

Distrito Escolar de Hermiston

**Manual de Calificación Basado en Estándares
para Escuelas Primarias**

Definición de Términos

Evaluación – Una medida del rendimiento de los estudiantes con respecto a un estándar en particular (habilidad/concepto).

Evaluación Formativa – Un control del progreso del estudiante con el aprendizaje. Una evaluación diseñada para informar el diseño de la instrucción y las intervenciones futuras al verificar el progreso del aprendizaje de los estudiantes en relación con un estándar.

Evaluación Sumativa – Un *resumen* del aprendizaje del estudiante. Una evaluación diseñada para medir el grado en que los estudiantes han alcanzado un estándar en particular al finalizar la instrucción.

Evaluación a la Proximidad de Dominio – Una evaluación que destaca el grado en que el estudiante ha alcanzado la maestría de un estándar de su nivel de grado. Estas evaluaciones tienen principalmente implicaciones para diferenciar la instrucción a nivel de grado y la preparación para aumentar los niveles de complejidad cognitiva con los estándares de nivel de grado.

Evaluación de Déficit – Una evaluación que destaca las habilidades/conceptos que a los estudiantes les falta de previas experiencias de aprendizaje (por ejemplo, previos niveles de grado). Estas evaluaciones principalmente tienen implicaciones para la intervención e informan al maestro sobre qué déficits necesitan tomar en cuenta a medida que ocurre la instrucción con los estándares de nivel de grado.

Retroalimentación – Información dada al estudiante que identifica áreas específicas de fortaleza, así como oportunidades de crecimiento en su aprendizaje.

Rúbrica – Parámetros que se usan para categorizar el rendimiento del estudiante en un estándar. Una rúbrica de calidad contiene parámetros específicos que el estudiante y el maestro pueden usar para comunicar que nivel de maestría el estándar requiere, así como también ayudar a identificar donde se encuentra el estudiante en su jornada hacia la maestría.

Basado en Estándares – Enseñanza que da prioridad al diseño de la instrucción, lecciones, y evaluaciones en torno a los estándares de nivel de grado.

Calificación Basado en Estándares – Sistema de calificación que utiliza la evidencia del rendimiento del estudiante en los estándares específicos de las materias, como son medidas por la rúbrica, para calcular un puntaje/calificación.

Estándar Prioritario – Un subconjunto de estándares de instrucción seleccionados por su amplitud e importancia en el aprendizaje del estudiante. Un estándar prioritario que abarca uno o más de un estándar de apoyo. Cuando los estudiantes alcanzan el dominio de un estándar prioritario es probable que hayan dominado los estándares de apoyo correspondientes. Los estándares prioritarios se seleccionan por su *permanencia* (conocimiento que perdure a pesar del tiempo), *preparación* (su importancia para el aprendizaje del siguiente nivel), *capacidad de influir* (conocimiento/habilidades que se pueden aplicar en múltiples contextos) y su presencia en *evaluaciones* externas (enfáticas en evaluaciones sumativas de alto nivel).

Objetivo de Aprendizaje – Una meta concreta (objetivo) escrita en un vocabulario fácil de entender por los estudiantes que claramente describe lo que los estudiantes aprenderán (contenido/estándar).

Criterios de Éxito – Criterios compartidos con el estudiante mediante los cuales el maestro puede determinar si el estudiante ha alcanzado el objetivo de aprendizaje (objetivo/meta). Los estudiantes deben ser capaces de explicar el grado en que han cumplido con los criterios de éxito.

Introducción

El Distrito Escolar de Hermiston tiene el placer de trabajar con una comunidad de diversos estudiantes de diversos orígenes. Creemos que tal diversidad es valiosa y la podemos utilizar para alcanzar niveles aún mayores de éxito con nuestros estudiantes, el personal y la comunidad. Creemos que todos los estudiantes merecen cruzar el escenario cuando se gradúan de la preparatoria con todas las oportunidades posibles para la universidad o profesión que estén disponibles para ellos. Se lo debemos a nuestros hijos brindarles las experiencias educativas que necesitan para adquirir el conocimiento y las habilidades que les permitan vivir la vida que ellos elijan, no una que se haya decidido por ellos debido a la falta de una educación de calidad. Este es el resultado equitativo “enfocados en el futuro” por el que todos luchamos. Para lograr eso, creemos que es fundamental que garanticemos un sistema de alineación en todas nuestras escuelas primarias. Creemos en la idea de ser un distrito escolar en lugar de un distrito de escuelas. Con esto en mente, nos mantenemos firmes en la idea de que, independientemente de la escuela a la que asista un estudiante en el Distrito Escolar de Hermiston, nos aseguramos de que reciba una educación equitativa construida sobre la base de prácticas educativas efectivas basadas en la investigación.

Una de esas prácticas es priorizar los estándares de contenido estatales y nacionales de nivel de grado en las experiencias educativas diseñadas para los estudiantes. Debido a la gran cantidad de estándares en cada área de contenido, es necesario priorizar ciertos estándares que brinden a los maestros y estudiantes un nivel de comprensión para garantizar que al dominar esos estándares “prioritarios,” los estudiantes habrán dominado todos los estándares. Llamamos a estos estándares nuestros estándares prioritarios. Con el fin de crear una mayor claridad en todo nuestro sistema educativo para maestros, estudiantes y padres con respecto al nivel de rigor requerido de los estudiantes para dominar los estándares prioritarios, identificar dónde se encuentra cada estudiante en su camino hacia el dominio de cada estándar prioritario y comunicar de manera efectiva tal progreso a todas las partes interesadas, es imperativo que adoptemos un sistema de calificación diseñado para tal fin.

Los sistemas de calificación tradicionales rara vez son útiles para los estudiantes, maestros y padres. En el sistema de calificación tradicional, los estudiantes obtienen calificaciones como una calificación con letras de la A a la F, o alguna variación de calificación con letras (un símbolo de más, una palomita, o símbolo de menos). Debido a la falta de claridad inherente a un sistema de calificación tradicional, los estudiantes y los padres no reciben información práctica que les permita comprender las fortalezas académicas particulares y las áreas de crecimiento que el niño posee en relación con dominio del nivel de grado. En el mejor de los casos, las calificaciones tradicionales ayudan a los padres a ver si su hijo está, “haciendo lo que el maestro quiere que haga.” En el peor de los casos, las calificaciones tradicionales estratifican a los estudiantes en categorías de estudiantes exitosos y no exitosos, lo que hace que los que no tienen éxito asuman que no son estudiantes académicamente capaces. No es sorprendente que tales prácticas de calificación perjudiquen a los estudiantes de poblaciones desatendidas y empeoren las brechas de rendimiento entre los grupos demográficos (Feldman, 2018; Alex, 2022). También es preocupante la presión que sienten con frecuencia los estudiantes que tienen éxito académico en este sistema para mantener su posición arbitraria como uno de los niños “inteligentes.”

Los sistemas de calificación tradicionales limitan la agencia de padres y estudiantes (Dueck, 2014). Posteriormente, el maestro se posiciona como el factor decisivo con respecto a qué tan bien les está yendo a los estudiantes. Incluso si un maestro usó criterios específicos para determinar las calificaciones, tales criterios rara vez son comunes en todo un distrito escolar, una escuela o incluso en los salones de clase del mismo grado en la misma escuela (Reeves, 2004), lo que resulta en un “revoltijo” de prácticas de calificación (Guskey, 2015). De acuerdo con Marzano (2000) y Townsley y Buckmiller (2016), con más de 100 años de uso del sistema de calificación tradicional en las escuelas estadounidenses, no existe ninguna investigación significativa que respalde su uso.

Puede surgir la pregunta de por qué se instituyó el sistema de calificación tradicional en primer lugar. La respuesta a eso puede atribuirse en parte a los objetivos de la era industrial en los Estados Unidos durante el siglo 20. El objetivo del sistema educativo en las escuelas estadounidenses en este tiempo era identificar a los

estudiantes que eran “buenos en la escuela” y podrían avanzar a la educación superior de aquellos estudiantes que serían más adecuados para trabajar en el sector industrial. (Townasley y Buckmiller, 2016).

Pasar a una era de educación pública que cree que todos los estudiantes pueden alcanzar altos niveles de competencia académica ha sido un cambio sísmico para las escuelas públicas estadounidenses en el siglo 21. La calificación tradicional es un sistema que no encaja en este nuevo paradigma (Craig, 2011), continuamente tirando de las instituciones educativas que trabajan con una mentalidad del siglo 21 “todos los estudiantes pueden alcanzar altos niveles de aprendizaje” en un modelo del siglo 20 “algunos estudiantes son adecuados para altos niveles de aprendizaje y algunos están destinados al trabajo manual.”

La clasificación basada en estándares es uno de esos modelos de calificación del siglo 21. Con todos los maestros enfocados en los mismos estándares de prioridad, pueden comunicar objetivos de aprendizaje específicos a padres y estudiantes. Con rúbricas en la mano que delinear claramente la progresión del aprendizaje y la evidencia requerida para que los estudiantes demuestren el dominio de cada estándar prioritario, todas las partes interesadas están empoderadas con información práctica que pueden usar para enfocarse intencionalmente en las necesidades de aprendizaje de los estudiantes. (Scriffiny, 2008). El sistema basado en estándares pone al estudiante directamente en el asiento del conductor de su aprendizaje y empodera a los padres con información para apoyar el progreso del aprendizaje de su hijo. Una libreta de calificaciones basada en estándares, entonces, no es más que una herramienta para comunicar el progreso de los estudiantes en relación con esos estándares de prioridad. Powerschool (2023) ha publicado una infografía que resume algunas de las distinciones clave entre la calificación tradicional y la basada en estándares.



Sistema de Calificación Tradicional

- Utiliza un sistema basado en porcentajes o calificaciones de letras
- Se apoya en evaluaciones como cuestionarios, tareas, ensayos, etc.
- A menudo incluye crédito adicional, ceros y puntajes en grupo
- Calificación general = promedio de muchos elementos diferentes
- Una sola calificación dada por curso (matemáticas, ciencias,

?



Sistema de Calificación Basado en Estándares

- Enfocado en los objetivos de aprendizaje y en el desarrollo de acuerdo con los estándares
- Mide el logro solamente, sin crédito adicional o ceros
- Solo se registran los elementos destinados a medir el logro
- La calificación se basa en la evidencia más reciente de aprendizaje
- El aprendizaje basado en estándares se puede combinar con un grado tradicional.

“Los maestros de todos los niveles deben ser capaces de defender las calificaciones que asignan y deben tener evidencia para respaldar sus decisiones. Para servir como comunicación significativa, las calificaciones deben ser justas, precisas y confiables. Es más probable que lo sean cuando los profesionales reflexivos están de acuerdo en el propósito de las calificaciones, miran la evidencia que tienen y luego deciden la calificación que mejor resume esa evidencia.” (Guskey & Jung, 2016, p. 54)

Esta guía está destinada a ser un recurso para los estudiantes, familias y educadores del Distrito Escolar de Hermiston. El objetivo para el resto de esta guía es aclarar una comprensión de la calificación basada en estándares, específicamente las formas en que el Distrito Escolar de Hermiston está implementando este sistema de calificación del siglo 21.

Principios de la Calificación Basada en Estándares

Las Escuelas Públicas de Federal Way en el estado de Washington publicaron seis principios de calificación basada en estándares que han ayudado a guiar nuestro trabajo aquí en Hermiston.

Principio 1 - Las calificaciones y los informes se basarán en objetivos de aprendizaje y estándares de desempeño claramente especificados.

Principio 2- La evidencia utilizada en la calificación será válida.

Principio 3 - La calificación se basará en criterios establecidos.

Principio 4 - No todos los rendimientos deben incluirse en las calificaciones.

Principio 5 - Calificar un estándar basado en el trabajo más reciente.

Principio 6 - Informar el logro y otros factores por separado.



Echemos un vistazo a cada uno de estos principios para articular cómo hacemos que eso sea una realidad en el Distrito Escolar de Hermiston para los estudiantes de la primaria. Esencialmente, como cada uno de los seis principios es ejecutado en nuestro modo de calificar en nuestro trabajo de calificación basado en estándares.



Principio 1 - *Las calificaciones y los informes se basarán en objetivos de aprendizaje y estándares de desempeño claramente especificados.*

En Hermiston, hemos formado tres comités separados por materias específicas compuestos por maestros, entrenadores de instrucción y administradores. Cada comité identifica estándares prioritarios para su materia y grado. Por ejemplo, nuestro Comité de Artes Lingüísticas en Inglés tiene representantes de todo el distrito escolar en cada nivel de grado. Los miembros del comité de nivel de grado pasan por un proceso cada año para identificar y revisar los estándares de mayor prioridad para que los estudiantes dominen al final de ese nivel de grado. Lo mismo ocurre con nuestros Comités de Matemáticas y el Desarrollo del Idioma Inglés. Los estándares prioritarios que identifica cada nivel de grado se eligen de la colección de estándares de nivel de grado adoptados por el Departamento de Educación de Oregon (siglas en inglés ODE). Es imposible para cualquier maestro enseñar todos los estándares para cada materia en cada grado. Simplemente no hay suficiente tiempo en el día. Incluso si lo hubiera, muchas normas suplementan y/o complementan otras normas. Esto nos lleva a un proceso de identificación de qué estándares son los más completos para que nuestros estudiantes aprendan. Es decir, qué estándares serían los más importantes para que un estudiante domine de tal manera que al dominar ese estándar más completo, hayan dominado los estándares más pequeños y suplementarios. Esta es una distinción

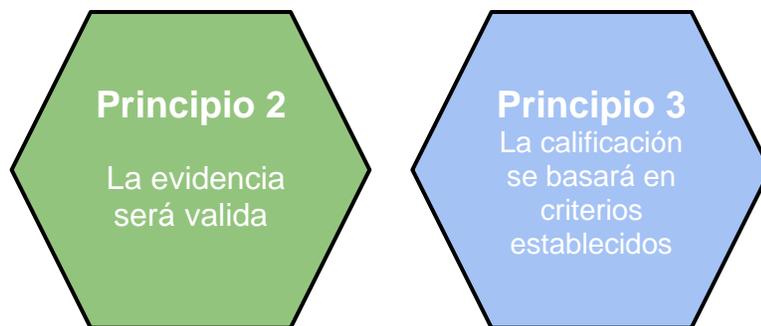
importante porque cada vez que un distrito escolar se somete al proceso de identificar estándares prioritarios para el aprendizaje de los estudiantes, se comparte legítimamente la preocupación de que el distrito escolar puede estar reduciendo el plan de estudios, limitando así el alcance de lo que los estudiantes tienen la oportunidad de aprender.

Otra pregunta que puede hacerse es cómo identifica cada comité qué normas son prioritarias y cuáles son suplementarias. Para esto, hemos utilizado un proceso desarrollado por Ainsworth (2013) y aprendido a través de nuestra asociación con el Centro Internacional para el Liderazgo en la Educación (International Center for Leadership in Education). Los miembros del comité evalúan cada estándar de acuerdo con cuatro criterios. Los criterios son:

1. Permanencia - ¿El dominio de este estándar proporcionará a los estudiantes el conocimiento y las habilidades que serán de valor más allá del presente?
2. Capacidad de Influir - ¿El dominio de este estándar tendrá aplicación a otros estándares dentro del área de contenido y cruzará a otras áreas de contenido?
3. Preparación - ¿El dominio de este estándar proporcionará a los estudiantes el conocimiento y habilidades esenciales que son necesarios para el éxito en el futuro?
4. Evaluaciones Externas - ¿El dominio de este estándar preparará a los estudiantes para los conceptos y habilidades que es más probable que encuentren en las pruebas estandarizadas anuales, los exámenes de ingreso a la universidad y los exámenes de competencia ocupacional?

Se debe señalar que ningún criterio es más importante que otro. Por ejemplo, el hecho de que se incluya un estándar de prioridad en la evaluación estatal de fin de año no significa que sea automáticamente un estándar de prioridad. Si ese estándar tuviera una mínima capacidad de influir y mínima permanencia, entonces no se seleccionaría como un estándar prioritario. Esencialmente, los cuatro criterios se aplican a cada estándar. Los estándares que mejor ejemplifican se alinean con los cuatro criterios son los seleccionados por cada comité.

Con los estándares prioritarios seleccionados, los maestros ahora están armados con los estándares específicos para alinear sus materiales del plan de estudio. Esto también crea una alineación en los seis edificios de las escuelas primarias, mientras que al mismo tiempo permite la autonomía y flexibilidad de los maestros para abordar esos estándares. Al entrar en múltiples salones de clase de 3er grado en el transcurso de una semana, verá que se enseñan estándares similares, pero los recursos que cada equipo de grado da prioridad para ayudar a los estudiantes a alcanzar el dominio de ese estándar dependen del maestro y el equipo. No todos los estudiantes son iguales. Los estudiantes requieren diversos recursos, métodos, estructuras de aprendizaje, etc...con el fin de dominar ciertos conceptos. Los maestros usan sus conocimientos del desarrollo infantil, el contenido, la pedagogía para reunir las necesidades de sus estudiantes. ¡Ese es el arte de enseñar!



Principio 2 y 3 - *La evidencia usada en la calificación será válida; La calificación se basará en criterios establecidos.*

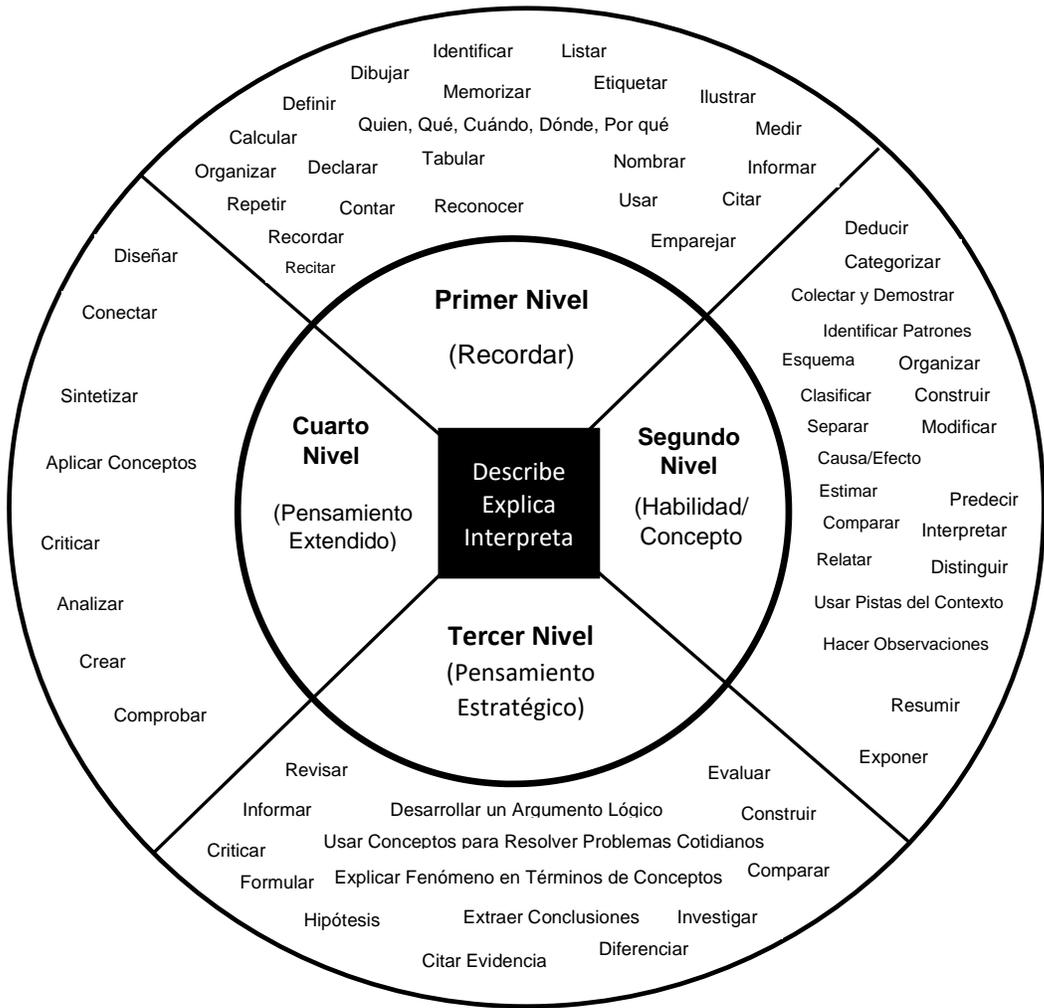
Como sistema, si no nos aseguramos de que los maestros tengan una métrica común para evaluar el dominio de los estudiantes, entonces hemos hecho poco para garantizar la equidad en todo nuestro sistema en el rigor de las evaluaciones que los maestros usan para determinar el dominio de los estándares de prioridad de los estudiantes. Por ejemplo, uno de nuestros estándares prioritarios de Artes Lingüísticas en Inglés de 5º grado es que los estudiantes integren información de varios textos sobre el mismo tema para escribir o hablar sobre el tema. Sin una métrica común para que los maestros usen, un estudiante en la Escuela A podría recibir una evaluación en la que se le pide que lea dos artículos sobre el reciclaje y luego responda una evaluación de opción múltiple de 10 preguntas que compare y contraste la información

en cada artículo. Al mismo tiempo, se le pide a un estudiante de la Escuela B que investigue y analice múltiples fuentes sobre los pros y los contras del reciclaje para luego redactar un ensayo de varios párrafos que justifique su posición sobre si el gobierno federal debería aprobar una legislación que ordene un programa de reciclaje en cada municipio de los Estados Unidos. Ambas evaluaciones evaluarían, en cierta medida, el estándar identificado. Sin embargo, está claro que cada evaluación varía significativamente en la evidencia requerida para demostrar el dominio.

En el Distrito Escolar de Hermiston, cada comité de nivel de grado ha creado una rúbrica de 4 puntos para cada estándar de prioridad. Esta rúbrica se alinea con la Profundidad del Conocimiento (siglas en inglés DOK), descrita en los Estándares del Estado de Oregon. Este es el DOK requerido de los estudiantes para demostrar dominio. En el Distrito Escolar de Hermiston, usamos la Profundidad de Conocimiento de Webb (2006) para determinar el nivel de complejidad en cada estándar (ver imagen a continuación). Esto ayuda a los maestros a comprender qué tipo de experiencias de aprendizaje necesitan los estudiantes para alcanzar la profundidad del conocimiento requerido de cada estándar de prioridad en particular. Cada estándar tiene un nivel DOK que el estudiante debe alcanzar para demostrar dominio con ese estándar. Por ejemplo, si un estudiante puede recordar información con éxito, puede etiquetar una gráfica, emparejar artículos, etc. con un estándar de nivel DOK 1, diríamos que el estudiante tiene el dominio. Sin embargo, digamos que un estándar requiere un nivel de comprensión DOK 3. Ahora el maestro debe proveer instrucción a un nivel mucho más profundo y, posteriormente, diseñar experiencias de aprendizaje que involucren a los estudiantes en un pensamiento significativamente más avanzado. En este caso, se les puede pedir a los estudiantes que desarrollen un argumento, formulen una hipótesis dándoles un conjunto de condiciones, investiguen un fenómeno, etc. Esencialmente, necesitan usar el conocimiento de nivel 1 que poseen para un estándar y aplicarlo en un contexto más riguroso. Si un estándar solicita el nivel de comprensión DOK 3 y un estudiante no puede completar ese tipo de tareas de nivel DOK 3, independientemente de cuánto pueda recordar, entonces no pueden mostrar dominio hasta que puedan hacerlo. Esto hace que la tarea de escribir rúbricas para cada estándar de prioridad sea

aún más importante para que los maestros tengan una guía clara y los estudiantes reciban oportunidades de aprendizaje equitativas.

Niveles de la Profundidad del Conocimiento (siglas en inglés DOK)



Con una comprensión del DOK requerido de cada estándar de prioridad, los comités se propusieron crear rúbricas específicas de tareas que mostrarán claramente lo que los estudiantes deben poder hacer para demostrar un dominio para cada estándar de prioridad. Cada rúbrica está en una escala del 1 al 4. La escala 1-4 se separa así:

- 4 - Supera el estándar
- 3 - Cumple con el estándar
- 2 - Casi cumple con el estándar
- 1 - No cumple con el estándar

Un puntaje de 1 estaría muy por debajo del estándar. Un puntaje de 2 significaría que el estudiante está cerca de cumplir con el estándar y puede necesitar algo de tiempo, práctica y oportunidad adicionales. Un puntaje de 3 significa que el estudiante ha alcanzado la expectativa de nivel de grado de fin de año para ese estándar. Un puntaje de 4 mostraría que el estudiante está por encima de las expectativas de fin de año para ese estándar. Es importante tener en cuenta esta escala al hacer juicios sobre el aprendizaje de los estudiantes. Esperamos que la mayoría de los estudiantes en la primera parte del año escolar sean de nivel 1 o 2 para la mayoría de los estándares. Si muchos estudiantes tenían 3 o 4 años en la primera parte del año escolar, entonces quedaría poco para que los estudiantes aprendieran en ese nivel de grado. A medida que avanza el año y los estudiantes tienen la oportunidad de aprender esos estándares, sus puntajes de rúbrica aumentarán. Un puntaje de 3 al final del año escolar es un éxito, ya que muestra que el estudiante ha alcanzado el resultado de nivel de grado para el estándar dado.

Las rúbricas a menudo sufren de un lenguaje subjetivo y vago que crea una delineación pobre entre los niveles de competencia. Considere la siguiente rúbrica:

1	2	3	4
Con orientación y apoyo del maestro, éxito parcial en las habilidades de nivel 2.	Con indicaciones y apoyo, aplica estrategias de lectura aprendidas o se aplica solo a textos por debajo del nivel.	Aplica de manera consistente e independiente las estrategias de lectura aprendidas a los textos a nivel de grado.	Además de un puntaje de 3.0 en rendimiento, el estudiante demuestra en profundidad interferencias y aplicaciones que van más allá de lo que se enseñó, por ejemplo: Aplica estrategias de lectura de manera independiente y consistente a textos por encima del nivel de grado.

Podemos ver que la rúbrica está evaluando el grado en que los estudiantes usan estrategias de lectura. Independiente aplicando las estrategias de lectura para comprender el texto es una habilidad crítica que todos los niños deben poseer. Sin embargo, la rúbrica no nos da una idea de qué estrategias de lectura estamos evaluando. Además, ¿qué quieren decir con “indicación y apoyo” o “aplica de manera consistente e independiente?” ¿Qué es una habilidad de “nivel 2?” ¿Cuál es la diferencia entre “orientación del maestro” e “indicaciones y apoyo?” Finalmente, ¿cómo demuestra un estudiante que ha ido "más allá de lo que se le enseñó"? El escritor de esta rúbrica a lo mejor sabe exactamente lo que querían decir en cada una de esas declaraciones. Desafortunadamente, cualquiera que no esté al tanto del pensamiento que se pensó en la creación de esta rúbrica debe interpretar cada uno de esos elementos por sí mismo. Todos nosotros tenemos nuestras propias interpretaciones de lo que creemos que significan esos términos. Ese es el problema. Existiría una variación significativa en las experiencias de aprendizaje para los estudiantes de Hermiston si se crearan rúbricas similares para nuestros estándares prioritarios. Esta es la razón por la cual los miembros del comité escribieron rúbricas específicas del trabajo/tarea. Una rúbrica específica de la tarea proporciona a los maestros un lenguaje concreto que describe los tipos de experiencias de aprendizaje, trabajos y evaluaciones que los estudiantes necesitan para demostrar que han alcanzado el nivel de competencia requerido en cada estándar. A continuación, una muestra de una rúbrica específica de la tarea para uno de nuestros estándares prioritarios de Artes Lingüísticas en Inglés de 3er grado.

	1	2	3	4
<p>3.RI.9 Comparar y contrastar los puntos más importantes y los detalles clave presentados en dos textos sobre el mismo tema.</p>	<p>Dados dos textos sobre el mismo tema, el niño es:</p> <ul style="list-style-type: none"> • incapaz de identificar 2-3 ideas principales compartidas <p>O</p> <ul style="list-style-type: none"> • no puede apoyar 2-3 ideas principales compartidas con 1 a 2 detalles clave de ambos textos 	<p>Dados dos textos sobre el mismo tema, el niño puede:</p> <ul style="list-style-type: none"> • identificar 2-3 ideas principales compartidas <p>MÁS uno de los siguientes:</p> <ul style="list-style-type: none"> • Apoyar 2-3 ideas principales compartidas con 1 a 2 detalles clave de ambos textos • explicar cómo cada idea principal compartida es similar/ diferente en ambos textos comparando / contrastando los detalles clave que los respaldan 	<p>Dados dos textos sobre el mismo tema, el niño puede:</p> <ul style="list-style-type: none"> • Identificar 2-3 ideas principales compartidas • Apoyar 2-3 ideas principales compartidas con 1-2 detalles clave de ambos textos • explicar cómo cada idea principal compartida es similar / diferente en ambos textos comparando/ contrastando los detalles clave que los respaldan 	<p>Dados dos textos sobre el mismo tema, el niño puede:</p> <ul style="list-style-type: none"> • Completar todos los criterios en 3 <p>MÁS uno de los siguientes:</p> <ul style="list-style-type: none"> • analizar y explicar cuál de los dos textos aborda el tema de manera más completa comparando / contrastando los 1-2 detalles clave que respaldan las ideas principales de cada texto. • analizar y explicar cómo ambos textos se complementan entre sí demostrando cómo los 2-3 puntos / 1-2 detalles clave en un texto complementan la información en el otro.

Estas rúbricas específicas de la tarea se convierten en los criterios por los cuales los maestros califican a los estudiantes, asegurando así la coherencia en todo el sistema de primaria y, lo que es más importante, la transparencia y la equidad para los estudiantes y los padres. Los puntajes que los estudiantes obtienen de estas evaluaciones reflejarán su nivel actual de dominio con el estándar de prioridad y, posteriormente, serán los puntajes que los maestros ingresan en el libro de calificaciones. A medida que se generan informes de progreso (final del trimestre) y las Boletas de Calificaciones (final del semestre) para compartir con los estudiantes y los padres, estas “calificaciones” se consideran para determinar un puntaje actual para cada estudiante en relación con cada estándar de prioridad.



Principio 4 - *No todo el rendimiento deben incluirse en las calificaciones.*

Como se mencionó en el Principio 1, cada maestro puede determinar cómo reunir mejor las necesidades de los estudiantes en su salón de clase. Algunos maestros pueden ofrecer oportunidades de práctica adicionales para una habilidad a modo de tarea. Algunos pueden sentir que las verificaciones de progreso disponibles en los planes de estudios son una buena indicación de cómo están progresando los estudiantes. Otros podrían creer que los estudiantes en su salón de clase necesitan algo completamente diferente para mostrar su progreso hacia el dominio de un estándar. Esto puede variar entre los maestros, e incluso en su propia clase a medida que avanzan en el año escolar. Entonces, ¿qué significa esto? Significa que se debe hacer una distinción con respecto a lo que se cuenta por motivos de una calificación. Para esto, una diferenciación rápida entre la evaluación formativa y sumativa es útil.

Evaluación Formativa – Un control del progreso del estudiante con el aprendizaje. Una evaluación diseñada para *informar* el diseño de la instrucción y las intervenciones futuras al verificar el progreso del aprendizaje de los estudiantes en relación con un estándar.

Evaluación Sumativa – Un *resumen* del aprendizaje del estudiante. Una evaluación diseñada para medir el grado en que los estudiantes han alcanzado un estándar en particular al finalizar la instrucción.

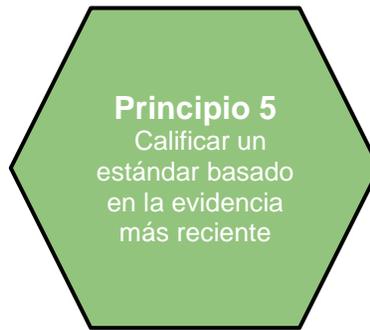
Las dos definiciones anteriores son útiles porque podemos ver qué tipo de evaluación sería mejor por motivos de una calificación y qué tipo de evaluación no sería tan útil. La pregunta que debe hacerse para determinar esto es qué evaluación, formativa o sumativa, proveerá la comprensión más completa de qué tan bien un estudiante ha aprendido un estándar en particular. Si adivinaste la evaluación sumativa, ¡tienes razón!

Las personas de las Escuelas Públicas de Federal Way (2023) tienen una explicación fantástica de esto:

“Las Evaluaciones Formativas (de Práctica) proveen información sobre lo que el estudiante entiende y lo preparan para las evaluaciones sumativas de manera muy similar a las prácticas de un equipo deportivo antes de un juego final. Esto promueve la comunicación bidireccional entre el maestro y el estudiante para desarrollar la responsabilidad y promover el crecimiento a través del aprendizaje continuo. Rara vez se tiene en cuenta en una calificación final. Por lo general, se usa para la práctica y generalmente no se incluye en los grados. Sin embargo, hay circunstancias en las que lo formativo puede convertirse en sumativo e incluirse en las calificaciones” (las Escuelas Públicas de Federal Way, 2023).

La última línea de la cita es importante. Las evaluaciones formativas pueden convertirse en evaluaciones sumativas. Cada equipo de nivel de grado decide eso. Un equipo de 2º grado podría revisar una evaluación formativa común que dieron recientemente y decidir que la evaluación fue una medida de calidad del desempeño del estudiante en relación con un estándar de prioridad. Por lo tanto, acuerdan mutuamente que la evaluación formativa se utilizará como un puntaje para informar el dominio del estudiante para ese estándar. Siempre y cuando el equipo esté de acuerdo y haya consistencia en cómo se midió a los estudiantes, entonces esto es perfectamente aceptable.

También se debe destacar lo que no se ha mencionado en esta sección sobre lo que se incluye en la calificación. El trabajo diario, exámenes sorpresa, tareas, etc... son todas herramientas razonables que los maestros probablemente usen en el curso de instruir a los estudiantes. Esas herramientas pueden ser útiles para proveer a los estudiantes práctica adicional y oportunidades para aprender. También pueden ser útiles para los maestros mientras siguen el progreso del aprendizaje de los estudiantes. Sin embargo, no son dignos de ser incluidos en un libro de calificaciones basado en estándares porque no reflejan la evidencia completa requerida para que un estudiante demuestre su nivel de dominio en un estándar. Más sobre esto en el Principio 5.



Principio 5 - Califique un estándar basado en el trabajo más reciente.

Considere la siguiente tabla que informa una serie de puntajes de evaluación para tres estudiantes diferentes que están trabajando para obtener su certificación de paracaidismo. Esta certificación permitirá al individuo dirigir expediciones de paracaidismo para miembros del público.

	Eval 1	Eval 2	Eval 3	Eval 4	Eval 5	Eval 6	Eval 7	Eval 8	Eval 9	Promedio
Estudiante 1	100	90	100	90	80	70	80	60	50	80
Estudiante 2	80	90	70	80	60	80	60	70	40	70
Estudiante 3	10	20	30	40	50	60	70	80	90	50
Puntaje de Aprobación	70	70	70	70	70	70	70	70	70	

En un sistema de calificación tradicional, el estudiante que ha progresado continuamente a lo largo del curso y actualmente está demostrando la mayor cantidad de conocimientos/habilidades relacionadas con el paracaidismo es el estudiante que recibiría una F (Estudiante 3). La pregunta aquí es, ¿con quién preferirías saltar de un avión? Claramente el mismo estudiante “reprobado.” Ciertamente esperamos que esta escuela de paracaidismo use un sistema de calificación basado en estándares donde la certificación se basa en si los estudiantes demuestran o no los conocimientos y habilidades necesarios en lugar de un promedio de todas sus evaluaciones desde el inicio del curso hasta el final. De hecho, así es como funcionan la mayoría de las certificaciones. No certificaríamos a un cirujano cardíaco en función de lo bien que se desempeñó en sus primeras evaluaciones de sus habilidades quirúrgicas en su primer año de la escuela de medicina. La junta de licencias médicas basaría su decisión de otorgar una licencia para practicar medicina en una evaluación exhaustiva de sus

habilidades actuales. La calificación basada en estándares es la misma. Considere el siguiente extracto de Wormeli (2011);

“LSAT. MCAT. Praxis. Examen SAT. Examen de Barra. Examen de CPA. Licencia de conducir. Licencia de piloto. Examen de certificación de mecánico de automóviles. Cada una de estas evaluaciones refleja las responsabilidades del mundo laboral a nivel de adulto que nuestros estudiantes enfrentarán algún día. Muchas de ellas son de alto riesgo: la vida de las personas depende de la validez de estas pruebas como medidas precisas de la competencia del individuo. Todos ellos se pueden rehacer una y otra vez para obtener crédito completo. Los abogados que finalmente aprueban el examen de la barra en su segundo o tercer intento no se limitan a ejercer la abogacía solo los martes o solo bajo el cuidado de un socio experimentado durante el resto de sus carreras. Si una evaluación de competencia es válida, el logro de sus puntajes de aprobación le otorga al individuo evaluado plenos derechos y privilegios sobre los mismos” (p.2).

Si creemos que los estudiantes adquieren conocimientos y habilidades entre más oportunidades y tiempo tienen para aprender, no tiene mucho sentido usar su desempeño en una evaluación previa de un estándar de prioridad en un cálculo de su calificación. En cambio, usaríamos su desempeño más reciente en una evaluación exhaustiva de esa norma de prioridad. Nos preocupa su nivel de comprensión hoy, no su nivel de comprensión el mes pasado. De hecho, todo el sistema educativo debe basarse en las habilidades que los estudiantes poseen ahora. Al buscar diferenciar la instrucción, proveer modificaciones, proveer adaptaciones, diseñar lecciones de intervención, etc. revisamos los datos más recientes de cada estudiante. La calificación basada en estándares no es diferente.

Mirándolo desde otro ángulo, no es difícil imaginar por qué muchos estudiantes se desconectan del aprendizaje a medida que avanza el año escolar. Imagine al Estudiante 3 nuevamente desde el ejemplo del paracaidismo. ¿Qué pasaría si los instructores de paracaidismo dijeran que la certificación de paracaidismo se basaba en un promedio de desempeño y no en el desempeño al final del curso? Después de las primeras evaluaciones, el Estudiante 3 se daría cuenta de que casi no tiene posibilidades de obtener un puntaje de aprobación. Imaginar que el Estudiante 3

abandone el curso es un resultado probable. En un sistema de educación obligatoria como el nuestro, a los estudiantes no se les permite abandonar la escuela. En cambio, se desconectan, desarrollan percepciones negativas de su capacidad para aprender y, a menudo, exhiben problemas de comportamiento (Dueck, 2014). Una vez que esto comienza para un estudiante, es un ciclo difícil de romper ya que su creencia en su capacidad académica continúa disminuyendo. Hay que asegurar que los estudiantes sepan que cada día es una nueva oportunidad para aprender y mostrar lo que saben al evaluar su aprendizaje en función de lo que pueden hacer hoy y no de lo que no pudieron hacer ayer, envía un mensaje positivo de que el éxito académico es un viaje, no un destino.

Hay una excepción al principio de la “evidencia más reciente” cuando se califica el nivel de dominio de un estudiante en un estándar de prioridad, especialmente cuando se ingresa el puntaje en un informe de progreso o boleta de calificaciones. Es posible que un maestro tenga una evaluación que no sea la más reciente, pero sea un ejemplo más completo del nivel de dominio actual de un estudiante. Por ejemplo, un maestro puede haber dado recientemente una evaluación de opción múltiple que tenía varias preguntas que requerían que los estudiantes demostraran su capacidad para identificar la idea principal y los detalles clave de una historia. Sin embargo, la semana anterior, los estudiantes leyeron un artículo y completaron un organizador gráfico donde tuvieron que escribir la idea principal de cada sección de la historia con sus propias palabras e incluir 3 detalles clave que respaldan cada idea principal. Al decidir qué evaluación usar como una indicación del nivel actual de comprensión de cada estudiante, el maestro podría argumentar que la evaluación del organizador gráfico es una mejor representación. El punto aquí es que hay cierta flexibilidad al determinar los puntajes para un informe de progreso o boleta de calificaciones entre usar la evidencia más reciente y/o la más completa. Sin embargo, algunos de ustedes, lectores inteligentes, pueden estar pensando que, si un maestro diseña sus tareas de evaluación de acuerdo con las rúbricas mencionadas en los principios 2 y 3, esto no se convierte en un problema. Si tuvieras ese pensamiento, estarías en lo correcto.



Principio 6- *Informar el logro y otros factores por separado.*

La mayoría de las boletas de calificaciones de los estudiantes incluyen calificaciones de rendimiento académico en materias básicas (lectura, escritura, matemáticas, educación física, música, ciencias, etc.), así como calificaciones de ciudadanía y comportamiento. Las categorías de ciudadanía y comportamiento incluyen cosas como seguir instrucciones, completar el trabajo a tiempo y ser respetuoso. Estas categorías con frecuencia se marcan con un símbolo de más, una palomita, o símbolo de menos, siendo el símbolo de más muy bueno y el de menos no tan bueno. Las calificaciones que reciben los estudiantes para estas categorías son para comunicar a los padres qué tan bien su hijo está demostrando habilidades de trabajo y sociales en el salón de clases. Estas son ciertamente cosas importantes que un padre debe saber. Pero, no pertenecen en una boleta de calificaciones basada en estándares.

Este principio puede ser uno de los principios más desafiantes para nosotros, aunque solo sea por la razón de que se ha hecho de esta manera por mucho tiempo. Es tan común que sería difícil encontrar demasiados ejemplos de distritos escolares que no hacen esto. Entonces, ¿por qué el Distrito Escolar de Hermiston está rompiendo con esta larga tradición? La información para un padre sobre qué tan bien su hijo está exhibiendo el trabajo adecuado y las habilidades sociales en el salón de clases es tan importante que esperar hasta el momento del informe de progreso o la boleta de calificaciones para compartir eso degrada la utilidad de dicha información. Si un niño tiene dificultades sociales o de comportamiento, los padres deben de estar recibiendo esa información de inmediato para que, como educadores, podamos asociarnos con nuestras familias para intervenir por el niño. Pocos estarían en desacuerdo con eso. Si el maestro se comunica con frecuencia con los padres cuando surgen inquietudes, marcar un símbolo de más, una palomita o un símbolo de menos en el informe de progreso y/o la boleta de calificaciones se vuelve innecesario y distrae la atención del propósito de la boleta de calificaciones. Lo mismo se aplica a los estudiantes que muestran con éxito habilidades adecuadas de

trabajo y sociales. Como educadores debemos comunicarnos con las familias de estos estudiantes para que puedan continuar apoyando el éxito de su hijo/a.

En ambos casos, la comunicación regular con las familias sobre el comportamiento y el éxito social de sus hijos demuestra que vemos a nuestras familias como socios en la educación. Como resultado, en el Distrito Escolar de Hermiston, la ciudadanía y el comportamiento no son categorías incluidas en el sistema de boletas de calificaciones basado en estándares. En cambio, es una característica de nuestra comunicación regular con los estudiantes y sus familias.

Boletas de Calificaciones e Informes de Progreso

Períodos de Informes

Para tener las calificaciones en mente, el año escolar se divide en cuatro términos. Cada término es un momento en que los maestros comunican el dominio de los estudiantes para los estándares prioritarios enseñados hasta ese momento en el año escolar.

Término 1 - Progreso 1 (P1): Principio del año – Finales de octubre

Término 2 - Semestre 1 (S1): Principio de noviembre – Finales de enero

Término 3 - Progreso 2 (P2): Principio de febrero – Finales de marzo

Término 4 – Semestre 2 (S2): Principio de abril – Final del año escolar

El tiempo establecido para cada término es una aproximación, ya que puede haber una ligera variación en cuanto a cuándo comienza y finaliza un término en particular. El Término 1 coincide con las conferencias de padres en el otoño. Este período de informe le llamamos Progreso 1 (P1). Los maestros compartirán un informe de progreso de P1 con los padres en las conferencias en el otoño, compartiendo el nivel actual de dominio de sus hijos en los estándares enseñados hasta ese momento. Al finalizar el Término 2 es el final del primer semestre escolar, el tiempo en el cual los maestros envían a casa la boleta de calificaciones que muestra las calificaciones del Semestre 1 (S1). El término 3 es el segundo período de informes de progreso y también coincide con las conferencias en la primavera a finales de marzo. Los padres tendrán la oportunidad de revisar un informe del progreso de su hijo en esta

conferencia. Finalmente, el Término 4 concluye el año escolar con la boleta de calificaciones del Semestre 2 (S2) que se envía a casa.

Rúbricas

En los principios 2 y 3, se habló de la importancia de las rúbricas específicas. Estas rúbricas forman la columna vertebral de cómo los maestros diseñan evaluaciones para evaluar cada estándar de prioridad y los criterios por los cuales los maestros evalúan el entendimiento de los estudiantes. Un puntaje en el libro de calificaciones para cualquier estándar de prioridad en particular siempre debe evaluarse con la rúbrica para ese estándar de prioridad. Las rúbricas de estándares de prioridad garantizan que los estudiantes de todo el Distrito Escolar de Hermiston sean evaluados equitativamente. Ya sea que un estudiante asista a la escuela A o la escuela B, un puntaje de 2 para un estándar de prioridad significa lo mismo. Esto asegura que nuestros estudiantes y padres puedan evaluar la validez de los logros académicos de sus hijos. También permite a los educadores de todo el distrito escolar comprender las tendencias en el rendimiento de los estudiantes.

Estándares de prioridad del área de contenido

Se han identificado estándares prioritarios en Artes Lingüísticas en Inglés (siglas en inglés ELA), Matemáticas, y el Desarrollo del Idioma Inglés (siglas en inglés ELD). En cada una de estas áreas de contenido, se enumeran varios estándares. Cada uno de esos estándares prioritarios recibirá un puntaje de 1-4 en el informe de progreso/boleta de calificaciones. Algunos estándares se enseñarán más adelante en el año escolar. Cuando eso suceda, el maestro ingresará N/A en la célula para que el padre y el estudiante sepan que los estudiantes no han sido evaluados en ese estándar porque aún no han tenido la oportunidad de aprender ese estándar. Para ELD, no todos los estudiantes recibirán un puntaje en el informe de progreso/boleta de calificaciones. Solo los estudiantes que reciben servicios de desarrollo del idioma inglés tendrán puntajes para esos estándares prioritarios.

Puntuaciones Para Otras Áreas de Contenido

ELA, ELD y Matemáticas no son las únicas materias importantes que los estudiantes deben aprender. Las Ciencias Naturales, Ciencias Sociales, Música, Salud y Educación Física también son de vital importancia para una educación integral para nuestros estudiantes. Sin embargo, aún no se han identificado estándares prioritarios en estos grados. Están Trabajando actualmente en Música y Educación Física. Pronto comenzaremos a buscar la posibilidad de identificar estándares de prioridad y rúbricas en Ciencias Naturales, Salud, y Ciencias Sociales. Hasta ese momento, los estudiantes recibirán una de las siguientes designaciones de puntaje en cada una de las áreas de contenido mencionadas anteriormente.

- + - Supera las Expectativas del Nivel de Grado
- = - Cumple con las Expectativas del Nivel de Grado
- - No Cumple con las Expectativas del Nivel de Grado

Los maestros de estas áreas de contenido continuarán diseñando experiencias de aprendizaje para sus estudiantes según los estándares estatales y nacionales para esa área de contenido. Además, continuarán diseñando evaluaciones que se alineen con el nivel de complejidad que requieren los estándares académicos que enseñan. La diferencia clave es que no habrá un conjunto específico de estándares para esa área de contenido en la que todos los maestros se enfoquen, ya que tampoco habrá un conjunto específico de criterios que estén alineados en todos los salones de clase para evaluar la comprensión de los estudiantes. Esto no significa que los estudiantes no estén recibiendo grandes experiencias de aprendizaje y que los maestros no estén evaluando su aprendizaje de manera efectiva en estas áreas de contenido, solo significa que es probable que haya alguna variación en los estándares en los que se enfocan los maestros y los puntajes que reciben los estudiantes en función de esos estándares.

Boleta de Calificaciones/Puntajes del Informe de Progreso

En la explicación de los principios 2 y 3, se mostró la escala que se utiliza para calificar cada estándar de prioridad.

- 4 - Supera el Estándar
- 3 - Cumple con el Estándar
- 2 - Casi Cumple con el Estándar
- 1 - No Cumple con el Estándar

Cada rúbrica estándar de prioridad usa esta escala. Es fundamental entender que cada rúbrica fue escrita para los resultados de aprendizaje a nivel de grado de fin de año. No se debe esperar que un estudiante logre puntajes de 3 y 4 en cada estándar de prioridad en los primeros dos términos del año escolar. Debería tomar a la mayoría de los estudiantes la mayor parte del año para lograr tal dominio. Un puntaje de 3 muestra que el estudiante ha cumplido con el estándar para esa calificación. Un puntaje de 4 significa que tienen un nivel de dominio más allá de su nivel de grado actual. Nuestro objetivo es que al final del año escolar, nuestros estudiantes hayan alcanzado el nivel 3 o 4 en la mayoría de los estándares. Digamos que un estudiante obtuvo los 3 y 4 en su primer informe de progreso a fines de octubre. Eso significaría que han dominado todos los estándares para su grado. Ese maestro tendría poco que enseñarles, en cuanto al nivel de grado, durante el resto del año escolar. Potencialmente puede existir una dificultad para algunos estudiantes y padres sin esta comprensión de cómo funciona la calificación basada en estándares. Si un estudiante y/o padre está acostumbrado a recibir A y B en su boleta de calificaciones, pueden concluir erróneamente que de repente están teniendo dificultad en la escuela. Este no es el caso. Una vez más, deberíamos ver a la mayoría de los estudiantes con 1 y 2 en los primeros trimestres, pasando a 3 y 4 en los últimos dos trimestres.

Estudiantes que Reciben Servicios de Educación Especial

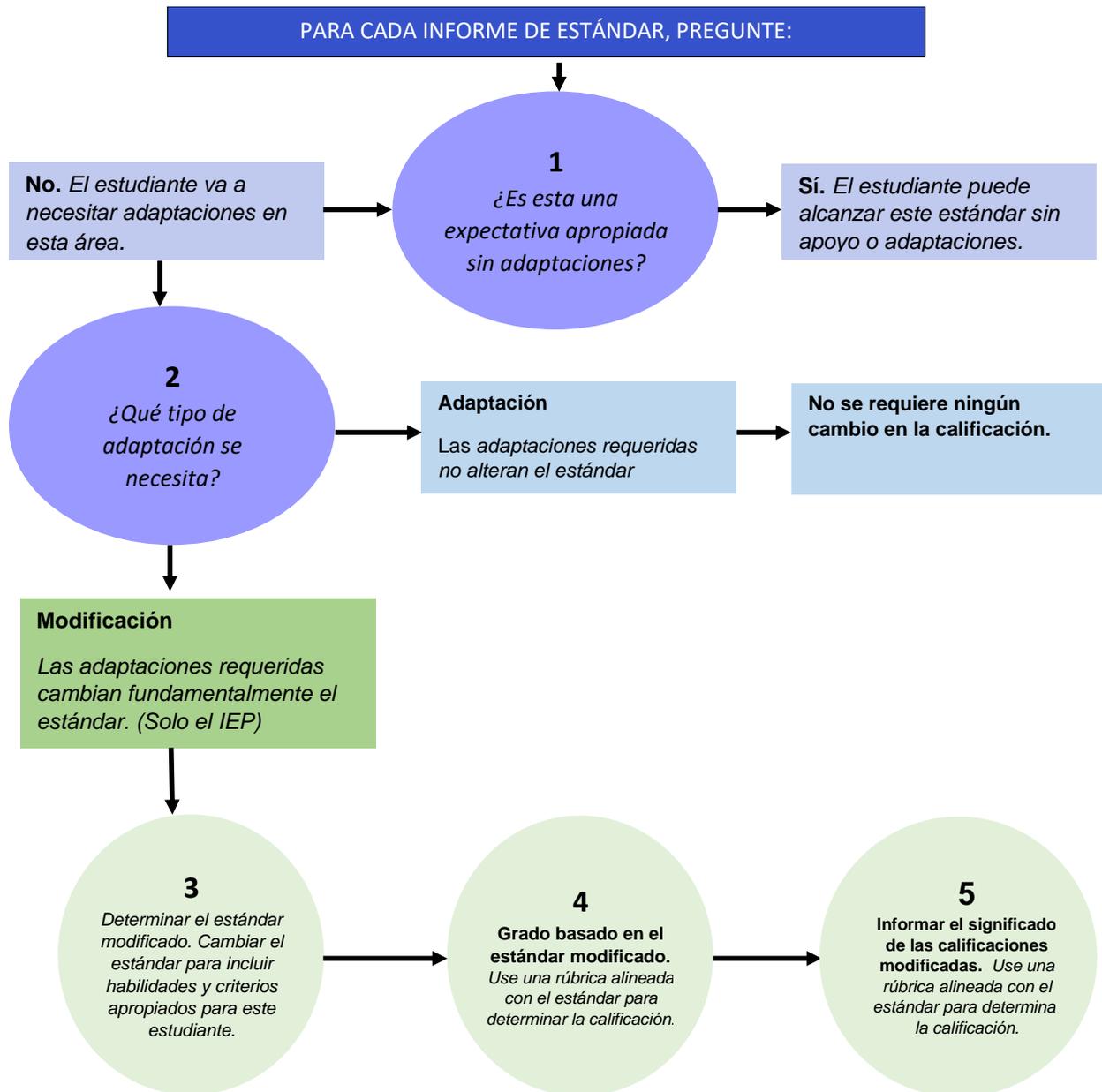
Los estudiantes que reciben adaptaciones, según su Plan de Educación Individualizado (siglas en inglés IEP) o 504, continuarán trabajando para cumplir con la competencia académica de nivel de grado en cada estándar prioritario. Como tal, recibirán el mismo informe de progreso/boleta de calificaciones que refleja los puntajes estándar de prioridad para el nivel de grado en el que se colocan, tal como lo haría

cualquier estudiante. Todos los estudiantes, independientemente de si están recibiendo servicios de educación especial, necesitan apoyo y alguna adaptación para tener éxito. La designación como estudiante que recibe servicios de educación especial no cambia las expectativas que ellos, sus padres y el Distrito Escolar de Hermiston tienen para su aprendizaje. Simplemente cambia las formas en que acomodamos y apoyamos su aprendizaje. Sin embargo, hay una excepción. Los estudiantes que reciben un plan de estudios modificado según su IEP, tendrán un sistema de informes que refleja su crecimiento hacia sus metas individuales establecidas en su IEP. Este informe se incluye como apéndice a la boleta de calificaciones. Jung (2018) tiene un diagrama útil que explica cómo funciona esto.

Modelo Diferenciado de Evaluación y Calificación



Una gráfica alternativa que representa una progresión similar fue desarrollada por Guskey y Jung (2007).



Referencias

- Ainsworth, L. (2013). *Prioritizing the common core: Identifying specific standards to emphasize the most*. Englewood, CO: The Leadership and Learning Center. American Speech-Language-Hearing Association. (n.d.).
- Alex, P. (2022). Time to Pull the Plug on Traditional Grading? Supporters of mastery-based grading say it could promote equity. *Education Next*, 22(4), 38-43.
- Craig T. A. (2011). *Effects of standards-based report cards on student learning*. (Doctoral dissertation). Retrieved from <https://repository.library.northeastern.edu/files/neu:1127>
- Dueck, M. (2014). *Grading smarter, not harder: Assessment strategies that motivate kids and help them learn*. Alexandria, VA: ASCD
- Federal Way Public Schools. (2023). *SBGR: Standards-Based grading and reporting: Staff handbook*. Retrieved June 22, 2023 from <https://www.k12.wa.us/sites/default/files/public/ossi/k12supports/pubdocs/SBGR%20Staff%20Handbook.WEB.pdf>
- Feldman, J. (2018). *Grading for equity: What it is, why it matters, and how it can transform schools and classrooms*. Corwin Press.
- Guskey, T. (2015). *Challenging the conventions of grading and reporting: On your mark*. Bloomington, IN. Solution Tree Press.
- Guskey, T.R., Jung, L.A. (2007). Standards-Based grading and reporting: A model for special education. *Teaching Exceptional Children*, 40(2), 48-53.
- Guskey, T. R., & Jung, L. A. (2009). Grading and reporting in a standards-based environment: Implications for students with special needs. *Theory Into Practice*, 48(1), 53-62.
- Jung, L. A. (2018). *Supporting and measuring growth for students with learning differences*. Alexandria, VA: ASCD.
- Marzano, R. (2000). *Transforming classroom grading*. Alexandria, VA: ASCD.

Orland School District 135. (2018). Standards-Based grading and handbook. Retrieved June 22, 2023 from <https://www.orland135.org/site/handlers/filedownload.ashx?moduleinstanceid=9366&dataid=11049&FileName=Standards-Based%20Grading%20Handbook%20%20WebVersion.pdf>

Powerschool. (2023). Everything you need to know about standards-based grading. Retrieved June 29, 2023 from <https://www.powerschool.com/blog/everything-you-need-to-know-about-standards-based-grading/>

Reeves, D. (2004). Making standards work: How to implement standards-based assessments in the classroom, school, and district. Englewood, CO: Advanced Learning Press.

Scriffiny, P. L. (2008). Seven reasons for standards-based grading. *Educational Leadership*, 66(2), 70-74.

Townsley, M., & Buckmiller, T. (2016). What Does the Research Say about Standards-Based Grading? A Research Primer. Online Submission.

Webb, N. L. and others. (2006). Web alignment tool. Wisconsin Center of Educational Research. University of Wisconsin-Madison.

Wormeli, R. (2011). Redos and retakes done right. *Educational Leadership*, 69(3), 22-26.