



# Factores que inciden en la calificación que los docentes asignan a los alumnos

Factors that affect the grade that teachers assign to students

Fatores que afetam a classificação que os professores atribuem aos alunos

ARTÍCULO DE REVISIÓN



José Pérez García   
resiljpg@gmail.com

María de los Ángeles López Ortega   
alopez25@cife.edu.mx

Centro Universitario CIFE. Cuernavaca, México

Escanea en tu dispositivo móvil  
o revisa este artículo en:

<https://doi.org/10.33996/revistahorizontes.v6i25.445>

Artículo recibido el 30 de abril 2022 | Aceptado el 21 de junio 2022 | Publicado el 30 de septiembre 2022

## RESUMEN

El propósito de este artículo fue determinar las prácticas de evaluación que sustentan la calificación asignada por los profesores a los alumnos y que tan válida y confiable es la información que sustenta la calificación. Se utilizó como metodología el análisis documental para identificar la literatura especializada sobre los factores que influyen en la calificación que los profesores asignan a los alumnos. Se analizaron 27 estudios y se encontró predominio de prácticas no alineadas a los programas de estudio, calificaciones basadas en una mezcla de factores cognitivos y no cognitivos y ausencia de formación en evaluación. Se concluye que es cuestionable la confiabilidad y validez de las calificaciones asignadas por los profesores, es necesario fortalecer la literacidad en evaluación de los docentes y directivos escolares, entrenamiento en prácticas de calificación efectivas y promover la evaluación formativa.

**Palabras clave:** Calificaciones escolares; Evaluación del aprendizaje; Evaluación formativa; Literacidad; Profesores

## ABSTRACT

The purpose of this article was to determine the evaluation practices that support the grade assigned by teachers to students and how valid and reliable is the information that supports the grade. Documentary analysis was used as a methodology to identify the specialized literature on the factors that influence the grade that teachers assign to students. Twenty-seven studies were analyzed and a predominance of practices not aligned with the study programs, grades based on a mixture of cognitive and non-cognitive factors and lack of training in evaluation were found. It is concluded that the reliability and validity of the grades assigned by teachers is questionable, it is necessary to strengthen the assessment literacy of teachers and school administrators, training in effective grading practices and to promote formative assessment.

**Key words:** School grades; Learning assessment; Formative assessment; Literacy; Teachers

## RESUMO

O objetivo deste artigo era determinar as práticas de avaliação que sustentam a classificação dos professores dos alunos e quais as práticas válidas e confiáveis são as informações que sustentam a classificação. A análise documental foi utilizada como uma metodologia para identificar a literatura especializada sobre os fatores que influenciam as notas que os professores atribuem aos alunos. Analisamos 27 estudos e encontramos uma predominância de práticas que não estavam alinhadas com o programa, notas baseadas em uma mistura de fatores cognitivos e não cognitivos, e uma falta de treinamento na avaliação. Conclui-se que a confiabilidade e validade das notas atribuídas pelos professores é questionável, é necessário fortalecer a alfabetização dos professores e administradores escolares, treinamento em práticas de classificação eficazes e promover a avaliação formativa.

**Palavras-chave:** Classificação escolar; Avaliação da aprendizagem; Avaliação formativa; Alfabetização; Professores

## INTRODUCCIÓN

Una de las responsabilidades que tiene el profesor en el aula es la de recopilar evidencias que sustenten en qué nivel se han logrado los aprendizajes esperados de un programa de estudios y asignar la calificación correspondiente a los alumnos. Este proceso de evaluación y calificación es relevante por las consecuencias que tiene ya sea de manera inmediata o cuando trasciende más allá de un ciclo escolar, porque de ello depende la continuación de estudios en grados superiores (Tierney, 2015); la calificación que se asienta en una boleta determina si el alumno es promovido de grado académico, si aparece en un cuadro de honor, si necesita clases de recuperación o si es aceptado en una institución superior (Guskey y Link, 2018); es decir, tiene un impacto en cuanto a las oportunidades futuras personales y educativas (Reeves, 2016).

La calificación, según la RAE (Real Academia Española, 2020), es la “puntuación obtenida en un examen o en cualquier tipo de prueba”. Para Brookhart y colaboradores (2016), son los símbolos que se asignan a los alumnos en las boletas de calificaciones, referidos a trabajos individuales o medidas compuestas de su desempeño. Haladyna y State (2019) la identifican como una variable cuantitativa, comúnmente expresada en número, que refleja el logro de desempeño de los alumnos. Un elemento adicional es que los números, letras o porcentajes asignados como calificación están basadas en estándares (Chen y Bonner, 2017).

A pesar de la relevancia de las calificaciones para la educación y la investigación, como señala Chen y Bonner (2017), no se ha documentado de

manera suficiente el proceso de toma de decisiones que explica cómo toma dicha decisión un docente y, es necesario conocer cómo están evaluando y calificando los profesores a sus alumnos para identificar en qué medida están considerando las pautas establecidas en los documentos oficiales como plan y programas de estudio (Gómez y Jakobsson, 2015), u otros documentos que orientan y regulan este proceso; o si sus prácticas están influenciadas por creencias personales sobre el aprendizaje, así como creencias y conocimientos sobre la calificación (Chen y Bonner, 2017). En la medida que exista una concordancia entre las calificaciones de los estudiantes y su desempeño, tendremos calificaciones efectivas y contaremos con prácticas válidas y confiables de calificación cuando estas estén alineadas con los criterios de evaluación (Svennberg et al., 2018).

La presente investigación tiene como propósito determinar las prácticas de evaluación que sustentan la calificación asignada por los profesores a los alumnos y que tan válida y confiable es la información que sustenta la calificación.

## METODOLOGÍA

Se realizó un estudio de revisión documental. En este tipo de estudio se identificó y sistematizó de manera cuidadosa la literatura relevante sobre un campo de conocimiento, metodología o enfoque teórico (Palmatier, Houston, y Hulland, 2018); además, los resultados de la búsqueda de publicaciones sobre el tema son reproducibles (Linares-Espinós, et al., 2018).

Para la búsqueda de la literatura se atendió a las tres recomendaciones que da Palmatier et al., (2018): 1) Reunir un cuerpo de investigaciones adecuado sobre el tema para que la síntesis sea valiosa; 2) debe estar bien ejecutado, con técnicas de análisis y literatura adecuada y 3) se deben ofrecer nuevas perspectivas a partir de la comparación sistemática de los estudios.

La búsqueda de la información se realizó en cuatro bases de datos especializadas: ERIC, Google Académico, SCOPUS y Web of Science y se utilizaron como criterios de inclusión que fueran publicaciones comprendidas en el periodo 2015-2020, publicadas en revistas indexadas y que correspondieran al ámbito de la educación básica y/o educación media superior. Se excluyeron publicaciones correspondientes a ponencias y tesis de grado, así como las referidas a campos diferentes al educativo.

En ERIC se utilizó la ecuación de búsqueda: (assessment AND grading) AND (secondary education) y se filtró el resultado para seleccionar artículos de revista de los años 2016 a 2020, en inglés. Se obtuvieron 56 resultados. En Google Académico se utilizó la ecuación de búsqueda: allintitle: "evaluación y calificación" + "educación" y se utilizó como filtro los años

2015-2020; se obtuvieron 12 resultados. En la base SCOPUS se utilizó la ecuación de búsqueda: (TITLE (assessment) AND TITLE-ABS-KEY (grading) AND TITLE-ABS-KEY (school)) AND (LIMIT-TO (PUBSTAGE, "final")). Se delimitó la búsqueda al periodo 2015-2020, artículos publicados en revistas indexadas. Se obtuvo como resultado 53 referencias. Adicionalmente se realizó una búsqueda relacionada con la práctica de la evaluación, con la ecuación: (TITLE (teachers) AND TITLE (assessment) AND TITLE (practice)), esta búsqueda se delimitó al periodo 2016-2020, artículos publicados en revistas indexadas y se localizaron 44 referencias. Finalmente, en la base Web of Science se buscó con la ecuación: TÍTULO: (assessment) AND TÍTULO: (teachers) AND TÍTULO: (practice). Se delimitó la búsqueda para artículos de revista arbitrada en inglés y/o español, en los años 2017-2020. Se encontraron 69 documentos. También se realizó una búsqueda de artículos sobre literacidad en evaluación, con la ecuación de búsqueda: TÍTULO: (assessment) AND TÍTULO: (teachers) AND TÍTULO: (literacy) se delimitó la búsqueda a los cinco años recientes y se localizaron 33 referencias (ver Tabla 1).

**Tabla 1.** Fuentes documentales identificadas.

Base de datos	Término de búsqueda	Salida
Google Académico	allintitle: "evaluación y calificación" + "educación"	12
ERIC	(Assessment and grading) and (secondary education)	43
Scopus	(TITLE (assessment) AND TITLE-ABS-KEY (grading) AND TITLE-ABS-KEY (school)) (TITLE (teachers) AND TITLE (assessment) AND TITLE (practice))	53 44
	(TITLE (teachers) AND TITLE (assessment) AND TITLE (literacy))	55
Web of Science	TÍTULO: (assessment) AND TÍTULO: (teachers) AND TÍTULO: (literacy)	33
	TÍTULO: (assessment) AND TÍTULO: (teachers) AND TÍTULO: (practice)	69

Todas las referencias se exportaron al gestor Mendeley, en donde se procedió a identificar duplicados para su eliminación. Posteriormente se realizó un cribado (Linares-Espinós, Hernández, Domínguez-Escrig, Fernández-Pello, Hevia, Mayor y Ribal, 2018) a partir de los títulos y resúmenes o con la revisión del texto completo cuando en el resumen no se tuvo información suficiente para decidir su inclusión. Finalmente, las referencias aceptadas de la búsqueda fueron 26 y se agregó un capítulo de libro citado con frecuencia en el tema.

Para la revisión de la información recopilada se establecieron cuatro categorías de análisis; 1) Evaluación y evaluación formativa; 2). Calificación; 3). Práctica de evaluación y 4). Literacidad en evaluación.

## **DESARROLLO Y DISCUSIÓN**

### **Evaluación y evaluación formativa**

En la revisión de la literatura se encontró que cinco investigaciones aluden a la relación del docente con la evaluación en general y la evaluación formativa en particular.

En una investigación se señala que los maestros tienen conocimiento o conocimiento parcial de los principios de evaluación; sin embargo, evidencian conflicto entre las recomendaciones de los expertos, la realidad en el aula y las políticas escolares (Chen y Bonner, 2017). Esta idea es reforzada también por la investigación de Mellati y Khademi (2018), quienes encontraron que los docentes con una baja literacidad en evaluación no usan la evaluación formativa, argumentando

limitaciones de tiempo, conocimiento y salario.

En otro estudio se exploró la actitud de maestros de primaria hacia la evaluación formativa (Ahmedi, 2019) y se encontró que los docentes manifiestan una actitud positiva hacia la evaluación formativa, sin embargo, no la implementan en el aula. En el mismo sentido, en un tercer estudio, se encontró una situación contradictoria pues aunque los maestros indicaron tener nociones de la evaluación formativa, en el aula aplican una evaluación sumativa (Fernández-Ruiz, y Panadero, 2020).

La cuarta investigación, es una revisión sistemática realizada por Schildkamp y colaboradores (2020), quienes identifican varios prerequisites que influyen en el uso de la evaluación formativa por parte del maestro. Dichos autores los agrupan en tres categorías: 1). Conocimiento y habilidades: acopio de información, construcción de instrumentos, análisis e interpretación de información, conocimiento pedagógico. 2). Factores psicológicos: actitudes y creencias, autonomía para tomar decisiones, presión social y 3). Factores sociales: colaboración con colegas e involucrar a los alumnos.

### **Práctica de evaluación**

En un primer grupo se identificaron siete estudios que dan cuenta de los instrumentos, técnicas y/o métodos de evaluación utilizados por los maestros. Utilizan pruebas o exámenes escritos (Arsyad, 2020; Gómez y Jakobsson, 2015; Fernández-Ruiz y Panadero, 2020; Narathakoon et al., 2020). También utilizan pruebas orales, trabajo individual y grupal, evaluación

formativa y sumativa (Isnawati y Saukah, 2017); retroalimentación verbal y no verbal (Seden y Svaricek, 2018); autoevaluación y autocalificación (Fraile Ruiz, Pardo, y Panadero, 2018); proyectos individuales y por equipo, evaluación del desempeño y, en menor medida, preguntas tipo ensayo (Cheng y Sun, 2015)

En un segundo grupo de estudios se ubicaron aquellos donde se identificaron los criterios utilizados por los maestros para evaluar. En la mayoría de los estudios se encontró que los maestros utilizan tanto factores cognitivos como no cognitivos (Arsyad Arrafii, 2020; Ashraf, y Zolfaghari, 2018; Cheng y Sun, 2015; Collazo, 2007; Guskey y Link, 2018; Marmeleira et al., 2020; Michael et al., 2016).

Además de los dos criterios anteriores, también señalaron la política escolar (Brookhart et al., 2016; y Isnawati y Saukah, 2017). Por su parte, en el estudio de Chiekem (2015), se encontraron sólo factores no cognitivos y, finalmente en el estudio de Seden y Svaricek (2018) el criterio más relevante fue la política escolar.

Un tercer grupo de investigaciones se refiere a dos estudios que tratan sobre la práctica de evaluación y qué tanto está alineada al programa. En uno de ellos se encontró que 74% de los maestros utilizó estándares para la evaluación y calificación de los alumnos (Michael et al., 2016); en el otro se encontró que la evaluación del aprendizaje está alejada de lo que establecen los principios teóricos y empíricos de la evaluación formativa (Fernández-Ruiz y Panadero, 2020).

## Literacidad en evaluación

En relación con los estudios referidos a la formación de los maestros en evaluación; la investigación de Michael et al., (2016) con docentes de educación física del estado de California, en EE.UU., encontró que 74% de los profesores refirió utilizar los estándares establecidos por el estado para evaluar y calificar a sus alumnos y quienes no los utilizaron, en su mayoría (91%) no tenían una formación sobre el tema. En el estudio de Collazo (2007), se reportó que 31% había aprobado un curso universitario específico de evaluación y, finalmente en el estudio de Cheng y Sun (2015), se reportó que 32% de los docentes tenía una formación en evaluación en lenguas.

En otro grupo de estudios se encontró que reportan escasa formación en evaluación en los profesores. Por ejemplo, en la investigación de Lam (2019), se encontró que los profesores tenían perfiles de alta calificación académica y una experiencia laboral sólida, pero un nivel básico de literacidad en evaluación. En otros estudios se reportó que los maestros sólo tienen una formación en un programa para docentes (Link, 2018), formación en el campo de la enseñanza del inglés (Tierney, 2015); otros reportaron que los profesores tienen poca o alguna formación (Chen y Bonner, 2017; Cheng et al., 2020; Guskey y Link, 2018); o que es incierta (Fernández-Ruiz, y Panadero, 2020); en otro estudio se reportó que tienen una formación en evaluación pero es inconsistente y poco útil (Battistone, Buckmiller, y Peters, 2019); otra investigación reportó

que 8 maestros de una muestra de 35 recibió entrenamiento en evaluación (Cheng, et al., 2020). En otros casos, recibieron una capacitación general en la implementación del plan de estudios (Isnawati y Saukah, 2017) y, finalmente Brookhart et al., (2016), establecieron en su estudio de revisión sistemática que la mayoría de los profesores no tienen una formación en el campo de la evaluación.

### **Calificación**

De acuerdo a los estudios revisados, algunas de las prácticas de evaluación se caracterizan por asignar ceros a trabajos incompletos y restar puntos por entregas tardías (Link, 2018) o utilizan el promedio y mecanismos de ajuste como redondeo, libretas, oportunidades adicionales, notas de clase diarias o eliminación de algunas notas (Collazo, 2007). En la investigación de Borghouts et al., (2017), el esfuerzo fue el factor más considerado para calificar; adicionalmente tomaron en cuenta las tácticas de juego, participación, conducta social y habilidades técnico/motoras.

En otro estudio de Guskey y Link (2018), se reportó que factores no cognitivos como participación en clase, hábitos de trabajo, limpieza, puntualidad en las entregas determinan del 10 al 20% de la calificación. En esta misma idea Seden y Svaricek (2018) reportaron en su estudio que la mayoría de los maestros utilizaron métodos obsoletos y contrarios a lo que establece la investigación de la evaluación en el aula.

En otra investigación se encontró que, además del rendimiento académico de los alumnos, los

maestros consideraron el esfuerzo, hábitos de estudio, participación, comportamiento para calificar y también influyó la política escolar (Isnawati y Saukah, 2017). Adicionalmente, hay casos en los que se alteraron calificaciones, por ejemplo Tierney (2015) reportó en su estudio el caso de 8 alumnos en los que se alteró la calificación por diferentes motivos: dar oportunidad a los estudiantes, motivos compasivos, y dar lecciones de vida. En un estudio se encontró que los maestros establecieron criterios de calificación y los dieron a conocer a sus alumnos (Lopez-Pastor, 2016).

### **Discusión**

En relación con los referentes que tienen los docentes para asignar las calificaciones, están los programas de estudio, la normatividad vigente referida a la evaluación y calificación, además de la formación y/o capacitación.

La revisión de la literatura en el presente estudio mostró que existen diferentes factores que determinan la calificación que el maestro asigna a los alumnos y existe algo que Guskey (2015) denomina ingenuidad compartida, en el sentido que el docente en la mayoría de los casos utiliza su idiosincrasia, sus métodos no están alineados con lo que establecen los estándares para el aprendizaje, además de que pocos maestros y directivos saben qué implementar para lograr un cambio en este tema. Lo anterior se confirma con lo que establece Guskey y Link, (2018) al señalar que las políticas de calificación de los maestros rara vez reflejan las recomendaciones de la investigación o están alineadas a enfoques

basados en estándares. En Suecia por ejemplo, como menciona Gómez y Jakobsson, (2015), después de 20 años de un cambio curricular, los docentes de ciencias siguen utilizando prácticas de evaluación tradicional en el aula.

La formación en el campo de la evaluación, sin embargo, no es factor determinante. Los participantes del estudio de Battistone et al., (2019), reconocieron haber recibido alguna formación sobre el tema, pero fue poco útil, pues no los involucró en el proceso de pensar y diseñar evaluaciones. En cambio, en el trabajo realizado por (López-Pastor et al., (2016) los participantes valoran positivamente su participación en un proceso de evaluación formativa al considerarlo como una experiencia con altas posibilidades de ser transferida a su práctica docente. Esto implica que cuenta la formación, pero más específicamente la metodología utilizada en dicho proceso.

Aquí puede resaltarse la necesidad de procesos de formación dirigidos a los futuros maestros o de capacitación a los que están en servicio utilizando por ejemplo la metodología de la evaluación socioformativa, en donde se parte de identificar y resolver problemas de contexto y la colaboración (Tobón, 2017). La capacitación o formación profesional debe apoyar a los maestros para que logren alinear su práctica de evaluación con lo que establece el currículo (Battistone, Buckmiller y Peters, 2019); debe contemplarse también el modelo propuesto por Schildkamp et al., (2020), que atiende tres factores importantes de la formación docente: social, psicológico y conocimientos.

Los docentes suelen considerar como algo complejo el proceso de calificación y esto puede deberse a carencias en la formación o capacitación, pero también de acuerdo con Collazo (2007), a la falta de sistematización de la información recopilada del estudiante y mezclar en una misma nota asuntos de diversa índole. Sin embargo, es aún limitada la investigación sobre la preparación que reciben los maestros en formación relacionada con la evaluación y la calificación (Battistone, Buckmiller y Peters, 2019).

En cuanto a las prácticas de evaluación que sustentan la calificación asignada por los docentes; para tomar la decisión ha realizado un proceso de acopio de información pertinente que refleja el desempeño del alumno; para ello utilizó instrumentos, técnicas, métodos, estableció criterios de evaluación y calificación; sin embargo, existen prácticas de evaluación que Reeves (2016) llama tóxicas porque tienen un impacto negativo en el rendimiento académico y conductual de los alumnos y se caracterizan por el uso del cero en una escala de cien puntos y el uso del promedio para calcular calificaciones finales.

También se debe señalar que en la mayoría de los estudios revisados la información no proviene de observación directa de la práctica docente y las respuestas obtenidas a través de un auto reporte pueden estar influidas por la deseabilidad social (Fernández-Ruiz y Panadero, 2020) y su práctica actual de evaluación y calificación puede ser diferente a lo que reportan (Michael et al., 2016). Por ejemplo, en la investigación de Ahmedi, (2019) se

encontró que existía en los docentes una actitud positiva hacia la evaluación formativa pero no la implementaban en el aula o, como se encontró en el estudio de Fernández-Ruiz y Panadero (2020) donde se señala que los docentes dicen tener nociones de la evaluación formativa, pero en su práctica aplican una evaluación sumativa. Hay disonancias entre las creencias y las prácticas de evaluación de los profesores como señalan Widiastuti y colaboradores (2020) y estas pueden deberse entre otros factores a la carga de trabajo del maestro, condiciones del salón de clase y limitaciones de tiempo.

Se ha señalado que la práctica de evaluación y calificación realizada por el docente no atiende las recomendaciones de la investigación y tampoco está alineada con los programas de estudio, tanto en el uso de instrumentos, técnicas y métodos, como en los criterios. La evaluación del aprendizaje que se realiza está alejada de lo que establecen los principios teóricos y empíricos de la evaluación formativa (Fernández-Ruiz y Panadero, 2020); pero aun teniendo conocimiento de evaluación y calificación, su práctica puede seguir igual porque existen otras condicionantes como la presión social y la política escolar; esto refuerza la idea de fortalecer en los docentes, además de los conocimientos, los factores psicológicos y sociales que señala (Schildkamp, et al., 2020).

En cuanto a los planteamientos que establecen el significado de la confiabilidad y validez de la información que sustenta la calificación, están expresados en dos estudios (Brookhart, 2017; Haladyna, 2019). La validez se entiende como la precisión en que la calificación

refleja el nivel de desempeño académico logrado por el alumno y la certeza de que efectivamente refleja lo que dijimos que calificamos y no otra cosa. Por su parte, la confiabilidad se refiere a la confianza que se tiene en la consistencia de esa medida.

Para que la calificación sea válida Haladyna (2019) sugiere que deben darse las siguientes condiciones: 1. Se identificaron los temas y se compartieron con los alumnos, 2) se proporcionó una instrucción que dio oportunidad al alumno de adquirir los conocimientos y habilidades establecidos, 3) los criterios de calificación representaron de manera justa y razonable el dominio, 4) se compartieron los principios de calificación con los alumnos y 5) los estándares de calificación fueron justos y se dieron a conocer el primer día de clases.

A partir de estos referentes se puede afirmar que es cuestionable la confiabilidad y validez en las calificaciones asignadas por los docentes, en los estudios revisados; excepto en el estudio de López-Pastor et al., (2016). El resto de los estudios revisados dan cuenta de calificaciones obtenidas a partir de una mezcla de factores cognitivos y no cognitivos; algunos criterios utilizados como asistencia, participación en clase, trabajos o créditos extra, esfuerzo y comportamiento, son denominados por Haladyna (2019) controversiales e inapropiados. Incluso en algunos casos se realizan prácticas tóxicas de evaluación (Reeves, 2016).

Para atender esta situación se debe promover en mayor medida el trabajo colaborativo ya que la mayoría de los docentes como señala Brookhart et al., (2016), además de no contar con una



capacitación para calificar a sus alumnos, realizan esta labor en solitario, rara vez realizan trabajo con sus colegas para verificar confiabilidad y validez de sus calificaciones. También los directivos que acompañan orientan y supervisan la actividad docente necesitan conocer como implementar cambios para modificar las prácticas de calificación.

La investigación apoya que se debe calificar basándose solamente en los estándares de desempeño establecidos en los programas de estudio y, en una evaluación separada considerar el esfuerzo y el comportamiento (Brookhart, 2017); se pueden evaluar además otros factores que inciden en el desarrollo integral del alumno pero estos se reportan por separado. Atendiendo a la validez de la calificación ésta se debe asignar considerando criterios alineados con el programa de estudio.

### CONCLUSIÓN

Para finalizar el estudio determinó que la tendencia en la práctica docente es el establecimiento de una evaluación formativa que involucre al alumno y la calificación basada en estándares alineados con los programas de estudio vigentes. Para poner en práctica los principios de evaluación y calificación establecidos en las reformas curriculares se deben acompañar con asesorías a docentes y directivos mediante estrategias que logren un impacto efectivo en el aula. Además, resulta cuestionable la validez y confiabilidad de las calificaciones asignadas a los alumnos, y cambiar las prácticas de evaluación y calificación de los maestros tiene el potencial de

beneficiar el éxito académico de los alumnos.

En futuras investigación se puede considerar la revisión de listas de evaluación o libro de evaluación del docente para verificar, si existe alguna ponderación en cuanto a los criterios de evaluación y en qué medida estos se alinean a lo establecido en el programa de estudios. Puede ser también de interés conocer qué formación tienen los maestros en el campo de la evaluación y calificación, sobre todo en el contexto peruano, así como investigar la participación de directivos y supervisores que deben tener un papel decisivo en la asesoría del docente y en la definición de políticas de calificación en la escuela alineadas a lo establecido en los programas educativos y lo que indica la normatividad sobre la evaluación y calificación.

### REFERENCIAS

- Ahmedi, V. (2019). Teachers' attitudes and practices towards formative assessment in primary schools. *Journal of Social Studies Education Research*, 10(3), 161–175. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85073470000&partnerID=40&md5=37a5f12686df73458db628cf4b211943>
- Arsyad Arrafii, M. (2020). Grades and Grade Inflation: Exploring Teachers' Grading Practices in Indonesian EFL Secondary School Classrooms. *Pedagogy, Culture and Society*, 28(3), 477–499. <https://doi.org/10.1080/14681366.2019.1663246>
- Ashraf, H., y Zolfaghari, S. (2018). EFL Teachers' Assessment Literacy and Their Reflective Teaching. *International Journal of Instruction*, 11(1), 425–436. <https://doi.org/10.12973/iji.2018.11129a>
- Battistone, W., Buckmiller, T., y Peters, R. (2019). Assessing Assessment Literacy: Are new teachers prepared to assume jobs

- in school districts engaging in grading and assessment reform efforts? *Studies in Educational Evaluation*, 62, 10–17. <https://doi.org/10.1016/j.stueduc.2019.04.009>
- Borghouts, L. B., Slingerland, M., y Haerens, L. (2017). Assessment Quality and Practices in Secondary PE in the Netherlands. *Physical Education and Sport Pedagogy*, 22(5), 473–489. <https://doi.org/10.1080/17408989.2016.1241226>
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., y Welsh, M. E. (2016). A Century of Grading Research: Meaning and Value in the Most Common Educational Measure. *Review of Educational Research*, 86(4), 803–848. <https://doi.org/10.3102/0034654316672069>
- Brookhart, S. M. (2017). *How to use grading to improve learning*. ASCD.
- Chen, P. P., y Bonner, S. M. (2017). Teachers' Beliefs About Grading Practices and a Constructivist Approach to Teaching. *Educational Assessment*, 22(1), 18–34. <https://doi.org/10.1080/10627197.2016.1271703>
- Cheng, L., y Sun, Y. (2015). Teachers grading decision making: Multiple influencing factors and methods. *Language Assessment Quarterly*, 12(2), 213–233. <https://doi.org/10.1080/15434303.2015.1010726>
- Cheng, L., DeLuca, C., Braund, H., Yan, W., y Rasooli, A. (2020). Teachers' grading decisions and practices across cultures: Exploring the value, consistency, and construction of grades across Canadian and Chinese secondary schools. *Studies in Educational Evaluation*, 67, 100928. <https://doi.org/10.1016/j.stueduc.2020.100928>
- Chiekem, E. (2015). Grading practice as Valid Measures of Academic Achievement of Secondary Schools Students for National Development. *Journal of Education and Practice*, 6(26), 24–29. <https://files.eric.ed.gov/fulltext/EJ1077389.pdf>
- Collazo, A. A. (2007). Una exploración a la intrincada experiencia de calificar el desempeño estudiantil. *Revista Pedagogía*, 40(1), 79–116.
- Fernández-Ruiz, J., y Panadero, E. (2020). Comparison between conceptions and assessment practices among secondary education teachers: more differences than similarities. *Infancia y Aprendizaje*, 43(2), 309–346. <https://doi.org/10.1080/02103702.2020.1722414>
- Fraile Ruiz, J., Pardo, R., y Panadero, E. (2018). Autoevaluación Y Autocalificación En El Grado En Ciencias De La Actividad Física Y Del Deporte: Estudio Censal De Las Guías Docentes. *Profesorado, Revista de Currículum y Formación Del Profesorado*, 22(3). <https://doi.org/10.30827/profesorado.v22i3.7997>
- Gómez, M., y Jakobsson, A. (2015). Science Teachers' Assessment and Grading Practices in Swedish Upper Secondary Schools. *Journal of Education and Training*, 2(2), 1. <https://doi.org/10.5296/jet.v2i2.7107>
- Guskey, T. R., y Link, L. J. (2018). Exploring the Factors Teachers Consider in Determining Students' Grades. *Assessment in Education: Principles, Policy and Practice*, 26(3), 303–320. <https://doi.org/10.1080/0969594X.2018.1555515>
- Guskey, T. R. (2015). *On your mark: Challenging the conventions of grading and reporting*. Solution Tree.
- Haladyna, T. M. (2019). Assigning a Valid and Reliable Grade in a Course. *IDEA*, 79(May), 1–15. <https://files.eric.ed.gov/fulltext/ED598950.pdf>
- Isnawati, I., y Saukah, A. (2017). Teachers' Grading Decision Making. In *TEFLIN Journal: A publication on the teaching and learning of English*. <http://dx.doi.org/10.15639/teflinjournal.v28i2/155-169>
- Lam, R. (2019). Teacher assessment literacy: Surveying knowledge, conceptions and practices of classroom-based writing assessment in Hong Kong. *System*, 81, 78–89. <https://doi.org/10.1016/j.system.2019.01.006>

- Linares-Espinós, E., Hernández, V., Domínguez-Escrig, J. L., Fernández-Pello, S., Hevia, V., Mayor, J., y Ribal, M. J. (2018). Methodology of a systematic review. *Actas Urológicas Españolas*, 42(8), 499–506. <https://doi.org/10.1016/j.acuro.2018.01.010>
- Link, L. (2018). Teachers' Perceptions of Grading Practices: How Pre-Service Training Makes a Difference. *Journal of Research in Education*, 28(1). <https://files.eric.ed.gov/fulltext/EJ1168160.pdf>
- Lopez-Pastor, V. M., Perez-Pueyo, A., Barba, J. J., y Lorente-Catalan, E. (2016). Students' perceptions of a graduated scale used for self-assessment and peer-assessment of written work in pre-service physical education teacher education (PETE). *Cultura\_ciencia\_deporte*, 11(31), 37–50. <https://doi.org/10.12800/ccd.v11i31.641>
- Marmeleira, J., Folgado, H., Martínez Guardado, I., y Batalha, N. (2020). Grading in Portuguese secondary school physical education: assessment parameters, gender differences and associations with academic achievement. *Physical Education and Sport Pedagogy*, 25(2), 119–136. <https://doi.org/10.1080/17408989.2019.1692807>
- Mellati, M., y Khademi, M. (2018). Exploring teachers' assessment literacy: Impact on learners' writing achievements and implications for teacher development. *Australian Journal of Teacher Education*, 43(6), 1–18. <https://doi.org/10.14221/ajte.2018v43n6.1>
- Michael, R. D., Webster, C., Patterson, D., Laguna, P., y Sherman, C. (2016). Standards-Based Assessment, Grading, and Professional Development of California Middle School Physical Education Teachers. *JOURNAL OF TEACHING IN PHYSICAL EDUCATION*, 35(3), 277–283. <https://doi.org/10.1123/jtpe.2015-0162>
- Narathakoon, A., Sapsirin, S., y Subphadoongchone, P. (2020). Beliefs and classroom assessment practices of english teachers in primary schools in Thailand. *International Journal of Instruction*, 13(3). <https://doi.org/10.29333/iji.2020.13310a>
- Palmatier, R. W., Houston, M. B., y Hulland, J. (2018). Review articles: purpose, process, and structure. *Journal of the Academy of Marketing Science*, 46(1), 1–5. <https://doi.org/10.1007/s11747-017-0563-4>
- Real Academia Española. (2020). *calificación | Definición | Diccionario de la lengua española | RAE - ASALE*. Diccionario de La Lengua Española. <https://dle.rae.es/calificación>
- Reeves, D. (2016). *Elements of grading. A guide to effective practice* (S. T. Press (ed.); Second Edi).
- Schildkamp, K., van der Kleij, F. M., Heitink, M. C., Kippers, W. B., & Veldkamp, B. P. (2020). Formative assessment: A systematic review of critical teacher prerequisites for classroom practice. *International Journal of Educational Research*, 103. <https://doi.org/10.1016/j.ijer.2020.101602>
- Seden, K., y Svaricek, R. (2018). Teacher subjectivity regarding assessment: Exploring english as a foreign language teachers' conceptions of assessment theories that influence student learning [Subjektivnost učiteljev pri ocenjevanju: Raziskovanje pojmovanj teorij ocenjevanja, ki vplivaj. *Center for Educational Policy Studies Journal*, 8(3), 119–139. <https://doi.org/10.26529/cepsj.500>
- Svennberg, L., Meckbach, J., y Redelius, K. (2018). Swedish PE teachers struggle with assessment in a criterion-referenced grading system. *SPORT EDUCATION AND SOCIETY*, 23(4), 381–393. <https://doi.org/10.1080/13573322.2016.1200025>
- Tierney, R. D. (2015). Altered grades: A grey zone in the ethics of classroom assessment. *Assessment Matters*, 8(5), 5–30. <https://doi.org/10.18296/am.0002>

Tobón, S. (2017). *Evaluación socioformativa. Estrategias e instrumentos* (Issue October). <https://doi.org/10.24944/isbn.978-1-945721-26-7>

Widiastuti, I. A. M. S., Mukminatien, N., Prayogo, J. A., y Irawati, E. (2020). Dissonances between teachers' beliefs and practices of formative

assessment in EFL classes. *International Journal of Instruction*, 13(1), 71–84. <https://doi.org/10.29333/iji.2020.1315a>