA EVALUACIÓN DEL DESEMPEÑO?

La evaluación de algunos tipos de aprendizaje requiere la observación del desempeño de los estudiantes. Los ejemplos incluyen tocar un instrumento musical, bailar, utilizar instrumentos científicos, manejar equipos, desarrollar rutinas de gimnasia, resolver problemas en grupo, aplicar habilidades de manejo de pelota y participar en pequeñas discusiones grupales.

La evaluación del desempeño es la evaluación de los estudiantes *cuando se involucran en una actividad*. Es la evaluación inmediata de una presentación, comportamiento e interacción.

1.1. Resultados de “realización”

Los documentos curriculares incluyen muchos resultados de aprendizaje que describen procesos, actividades o presentaciones. Se puede describir estos resultados como resultados de “realización”.

La evaluación del desempeño, como se define aquí, es la evaluación que se basa en la observación y el juicio de las actividades mientras van sucediendo.

Para evaluar estos resultados es necesario observar al estudiante en actividad. Este tipo de evaluación es especialmente importante en algunas áreas de aprendizaje, como Artes y Educación Física, y en algunas facetas de aprendizaje, como la parte oral en Lenguaje.

Otras definiciones

Algunas veces se define de forma más libre la evaluación del desempeño como cualquier evaluación que requiere que los estudiantes construyan o generen una respuesta. Por ejemplo, en ocasiones, las respuestas cortas y ensayos son denominados *evaluación del desempeño*. Esto no es a lo que nosotros nos referimos.

En ocasiones, el término *evaluación del desempeño* se usa para referirse a una evaluación “auténtica” del mundo real, del desempeño *in situ*. Esto tampoco es a lo que nosotros nos referimos.

Uso de herramientas y equipos

Algunos resultados en el Currículo Normativo de Ontario no se relacionan directamente con el contenido de facetas de aprendizaje particulares, sino que tratan con *procesos* generales.¹ Por ejemplo, se espera que los estudiantes de 3.^{er} grado:

- Utilicen una variedad de herramientas simples y materiales manipulables para extender los sentidos en investigación práctica.
- Manipulen materiales concretos de manera espontánea y con seguridad.

Se espera que los estudiantes de 9.^{no} grado:

- Utilicen de manera segura y analicen una variedad de herramientas, maquinaria y materiales de manipulación para respaldar un diseño grupal y personal, construyan, experimenten, recolecten información, investiguen y realicen estudios independientes

Documentos curriculares de Ontario

¹ Ministry of Education and Training (1993). *The Ontario Common Curriculum, Grades 1-9*.

Conducta saludable

El cuadro de Documentos Nacionales de Estados Unidos contiene estándares para diez disciplinas. Muchas incluyen resultados de realización. Por ejemplo, para lograr el Estándar 5 al final del 4.º grado en Educación Sanitaria, los estudiantes deben poder demostrar los siguientes comportamientos:

- Formas saludables para expresar necesidades, deseos y sentimientos.
- Formas para comunicar cuidado, consideración y respeto hacia uno mismo y hacia los demás.
- Habilidades para escuchar atentamente a los demás y así construir y mantener relaciones saludables.
- Habilidades para mejorar la salud.
- Estrategias no violentas para resolver conflictos.²

Estándares Nacionales de Estados Unidos

Para evaluar estos resultados, los maestros deben observar los desempeños de los estudiantes, en lugar de los “productos” de su trabajo.



² *Education Week* (12 de abril de 1995). Washington, Special Report “Struggling for Standards”, p. 59.

Movimiento

Los Perfiles Curriculares para Escuelas Australianas detallan resultados para ocho áreas clave de aprendizaje. Muchos de estos resultados son resultados de "realización".

Educación Física y Sanitaria

"Demuestra acciones coordinadas del cuerpo, representando y modificando secuencias de movimiento".

Por ejemplo, se muestra cuando los estudiantes:

- Adaptan la velocidad, fuerza, flujo, dirección y altura del movimiento para mejorar el desempeño.
- Varían los patrones de movimiento al correr carreras cortas, con vallas y con distancias largas.³

Actuación y trabajo grupal

Estudios de Sociedad y Ambiente

"Expresa una visión personal del significado de la información".

Por ejemplo, se muestra cuando los estudiantes:

- Interpretan el papel de un personaje de un texto para mostrar las respuestas de los personajes en situaciones similares.

"Identifica las causas de conflicto y la falta de resultados en un trabajo de grupo y negocia las soluciones".

Por ejemplo, se muestra cuando los estudiantes:

- Resuelven conflictos y problemas en un grupo sin recurrir a una autoridad.
- Ponen de lado sentimientos personales para completar tareas (trabajan con sus compañeros, no necesariamente amigos, para compartir habilidades y conocimientos).⁴

³ Curriculum Corporation (1994). *Health and Physical Education – A Curriculum Profile for Australian Schools*. Carlton: Curriculum Corporation, p. 86.

⁴ Curriculum Corporation (1994). *Studies of Society and Environment – A Curriculum Profile for Australian Schools*. Carlton: Curriculum Corporation, pp. 37 y 85.

Danza

Las Artes

“Experimenta con ideas y explora sentimientos para encontrar soluciones satisfactorias a tareas”.

Por ejemplo, se muestra cuando los estudiantes:

- Unen formas y movimientos para explorar temas, ideas o sentimientos; y
- Realizan movimientos expresando emociones o estados de ánimo (miedo, molestia, alegría) y enlazan varias de las formas para hacer una secuencia de danza.⁵

Perfiles Curriculares para Escuelas Australianas

Las tarjetas de reportes suelen reflejar la importancia que profesores y escuelas asignan a los resultados de aprendizaje que requieren la observación del desempeño del estudiante.



⁵ Curriculum Corporation (1994). *The Arts – A Curriculum Profile for Australian Schools*. Carlton: Curriculum Corporation, p. 70.

Utilizando instrumentos

Matemáticas

“Mide y construye cosas utilizando una variedad de escalas y estrategias para realizar medidas que son más precisas que las que permite el equipo disponible”.

Por ejemplo, se muestra cuando los estudiantes:

- Utilizan instrumentos de forma correcta y acertada (coloca una jarra de medida en una superficie nivelada y lee con marcadores a nivel del ojo para reducir error de lectura).⁶

Desempeño musical

Las Artes

“Practica, presenta y promociona trabajos musicales en formas apropiadas para audiencias particulares”.

Por ejemplo, se muestra cuando los estudiantes:

- Ensayan y ponen en práctica trabajos musicales como miembros de un conjunto demostrando la habilidad de mantener una parte independiente y mostrar sensibilidad y balance dentro del conjunto, mientras representan tanto como conductor, líder o parte del grupo.⁷

Perfiles Curriculares para Escuelas Australianas

1.2. Contextos y propósitos

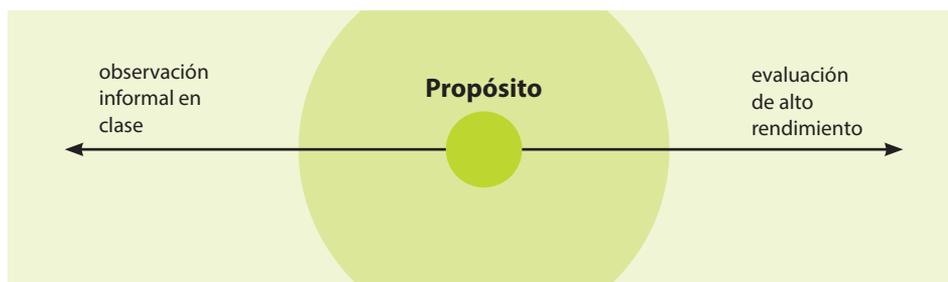
Hay muchos contextos y propósitos para la evaluación de resultados de “realización”, lo que significa que hay muchos tipos de evaluación de desempeño. Por ejemplo, un profesor se involucra en la evaluación del desempeño cuando, al discutir un problema de matemáticas con un estudiante, observa y registra la respuesta reflejando una nueva comprensión: “¡Ahora entiendo por qué $\frac{1}{4}$ es 25%! ”.

⁶ Curriculum Corporation (1994). *Mathematics – A Curriculum Profile for Australian Schools*. Carlton: Curriculum Corporation, p. 90.

⁷ Curriculum Corporation (1994). *The Arts – A Curriculum Profile for Australian Schools*. Carlton: Curriculum Corporation, p. 114.

Los contextos en el salón de clase para la evaluación del desempeño también incluyen evaluaciones de “eventos” más estructurados; por ejemplo: un juego de roles en Idiomas, un pequeño grupo de representación dramática, o evaluación de Educación Física en función de su habilidad para lanzar la pelota. Y se realiza una evaluación del desempeño de tipo más formal y estructurado como parte de una evaluación de alto rendimiento en danza y música.

Es útil pensar sobre estos diferentes tipos de evaluación de desempeño junto con un cuadro de “propósito”, tal como se muestra a continuación, con evaluación formativa y diagnosticada en el salón de clases en un extremo y evaluación sumativa de alto rendimiento en el otro. Las observaciones informales en el salón de clases están hacia la izquierda del cuadro. En el centro del cuadro se encuentran los eventos de evaluación planeados por los profesores en el salón de clase. Por ejemplo: una reunión profesor-estudiante o la observación de un grupo en Ciencia. En la parte externa derecha se fijan las evaluaciones de alto rendimiento.



La forma en la que se planean estos diferentes tipos de evaluaciones, la forma en la que se califica el desempeño del estudiante, y la forma en la que se utiliza el resultado para inferir la ubicación de un estudiante en un mapa de progreso, dependen del contexto y el propósito de la evaluación.

1.3. Planear evaluaciones de desempeño

La utilidad de la evidencia brindada por la evaluación de desempeño para inferir el nivel de logro del estudiante depende de:

- Relevancia y cobertura (un tema de validación).
- Cantidad de evidencia (un tema de fiabilidad).

Relevancia y cobertura

Relevancia describe el grado en el que la evidencia se dirige al conocimiento, habilidad y comprensión, o a los resultados de un área de aprendizaje. Es probable que los profesores, en el salón de clases, se encuentren más interesados en asegurar que las observaciones del desempeño del estudiante se enfoquen en metas de enseñanza y en que haya una relación entre las tareas de desempeño planeadas y los resultados que serán evaluados.

Cobertura describe el grado en el cual la evidencia tiene como muestras diferentes resultados en un área de aprendizaje. Normalmente, los profesores observan distintos desempeños del estudiante en sus intentos por recolectar evidencia sobre una diversidad de resultados. Es probable que los profesores también se encuentren interesados en desarrollar formas para registrar eficientemente información anecdótica sobre el desempeño del estudiante, con el fin de brindar evidencia sistemática para guiar las instrucciones y brindar una retroalimentación a los estudiantes y a los padres.

Cantidad de evidencia

A pesar de que la evaluación del logro del estudiante se basa en observaciones de su desempeño en tareas específicas, el propósito e interés siempre estarán en el conocimiento, la habilidad y la comprensión requeridos para desarrollar esas tareas. La intención es inferir el logro de los estudiantes en un área de aprendizaje a partir de su desempeño en una tarea específica, lo que significa generalizar desde una muestra de desempeño limitado.

Esta observación lleva a la pregunta sobre la cantidad de evidencia (número de tareas) requerida para brindar un estimado lo suficientemente acertado

del nivel general de logro del estudiante en un área de aprendizaje. Dicha pregunta es particularmente importante en contextos de alto rendimiento, en los cuales el desempeño del estudiante influye en la admisión a cursos, ofertas de becas o diplomas.

1.4. Calificación del desempeño

Cuando se califica el desempeño para inferir el nivel de logro del estudiante, se deben considerar dos características de evaluación:

- El método de calificación.
- La comparación de calificaciones (fiabilidad entre evaluadores).

Método de calificación

El desempeño se califica holística o analíticamente. Los profesores hacen calificaciones *holísticas* del desempeño de un estudiante cuando le dan una sola calificación basados en una impresión general de su desempeño.

Los profesores hacen una calificación *analítica* del desempeño de los estudiantes cuando puntúan diferentes aspectos del desempeño.

Una representación de danza, por ejemplo, puede ser calificada por el contenido y las características de la presentación: ideas, organización, ritmo brindado, volumen, expresiones faciales y contacto con la audiencia. Los profesores pueden reportar estas puntuaciones por separado o utilizarlas para hacer, en definitiva, una evaluación final del desempeño.

Comparación de las calificaciones

Una característica fundamental de la evaluación del desempeño es su fiabilidad en las calificaciones. Dos personas que ven la misma presentación

pueden calificar el desempeño de forma diferente. Para asegurar que las evaluaciones sean justas —lo que significa que la puntuación del desempeño de un estudiante no dependa de quién lo evalúe—, es importante minimizar las diferencias entre evaluadores, en otras palabras, asegurar la fiabilidad entre evaluadores.

Cuando se juzga el desempeño para inferir el nivel de logro del estudiante, hay que considerar dos características de la evaluación:

- el método de calificación; y
- la comparabilidad de las calificaciones (fiabilidad entre evaluadores).

En el contexto del salón de clase, si otro profesor realiza alguna calificación del desempeño de un estudiante, no es muy importante si se utiliza la misma guía de evaluación o no. Sin embargo, en los contextos de alto rendimiento, la fiabilidad entre evaluadores es crucial.

Normalmente, mientras mayor sea el requerimiento para comparar las evaluaciones de desempeño, más precisas serán las especificaciones del criterio de evaluación. En una situación de alto rendimiento, como la evaluación de final de año, el criterio para la evaluación puede estar muy especificado. Los evaluadores pueden estar capacitados para llevar a cabo el proceso de evaluación monitoreando cuidadosamente, para asegurar un alto nivel de concordancia entre evaluadores.

1.5. Estimar y reportar las ubicaciones en un mapa de progreso

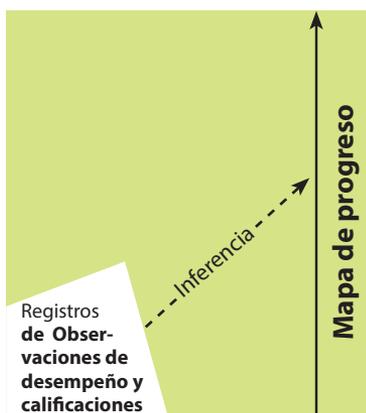
En la evaluación de desarrollo (ver *Evaluación de desarrollo: recursos de evaluación del Consejo Australiano para la Investigación Educativa 2*), los profesores monitorean el progreso del estudiante sobre un mapa preconstruido de habilidades, conocimiento y comprensión. Los profesores hacen observaciones sobre el desempeño del estudiante en contextos relevantes al área de

aprendizaje y a los resultados descritos en el mapa. Estas observaciones son la “evidencia” utilizada para estimar el nivel de atención de los estudiantes. Mientras mayor sea el número de observaciones, más confiable será la conclusión sobre el nivel de desempeño de cada estudiante.

La evaluación del desarrollo requiere una decisión final (inferencia) sobre la ubicación de un estudiante en un mapa de progreso basándose en la evidencia disponible. La forma en la que se realiza esta inferencia normalmente está determinada por el propósito de la evaluación de desempeño. Mientras más alto sea el rendimiento, mayor será el requerimiento de comparación, y las especificaciones en el “proceso de inferencia” serán más ajustadas.

Estimaciones subjetivas

Las evaluaciones de desempeño son una fuente de evidencia desde una variedad de métodos que utilizan los profesores para observar, registrar y recoger evidencia. Un profesor de Educación para la Salud puede hacer una inferencia sobre el logro de un estudiante en “movimiento humano”, por ejemplo, utilizando una combinación de evidencia de registros anecdóticos de su desempeño en clase, evaluaciones escritas y observaciones en tareas desarrolladas. En algunas instancias, la ubicación de un estudiante en un mapa de progreso puede ser estimada completamente desde su desempeño en una tarea asignada.



Cuando el estimado de nivel de logro del estudiante se realiza subjetivamente, puede que solo haya una conexión faltante entre la evidencia disponible (observaciones y calificaciones de desempeño) y el resultado estimado de la ubicación del estudiante en un mapa de progreso.

Estimaciones objetivas

En situaciones de alto rendimiento, donde son deseables altos niveles de comparación entre evaluadores, la manera en que las calificaciones se combinan para estimar la ubicación de un estudiante en un mapa de progreso pueden ser prescritas con firmeza, a fin de asegurar que la estimación sea realizada de la misma forma por diferentes evaluadores (objetivamente). Por ejemplo, la conclusión puede ser numérica a partir de un patrón de evaluación de calificaciones sobre un criterio definido cuidadosamente.

Los siguientes tres capítulos exploran diferentes tipos de evaluación de desempeño:

"Observaciones informales en el salón de clases" describe la evaluación día a día en el salón de clase de resultados de "realización" y dos estrategias para mejorar la calidad de información:

- Enfocarse en las observaciones.
- Registrar las observaciones eficientemente.

"Eventos de evaluación del desempeño en el salón de clases" describe estrategias para planear, calificar y registrar presentaciones. Por ejemplo:

- La autocalificación del desempeño en un problema de colaboración resolviendo una tarea.
- La autocalificación de participación en una tarea grupal.
- La puntuación analítica de una presentación musical.

- La puntuación holística del lenguaje oral.
- La puntuación analítica de una presentación dramática.

“Cuando la comparación importa” considera el uso de evaluación del desempeño en contextos en los que la comparación de evaluaciones es particularmente importante. Los elementos en este contexto incluyen:

- Coincidencia curricular.
- Legitimidad.
- Generalización.
- Estandarización.
- Fiabilidad entre evaluadores.

Para discusión

“Una evaluación que requiera que el alumno realice una actividad compleja o interesante no es necesariamente una buena evaluación. Una evaluación confiable mide más allá de tareas específicas que se pide completar al estudiante. El resultado de una buena evaluación identifica lo que los estudiantes pueden realizar con un amplio conocimiento o dominio de habilidades. Las habilidades que muestran los estudiantes en la situación de evaluación debe ser transferida a otras situaciones y a otros problemas.”⁸

⁸ Herman, J.; Aschbacher, P. y Winters, L. (1992). The National Center for Research on Evaluation, Standards and Student Testing (CRESST). *A Practical Guide to Alternative Assessment*. Alexandria, VA: Association of Supervision and Curriculum Development, p. 9. Copyright © by The Regents of the University of California and supported under the Office of Educational Research and Improvement (OERI), US Department of Education.

2. OBSERVACIONES INFORMALES EN EL SALÓN DE CLASES

La evaluación del desempeño es la evaluación de estudiantes involucrados en una actividad, ya sea utilizar aparatos científicos, discutir una historia o arrojar una jabalina. Hay muchos contextos y propósitos para evaluar el desempeño: desde la evaluación informal continua en el salón hasta evaluaciones formales de alto rendimiento. Este artículo se enfoca en la evaluación informal del desempeño: la evaluación que hacen los profesores en el salón *mientras van enseñando* en lugar de hacerlo en “eventos” planeados de evaluación.

Las observaciones del comportamiento del estudiante son necesarias para algunos tipos de resultados que no pueden ser evaluados inmediatamente mirando el producto del trabajo del estudiante. Por ejemplo, evaluar la habilidad de un estudiante para trabajar eficientemente en equipo, o la habilidad que utilizan en una técnica particular en ciencias, requiere observaciones del desempeño del estudiante. Otros ejemplos incluyen: utilizar instrumentos matemáticos, demostrar técnicas de atletismo y mantener posiciones en danza.

Los profesores tienen muchas oportunidades de observar el desempeño de un estudiante. Ellos observan la participación de estos en actividades de clase, cuando intervienen en las discusiones dentro del aula o interactúan con otros estudiantes. La mayoría de la información obtenida de esta forma se usa inmediatamente para informar sobre la práctica, para responder a las necesidades específicas de los estudiantes y para agregar una imagen general del progreso del estudiante como aprendiz. Pero el proceso es mayormente aleatorio, y depende de la oportunidad de la observación.

La utilidad y la calidad de la información obtenida desde una observación día a día del desempeño de estudiantes puede ser mejorada al:

- Enfocar las observaciones en resultados importantes de aprendizaje.
- Registrar observaciones sistemáticamente.

2.1. Enfocar las observaciones

No es práctico registrar los comentarios de cada alumno, cada día, sobre todo lo que dice o hace.

Las observaciones deben estar enfocadas en *eventos significativos que puedan relacionarse con resultados de aprendizaje relevantes*. Eventos “significativos” pueden ser:

- Comportamiento inesperado del estudiante.
- Ejemplos claros de comprensión o falta de comprensión.
- Evidencia clara concerniente a un resultado particular de aprendizaje.

Centrar la atención

Una técnica para enfocar las observaciones en resultados importantes de aprendizaje es “centrar la atención”. Se denomina así al proceso de observar solo algunos aspectos del comportamiento del estudiante. Los ejemplos de centrar la atención en el lenguaje oral y en el comportamiento psicomotor se muestran en las páginas 25 y 26.

Para obtener más información posible de una observación diaria, los profesores necesitan desarrollar un sistema de registro que sea:

- organizado;
- fácil de usar, y
- eficiente.

También necesitan:

- dar acceso rápido a la información, y
- permitir crecimiento que se vea en el tiempo.

En ocasiones, los profesores escogen enfocar los resultados de diferentes áreas de aprendizaje en la misma actividad. Por ejemplo, un profesor puede tener como propósito un resultado en Trabajo Científico y uno en Lenguaje Oral y Escrito, en una tarea de ciencia que requiere de los estudiantes trabajar en pequeños grupos utilizando equipos para investigar un problema.

Cuando se centra la atención en el comportamiento, los profesores a menudo enfocan sus observaciones en diferentes etapas del proceso. Por ejemplo, si los estudiantes están aprendiendo una serie de habilidades aisladas antes de combinarlas para un desempeño complejo (por ejemplo, una rutina de

Enfocar las observaciones		Centrar la atención		
Terry Spottiswood ha desarrollado una lista para observación del lenguaje oral que se enfoca solo en el comportamiento de cuatro estudiantes. Ella observa para tener evidencia de estos cuatro comportamientos en la interacción de un grupo pequeño. ¹				
Nombre	Indica desacuerdo en una forma positiva.	Apoya las declaraciones de los demás.	Hace preguntas para obtener información.	Escucha sin interrumpir.
Georgio	✓		✓	✓
Dean			✓	
Sam	✓	✓	✓	✓
Wai-Ming	✓	✓	✓	

¹ Terry Spottiswood, Waldau Primary School, Victoria, Australia.

Enfocar las observaciones	Centrar la atención en el comportamiento psicomotor
<p>Roy Killen sugiere hacer una lista para observaciones enfocadas en comportamientos psicomotores que incluya la secuencia de pasos que se espera que los estudiantes sigan en una actividad en el salón de clase.²</p>	
Utilizando una barra de medición	Hecho
<ul style="list-style-type: none"> • Ubicar objetos de pesos desconocidos al lado izquierdo de la balanza. 	<input type="checkbox"/>
<ul style="list-style-type: none"> • Estimar el peso y los lugares apropiados para poner el peso en el lado derecho de la balanza. 	<input type="checkbox"/>
<ul style="list-style-type: none"> • Subir el brazo de la balanza y revisar el balance. 	<input type="checkbox"/>
<ul style="list-style-type: none"> • Notar si el peso fue muy grande o muy pequeño. 	<input type="checkbox"/>
<ul style="list-style-type: none"> • Bajar la balanza. 	<input type="checkbox"/>
<ul style="list-style-type: none"> • Agregar o disminuir pesos (como sea apropiado) y volver a revisar el balance. 	<input type="checkbox"/>
<ul style="list-style-type: none"> • Mientras se acerca al equilibrio, agregue y disminuya pesos. 	<input type="checkbox"/>
<ul style="list-style-type: none"> • Lleve la balanza cerca al equilibrio con la adición o substracción de cada peso (en vez de tratar de lograr el balance en una forma fallo-error). 	<input type="checkbox"/>
<ul style="list-style-type: none"> • Cuando se logre el equilibrio, agregue pesos para tener el peso correcto del objeto. 	<input type="checkbox"/>

gimnasia), un profesor puede centrar la atención del desempeño en una o dos habilidades aisladas tanto como en aspectos del desempeño final.

También puede ayudar “centrar la atención” en estudiantes particulares en lugar de tratar de observar a todos los estudiantes en todas las lecciones. Por ejemplo, los profesores pueden escoger observar cinco estudiantes cada día.

² Killen, R. (1993). *Assessment and Evaluation*. Newcastle: University of Newcastle, Faculty of Education, p. 94.

Trabajar sistemáticamente en las clases de este modo ayuda a alertar a los profesores sobre los estudiantes “invisibles”, que tienen poca oportunidad de ser vistos, para poder registrar sus actividades.

Cuestionamiento enfocado

El cuestionamiento enfocado es una técnica que permite reunir información adicional mientras se centra la atención en los estudiantes. Los profesores utilizan el cuestionamiento enfocado para mejorar su percepción de la captación conceptual de los estudiantes en una tarea: ¿por qué tratan de resolver un problema en una forma en particular? ¿Cómo sus descubrimientos están relacionados con sus predicciones y con otras experiencias que han tenido? ¿Qué tratan de comprender sobre las habilidades requeridas para completar una serie compleja de movimientos?

2.2. Registrar observaciones

Para obtener la mayor información posible de la observación cotidiana, los profesores necesitan obtener un sistema de registros que sea:

- Organizado.
- Fácil de usar.
- Eficiente.

También es necesario:

- Dar un acceso rápido a la información.
- Permitir crecimiento que se vea en el tiempo.

Hay muchas formas de registrar las observaciones en un salón de clase, incluidos ficheros, registros computarizados, cuadros y hojas de observación.

Los profesores necesitan probar varios métodos, combinarlos o inventar los suyos antes de encontrar el sistema que mejor encaje con ellos, su área de estudio y su salón de clase.

Ficheros

Los profesores han desarrollado diversas formas de usar ficheros para registrar evaluación informal del desempeño. Ann Beyer sugiere los siguientes:³

- i. Conservar las fichas en un fólдер. Utilizar una ficha para cada estudiante en cada área de aprendizaje. Escribir la fecha y las observaciones directamente en las fichas. Cuando se llenen las fichas, colocar encima una adicional.
- ii. Mantener el conjunto de fichas juntas para cada estudiante en un fólдер. Las fichas de diferentes colores pueden indicar diferentes facetas de un área de aprendizaje.
- iii. Pegar las fichas en una sola hoja para mantenerlas en una pizarra para un uso más fácil. Una vez que la tarjeta esté llena, archívela para una referencia futura.

Registros computarizados

Si los profesores tienen acceso a la computadora en el salón de clase, pueden ingresar las observaciones directamente en la computadora. Utilizando copiar y pegar, editan las funciones que pueden transferir en sus observaciones para archivar a las estudiantes individualmente. Estos archivos pueden ser importados a un archivo de reportes en otra fecha.

³ Beyer, A. (1993). "Assessing Students' Performance Using Observations, Reflections, and Other Methods". En Webb, N. y Coxford, A., *Assessment in the Mathematics Classroom*. Virginia: National Council of Teachers of Mathematics, Inc., p. 119.

Listas de anotaciones en clase

Las listas de anotaciones en clase pueden usarse para registrar las observaciones del comportamiento del estudiante.

Hojas o cuadros de observación

Las observaciones del desempeño del estudiante también pueden registrarse en hojas o cuadros en clase. Pueden usarse para un período particular en una tarea en particular. Algunos profesores pegan los cuadros dentro de un fólter para mantener su privacidad. Otros los ponen en la pared de la clase para que los estudiantes puedan ver los resultados de sus evaluaciones.

Algunos profesores cortan la hoja de observaciones al final de cada semana y transfieren entradas individuales en el fólter de registro de cada estudiante. Otros han refinado el sistema para un registro más fácil y más eficiente: escriben sus observaciones en *post-its* en los cuadros (ver página 31) y luego pueden transferirlas fácilmente a un cuaderno de registro.



Registro de observaciones
Cuadro de observaciones

Kevin Olsen y sus colegas describen el uso de un profesor de hojas A3 para cuadros de observaciones.⁴ Al final de cada semana, el profesor pone fecha a las entradas, las recorta y hace que el estudiante las pegue en una tarjeta en su fólder. Es muy importante que el comentario ubicado en un cuadro de este tipo *esté separado*; eso significa que los comentarios sobre un estudiante significarán algo luego de que se hayan separado del contexto que rodea los comentarios. Las siguientes observaciones corresponden al área de comprensión matemática, pero este sistema puede usarse para registrar observaciones del desempeño de los estudiantes en cualquier área de aprendizaje.

Reina 20/05: Animales, patrones de objetos naturales, ordinales "... y la oveja es la primera y la gallina es la segunda...?" 21/05: Secuencia temporal - El uso de ordinales ocurre con frecuencia en el día; utiliza primero y segundo para describir secuencia - Registrado oralmente y en dibujos.	Xenia	Antonia	Kate
Chris	Conrado	Juan 17/05: Patrones - Ama los colores. Los usa y utiliza ordinales "Primero este color aquí; y el azul es el siguiente...". Precisión al combinar colores.	Rose

... y los otros 23 niños.

⁴ Olsen, K.; Adams, G.; Grace, N. y Anderson, P. (1994). *Using the Mathematics Profile*. Carlton: Curriculum Corporation, p. 84.

Registro de observaciones

Post-its

Bonnie Campbell Hill ha refinado el sistema de cuadros para un registro más eficiente y más fácil utilizando notas adhesivas.⁵

Los profesores, utilizando su sistema, copian un formato con bloques del tamaño de pequeños post-its y el espacio para colocar el nombre del estudiante. Los profesores colocan el nombre del estudiante y luego hacen varias copias del formato. Se colocan post-its en blanco en los espacios bajo cada nombre. El profesor lleva el formato en una pizarra y coloca las notas con observaciones directamente en los post-its. Estas tienen fecha. En intervalos regulares, los post-its se transfieren a la hoja de registro de cada estudiante. Estos son guardados juntos en un fólter por separado (el mismo proceso se puede seguir con hojas de etiquetas, que funcionan mejor porque se pegan más firmemente; sin embargo, solo pueden extraerse una vez.) Las siguientes observaciones abarcan desde actitudes generales hacia las matemáticas hasta comportamientos de escritura. Este sistema puede emplearse para enfocar las observaciones en comportamientos en cualquier área de aprendizaje.

<p style="text-align: center;">Alicia</p> <p>19/01: Empezó a leer capítulos más difíciles de libros, de manera competente y con mínima ayuda... ¡Es difícil, pero ella puede hacerlo!</p>	<p style="text-align: center;">Braulio</p> <p>11/01: Hablamos acerca de dejar los videojuegos y esforzarse al máximo; no puede avanzar porque está atrapado en la dinámica de copiar con un amigo. Mostró remordimiento, estuvo de acuerdo conmigo, y se comprometió a trabajar más.</p>	<p style="text-align: center;">David</p> <p>15/01: Necesita ir al baño con mucha frecuencia. Él también lo notó; dijo que hablará con sus padres al respecto.</p>	<p style="text-align: center;">Daniel</p> <p>15/01: Toma con seriedad la responsabilidad de su propio aprendizaje. Se quedó en el aula durante el recreo y terminar le tomó solamente 5 minutos. Bien hecho.</p>
<p style="text-align: center;">Elizabeth</p> <p>12/01: ¡Toda la mañana distraída! Estuvo moviéndose de un lado a otro, interrumpiendo, desconectada de las tareas, haciendo ruidos y otros sonidos. Fue necesario repetir las instrucciones. Conversamos sobre el comportamiento respetuoso.</p>	<p style="text-align: center;">Randy</p> <p>19/01: Llegó puntual todos los días de la semana... ¿Tal vez ya se superó el problema de las tardanzas?</p>	<p style="text-align: center;">Susana</p> <p>11/01: Una gran ayuda para poner al día a Tony que volvía de estar enfermo. Buena "profesora"; explica los temas con sus propias palabras, y está lista para ayudar a su "alumno".</p>	<p style="text-align: center;">Nancy</p>

⁵ Campbell Hill, B. y Ruptic, C. (1994). *Practical Aspects of Authentic Assessment*. Norwood MA: Christopher-Gordon Publishers, Inc., pp. 55-60.

Enfocar y registrar observaciones

Pegar post-its y enfocar hojas de registro

El sistema de post-its puede combinarse con una hoja de registro diseñada para enfocar las observaciones en resultados particulares de aprendizaje. Las observaciones hechas en post-its pueden transferirse a una hoja de registro que puede ser usada por varios estudiantes. Este ejemplo muestra cómo las observaciones de desempeño de los estudiantes pueden registrarse en relación con cuatro resultados enfocados en ciencia.⁶

Clase		Actividad / proyecto		
	Resultados enfocados			
Estudiante	Planifica la investigación	Conduce la investigación	Utiliza la ciencia	Actúa responsablemente
Beth	15/6 Utiliza mapas de conceptos para hacer un plan		17/6 Utiliza apropiadamente los equipos	17/6 Ayuda a otros estudiantes que tienen problemas con el mechero Bunsen.
James		17/6 Fija dos secuencias alternativas	17/6 Hace una comparación metodológica	



⁶ Adaptado de Deleuil, I. y Malcolm, C. (1994). *Using the Science Profile*. Carlton: Curriculum Corporation, p. 73.

Observaciones enfocadas y registradas

Lista de clase

Leon Deleuil ha desarrollado una lista de verificación para registro de observaciones en el desempeño del estudiante.⁷ En este ejemplo, los estudiantes crean los instrumentos musicales que luego tocan. Los aspectos del desempeño de cada estudiante (por ejemplo, el criterio de evaluación) en el que León se enfoca son mostrados al inicio del cuadro. Algunos de los criterios pueden ser evaluados en un contexto de evaluación estructurada; por ejemplo, el profesor puede fijar un tiempo para que el estudiante toque el instrumento para evaluar el criterio "obtiene el sonido". El criterio "seguridad con herramientas", sin embargo, será evaluado mientras el estudiante crea el instrumento. Los resultados meta son de dos niveles y de dos tipos del Perfil de Ciencias Australiano.

Nombre	Instrumento	Obtiene el sonido	Cambia la nota	Llama la atención	Seguridad con las herramientas
Beth	Zampoña	✓✓✓✓ ✓	✓✓✓✓ ✓	✓✓✓✓ ✓	✓✓✓✓ ✓
Brian	Xilófono de botellas	✓✓✓✓ ✓	✓✓✓✓ ✓	✓✓✓✓ ✓	
Buckie	Guitarra	✓✓✓✓ ✓	✓✓✓✓ ✓	✓✓✓✓ ✓	Necesita ser más cuidadoso con el cuchillo.
Choo	Flauta	✓✓✓✓ ✓	✓✓✓✓ ✓	✓✓✓✓ ✓	✓✓✓✓ ✓

⁷ Deleuil, I. y Malcolm, C. (1994). *Using the Science Profile*. Carlton: Curriculum Corporation, p. 47.

2.3. Estimación y reporte del logro

En la evaluación del desempeño, el crecimiento se monitorea sobre el mapa de progreso que describe el desarrollo dentro de un área de aprendizaje particular. Las observaciones enfocadas en el salón de clase brindan a los profesores evidencia que puede usarse para monitorear y reportar el crecimiento del estudiante sobre dicho mapa. Mientras mayor sea el número de observaciones, mayor es la evidencia que el profesor tendrá para estimar el nivel de logro del estudiante.

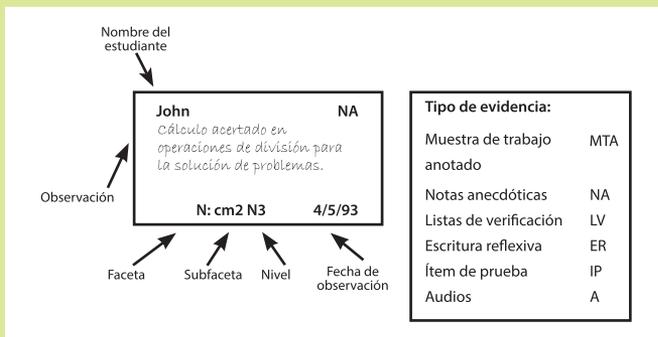
Los profesores que emplean el método Bainbridge Island para registrar, reportar y hacer seguimiento al crecimiento del estudiante en Escritura hacen un solo estimado final del nivel de logro del estudiante, indicando el nivel estimado con la fecha en la hoja de seguimiento (ver *Portafolios: Recursos de evaluación del Consejo Australiano para la Investigación Educativa 7*, pp. 44-45). Este sistema puede adaptarse para observaciones en otras áreas de aprendizaje. En el cuadro de la página 36, se muestra un ejemplo de una adaptación para las observaciones de desarrollo en actuación.

Las observaciones enfocadas en el salón de clase brindan a los profesores evidencia que puede usarse para monitorear y reportar el crecimiento del estudiante sobre dicho mapa. Mientras mayor sea el número de observaciones, mayor es la evidencia que el profesor tendrá para estimar el nivel de logro del estudiante.



Estimación y reporte de ubicación en un mapa de progreso

Lorelle Holcroft y Robyn Coates han desarrollado un sistema para hacer seguimiento al crecimiento de un estudiante sobre los niveles de Estándares de Desempeño del Estudiante en Queensland.⁸ Su sistema no permite específicamente evaluar el desempeño. Cada pieza de evaluación evidenciada tiene fecha y utiliza dos códigos. El primero indica el tipo de evidencia (ejemplo: MTA = ejemplo de trabajo anotado). El segundo es el código para estandarizar el desempeño del estudiante, indicando la faceta, la subfaceta y el trabajo.



Luego, estos registros se transfieren a una Hoja de Resumen que junta toda la evidencia disponible: tipo de evidencia (herramienta de registro), fecha de evidencia, faceta, subfaceta y nivel de desempeño para cada observación. Después, los profesores usan esta información para hacer una calificación final del logro de los estudiantes.

Subfaceta	Nivel 1	Nivel 2	Nivel 3	Nivel 4
Cálculo mental N: cm	1.15 Utiliza conteo y otras estrategias para resolver preguntas que involucran números pequeños formuladas oralmente.	2.15 Estima y calcula mentalmente, incluyendo sumas y restas de números hasta 10 y hace extensiones basándose en el valor.	3.15 Estima y calcula mentalmente, incluyendo sumas (hasta 100), restas de números de dos dígitos y multiplicación de números hasta 10.	4.15 Estima y calcula mentalmente, incluyendo adición y sustracción de la mayoría de números de dos dígitos y multiplica números hasta 10.
Herram. de registro			NA	
Fecha de evidencia			4/5/93	
Cálculo escrito N: ce		2.16 Utiliza una variedad de estrategias, incluyendo reagrupación para ayudarse en la suma y resta de los números que no puede completar mentalmente.	3.16 Utiliza métodos de comprensión escrita para sumar y restar cualquier número y monto de dinero y multiplicar y dividir todos los números hasta 10.	4.16 Utiliza métodos para sumar, restar y dividir todos los números y medidas (dos dígitos).
Herram. de registro		IP	MTA	
Fecha de evidencia		2/93	3/93	

⁸ Holcroft, I. y Coates, R. *Monitoring Learning in Mathematics*. Brisbane: Mount Coot-tha School Support Centre.

Estimar y reportar las ubicaciones en un mapa de progreso

El ejemplo aquí mostrado es la adaptación del método de Bainbridge Island para registrar el crecimiento del estudiante en escritura (ver *Portafolios: Recursos de evaluación del Consejo Australiano para la Investigación Educativa 7*, pp. 44-45). El criterio se toma para los niveles 3, 4 y 5 de la Faceta de Actuación del Perfil Curricular de Artes para escuelas australianas.

	Nivel 3	Nivel 4	Nivel 5
	<ul style="list-style-type: none"> Explora ideas y sentimientos a través de la actuación. Explora y utiliza varios elementos de actuación y utiliza habilidades específicas, técnicas y procesos, como improvisación y juego de roles. Planea y presenta actuaciones para una audiencia en particular o con algún propósito. 	<ul style="list-style-type: none"> Experimenta con ideas y explora sentimientos para encontrar soluciones satisfactorias a las tareas. Selecciona, combina y manipula elementos de actuación, utilizando diversas habilidades, técnicas y procesos. Muestra diversas habilidades para presentar una actuación para distintas audiencias y propósitos. 	<ul style="list-style-type: none"> Utiliza puntos de inicio como observaciones, experiencia e investigación para expresar ideas y sentimientos. Estructura la actuación organizando elementos de actuación y aplicando habilidades apropiadas, técnicas y procesos. Planea, selecciona y modifica presentaciones para ocasiones particulares tomando en cuenta factores como propósito, espacio, materiales y equipo.
1 Karley A.	9/92	1/93 →	
2 Mackenzie B.	9/92	1/93	
3 Paul B.	9/92	← 1/93	
4 Zach C.	9/92	1/93	
5 Lucas C.	9/92	1/93	
6 Win D.		9/92	
7 William D.	9/92	1/93	1/93
8 Tiffany E.			11/92
9 Clayton H.		11/92 →	
10 Nicole H.		12/92	1/93
11 Julie H.	9/92	2/93	
12 Molly J.	9/92	1/93	
13 Kirstin L.		9/92	2/93
14 Michelle M.			9/92
15 J.J. M.	9/92	9/93	¡NO HAY CRECIMIENTO!
16 Robbie O.	9/92	2/93	
17 Marian S.		9/92	2/93
18 Ryan S.	9/92		2/93 →
19 Patrick S.	9/92		2/93
20 Alex W.	9/92		2/93
21			CON AYUDA
22			
23			

3. "EVENTOS" DE EVALUACIÓN DE DESEMPEÑO EN EL SALÓN DE CLASES

En ocasiones, los profesores hacen que los estudiantes realicen presentaciones con la intención de evaluar su desempeño. Algunos ejemplos incluyen una presentación de una composición musical original, un juego de roles en una conversación en otro idioma, una producción actoral o de danza, una evaluación de técnicas de atletismo y una presentación oral de los resultados de una investigación de historia local.

Las evaluaciones de desempeño de este tipo son diferentes de las observaciones informales en el salón hechas en el transcurso de la enseñanza (ver las páginas 23 a 36).

También difieren de las evaluaciones estructuradas diseñadas para contextos de alto rendimiento, donde se hacen esfuerzos especiales para asegurar la comparación de evaluaciones de estudiante a estudiante, evaluador a evaluador y escuela a escuela (ver páginas 60 a 75).

Los profesores utilizan eventos de evaluación del desempeño en lugar de observaciones informales para recoger información que es difícil de ver y recolectar eficientemente en el momento adecuado.

Esta es la ocasión de hacer evaluaciones *planeadas* para la observación de estudiantes *involucrados en una actividad*. El profesor selecciona un contexto para evaluar y decidir, por adelantado, un método para calificar el desempeño de los estudiantes. Estas calificaciones brindan evidencia sobre los resultados alcanzados en un área de aprendizaje.

El criterio de evaluación ha sido desarrollado para enfocar las observaciones de los profesores en tipos particulares de evidencia y utiliza un sistema de registro de observaciones.

3.1. Planeación de eventos de evaluación de desempeño

La utilidad de una presentación o desempeño como un recurso de evidencia sobre resultados alcanzados en un área de aprendizaje depende de:

- Cuán apropiada es la evidencia para dirigir las metas instruccionales, o “resultados”, de un área de aprendizaje (relevancia).
- Cuán apropiada es la tarea de desempeño para ejemplificar el rango de resultados en área de aprendizaje (cobertura).
- Cuán apropiada es la tarea para permitir que los estudiantes muestren sus fortalezas y debilidades.
- Cuán justa es la tarea para los estudiantes de diferentes lenguajes, culturas e historial sociocultural.

Los “eventos” de evaluación de desempeño en el salón de clases son evaluaciones *planeadas* para la observación de estudiantes *involucrados en una actividad*. El profesor selecciona un contexto para evaluar y decidir, por adelantado, un método para calificar el desempeño de los estudiantes. Estas calificaciones brindan evidencia sobre los resultados alcanzados en un área de aprendizaje.

El manejo de tareas de desempeño en el salón de clases también debe tomarse en cuenta. Una tarea que ofrece excelente información sobre el nivel del logro del estudiante, pero requiere que la atención del profesor se dirija a uno solo de ellos por un largo período, puede ser impracticable en el contexto del salón.

Dirigir resultados

Cuando se evalúa el desempeño del estudiante, es importante aclarar la dirección de los resultados. Las tareas pueden estar apropiadamente

estructuradas y las observaciones enfocadas para la recolección de evidencia, incluyendo resultados de diferentes niveles en un mapa de progreso. Un evento de evaluación es más útil cuando permite la recolección de evidencia sobre un rango de resultados.

Planeamiento de eventos de evaluación de desempeño

Al completar la unidad de trabajo, los estudiantes de Terry Spottiswood desarrollaron una tarea de oratoria como evaluación final.¹ Ellos presentaron libros que habían escrito y reportaron el proceso a una audiencia con compañeros de clase, estudiantes menores y padres de familia. Terry enfocó su evaluación en los resultados observados desde estructuras "lingüísticas" y características de subfacetas de la faceta de lenguaje oral y comprensión auditiva del perfil de Inglés. ¿Pueden los estudiantes controlar "la mayoría de las estructuras lingüísticas y características del lenguaje hablado para interpretar significados y desarrollar y presentar ideas e información en situaciones familiares"? Terry busca que las habilidades de los estudiantes:

- Despierten y mantengan el interés de la audiencia.
- Utilicen oraciones completas en el lenguaje hablado.
- Varíen el estilo, el tono y el ritmo.
- Se sobrepongan al nerviosismo.

Ella realiza una evaluación de inmediato utilizando un sistema de vistos para indicar el grado de competencia: un visto: *bueno*; dos vistos: *muy bueno*; tres vistos: *excelente*; cuatro vistos: *extraordinario*. Terry también deja espacio en su hoja de registro para otros comentarios.

Dirigir resultados

Nombre	Despierta y mantiene interesada a la audiencia	Utiliza oraciones completas	Varía estilo, tono y ritmo	Se sobrepone al nerviosismo	Otros comentarios
Jody	✓✓	✓✓	✓✓	-	
Nam Huong	✓✓	-	✓✓	✓	gran mejora muy nervioso
Tom	✓✓✓✓	✓✓	✓	-	maravilloso
Gino	✓	✓	✓	-	bien secuenciado buenos puntos

¹ Terry Spottiswood, Waldau Primary School, Victoria.

3.2. Asegurar una profunda evidencia

Los profesores usan presentaciones y desempeños como oportunidades para recolectar información sobre las habilidades y talentos ocultos de los estudiantes. A menudo, los temas particulares en que los estudiantes realizan sus presentaciones, piezas de música en particular o rutinas particulares de danza que presentan son menos importantes que las habilidades y talentos ocultos sobre los cuales el desempeño brinda evidencia. Mientras más oportunidades tenga el profesor de observar el desempeño de los estudiantes, más confiable será la calificación de su habilidad oculta. Las observaciones de desempeño del estudiante necesitan ocurrir en un *número de ocasiones*.

Las observaciones de desempeño del estudiante necesitan ocurrir en un *número de ocasiones* y en un *rango de contextos*.

Las observaciones de desempeño necesitan ocurrir en un *rango de contextos*. Un profesor que evalúe el lenguaje oral puede observar al estudiante en diferentes tipos de desempeños orales, por ejemplo, discusiones de pequeños grupos o discursos formales con distintas audiencias. Antes de ello, el profesor debe estar preparado para realizar una calificación sobre el nivel de rendimiento oral del estudiante.

Pensando sobre este rango de contextos, los profesores necesitan considerar si las evaluaciones corresponden al *mismo tipo*

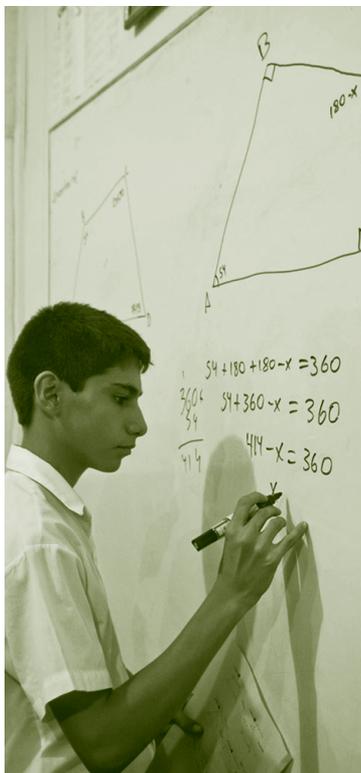
de habilidad o talento oculto. ¿Tiene sentido inferir algo en general sobre la habilidad de un estudiante para "lanzar la pelota" a partir de desempeños en contextos tan diferentes como básquetbol, tenis y fútbol, por ejemplo?

Considerando la fiabilidad

También es importante tener conciencia de la forma en que el género, historial cultural y personalidad pueden influir en el desempeño del estudiante. Los profesores necesitan considerar el contexto de eventos de desempeño y criterios de evaluación para asegurarse de que los indicadores de logro sean justos.

Planeamiento de eventos de evaluación de desempeño	Considerando fiabilidad
<p>El desempeño de los estudiantes puede verse influido por características de la situación de la evaluación irrelevantes para los resultados que van a ser evaluados.</p> <p>Algunos ejemplos:</p> <p><i>Género</i></p> <p>¿Las niñas se desenvuelven igual de bien en actividades físicas si están jugando en equipos mixtos?</p> <p>Steve Quartermaine encontró que generalmente las niñas muestran menor nivel de habilidades físicas cuando son evaluadas en equipos mixtos.²</p> <p><i>Historial cultural</i></p> <p>¿Algunas prácticas culturales ponen en desventaja a los estudiantes en evaluaciones desarrolladas?</p> <p>Una evaluación de desempeño oral puede requerir "mantener contacto visual" como criterio de evaluación. Esto puede resultar inapropiado para algunos grupos culturales en los cuales no se suele hacer contacto visual.</p>	

² Conversación con Steve Quartermaine, *Monitoring Standards in Education*, Department of Education, Western Australia, 1995.



Permitir un rango de niveles de desempeño

Los eventos de evaluación de desempeño deben ser lo suficientemente indefinidos como para permitir a los estudiantes de bajo rendimiento que experimenten el éxito, mientras que a otros se les exige —y se les permite— demostrar altos niveles de logro.

Un evento de evaluación es más útil cuando permite recoger la evidencia a partir de diversos niveles de resultados.

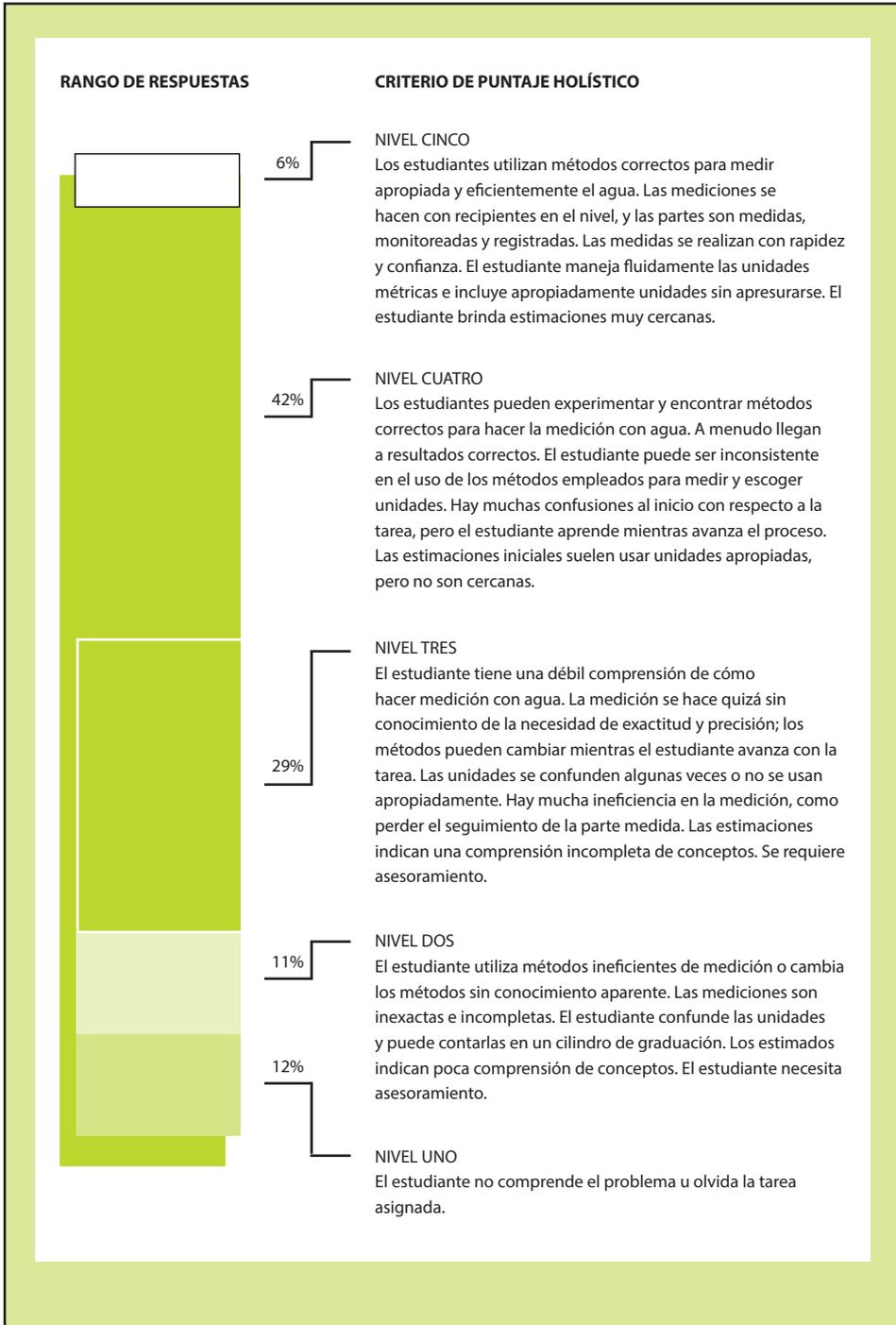
En el ejemplo de abajo, las habilidades de los estudiantes de 6.^{to} grado se evalúan según su capacidad y volumen a partir de cinco niveles de desempeño.

Planeamiento de eventos de evaluación de desempeño

Permitir a los estudiantes que muestren diversos niveles de desarrollo

La Junta de Educación de Toronto tiene como tarea “purificar el agua”. La tarea es lo suficientemente indefinida como para que pueda lograrse de distintas formas que reflejen diferentes niveles de desempeño.³ La Junta define cinco niveles de desempeño en la tarea y muestra el porcentaje de estudiantes que desarrollan cada nivel.

³ Toronto Board of Education, Benchmark “Pouring Water” Task.



3.3. Calificación y registro

El criterio para calificar el desempeño del estudiante y la forma en la cual se registran las calificaciones dependen del área de aprendizaje, de los resultados dirigidos y del contexto de evaluación inmediata.

Los ejemplos en las páginas 44 a 54 muestran el criterio de evaluación y los métodos de registro usados en diferentes tipos de eventos para la evaluación de desempeño. En estos ejemplos, los profesores, los compañeros o el mismo estudiante han realizado las calificaciones, ya sea de forma analítica u holística.

Calificación y registro	Calificación analítica (autoevaluación)																					
<p>Lorelle Holcroft y Robyn Coates sugieren que los estudiantes evalúen su competencia y comprensión sobre un conjunto de tareas y participaciones en una actividad de solución de problema de grupo.⁴ Esta autoevaluación puede formar la base para una futura reunión del profesor.</p>																						
<p>Nombre: _____</p> <p>Actividad: _____</p> <p>Trabajé en un grupo de personas.</p> <p>Los otros miembros de mi grupo fueron:</p> <p>Mientras resolvía el problema, yo:</p> <ul style="list-style-type: none"> • Me sentí confiado • Me divertí con lo que estaba haciendo • Hice sugerencias • Escuché lo que otros en mi grupo decían • Participé en discusiones de grupo • Ayudé a otros en mi grupo • Me rendí fácilmente / Seguí intentando 																						
	<table border="1"> <thead> <tr> <th colspan="2">SÍ</th> <th>NO</th> </tr> </thead> <tbody> <tr> <td>PARA NADA</td> <td>A VECES</td> <td>SIEMPRE</td> </tr> <tr> <td>PARA NADA</td> <td>A VECES</td> <td>SIEMPRE</td> </tr> <tr> <td>NUNCA</td> <td>A VECES</td> <td>CUANDO ERA APROPIADO</td> </tr> <tr> <td>NUNCA</td> <td>A VECES</td> <td>SIEMPRE</td> </tr> <tr> <td colspan="2">NUNCA</td> <td>CUANDO PUDE</td> </tr> <tr> <td colspan="2">ME RENDÍ FÁCILMENTE</td> <td>SEGÚÍ INTENTANDO</td> </tr> </tbody> </table>	SÍ		NO	PARA NADA	A VECES	SIEMPRE	PARA NADA	A VECES	SIEMPRE	NUNCA	A VECES	CUANDO ERA APROPIADO	NUNCA	A VECES	SIEMPRE	NUNCA		CUANDO PUDE	ME RENDÍ FÁCILMENTE		SEGÚÍ INTENTANDO
SÍ		NO																				
PARA NADA	A VECES	SIEMPRE																				
PARA NADA	A VECES	SIEMPRE																				
NUNCA	A VECES	CUANDO ERA APROPIADO																				
NUNCA	A VECES	SIEMPRE																				
NUNCA		CUANDO PUDE																				
ME RENDÍ FÁCILMENTE		SEGÚÍ INTENTANDO																				

⁴ Holcroft, L. y Coates, R. *Monitoring Learning in Mathematics*. Brisbane: Mount Coot-tha School Support Centre.

Cuando estaba trabajando en este problema, pude demostrarle a mi profesor que yo sabía...

Además de los miembros de mi grupo, yo hablé sobre este problema con...

En una escala de puntuación (A B C D E), por mi reporte, creo que debo recibir

○

Calificación y registro

**Calificación analítica
(autoevaluación)**

El proyecto de Evaluación de Aprendizaje de Connecticut Common Core ha desarrollado una autoevaluación para evaluar la participación en una actividad de grupo.⁵ Los estudiantes circulan sus autocalificaciones a cada una de las personas del grupo para su revisión y firma al final de la actividad. Si algún miembro del grupo está en desacuerdo con la autocalificación del estudiante, se discuten sus razones. Luego, el estudiante decide si cambiar o no el puntaje original.

Nombre del estudiante: _____

Número de código del estudiante: _____

	Marcar solo uno			
A. PARTICIPAR EN EL GRUPO	Casi siempre	Seguido	Algunas veces	Pocas veces
1. Participa en las discusiones del grupo sin apresurarse.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Hizo su parte justa del trabajo.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Trató de dominar el grupo, interrumpió a otros, habló demasiado.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Participó en las actividades del grupo.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

⁵ Collison, Judith (1992). "Using Performance Assessment". *The Arithmetic Teacher*, Vol. 39. N.º 6, p. 45.

B. MANTENERSE EN EL TEMA

	Casi siempre	Seguido	Algunas veces	Pocas veces
5. Prestó atención, escuchó lo que se dijo y atendió a lo que se hizo.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Hizo comentarios para que el grupo se centrara en el tema.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Se salió del tema o cambió la asignatura.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Permaneció en el tema.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

C. BRINDAR IDEAS ÚTILES

	Casi siempre	Seguido	Algunas veces	Pocas veces
9. Brindó ideas y sugerencias que ayudaron al grupo.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. Ofreció comentarios y críticas constructivas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. Influyó en las decisiones y planes del grupo.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. Brindó ideas útiles.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

D. DEMOSTRAR CONSIDERACIÓN

	Casi siempre	Seguido	Algunas veces	Pocas veces
13. Realizó comentarios positivos y favorables sobre los miembros del grupo y sus ideas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14. Brindó reconocimiento y crédito a los demás por sus ideas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15. Hizo comentarios hostiles o desconsiderados sobre un miembro del grupo.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16. Fue considerado con los demás.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

E. INVOLUCRARSE CON LOS DEMÁS

	Casi siempre	Seguido	Algunas veces	Pocas veces
17. Mantuvo a los demás involucrados haciendo preguntas, solicitando contribuciones, o exigiendo a los demás.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
18. Trató de hacer trabajar al grupo en conjunto para alcanzar acuerdos grupales.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
19. Consideró seriamente las ideas de los demás.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
20. Involucró a los demás.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

F. COMUNICACIÓN

	Casi siempre	Seguido	Algunas veces	Pocas veces
21. Habló claramente, fue fácil de escuchar y comprender.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
22. Expresó ideas claramente y de forma efectiva.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
23. Se comunicó claramente.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

G. EXPERIENCIA EN GENERAL

	Casi siempre	Seguido	Algunas veces	Pocas veces
24. Este grupo me ayudó a mejorar mi comprensión sobre los problemas y las formas de resolverlos a diferencia de si lo hubiera hecho por mi cuenta.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
25. Trabajar con el grupo fue una experiencia divertida.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Calificación y registro

**Calificación analítica
(autoevaluación)**

Ian Fox brinda un ejemplo de autoevaluación general para jóvenes gimnastas en la escuela.⁶ Los estudiantes califican su propio desempeño en una escala de cinco puntos, desde *bajo* a *alto*, en siete partes. Se les pide que comenten si les gusta la gimnasia y que reflexionen acerca de en qué necesitan mayor práctica. A los estudiantes se les podría pedir, sobre la base de estas evaluaciones en cada aparato, hacer una evaluación equilibrada sobre su desempeño general en gimnasia.

Gimnasia para niños

	BAJO				ALTO
Viga	<input type="checkbox"/>				
Trampolín	<input type="checkbox"/>				
Caballote	<input type="checkbox"/>				
Cuerdas	<input type="checkbox"/>				
Volantines	<input type="checkbox"/>				
Aros	<input type="checkbox"/>				
Colchoneta	<input type="checkbox"/>				

Firmado _____

Profesor: _____

⁶ Ian Fox, Buckland Beach Intermediate School, Auckland, Nueva Zelanda.

Calificación y registro**Calificación analítica (autoevaluación y evaluación del profesor)**

Ian Fox ofrece este ejemplo de una evaluación combinada realizada por el profesor y el alumno para el desempeño en música utilizando instrumentos de percusión.⁷ Los profesores y estudiantes puntúan el desempeño en cinco criterios utilizando una escala de tres puntos: *sigue intentando, tiene dominio, es competente*.

Habilidades musicales

	Sigue intentando		Tiene dominio		Es competente	
	estudiante	profesor	estudiante	profesor	estudiante	profesor
TONO		✓	✓			
TIEMPO					✓	✓
UTILIZA BAQUETAS					✓	✓
LEE MÚSICA					✓	✓
SUAVIDAD				✓	✓	

Calificación y registro**Calificación analítica (evaluación del profesor)**

Las reuniones y entrevistas estructuradas son otro tipo de eventos de evaluación que permiten calificar resultados de aprendizaje específicos. Una reunión sobre lectura estructurada, por ejemplo, brinda al profesor la oportunidad de escuchar a un estudiante leer en voz alta un texto seleccionado y evaluar su habilidad de lectura. Cindy Fulton ha desarrollado el siguiente cuadro para evaluar las estrategias de lectura del estudiante, fluidez, inflexión y comprensión de lectura.⁸ Un visto indica que la habilidad se muestra a veces; un signo más indica que la habilidad se muestra *constantemente* durante la reunión. Cuando el profesor utiliza esta evaluación en diferentes reuniones de lectura con un estudiante, puede documentar y ver la mejora a través del tiempo.

⁷ Ian Fox, Buckland Beach Intermediate School, Auckland, Nueva Zelanda.

⁸ Campbell Hill, B. y Ruptic, C. (1994). *Practical Aspects of Authentic Assessment*. Norwood, MA: Christopher-Gordon Publishers, Inc., p. 109.

Registro de lectura

Estudiante: _____

	Fecha: 8/12	Fecha: 17/12	Fecha:	Fecha:
	Título: <i>Dos pequeños perros</i>	Título: <i>Durmiendo</i>	Título:	Título:
Nivel Apropiado	<i>Frustración instruccional</i>	<i>Independiente de la instrucción</i>		
Estrategias: Idea completa				
Idea con dibujos	✓			
Patrones				
Palabras vistas	✓	✓		
Primera letra				
Decodificación	✓	✓		
Idea con contexto	✓	✓		
Saltar / regresar				
Volver a leer	✓	✓		
Lee con fluidez		✓		
Con inflexión		✓		
Comprensión literal	✓	+		
Comprensión interpretativa	✓	+		
Estrategia enseñada	<i>Contracciones Contexto</i>	<i>Contracciones Comillas</i>		
Comentarios				

CÓDIGOS: + consistentemente
✓ a veces

Calificación y registro**Calificación analítica
(evaluación del profesor)**

En este esquema de puntuación analítica para el desempeño de un grupo musical, se evalúan ocho aspectos de cada participación en el desempeño del estudiante. Cada aspecto (criterio) es calificado en una escala de cuatro puntos: *no evidente*, *limitadamente competente*, *competente*, *muy competente*. Se añade un espacio para que el profesor pueda hacer comentarios.

ASPECTOS DE RENDIMIENTO MUSICAL	No evidente	Limitadamente competente	Competente	Muy competente	Comentarios
1 Técnica de desempeño					
2 Control de presentación					
3 Calidad del tono					
4 Variación del tono					
5 Articulación y parafraseo					
6 Habilidad en presentación como miembro del grupo					
7 "Musicalidad" individual					
8 Presentación					

Calificación y registro**Calificación holística
(evaluación del profesor)**

Esta guía para evaluar el desempeño de los estudiantes en una tarea de discusión de un grupo pequeño DART (Recursos de Evaluación de Desarrollo para Profesores; DART, por sus siglas en inglés), brinda a los profesores niveles descritos de desempeño que coinciden con los niveles 2, 3, 4 y 5 del Perfil de Inglés para Colegios Australianos.⁹ Los profesores observan el trabajo individual de los estudiantes en un grupo pequeño y realizan una calificación final de su nivel de desempeño.

⁹ Forster, M.; Mendelovits, J. y Masters, G. N. (1994). *Developmental Assessment Resource for Teachers DART English*. Camberwell: Australian Council for Educational Research, p. 84.

Nivel 5

- Dirige y resume de manera implícita o explícita discusiones en grupo y coordina decisiones sobre la organización del desempeño.
- Sus puntos de vista son relevantes y persuasivos, y reflejan algún conocimiento de cómo los puntos de vista de diferentes personas se ven influenciados.
- Puede escoger las implicaciones de contribuciones de otro miembro de grupo y las articula.
- Coordina decisiones y organiza tareas donde es apropiado.

Nivel 4

- Toma el liderazgo al organizar la discusión del grupo, anima a los demás y dirige y resume el progreso.
- Ofrece sus propias opiniones enérgicamente (pero no necesariamente es dominante).
- Es articulado y persuasivo en la forma en la que se expresa por sí mismo.
- Tiene una función sustancial en la organización del desempeño en poesía, o copiando los roles de la televisión haciendo reportes, o en discusiones sobre espectáculos de mascotas.

Nivel 3

- Inicia algunas partes de la discusión y trata de hacer que los demás tomen parte.
- Ofrece sus propias opiniones sobre el tema que tratan de explicar.
- Escucha responsablemente otros puntos de vista.
- Su lenguaje tiene originalidad o precisión.

Nivel 2

- Comprende la tarea y coopera con el grupo.
- Sus puntos de vista sobre el tema son obtenidos y no voluntarios.
- Formula preguntas para comprender lo que dicen los demás.
- Utiliza un lenguaje simple y adecuado para comunicar sus ideas.

Calificación y registro

Calificación analítica (evaluación del profesor)

Neryl Jeanneret brinda el siguiente ejemplo de una escala de puntuación analítica para el desempeño de actuación.¹⁰ Se evalúan ocho aspectos utilizando una escala de cuatro puntos, desde *completamente en desacuerdo con que la oración sea descriptiva* hasta *conuerdo completamente con que la oración es descriptiva*.

- Escala de puntuación
- CD completamente en desacuerdo con que la oración sea descriptiva.
 - D en desacuerdo con que la oración sea descriptiva.
 - C conuerdo con que la oración es descriptiva.
 - CC conuerdo completamente con que la oración es descriptiva.

ASPECTOS DE DESEMPEÑO PARA ACTUACIÓN	CD	D	C	CC
1 La presentación está bien ensayada				
2 La presentación no es mecánica como la de un robot				
3 El ítem está dentro de las capacidades técnicas del actor				
4 La presentación es confiable				
5 El estilo/género es retratado exitosamente				
6 El actor tiene control de la escena				
7 La presentación es artísticamente exitosa				
8 Hay un compromiso original en la interpretación				



¹⁰ Killen, R. (1993). *Assessment and Evaluation*. Newcastle: University of Newcastle, Faculty of Education, p. 95.

Calificación y registro

Calificación analítica (evaluación del profesor)

El Procedimiento de Evaluación de Necesidad del Estudiante (SNAP, por sus siglas en inglés) ha sido desarrollado para apoyar a los profesores en las evaluaciones formales del trabajo escrito y oral de los estudiantes no angloparlantes.¹¹ Las actividades de evaluación del lenguaje oral para estudiantes de escuela secundaria incluyen discusiones en grupo, debates formales, entrevistas, conocimiento del procedimiento, argumentos orales, descripción oral y narrativa oral. En la evaluación del desempeño en la tarea de argumentación oral de la SNAP, los profesores califican mediante un criterio de rangos. Se brindan notas con explicaciones para ayudar en la evaluación.

Actividad de Evaluación de Lenguaje Oral para Secundaria			ARGUMENTACIÓN ORAL		
Nombre del estudiante:	Nivel, grado:	Clase:	Fecha:		
Tarea:	Contexto:				
Criterio (marque la casilla apropiada)	Muy competente	Competente	Competencia limitada	Sin evidencia de competencia	Comentarios
Estructura esquemática. <i>El estudiante:</i> <ul style="list-style-type: none"> Hace una declaración inicial, presentando un avance del tema Toma una posición declaratoria, por ejemplo una tesis Presenta argumentos relevantes para respaldar su posición Respalda los argumentos con evidencia apropiada Anticipa y refuta el punto de vista opuesto (opcional) Resume la evidencia y apela a la acción (si es apropiado) 					
Característica del lenguaje Participantes. <i>El estudiante:</i> <ul style="list-style-type: none"> Brinda información detallada sobre los participantes Utiliza vocabulario específico del tema 					
Procesos y circunstancias. <i>El estudiante:</i> <ul style="list-style-type: none"> Utiliza una variedad de verbos (procesos) Utiliza tiempo apropiado Utiliza un rango de circunstancias 					

¹¹ Mincham, L. (1992). *ESL Students Needs Assessment procedures R-12*. Documento interno. South Australian Education Department.

<p>Tenor (efecto comunicativo) El estudiante:</p> <ul style="list-style-type: none"> Utiliza correctamente la persona gramatical Utiliza el lenguaje en una forma personal e interactiva, pidiendo el uso de la palabra, o usa un lenguaje neutral/ impersonal si debate analíticamente Utiliza modalidades para expresar obligación (cuando apela a una acción) o certeza (cuando apela a la lógica) Se comunica claramente y de modo audible Explica términos poco comunes Responde a la audiencia Habla de forma fluida Utiliza el lenguaje corporal para establecer o mantener la comunicación 					
<p>Cohesión y complejidad del lenguaje. El estudiante:</p> <ul style="list-style-type: none"> Utiliza términos de referencia Utiliza diversas conjunciones Utiliza oraciones complejas 					
<p>Precisión El estudiante:</p> <ul style="list-style-type: none"> Utiliza la gramática de forma precisa y sin omisiones <ul style="list-style-type: none"> Verbos (procesos) Artículos Preposiciones Orden de las palabras/estructura de las oraciones Utiliza la concordancia de forma precisa Se aproxima a la pronunciación y entonación del castellano Se autocorrige 					
<p>Estrategias, habilidades y respaldo El estudiante:</p> <ul style="list-style-type: none"> Utiliza información recolectada, muestra habilidades de investigación Completa la tarea independientemente (es decir, con mínimo apoyo) Alude a imágenes para mejorar el significado (opcional) 					
Comentarios generales (incluyendo actitud)					
<p>Calificación global (rodear con un círculo) bajo 1 2 3 4 5 alto (es decir, habilidad de escribir apropiadamente para cumplir el propósito y la audiencia)</p>					

3.4. Estimar y reportar la ubicación en un mapa de progreso

Los eventos de evaluación del desempeño brindan evidencia que puede ser utilizada para estimar la ubicación de un estudiante en un mapa de progreso. Esta evidencia puede tomar una variedad de formas, incluyendo los registros de los profesores de los comportamientos demostrados durante un evento de evaluación, calificaciones analíticas, calificaciones holísticas de presentaciones completas y registros de puntuación de las respuestas del estudiante.

Cuando los profesores mantienen los registros de los comportamientos o resultados listados en diferentes niveles en el mapa de progreso, pueden usarlos para tomar decisiones finales sobre el nivel de logro general del estudiante (ver página 56).

Las decisiones sobre el nivel de logro del estudiante también pueden realizarse sobre la base de calificaciones analíticas y holísticas del desempeño, particularmente si los niveles de puntuación se alinean directamente con los niveles descritos en un mapa.

En el ejemplo de la página 57, el desempeño de los estudiantes en fútbol es calificado según dos criterios (es decir, de forma analítica), utilizando un juego de tres niveles de descripción para corresponder a los niveles 2, 3 y 4 del mapa de progreso de Educación Física en Australia Occidental.

Los niveles de logro de los estudiantes también pueden calcularse a partir de los registros de puntuación de sus desempeños en tareas asignadas. En el evento de evaluación de desempeño descrito en la página 58, los profesores asignan puntuación al desempeño de los niños en siete tareas por separado. Cada puntuación es de 0 a 3.

El puntaje total del niño (0 a 21) brindado estima su ubicación en un mapa de desempeño descrito para Habilidad Motriz Fina (ver página 59).

Estimación y reporte

Registros de comportamientos observados

Cindy Ruptic ha diseñado un formato para sus reuniones de lectura que incluye los Niveles de Lectura de Bainbridge Island (*Preconvencional, Emergente, En desarrollo, Inicial y En expansión*).¹² Durante la reunión, Cindy realiza notas detalladas en las habilidades de lectura del estudiante. Además, rodea con círculos los indicadores de desempeño. Luego, realiza una calificación final del desempeño de los estudiantes. En este ejemplo, Cindy ha encerrado en un círculo todos los indicadores de la etapa *En desarrollo*, y algunos en la etapa *Inicial*. La evaluación general de Cindy indica que este estudiante se encuentra en una etapa *En desarrollo*.

Fecha: 8/12 Título: Dos pequeños perros 16 páginas ¡Muy difícil!										
EVALUACIÓN DE LECTURA		NIVEL:		Frustración Instru <u>ccional</u> Independiente						
EVALUACIÓN DE LECTURA		Imprime conceptos; Nombra las letras; Se enfoca en las imágenes; Responde; Escoge/tiene favoritos								
EMERGENTE		Figuras para contar la historia; Nombres/palabras en contexto; Sonido de letras; Memoriza patrones; Ritmo								
EN DESARROLLO		<u>Lee solo</u> ; <u>Concepto de palabra</u> ; <u>Impresión + dibujos</u> ; <u>Palabras simples</u> <u>Patrón de libros</u> ; <u>Vuelve a contar la idea principal</u>								
INICIAL		Nivel aproximado; Susurra las palabras; <u>Se autocorrije</u> ; Puntuación; <u>Fónica</u> ; Estructura de la oración; <u>Contexto</u>								
EN EXPANSIÓN		Muchas estrategias; Capítulos cortos del libro; Expresiones; Vuelve a contar la historia; Conecta la experiencia								
Estrategias, Preguntas y Evidencia		Figuras	Patrones	Fónica	Palabras susurradas	Contexto	Errores	AC	CP	I
vacilante; trabajando a través del texto laboriosamente, pero con buena comprensión Del = de + el Al = a + el Donde + quiera = doquier		ratón madri- guera del conejo		ella él al doquier cerró		corrió correr	el (del) del (de él) Empieza a usar el contexto para autocorregirse	✓ ✓		
Estrategia/Tarea: <u>Contracciones (al, del...)</u> / utilizando pistas contextuales, contexto = con el texto ©1993 C. Ruptic, Bainbridge Island, WA. Códigos: AC = Autocorregido, CP = Corregido por el profesor, I = Ignorado										

¹² Campbell Hill, B. y Ruptic, C. (1994). *Practical Aspects of Authentic Assessment*. Norwood, MA: Christopher-Gordon Publishers, Inc., p. 113.

Estimación y reporte

Calificación analítica

Los profesores utilizan este Estándar de Monitoreo en Educación marcando dos aspectos de desempeño de cada estudiante.¹³ Los niveles descritos de desempeño corresponden a los niveles 2, 3 y 4 de las Declaraciones de Resultados de Australia Occidental, por eso ambos puntajes pueden usarse directamente para estimar el nivel de logro del estudiante en esta área curricular de Educación Física.

Habilidad en fútbol	Nivel de resultado		
	2	3	4
Dar un pase	Uso inconsistente de técnica correcta 1. Patea con la punta del dedo 2. Poca transmisión de peso 3. Se ve incómodo 4. Sin precisión	Técnica correcta y control sin presión 1. Patea con la parte interna del pie 2. Transmisión limitada de peso 3. Mantiene balance 4. Algunas veces sin precisión	Técnica correcta y control bajo presión 1. Utiliza los brazos para el balance 2. Transmisión eficiente de peso 3. Controla distancia y dirección 4. Es preciso
	Nivel de resultado		
	2	3	4
Recibir un pase	Rara vez controla la pelota 1. Pie rígido 2. Algunas veces controla la pelota sin presión	Controla la pelota con la técnica correcta sin presión 1. Le da con el pie 2. Mantiene balance 3. Controla la pelota en el piso	Controla la pelota con la técnica correcta bajo presión 1. La atrapa mientras corre 2. Controla la mayoría de los pases 3. Controla la pelota en el aire y en el piso

¹³ *Monitoring Standards in Education*, Department of Education, Western Australia.

Estimación y reporte**Respuestas calificadas**

Carol Mardell y Dorothea Goldenberg han desarrollado un instrumento de revisión conocido como *Indicadores de Desarrollo para la Evaluación del Aprendizaje* (DIAL, por sus siglas en inglés) que identifica a los niños en etapa pre jardín con necesidad de servicios de seguimiento y una evaluación de diagnóstico más detallada.¹⁴

Como parte del DIAL, a los niños se les administra un conjunto de tareas diseñadas para evaluar sus habilidades motrices finas (unir formas, construir con bloques, cortar tarjetas, copiar formas, copiar letras, tocar los dedos con el pulgar y aplaudir). Los evaluadores observan a cada niño y le asignan una puntuación de 0 a 3 en cada tarea (ver la puntuación para la tarea de tocar los dedos en el siguiente ejemplo).

Cada niño recibe un puntaje total que va del 0 al 21 en el juego de habilidades motrices finas. Este puntaje brinda un estimado de la posición del niño en un mapa de desarrollo de las habilidades motrices definidas por estas siete tareas (ver página 59). Por ejemplo, un niño con un puntaje de 15, ha logrado un nivel de habilidad motriz fina que le permite enlazar, al menos, algunas formas, construir torres de bloques de tres, cortar líneas rectas, y aplaudir dos veces.

Tocar los dedos

- a) El operador dice: "Está bien [nombre del niño], mira lo que yo hago".
- b) El operador toca cada dedo en una mano con el pulgar de la misma mano a un rango de uno por segundo.
- c) Luego el operador dice: "Ahora hazlo tú".
- d) El operador dice: "Ahora trata de hacerlo con tu otra mano", sin importar si el niño logró hacerlo con la primera mano.

Puntuación

- 3** El niño toca con el pulgar los dedos en secuencia (empezando en cualquier dirección) con ambas manos.
- 2** El niño toca con el pulgar los dedos en secuencia (empezando en cualquier dirección) con una mano.
- 1** El niño toca con el pulgar todos los dedos sin importar la secuencia con ambas manos.

¹⁴ Mardell, C. y Goldenberg, D. S. (1972). *DIAL Developmental Indicators for the Assessment of Learning*. Illinois: DIAL Inc.

4. CUANDO LA COMPARACIÓN IMPORTA

En situaciones de alto rendimiento, donde el desempeño de los estudiantes puede influir en la admisión a cursos, ofertas de becas o diplomas, es deseable la *legitimidad*, o comparación de estudiante a estudiante y de evaluador a evaluador.

En ocasiones, es importante que las evaluaciones del desempeño de los estudiantes sean comparables de estudiante a estudiante, de evaluador a evaluador y de escuela a escuela. El desempeño de los estudiantes de alto rendimiento puede influir en la admisión a cursos, ofertas de becas o diplomas. Los ejemplos incluyen exámenes orales de lenguaje al final del año, presentaciones instrumentales para la selección a un programa de música, evaluaciones para una beca de danza, evaluaciones para certificación en música. En estas situaciones de alto rendimiento se desea la *legitimidad*, o la comparación de estudiante a estudiante y de evaluador a evaluador.

Además, se han realizado esfuerzos especiales para asegurar la comparación de las evaluaciones del desempeño en programas que monitorean niveles de logro del estudiante a lo largo de sistemas educativos y del tiempo. Los ejemplos incluyen el uso de tareas de desempeño en ciencia en programas de evaluación nacional, como la Evaluación Nacional de Estados Unidos para el Progreso Educativo (NAEP, por sus siglas en inglés) y la de la Unidad de Evaluación de Desempeño de Reino Unido (APU, por sus siglas en inglés).

En situaciones en las que es importante la comparación, se debe dar una consideración especial a:

Planeamiento de evaluaciones de desempeño

- Coincidencia curricular: la medida en que las evaluaciones del desempeño abordan un amplio rango de metas de aprendizaje.
- Legitimidad: la medida en que las evaluaciones del desempeño son justas para todos los grupos de estudiantes.
- Estandarización: la medida en que las condiciones de evaluación son similares de estudiante a estudiante.
- Generalización: la medida en que las evaluaciones son generalizadas, más allá de las tareas particulares en las que se han basado.

Ejemplos de contextos en los que la evaluación del desempeño requiere una comparación de estudiante a estudiante, de evaluador a evaluador y de escuela a escuela:

Evaluaciones al final del año escolar en:

Estilos de danza

Presentación actoral individual

Presentación de música grupal

Examen oral de idiomas diferentes del inglés (LOTE, por sus siglas en inglés)

Evaluación musical

Campeonato de clavados

Audiciones musicales para admisión a la Escuela de Arte

Evaluaciones nacionales o estatales que involucran desempeño (por ejemplo, Estándares de Monitoreo en Educación Western Australia - Evaluación Oral, Evaluación Nacional de Estados Unidos para el Progreso Educativo y Evaluación Curricular Nacional de Reino Unido).

Cuando la comparación es importante, se debe tener especial consideración en:

- coincidencia curricular;
- legitimidad;
- generalización;
- estandarización;
- fiabilidad interevaluadores; y
- objetividad.

Calificación del desempeño

- **Fiabilidad interevaluadores:** la medida en que diferentes evaluadores son consistentes con el uso de su criterio de evaluación.

Resumen y reporte del logro

- **Objetividad:** la medida en que se especifican los procedimientos para hacer estimados finales de los niveles de logro del estudiante.

4.1. Planeamiento de la evaluación del desempeño

Cuando la comparación es importante, la evaluación del desempeño no solo brinda información sobre lo que los estudiantes saben y pueden hacer, sino que también sirve, por ejemplo, para tomar decisiones sobre qué calificaciones están recibiendo y si pueden recibir becas u obtener ingreso a cursos universitarios. Si se toma una decisión justa, entonces las evaluaciones de desempeño serán válidas y confiables.

Coincidencia curricular

Debido a que la evaluación del desempeño de alto rendimiento puede reducir —y posiblemente tergiversar— el aprendizaje

del estudiante, debe existir una correcta coincidencia entre las tareas de evaluación y el amplio rango curricular y metas de enseñanza. Un método para asegurar esta coincidencia es tener un panel de expertos que revise las tareas de evaluación.

Cuanto mayor sea el requisito de comparabilidad, más cuidadosamente:

- las tareas de evaluación están estructuradas;
- los criterios de evaluación están especificados; y
- se especifica el proceso para estimar los niveles de logro de los estudiantes.

Planeamiento de evaluación del desempeño

Coincidencia curricular

En el programa de Estándares de Monitoreo en Educación de Australia Occidental (MSE, por sus siglas en inglés) un grupo de profesores de educación física y especialistas en evaluación revisaron las evaluaciones de desempeño propuestas para asegurar la coincidencia entre las tareas y el rango de metas curriculares y de enseñanza. Las evaluaciones piloto se prueban cuidadosamente en grupos de estudiantes y profesores.

Legitimidad

Es importante que las tareas de evaluación del desempeño no beneficien o perjudiquen de manera injusta a estudiantes individuales o grupos particulares de estudiantes.

Los estudiantes deben tener los recursos físicos necesarios para completar las tareas de desempeño. Por ejemplo, una tarea que requiera hacer correr el agua o salidas de electricidad puede ser excluida de una evaluación de desempeño en ciencia, ya que puede poner en desventaja a estudiantes en algunos ambientes rurales.

Los estudiantes también deben ser evaluados en contextos que no les resulten poco familiares, pues ello podría afectar su desempeño de forma negativa. Por ejemplo, para hablar en público, un estudiante competente puede sentirse intimidado al tener que utilizar un micrófono por primera vez durante una evaluación oral.

Durante las evaluaciones de desempeño, los estudiantes son vulnerables a la retroalimentación inmediata de la audiencia de forma diferente que en otros tipos de evaluación (por ejemplo, evaluación escrita, de su portafolio o de un proyecto). Por esta razón, la evaluación del desempeño, en particular, debería ocurrir en ambientes de apoyo.

El desempeño puede verse afectado según quién esté realizando la evaluación (por ejemplo, los estudiantes con una tradición cultural particular pueden sentirse poco cómodos si son evaluados por adultos con otra tradición cultural), por quien se encuentre presente durante la evaluación (por ejemplo,

<p>Planeamiento de evaluación del desempeño</p> <p>En estudios internacionales, es importante que las tareas de evaluación no beneficien o perjudiquen de manera injusta a los estudiantes de diferentes grupos culturales, idiomáticos o socioeconómicos.</p> <p>Se proyectaron tareas de desempeño diseñadas para el IEA Tercer Estudio Internacional de Matemáticas y Ciencias (TIMSS, por sus siglas en inglés) por un panel internacional, para asegurar que las tareas sean igualmente significativas para estudiantes de diferentes procedencias. Las tareas piloto también fueron probadas cuidadosamente antes de ser aprobadas.</p> <p>Se excluyeron algunas tareas porque se requería acceso a recursos particulares, como agua caliente y fría y electricidad, lo que posiblemente no se encuentre disponible en algunas escuelas. Otras fueron excluidas si se notaba que ponían en desventaja a estudiantes urbanos (por ejemplo, las tareas que requerían conocimiento agrícola).</p>	<p>Asegurar legitimidad</p>

los estudiantes pueden tener una mala presentación ante compañeros poco empáticos), y por quien se encuentre incluido en la tarea de desempeño (por ejemplo, las niñas tienen un menor desempeño en contextos mixtos, a diferencia de contextos de un solo género).

La revisión de las tareas de evaluaciones propuestas y la prueba cuidadosa de los pilotos realizadas por expertos puede ayudar a asegurar la legitimidad y disminuir el margen de error.

Generalización

Aunque las evaluaciones del logro del estudiante se basan en observaciones del desempeño en tareas específicas, a menudo su propósito e interés se centra en algunas habilidades generales ocultas. El desempeño en un experimento práctico particular de ciencias, por ejemplo, puede ser de interés primario como indicador de habilidades generales del estudiante para utilizar procedimientos científicos.

Debido a que la evaluación del desempeño es compleja y consume gran cantidad de tiempo, con frecuencia es muy difícil observar a los estudiantes

Planeamiento de evaluación del desempeño	Generalización
<p>En un intento por asegurar qué conclusiones significativas puede mostrar el nivel de logro del estudiante desde sus desempeños en un número limitado de tareas, la Junta de Evaluadores de Música Australiana (AMEB, por sus siglas en inglés) muestra un amplio rango de desempeños del estudiante.</p>	
<p>Por ejemplo, en clase de violín para 5.^{to} grado, los estudiantes realizan cuatro estudios y piezas —elegidas de las cuatro listas—, más una lista extra de tres trabajos. Cada uno de los estudios y piezas requieren que los estudiantes toquen en diferentes estilos. Los evaluadores observan la técnica y comprensión de cada estilo. El criterio de evaluación incluye tono, entonación, musicalización, expresión y articulación.</p>	

en un número pequeño de tareas estandarizadas. Como resultado, hay cuestionamientos sobre:

- El número de tareas requeridas.
- La posibilidad de que el desempeño puede generalizarse a otras tareas similares es particularmente importante en evaluaciones de desempeño de alto rendimiento.

Los estudios de investigación muestran que en las evaluaciones de alto rendimiento, los estudiantes deben ser observados en más de una o dos tareas de desempeño. Por ejemplo, la investigación realizada por Shavelson, Baxter y Xiaohang Gao sugiere que sería necesario observar a los estudiantes en hasta quince ocasiones antes de obtener una medición precisa de algunos aspectos de rendimiento en ciencias.¹

Debido a que en ocasiones las tareas de desempeño complejas muestran áreas de conocimiento y habilidades particulares, se debe ser muy cuidadoso al dar un contexto más amplio (dominio) de desempeño en el cual la observación de desempeño de los estudiantes pueda generalizarse. Por ejemplo, ¿es posible concluir algo significativo sobre el nivel de logro del estudiante en educación física a partir de su desempeño en áreas tan diferentes como fútbol, tenis, básquetbol y atletismo?

Estandarización

La forma en la que una tarea de desempeño es presentada o explicada, así como el grado de ayuda que reciben los estudiantes para completar la tarea, puede afectar su desempeño.

¹ Shavelson, R. y Baxter, G. (1992). "What We've Learned About Assessing Hands on Science". *Educational Leadership*, Vol. 49, N.º 8, pp. 20-25; y Shavelson, R.; Baxter, G. y Xiaohang Gao (1993). "Sampling Variability of Performance Assessment". *Journal of Educational Measurement*, Vol. 30, N.º 3, pp. 215-232.

Planeamiento de evaluación del desempeño

Estandarización

El Tercer Estudio Internacional de Matemáticas y Ciencias (TIMSS, por sus siglas en inglés) brinda sesiones de capacitación para todas las personas encargadas de administrar las tareas y para observadores, para que la tarea del “circo” de ciencia práctica sea administrada en la misma forma en y a lo largo de todos los países. Esto significa que los estudiantes reciben la misma información introductoria e instrucciones, el mismo tipo de ayuda cuando se encuentran con problemas, y tienen el mismo tiempo para completar las tareas. El Manual de Administración describe la información que debe brindarse a los estudiantes y ofrece un texto introductorio sugerido.

Calificación de la evaluación del desempeño

Especificar criterio de evaluación

El criterio de evaluación para la tarea de ciencias del Reino Unido *investigar la flotación y el hundimiento* fue especificado de la siguiente manera:²

- Se da cuenta de la necesidad de usar las balanzas de cocina con un dial, porque no puede cuantificar la mínima diferencia de pesos entre los dos objetos utilizando las balanzas del salón de clases. Pesa los dos objetos hasta la división etiquetada más cercana, utilizando las balanzas tipo cocina con dial.
- Muestra un cuadro o tabla, o contribuye a su construcción, y registra acertadamente los descubrimientos para siete de ocho objetos.
- Brinda una descripción oral o escrita de lo que realizó, incluyendo cada una de las etapas en el orden correcto.
- Comunica de forma oral o escrita que tiene conocimiento de la diferencia entre una prueba *justa e injusta*.

² Department of Education and Science School Examinations Assessment Council (1991). *Assessment Record Booklet, Key Stage 1*. Lincoln: NFER-Nelson, p. 48.

En contextos en los cuales la comparación es importante, es necesario que exista la menor variación posible en los requerimientos hechos a los estudiantes y en las oportunidades que tienen para mostrar su rendimiento. Mientras mayor sea el requerimiento para una comparación, mayor será el cuidado en administrar las tareas con un desempeño más controlado o *estandarizado*.

Los evaluadores serán capacitados o se les brindarán instrucciones de administración estandarizadas para que los estudiantes reciban la misma información introductoria e instrucciones, los mismos tipos de ayuda cuando se encuentren con problemas y el mismo tiempo para llevar a cabo las tareas.

4.2. Calificación de evaluaciones del desempeño

Las evaluaciones del desempeño dependen del juicio humano. Dos personas que ven el mismo desempeño pueden reconocer diferentes cualidades y, por tanto, calificar de diferente forma.

Los evaluadores también pueden oscilar entre cualidades en el desempeño que no son centrales en la habilidad que está siendo evaluada. Por ejemplo, un gimnasta pequeño y delicado puede recibir una mejor calificación que un competidor físicamente más maduro por su atractivo físico y no por su desempeño superior.

Las calificaciones de desempeño pueden compararse a partir de:

- Definir cuidadosamente el criterio para evaluación del desempeño.
- Asegurar consistencia de criterio en la interpretación.
- Utilizar un número de calificaciones.
- Identificar y tomar en cuenta las diferencias sistemáticas sobre los evaluadores.

Especificar el criterio de evaluación

La evaluación del desempeño requiere calificar la calidad de —y, algunas veces, el proceso como se llegó a— una respuesta compleja. El criterio de evaluación debe estar bien concebido y ser definido de forma explícita, en un intento por minimizar la influencia de características irrelevantes del desempeño del estudiante o de la calificación del evaluador.

En ocasiones el criterio de evaluación es respaldado con ejemplos de desempeño de los estudiantes. Por ejemplo, los Puntos de Referencia de Toronto brindan a los profesores ejemplos en video sobre desempeños en el lenguaje oral.

Los profesores pueden practicar utilizando estos ejemplos, que están acompañados por comentarios detallados.

Calificación de la evaluación del desempeño	Diversos calificadores
<p>En la evaluación de danza de Victorian College of the Arts, cinco evaluadores observan el desempeño de cada estudiante. El desempeño de cada estudiante se califica en una escala de 0 a 100 a partir de los siguientes criterios:</p>	<ul style="list-style-type: none"> • precisión técnica; • habilidad rítmica y de musicalización; • uso de estilo y presentación; • habilidad para mantener ubicación y formación; y • habilidad para demostrar comprensión de su propio cuerpo (conocimiento cinético).
<p>Para calcular el puntaje en el examen del estudiante, se eliminaron los puntajes más altos y los más bajos. Los tres puntajes restantes fueron promediados. Si el resultado se encuentra entre dos números enteros, el puntaje es redondeado hacia arriba o hacia abajo, según el número entero más cercano.</p>	

Asegurar la consistencia de la interpretación

Para hacer calificaciones que puedan compararse, es necesario que los evaluadores desarrollen una comprensión compartida del criterio de evaluación. Brindar sesiones de capacitación en las que el evaluador revise y discuta ejemplos de desempeño del estudiante es un buen método para mejorar la comparación. Cuando los evaluadores tienen la oportunidad de trabajar juntos, es posible lograr una mejor consistencia en la interpretación del criterio a través de discusiones de los desempeños actuales.

En algunos programas de evaluación, pequeños grupos de evaluadores trabajan juntos por largos períodos, desarrollando una *cultura* compartida de evaluación.

Diversos evaluadores

Sin importar cuán claro sea el criterio de evaluación, siempre habrá preocupación por la subjetividad de las calificaciones. Una respuesta a esta preocupación es tener diversos evaluadores para juzgar el mismo desempeño y luego promediar sus calificaciones, posiblemente ignorando las puntuaciones más altas y las más bajas. A veces este método se usa para evaluar saltos de trampolín en natación, por ejemplo.

Adaptar las diferencias sistemáticas entre evaluadores

Otra forma de mejorar la comparabilidad de evaluaciones del desempeño es identificar y tomar en cuenta las diferencias sistemáticas entre los evaluadores. Las adaptaciones pueden realizarse considerando la tendencia del evaluador a ser demasiado severo o demasiado tolerante al momento de calificar el desempeño. Este proceso se llama, algunas veces, "control" de las diferencias entre los evaluadores (a diferencia de los procedimientos diseñados para asegurar que el evaluador tenga una similar comprensión del criterio y el proceso de evaluación).

Resumen y reporte del logro

Cuando los profesores de Reino Unido evalúan el desempeño de un estudiante utilizando la Evaluación Estandarizada de Tareas para “Explicación de Ciencia”, usan el criterio mostrado entre los niveles 1 y 3 de las metas logradas. Para estimar el nivel del estudiante en este mapa de progreso de Ciencias, los profesores usan la regla numérica ilustrada a continuación:³

Diversos calificadores

Nivel	Criterio logrado			
3			✓✓	✓✓✓✓✓
2			✓✓✓	
1	✓	✓✓		
Regla	Menos de dos criterios del nivel 1	Dos criterios del nivel 1	Cinco o más criterios de los niveles 2 y 3 mezclados	Seis o más criterios del nivel 3
Evaluación general	Trabaja hacia el nivel 1	Nivel 1	Nivel 2	Nivel 3

4.3. Resumir y reportar el rendimiento

Las calificaciones en las tareas de desempeño pueden usarse para realizar una evaluación general (final) del nivel de logro del estudiante en un área de aprendizaje. Los procedimientos del uso de las calificaciones del desempeño para hacer una evaluación general son paralelos a los procedimientos disponibles para el desempeño en el salón de clases (ver páginas 37 a 59).

El ejemplo en esta página muestra cómo las calificaciones de *criterio logrado* durante el desempeño de una tarea de ciencia son utilizadas en un programa

³ Department of Education and Science School Examinations Assessment Council (1991). *Assessment Record Booklet, Key Stage 1*. Lincoln: NFER-Nelson, p. 46.

de evaluación nacional para estimar los niveles en los cuales se encuentran trabajando los estudiantes.

Mientras los estudiantes se involucran en las tareas, los profesores puntúan el criterio según lo demostrado por ellos. Estos criterios se muestran según varios niveles de logro en las tareas de ciencia. Luego, los profesores usan una regla numérica para calcular el nivel general de logro del estudiante.

En algunas evaluaciones de desempeño de alto rendimiento, los evaluadores realizan calificaciones *holísticas* del desempeño que, posteriormente, se combinan con las demás para brindar un estimado de nivel de logro general del estudiante.

En ocasiones, los evaluadores califican diferentes aspectos de un desempeño del estudiante y luego combinan estas calificaciones *analíticas* para obtener una medida general (equilibrada) de logro. Por ejemplo, para estimar el nivel de logro para danza en alto rendimiento al final del año escolar, los profesores primero califican tres aspectos de presentación de danza de los estudiantes: coreografía, presentación y apreciación, cada una en una escala de 5 puntos, desde E hasta A.

En el momento de calificar la presentación, los profesores ponen atención a seis criterios (ver páginas 73-74): alineamiento, mecánica del cuerpo, conocimiento espacial, conocimiento rítmico, uso de estilo y comunicación en la presentación.

Sobre la base de estas calificaciones, los profesores luego obtienen un solo nivel de logro, desde *muy limitado* hasta *muy alto*, utilizando un proceso numérico. Por ejemplo, un estudiante calificado con una B (4 puntos) por coreografía, A (5 puntos) por presentación y B (4 puntos) por apreciación tendrá un acumulado de 13 ($4 + 5 + 4$), un nivel de logro *alto*. Ese proceso numérico se describe con detalle en la página 75.

Resumen y reporte de logro**Calificación analítica**

Para estimar el nivel del logro de estudiantes del último año de Danza en Queensland, los profesores primero evalúan tres aspectos de la presentación de la danza: coreografía, presentación y apreciación. Se usa una escala de cinco puntos, desde E hasta A, para cada una de estas tres evaluaciones. A continuación, se muestra la escala utilizada para evaluar la presentación. Los profesores pueden estimar un solo nivel de logro en danza utilizando el proceso numérico mostrado en la página 75.⁴

El estudiante muestra consistentemente:

A

- comprensión rigurosa de alineamiento en movimiento locomotor y no locomotor;
- fuerza, balance, coordinación y control en mecánica eficiente del cuerpo;
- alto desarrollo de conocimiento espacial;
- alto desarrollo de conocimiento rítmico;
- sensibilidad en el uso apropiado de estilo;
- energía adecuada y concentración al comunicarse con la audiencia durante la presentación.

El estudiante muestra la mayoría del tiempo:

B

- comprensión substancial de alineamiento en movimiento locomotor y no locomotor;
- fuerza, balance, coordinación y control en mecánica eficiente del cuerpo;
- buen desarrollo de conocimiento espacial;
- buen desarrollo de conocimiento rítmico;
- apropiado uso de estilo;
- energía adecuada y concentración al comunicarse con la audiencia durante la presentación.

El estudiante demuestra:

C

- una habilidad variable de alineamiento en movimiento locomotor y no locomotor;
- algo de fuerza, balance, coordinación y control en mecánica eficiente del cuerpo;
- conocimiento espacial en secuencias de movimiento simples;
- conocimiento rítmico en secuencias de movimiento simples;
- algún uso apropiado del estilo;
- inconsistente energía y concentración al comunicarse con la audiencia durante la presentación.

⁴ Board of Senior Secondary School Studies, Queensland. *Senior Dance*. Queensland: Department of Education.

El estudiante
demuestra
inconsistentemente:

D

- apropiado alineamiento del cuerpo;
- utiliza mecánica del cuerpo con fuerza, balance, coordinación y control;
- uso del conocimiento espacial en la mayoría de secuencias de movimiento;
- uso del conocimiento rítmico en la mayoría de secuencias de movimiento;
- uso de estilo apropiado;
- comunicación con la audiencia durante la presentación.

El estudiante rara
vez demuestra la
habilidad de:

E

- alineamiento del cuerpo;
- uso de la mecánica del cuerpo con fuerza, balance, coordinación y control;
- uso del conocimiento espacial en la mayoría de secuencias de movimiento;
- uso del conocimiento rítmico en la mayoría de secuencias de movimiento;
- uso de estilo apropiado;
- comunicación con la audiencia durante la presentación.

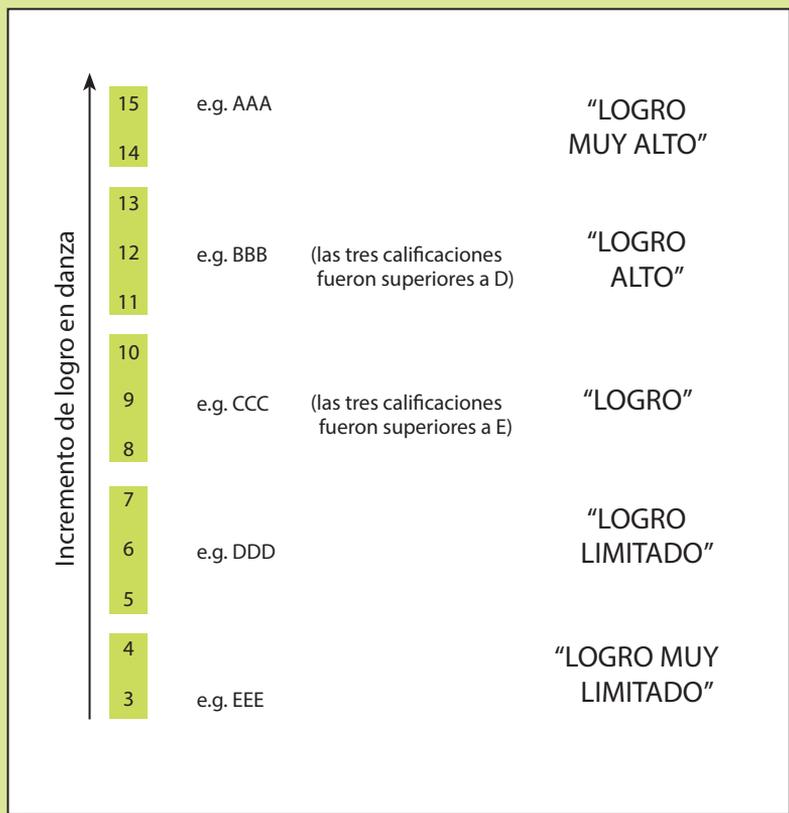


Resumen y reporte de logro

Calificación analítica

En Queensland, los profesores estiman un solo nivel de logro al final del año escolar en Danza, en una escala que va desde *logro muy limitado* hasta *logro muy alto*, utilizando el proceso numérico descrito a continuación. Los profesores obtienen un puntaje agregado de 3 a 15, al sumar los puntos otorgados en cada una de las tres áreas de danza (coreografía, presentación y apreciación).

Como máximo, un estudiante puede alcanzar cinco puntos por cada área. Por ejemplo, un estudiante con cuatro puntos (B) en coreografía, cinco puntos (A) en presentación y cuatro puntos (B) en apreciación tiene un agregado de 13 (4 + 5 + 4) y, por ello, un nivel de *logro alto*.



5. DISEÑANDO LA EVALUACIÓN DEL DESEMPEÑO

En la evaluación del desempeño, los profesores monitorean el progreso del estudiante en un mapa de progreso de desarrollo de habilidades, conocimientos y comprensión. Las evaluaciones del desempeño son un método que puede usarse para recolectar evidencia del logro del estudiante.

Hay muchos contextos para —y tipos de— evaluaciones de desempeño. Diferentes tipos de evaluación de desempeño brindan distintos tipos de evidencia que los profesores pueden usar para estimar los niveles de rendimiento del estudiante. Este artículo enlista los temas que los profesores deben considerar cuando diseñan una evaluación de desempeño. Los temas incluyen preguntas de propósito evaluador, métodos para calificar y registrar el desempeño del estudiante, y formas para estimar la ubicación de un estudiante en un mapa de progreso. Al final de este artículo, se incluye un resumen de listas de verificación del proceso de diseño.

5.1. Planeamiento de la evaluación del desempeño

¿Cuál es el propósito de la evaluación?

¿Evaluación formativa o sumativa? ¿Está interesado en hacer un monitoreo informal en desarrollo para propósitos de diagnóstico formativo o en desarrollar “eventos” de evaluación del desarrollo para propósitos de evaluación sumativa?

La forma como responda a estas preguntas determinará su método de evaluación del desempeño. Puede escoger utilizar la observación y registrar mediante estrategias como aquellas sugeridas en las páginas 23 a 36, o considerar el desarrollo de eventos de evaluación planeados. Puede escoger utilizar ambas estrategias: registrar evidencia de desempeño anecdótico y enfocar sus observaciones en resultados específicos con tareas planeadas ocasionalmente.

¿Qué metas curriculares o resultados recogerá como evidencia si piensa registrar observaciones informales del desempeño del estudiante?

¿Sus observaciones se enfocan en metas explícitas de enseñanza o resultados?

¿La evidencia que está recolectando es relevante?

Si usted está observando el desempeño de los “eventos” de evaluación:

¿Cuán bien mide la evaluación aquello para lo que fue diseñada (validación)?

¿Hay una coincidencia entre las tareas y el resultado de realización que está siendo evaluado?

¿Qué aspectos o detalles del desempeño observará? ¿Puede la evaluación ser utilizada para inferir un desempeño en otro contexto (generalización)?

¿Cuántas tareas utilizará?

¿Las tareas son justas para todos los estudiantes?

¿Las tareas permiten a los estudiantes mostrar sus fortalezas y debilidades? (¿son lo suficiente indefinidas?)

En resumen:

¿Cuál es el propósito de la evaluación del desempeño?

- ¿Formativo o sumativo?
- ¿Evidencia de qué meta curricular / resultado de mapa de progreso?
- ¿La evidencia es relevante?
- ¿Qué detalles se observarán?
- ¿Qué tareas y en qué cantidad?

5.2. Enfocar y registrar observaciones / Calificar el desempeño

Si está observando de manera informal el desempeño del estudiante:

¿En qué resultado se enfocará y cómo? ¿Observará una gran cantidad de resultados o se enfocará en un número pequeño?

¿A qué estudiantes observará y cómo? ¿Observará a un gran número de estudiantes o se enfocará en algunos pocos? ¿Utilizará técnicas particulares como “cuestionamiento enfocado” (ver página 27)?

¿Cómo registrará su observación? ¿Qué método utilizará (por ejemplo, fichas, registros computarizados, listas anotadas en clases, cuadros de observación, *post its*)? ¿Su método es eficiente? ¿Este le brinda información accesible para propósitos de enseñanza y para retroalimentación a estudiantes y padres de familia? ¿Le permite ver el crecimiento a lo largo del tiempo?

Si está observando eventos de evaluación del desempeño

¿Cómo calificará las presentaciones? ¿Realizará calificaciones holísticas o utilizará una guía de puntuación analítica?

¿Qué criterio de evaluación utilizará? ¿Es justo? (¿Ha considerado el género, la tradición cultural y de lenguaje del estudiante?). ¿Ellos se comunican fácilmente?

Si va a desarrollar tareas en un contexto donde la comparación del desempeño es particularmente importante, necesitará considerar cuán consistentemente la evaluación mide el desempeño del estudiante (confiabilidad). Necesitará controlar las condiciones de la evaluación, asegurar consistencia del criterio de interpretación (procedimientos de moderación y capacitación de evaluación) y utilizar un número de evaluadores (ver páginas 65-67).

En resumen:

¿Cómo enfocará y registrará las observaciones y cómo calificará los eventos de desempeño?

- ¿En qué resultados se enfocará y cómo?
- ¿A qué estudiantes observará y cómo?
- ¿Cómo registrará sus observaciones?
- ¿Cómo calificará el desempeño?
- ¿Qué criterio utilizará?
- ¿Cuán confiable es la evaluación?

5.3. Ubicar a los estudiantes en un mapa de progreso

¿Cómo estimará y reportará la ubicación de los estudiantes en un mapa de progreso?

¿Por nivel o ubicación en el cuadro? ¿Quiere colocar a un estudiante "en un nivel" y reportar este nivel, o ubicarlo en un continuo y reportar esta ubicación?

¿Estimación objetiva o subjetiva? ¿Quiere hacer un estimado subjetivo final o un estimado estadístico final de la ubicación del estudiante en un mapa de progreso?

En resumen:

¿Cómo calculará o reportará la ubicación de un estudiante en un mapa de progreso?

- ¿Por nivel o ubicación en el cuadro?
- ¿Con estimación objetiva o subjetiva?

6. UN RESUMEN DEL PROCESO DE DISEÑO DE LA EVALUACIÓN DEL DESEMPEÑO

Etapas de diseño de tareas para el desempeño	Estrategias de diseño de tareas
Decidir el propósito de evaluación del desempeño	<ul style="list-style-type: none"> • Describir el propósito de evaluación • Revisar estas descripciones sobre objetivos curriculares importantes y resultados de mapa de progreso en áreas de aprendizaje.
Decidir las metas curriculares o los resultados en los que se enfocará.	<ul style="list-style-type: none"> • Hacer una lista de metas o resultados. • Revisarlas sobre objetivos curriculares importantes y resultados de mapas de progreso en áreas de aprendizaje para revisar la relevancia.
Decidir las metas curriculares o los resultados en los que se enfocará.	<ul style="list-style-type: none"> • Decidir el número de "eventos" de desempeño. • Revisar los plazos definidos. • Revisar la generalización. • Revisar eficiencia, accesibilidad de información y habilidad para mostrar crecimiento sobre el tiempo. • Revisar legitimidad (género, cultura, lenguaje).
Decidir un método para enfocarse y registrar observaciones y calificar las tareas de desempeño.	<ul style="list-style-type: none"> • Decidir cómo enfocará los resultados. • Decidir a qué estudiantes observará. • Decidir cómo registrará sus observaciones. • Decidir cómo calificará el desempeño (criterio de evaluación). • Revisar para una fácil comunicación.
Decidir un método para estimar y reportar ubicaciones en un mapa de progreso.	<ul style="list-style-type: none"> • Describir el método para estimar la ubicación en un mapa de progreso. • Describir el método para reportar la ubicación en un mapa de progreso. • Comprobar estas descripciones sobre el propósito de la tarea y la audiencia.

Evaluación de desempeño Recursos de evaluación del Consejo Australiano para la Investigación Educativa 5

Durante la elaboración de estándares de aprendizaje en que estuvo abocado el SINEACE entre los años 2009-2015, se conoció la experiencia australiana de evaluación de los aprendizajes realizada por el Consejo Australiano para la Investigación Educativa (ACER). El SINEACE suscribió un convenio con el ACER para traducir los folletos que comprenden el Kit de Recursos de Evaluación, que fueron un valioso material durante el proceso; esto permite poner al alcance de los lectores de habla castellana el quinto número de la serie, el cual rescata la importancia de la evaluación de desempeño; es decir, la que se realiza cuando los estudiantes se involucran en una actividad.

Hay muchos contextos para —y tipos de— evaluaciones de desempeño, cada una de las cuales brinda distintos tipos de evidencia que los profesores pueden usar para estimar los niveles de rendimiento del estudiante. En la publicación se describen ampliamente los sistemas que los docentes pueden utilizar para registrar las observaciones.

SERIE DOCUMENTOS TÉCNICOS



SISTEMA NACIONAL DE EVALUACIÓN,
ACREDITACIÓN Y CERTIFICACIÓN
DE LA CALIDAD EDUCATIVA



PERÚ

Ministerio
de Educación

ISBN: 978-612-4322-25-9



9 786124 322259