

Teachers' Grading Practices: In Search for Clear Grading Criteria

Siti Zulaiha*

University of Muhammadiyah Prof. DR. HAMKA (UHAMKA), Jakarta, Indonesia

DOI: 10.22236/JER_Vol2Issue1pp15-23

Grading is a complicated process yet serves many significant roles such as judging students' achievement and providing parental feedback. However, studies suggest that grading practices are not always uniform in that non-achievement factors are also frequently included in determining grades. Thus it is important to review and further understand teachers' grading practices. This article explores the discussion and debates around grading practices both in the general education field as well as studies in the English as a foreign language (EFL) context. The discussion, in particular, focuses on the range of factors that influence teachers' grading practices. Finally, this article reflects on implications of grading practices and offers suggestions that have been commonly recommended by experts.

Keywords: grade, grading, achievement, non-achievement factors

Penilaian merupakan proses rumit namun memberikan peran penting untuk mengetahui perkembangan siswa dan memberikan umpan balik. Sejumlah penelitian menyarankan proses penilaian tidak harus seragam karena factor-faktor non-pencapaian sering diikutsertakan dalam menentukan nilai, sehingga penting untuk mengulas dan mengerti lebih jauh tentang praktek penilaian guru. Artikel ini menampilkan diskusi dan debat tentang masalah ini dalam kontek pendidikan secara umum dan pembelajaran bahasa Inggris sebagai bahasa asing. Pembahasan dalam artikel ini fokus pada faktor-faktor yang mempengaruhi praktek penilaian guru. Di bagian akhir, artikel ini memberikan refleksi praktek penilaian dan memberikan usulan yang sebagian besar telah direkomendasikan oleh para pakar.

* Corresponding author. Email: siti.zulaiha@uhamka.ac.id

INTRODUCTION

Grading is a major component of educational practice. Grades are frequently a significant issue among teachers, students, and parents. While some teachers may enjoy grading student achievement, there are certainly teachers who find the grading process distressing. From a student perspective, a student may frequently question a grade assigned by a teacher. Similarly, many parents question teachers as to why their children are assigned a particular grade or specifically a lower grade. Thus teachers' grading practices are under scrutiny in relation to what grade a teacher assigns actually means or does not mean (Brookhart, 1994; Brown, 2004; Cheng & Sun, 2015; Randall & Engelhard, 2010). In other words, what an assigned grade represents is commonly under scrutiny by parents. For example, does getting an 'A' in an English writing course reflect the actual student's writing skill? The answers to questions such as this will reveal the specific factors of teachers' grading practices.

The question about what should be comprised in grades has become the focus of ongoing debates. Some assessment experts and educators advise that grades should only be based on academic achievement (Gronlund, 1998; McMillan, 2008). However, other studies offer evidence that while teachers consider achievement factors when grading students, they also commonly include non-achievement factors when assigning the final grade to students (Cheng & Sun, 2015; Liu, 2013; Randall & Engelhard, 2009; 2010; Sun & Cheng, 2013; Yesbeck, 2011). In light of this, this article discusses teachers' grading practices in the general education field as well as in the context of EFL and clarifies the meaning of grading. The article further reviews several studies on teachers' grading practices and focuses on the factors influencing teachers' decision making in grading. Finally, reflections and implications are discussed.

MAKING MEANING OF GRADING

Grades have a great impact on students. At a glance, the process of assigning grades seems simple. Teachers assess students and give students letter or number grades based on students' performance. However, in reality, the grading practice is notably complex. In fact, researchers have long been attempting to clarify and comprehend the meaning and purposes of grading.

A grade is defined as a symbol of a students' achievement level of the learning goals in a particular reporting period (Frisbie & Waltman, 1992). Brookhart (1994) defines grading as "the reporting of student achievement or progress for a report card period or a term" (p. 120). Grades should just reflect academic performance towards learning goals (Randall & Engelhard, 2010). From these definitions, it is clear that a grade should accurately represent students' mastery of the course materials (Liu, 2013). Nonetheless, Wormeli (2006) states that many teachers use grades to communicate students' performance in both academic and other factors such as effort, conduct, and attitude.

Many experts advocate that a grade should represent a student's attainment in that teachers should only consider the achievement factor in their grading decision making. Gronlund (1998) for example, contends that grades should only include achievement factors as the inclusion of non-achievement factors will misrepresent the meaning of the grade as evidence of the student's achievement. Similarly, Linn & Miller (2005) question the appropriateness of the inclusion of another element in the final grade. Likewise, Brookhart (2004) recognizes the need for teachers to assess non-achievement variables. However, she

suggests that these variables must not be graded. Several researchers further argue that grading only on students' achievement will inform students, parents, and other stakeholders of a student's actual achievement and progress. Thus the inclusion of non-achievement factors may risk the inferences made of grades (Brookhart, 1993; Cheng & Sun, 2015; Zoeckler, 2007).

Literature shows that meaning is variedly defined, and that grading has multi purposes. There seems no congruence between experts' advice and teachers' perceptions and practices. Thus, how teachers determine grading will influence their grading practices and the validity of the grades they assign. Cheng and Wang (2007) noted that teachers' beliefs and their considerations of the purpose and consequences of grading become the basis of teachers' practices. In the next section, studies on teachers' grading practices are discussed to understand the complexity of this issue.

TEACHERS' GRADING PRACTICES: WHAT STUDIES TELL US

Research has documented that teachers' grading practices are significantly variable. Teachers consider a variety of factors and criteria in assigning grades to their students (Brown, 2004; Sun & Cheng, 2013). In the general field of education, extensive studies investigating teachers' grading practices have been done. Studies indicate that teachers' grading practices are not consistent with the recommended practices provided in the literature. Many teachers use achievement as well as non-achievement factors when assessing and grading students. The most commonly mentioned factors are effort, behavior, participation, study habits, and improvement as found in the following reviewed studies.

In 2001, McMillan & Lawson examined 213 secondary school teachers' practices representing urban, suburban, and rural schools. Their study investigated the factors that teachers used to grade students, the types of assessments used, the cognitive level of assessments, and the grades assigned. The results indicated that teachers used a variety of factors to grade students. The most important factor was academic performance, but effort, participation, and extra-credit work were also important components of elementary school teachers' grading decisions.

Conducted in a primary school context, McMillan, Myran, & Workman (2002) examined assessment and grading practices of over 900 grades responsible for grading class three to five students. The teachers in this research specified the various factors they used to grade students, and the grades awarded. The findings showed that teachers used an assortment of factors; placing most weight on academic performance as well as academic enabling behaviors such as effort and improvement, and much less emphasis on homework, comparisons with other students, grade distributions of other teachers, and borderline cases.

In a survey study carried out in both elementary and middle school, Randall and Engelhard (2009) reported similar results with teachers considering non-achievement factors such as behavior and effort when assigning final grades. Randall and Engelhard examined differences in the grading practices between elementary and middle school teachers. The 234 public school teachers were asked to respond to 53 scenarios to illustrate a variety of student characteristics that may influence the grades that teachers gave. Results indicated that, in general, elementary school teachers gave higher grades than did their middle school colleagues. Randall and Engelhard further concluded that grade-level differences in grading practices might be related to student behavior and effort.

In a later study involving 516 American public school teachers, Randall and Engelhard (2010) investigated the meaning of grades and aspects teachers consider when giving final grades with a focus on borderline cases. Randall and Engelhard found that in general although teachers followed the local grading policy, they tended to use non-achievement factors for several borderline cases.

In the teaching of English as a foreign language (EFL) context, very few studies have been done on teachers' grading practices. However, the findings of these studies show that the issue of non-achievement factors influencing grades is also common in the EFL context. Sun and Cheng (2013) surveyed 350 junior and senior school English language teachers in China to examine the meaning of grades teachers assigned to their students and the value judgments they made in grading. Regarding construct validity, the findings indicated that the teachers included achievement as well as non-achievement factors such as effort, habit, attitude, improvement, and motivation in assigning grades. Moreover, among the non-achievement factors mentioned the teachers in this context placed particular weight on effort. Recently, Cheng and Sun (2015) published their study on secondary school English language teachers in China focusing on the factors and types of assessment teachers used for grading. The results showed that the teachers placed greater weight on non-achievement factors (e.g. effort, homework, and study habits) when grading. Moreover, these teachers used multiple types of assessment including performance and project-based assessment, teacher self-developed assessment, as well as paper and pencil testing.

All of these studies together confirm that teachers consider a variety of factors when grading for different reasons. Researchers have documented that teachers have a desire to be fair to students (McMillan, 2001; Zoeckler, 2007). Teachers tended to use non-cognitive factors when dealing with borderline cases resulting in lower proficiency students receiving additional scores (Randall & Engelhard, 2010). Teachers were also concerned about the consequences of grades increasing students' confidence and good behavior (Brookhart, 1994). There was also evidence that teachers used non-achievement factors to motivate students. In short, teachers used grades not only as an indicator of students' achievement but also as a tool in classroom management and student motivation.

REFLECTIONS AND IMPLICATIONS

It is apparent that grading practices are not as simple as it would seem. The complexity of grading has been marked by the different beliefs and values held by teachers, researchers, and assessment experts. It is interesting to note, although studies on grading practices were conducted across various educational levels, subject matters, and countries, that studies yield similar finding that teachers consider both achievement and non-achievement factors when assigning the final grade. The most noteworthy issue is perhaps the fact that regardless of the concerns about validity and reliability of grading, the grading practices remain the same. In a recent extensive literature review of research in grading practices, Brookhart, Guskey, Bowers, McMillan, Smith, Smith, Stevens, and Welsh (2016), also confirm that over 100 years of research, "grades typically represent a mixture of multiple factors that teachers value" (p. 834).

Given this fact, the questions that should be considered now are how should we better approach the practice of grading? and What are factors to consider for a more accurate, reliable and thus more accountable grading? Several experts commonly agree to factors to be

considered to improve grading practices, yet, consideration of these factors is not without constraints. These factors are particularly relevant within an EFL context where it is particularly important to provide legitimate grading both in terms of outcomes and ongoing motivation of students.

One of the crucial factors to consider is context. Context involves social and cultural factors, variations between education systems, individual grading policies and the specific beliefs that stakeholders in a particular institution hold about grades and the grading process. When considered together, this combination of factors will influence teachers' grading practices. Studies have shown that one of the factors affecting teachers' grading practices is the local/institution grading policy (Zoeckler, 2007). Teachers, therefore, have to comply in their practices with the policy in their teaching context. Brown (2004) elucidates that each institution has its own grading beliefs and grading regulations. Brown further says that "some institutions, for example, mandate deductions for unexcused absences. Others require that only the final exam determines a course grade" (p. 285). Unfortunately, not many studies to date look into the beliefs of stakeholders in a variety of contexts, especially in Asian EFL settings and in particular in Indonesia. Most studies focus on teachers' practices, yet, teachers experience pressure from the principal, parents, and students when assigning grades. This includes the pressure to change grades (Tierney, 2015). Understanding such a context is essential.

Teachers' roles, moreover, as perceived by teachers and the local educational community, may also determine teachers' grading processes. A study conducted by Sun and Cheng (2013) of secondary EFL teachers in China showed that "the teacher's role in educating students as an all-rounded person is more emphasized, especially at the K-12 school level" (p.340). They further stated that teachers are not only required to nurture the "intellectual aspect" but also students' "emotional necessities". Thus the teachers' multi-roles as an educator, mentor and, counselor are embedded in teachers' grading practices where teachers consider non-achievement factors to hold themselves accountable for the roles they perceived.

However, studies indicate that although stakeholders hold different beliefs about grading that in fact these beliefs are not necessarily shared nor negotiated, parents and students view grades differently to teachers or administrators. Also, teachers do not share their grading practices with fellow teachers (Kain, 1996). Given this, Friedman and Frisbie (2000) advocate that teachers and administrators need to have a common understanding of what constitute a grade as a grade should reflect actual students' achievement and progress which are very important for parents to know. It is therefore of vital importance that stakeholders in one specific context share the beliefs and practices they hold about grading.

Another factor that is commonly proposed for improved grading practices is training on assessment and grading. Researchers have documented that one of the reasons causing invalid and unreliable grading is a lack of teacher training in essential grading practices (Brookhart et al., 2016; Brookhart, 1994; Stiggins, 1998). It is argued that training will help teachers to develop clear criteria and observe as well as interpret achievement and non-achievement factors accurately (Brookhart et al., 2016; McMillan, 2003). However, the content of the training should not solely focus on grading but also focus on how assessment and instruction including feedback are conducted in a way that will help students to improve their learning. Moreover, it is also important that school administrators are involved in such training.

Training is a great opportunity for both administrators and teachers to come together and thus have shared perceptions of grading practices for the school. In this way administrators can also support teachers in their practices. The need for such training also has implications for pre-service teachers Training Institution. The system within Teacher Training Institution should include and discuss a particular topic on language assessment and grading to consider the many complex issues research in this field has shown.

The next factors worthy of attention is the need to develop and communicate clear purposes and grading criteria. Studies reveal that many teachers do not have a set of clear criteria as the basis for assigning grades. Clear criteria will hold teachers accountable in grading. Regarding non-achievement factors, Brookhart (1993) acknowledges that there is the need for teachers to accommodate non-achievement factors in grading such as effort and achievement as a consequence of roles and responsibilities that are attached to teachers. Responding to this issue, some experts posit a solution for teachers to assess and grade achievement as well as non-achievement factors separately. This is true in the case of Indonesia.

The Indonesian 2013 Curriculum promotes character building and thus has three aspects of graduate competency standard, i.e., attitude, skills, and knowledge. In this framework, EFL teachers have to assess and grade students' language performance as well as students' attitude. Although the two factors are reported separately on the report card, this practice is not unproblematic. One still can question what counts as achievement and non-achievement factors? For example, does teachers' assessment of language performance purely reflects the performance factor only. Research evidence shows that teachers who actually state that they only consider the achievement factor were also found to occasionally include factors such as effort and attitude when grading (Hay & Macdonald, 2008).

Given this fact, the challenge then remains as to how to measure non-achievement factors validly and reliably. Brookhart (1994) argues that it should be made clear what constitutes effort, attitude, and achievement when grading. Nonetheless, Brown (2004) contends that teachers need to acknowledge the subjectivity of non-achievement factors by converting such factors into clear and assessable outcomes. The criteria then should be specified in a rubric. Czekanski and Wolf (2013) emphasizes the needs for an explicit rubric as it provides structured criteria and detailed expectations used for assessing and scoring students' overall achievement. Moreover, teachers should be encouraged to involve students in developing scoring criteria.

Many works of literature support the idea that students should have a voice in designing and establishing the criteria so that students will have a better comprehension of teachers' expectations and how to meet them as well as having control over their learning (Czekanski & Wolf, 2013). However, research shows that this rarely occurs. Interestingly, in an interview of 79 ESL/EFL University teachers in Canada, Hong Kong, and China, Cheng and Wang, (2007) found that although most teachers in the three contexts developed their rubrics for grading, they did not involve their students in creating the scoring criteria. Moreover, the institution grading policy should carefully consider the weight of non-achievement factors in the overall scoring criteria listed in the rubric. Brown (2004) proposes that non-achievement factors should receive a small amount of weight so that a grade mostly represents students' achievement. The criteria

and the weight that are presented in a rubric then need to be communicated to students early in the learning period so that students know what they will be assessed and graded on.

Meaningful and regular feedback has been seen as an influential factor in improving students' learning and achievement as well as grading practice. Grading and feedback are perhaps two of the prevailing approaches in which teachers communicate students' learning progress and achievement to students and other relevant stakeholders (Guskey, 1987). Research has found that feedback has impacts on students' motivation and learning (e.g., Hattie & Timperley, 2007). In the case of grading, students perceive grades as their responsibility if they are given regular feedback (Cheng & Wang, 2007). Moreover, frequent feedback helps students to change their attitudes and strategy before the final grade is assigned (Brown, 2004).

Nevertheless, studies show that there are factors that influence teachers in providing comprehensive feedback. One of them is class size. This applies in particular in the Chinese context as Cheng and Wang (2007) found in their study that most teachers had large classes and thus found that giving an individualized feedback was a real challenge. It is perhaps also true in Indonesian context where EFL teachers in one school have several classes to teach, and each class has a large number of students. In this case, teachers should explore and judge the best possible way of giving feedback in a manner that is in the students' best interest. Feedback, after all, is crucial to help students to make meaning and also hold students accountable for their learning and maybe of more importance than the actual grading itself in terms of student's motivation and learning (Wormeli, 2006).

Last but not least is the importance of using multiple assessment strategies before teachers assign the final grades. For example, in language assessment, "performance does not indicate actual competence..." (Brown, 2004, p. 117). Therefore, Brown further argues that to measure the actual competence, teachers should not use one single assessment strategy. The use of multiple assessment strategies such as standardized tests, teacher-made tests, portfolios, and journals provide a documentation of students performances that better reflect competence and hence provide a more reliable and valid grade.

In summary, grading practice is context-specific. Moreover, grading has to be seen as an integral part of instruction and assessment process. Therefore, the above factors need to be considered in light of the local context. The challenge for researchers now is to document more practices in a variety of contexts that focus not only on grades but also on how grades relate to assessment, instruction, and socio-cultural aspects that are specific to the teachers' teaching context (Brookhart et al., 2016). Studies of this nature are needed especially in Indonesia as studies on grading practices are currently scarce. The findings of future studies hopefully will provide us with a better understanding of the complexity of grading practices to develop better classroom practice.

CONCLUSIONS

Grading is unarguably a complex issue that must take into account the values and beliefs held by all stakeholders. Teachers define the meaning and purposes of grading differently, and this in itself leads to different practices that consider both achievement and non-achievement factors and thus do not necessarily comply with what experts have recommended. However, there is clearly a need for teachers to assign grades both in terms of ongoing reporting procedures and as part of classroom management tools. Therefore, there should be an effort to

help teachers clarify grading criteria that is grounded in courses provided in teacher training institutions. Factors such as the impact of the specific local institution and pressures from this system to moderate grades and the complexity of teachers' multiple roles must all be considered. A joint commitment to an improved grading practice via strategies such as the development of strong grading criteria to include multiple grading processes, a consideration of variables such as class size, grade level difference, and ongoing discussion about the means in which non-achievement factors can be adequately assessed will offer a sound opportunity to make a difference and is in the best interest of all students.

REFERENCES

- Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement*, 30(2), 123–142.
- Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education*, 7(4), 279–301.
- Brookhart, S. M. (2004). *Grading*. Upper Saddle River, NJ: Pearson-Merrill-Prentice Hall.
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., Mcmillan, J. H., Smith, J. K., Smith, L. F., ... Welsh, M. E. (2016). A century of grading research : Meaning and value in the most common educational measure, 86(4), 803–848.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. White Plains, NY: Pearson Education.
- Cheng, L., & Sun, Y. (2015). Teachers' grading decision making: Multiple influencing factors and methods. *Language Assessment Quarterly*, 12(12), 213–233.
- Cheng, L., & Wang, X. (2007). Grading, feedback, and reporting in ESL/EFL classrooms. *Language Assessment Quarterly*, 4(1), 85–107.
<https://doi.org/10.1080/15434300701348409>
- Czekanski, K. E., & Wolf, Z. R. (2013). Encouraging and evaluating class participation. *Journal of University Teaching and Learning Practice*, 10(1), 1–10. Retrieved from <http://eds.a.ebscohost.com/eds/detail/detail?sid=5e5b9762-5c31-4e67-ae98-07856c6880cb@sessionmgr4003&vid=0&hid=4111&bdata=JnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2l0ZQ==#AN=EJ1005280&db=eric>
- Friedman, S. J., & Frisbie, D. A. (2000). Making report cards measure up. *Education Digest*, 65(5), 45–50.
- Frisbie, D. A., & Waltman, K. K. (1992). Developing a personal grading plan. *Educational Measurement : Issues and Practice*, 11, 35–42.
- Gronlund, N. E. (1998). *Assessment of student achievement*. Boston: Allyn and Bacon Huba.
- Guskey, T. R. (1987). Rethinking mastery learning considered. *Review of Educational Research*, 57, 225–229.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Hay, P. J., & Macdonald, D. (2008). Mis)appropriations of criteria and standards- referenced assessment in a performance-based subject. *Assessment in Education*, 15, 153–168.
- Kain, D. L. (1996). Looking beneath the surface: Teacher collaboration through the lens of grading practices. *Teachers College Record*, 94(4), 569–587.
- Linn, R., & Miller, M. (2005). *Measurement and assessment in teaching*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Liu, X. (2013). Investigating factors influencing grading decisions among teachers of chinese to speakers of other languages.
- McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice*, 20(1), 20–32.

- McMillan, J. H. (2003). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Measurement: Issues and Practice*, 22(4), 34–43.
- McMillan, J. H. (2008). *Assessment essentials for standards-based education* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- McMillan, J. H., & Lawson, S. R. (2001). Secondary science teachers' classroom assessment and grading practices. Retrieved from ERIC database. (ED 450 158)
- McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *The Journal of Educational Research*, 95(4), 203–213.
- Randall, J., & Engelhard, G. (2009). Differences between Teachers' Grading Practices in Elementary and Middle Schools. *Journal of Educational Research*, 102(3), 175–185.
- Randall, J., & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education*, 26(7), 1372–1380. <https://doi.org/10.1016/j.tate.2010.03.008>
- Stiggins, R. J. (1998). Revitalizing classroom assessment: The highest instructional priority. *Phi Delta Kappan*, 68, 363–368.
- Sun, Y., & Cheng, L. (2013). Teachers' grading practices: meaning and values assigned. *Assessment in Education: Principles, Policy & Practice*, 30(March 2013), 1–18.
- Tierney, R. D. (2015). Altered grades : A grey zone in the ethics of classroom assessment, (5), 5–30. Retrieved from http://www.nzcer.org.nz/system/files/AM2015_8_005.pdf
- Wormeli, R. (2006). Accountability: Teaching through assessment and feedback, not grading. *American Secondary Education*, 34(3), 14–28.
- Yesbeck, D. M. (2011). Grading Practices: Teachers' considerations of academic and non-academic factors. *Grading Practices: Teachers' Considerations of Academic and Non Academic Factors*, (December). Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:No+Title#0>
- Zoeckler, L. (2007). Moral aspects of grading: A study of high school English teachers' perceptions. *American Secondary Education*, 35(2), 83–102.