



Research Article

© 2019 Eric Anane and Jephhtar Adu-Mensah.
This is an open access article licensed under the Creative Commons
Attribution-NonCommercial-NoDerivs License
(<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

Teachers Scoring and Grading of Students' Responses to Tasks: The Ghanaian Basic School Experience

Eric Anane

*Institute of Education,
University of Cape Coast
Corresponding Author*

Jephhtar Adu-Mensah

Association of African Universities

Doi: 10.2478/ajis-2019-0009

Abstract

This paper explored the scoring and grading practices of teachers in basic schools. The study sought to understand the various factors teachers take into account when scoring students' responses to tasks in the classroom. The stratified sampling procedure was used to select 278 primary and junior high school teachers from whom data were collected for the study. In the descriptive analysis, the results revealed that teachers considered students' behaviour such as punctuality, interest in subject and neatness of work presented, and number of questions attempted by a student, and awarded marks when grading. These academic and non-academic factors considered by teachers in this study during scoring and grading are consistent with the literature. In independent samples t-test analyses, primary school teachers did not differ significantly from their counterparts in the junior high schools in terms of factors they considered during the scoring and grading of students' responses to tasks in the classroom. This study adds a circumstantial data to the existing debate on teacher classroom assessment practices.

Keywords: *non-academic factors, scoring and grading practices, classroom assessment*

1. Introduction

In the current era of student answerability, together with high-stakes testing, schools have focused on the synchronisation of curricula and assessments. At present, assessment is viewed by experts as an invisible thread in the teaching and learning environment. Teacher scoring and grading practices is therefore seen as an essential part in the quest to improving students' learning. However, developing standardised grading practices is still under examination in many education systems. Measurement experts recommend that grading of educational constructs such as achievement should be accurate (Ebel & Frisbie, 1991). However, achieving this condition has proven a daunting task for both experienced and novice teachers, because, when a student's grade represents multiple components of abilities, it is very difficult to validate and report same in relation to students' performance.

In addition to Ebel and Frisbie's (1991) view that there is no test score free of error, researchers have recognised and identified some challenges with the assessment practices of teachers since the introduction of continuous assessment in schools (West Africa Examinations Council [WAEC], 1990 & 1993; Winger, 2005; Guskey, 2006; Wormeli, 2006; Dueck, 2014). For example, a study conducted by WAEC in 1990 and 1993 revealed that teachers appeared to be generous in the awarding of marks to their students. Furthermore, Allen, Gregory, Mikami, Lun,

Hamre, and Pianta (2013) and Anane (2011) found in their respective studies that teachers make conscious effort to make students get close to the maximum mark, especially with pass marks established by external examining bodies.

Grading practices continue to be contentious and misconstrued. The usefulness of classroom assessment and grading of students' responses to assignments and tasks has become progressively imperative in this era of teacher accountability (Liu, 2008; Bonesronning, 2004). There seems to be disconnect between grading purposes, practices, and policies, and the current time of accountability and seemingly high-stakes nature of national examinations might have contributed significantly to this situation. Students' formative assessment scores in class should, in theory, partially tie with performance on summative scores. However, this is often not the case in most Ghanaian classrooms (Anane, 2011). Measurement experts believe that educational reform is on the rise in the areas of curriculum development (both objective- and standards-based), the use of standardised assessments, and high-stakes testing policies. There is a greater need to further the reforms to include the upgrading of teachers' professional skills in classroom scoring and grading practices in order to eliminate the unwieldy assessment and grading practices of basic school teachers which often lead to ambiguity in test scores, thus, making the actual grades students obtain meaningless (Marzano, 2000).

A number of researches have been conducted concerning factors teachers consider when scoring and grading students' assignments and answers to questions in school situations (Liu, 2008; McMillan, 2008; Guskey, 2006; Guskey & Bailey, 2001). According to Guskey (2006), teachers use unreliable procedures when scoring and grading students' responses to tasks. In Guskey's (2006) study of understanding why and how teachers grade, the results indicated that instructors draw mainly from their own experiences as students in deciding the scoring and grading practices they use. Guskey (2006) recommended therefore that educators must make every effort to guarantee that grading practices are as clear, explicit, and impartial as possible. Furthermore, he believes educators need to desist from using personal beliefs and insentient biases as influences while scoring and assigning grades, ensuring that above all, their grading policies and practices will be fair and impartial.

Similarly, Winger (2005), Dueck (2014) and Wormeli (2006) believe that teachers include non-academic factors such as effort, responsibility and attitude when grading students, and these contribute to variation in teachers' grading practices. The unreliable procedures used by teachers when grading might stem from their scanty knowledge on issues relating to grading (Brookhart & Nitko, 2008; Stiggins, 2008). Such knowledge is scanty because they rely on traditional grading procedures, often replicating some poor grading practices they experienced as students (Guskey & Bailey, 2001). Similarly, most teachers consider the award of marks as a way of motivating students to do their best (Munk & Bursuck, 2004). In addition, though grading is considered as an important professional responsibility, it is perceived by most teachers as one of the energy-sapping activities to undertake, very confusing and not likely an activity they would undertake during teaching (Green & Emerson, 2007). These perceptions make scoring and grading practices a bane for teachers and other educators.

Scores and grades as a measure of a student's performance or learning achievement has strong and lasting effects on the students' attitude, persistence in school, and motivation to learn (Brookhart, 2004). Therefore, teachers' grading practices are expected to match recommended assessment practices. In Ghanaian classrooms, although, the School Based Assessment (SBA) clearly stipulates guidelines on how assessment activities should be done, teachers at times lack the knowledge; are in haste or use undefined and unreliable procedures when assessing students' learning (Kubiszyn & Borich, 2013; Guskey, 2006; Wormeli, 2006). Such practices lead to much erraticism in grading practices which offer little consistency across schools and within classrooms, even when schools and teachers have adopted the same grading guidelines and schemes (McMillan, 2008).

Teachers' grading practices are being examined closely to change to more answerability measurements and performance-based assessments, and determining how to score and grade the assessments that truthfully reflect student attainment. Given the diversity of grading practices that studies and literature report, there is therefore the need to conduct research to more fully

understand why such differences exist among teachers' grading practices, and how the practices relate to measurement theory. Inquiries may lead to understanding a summary of grading practices that teachers use to generate grades, such as the measurement procedures they use, their rules of evidence or the standards they apply (Liu, 2008; Black, Harrison, Lee, Marshall, & William, 2004; Stiggins, Frisbie & Griswold, 1989).

However, not much has been done in terms of research on teachers' scoring and grading practices; especially within the setting of basic school teachers in Ghana, and from a developing country's perspective. Again, reviews from literature (Liu, 2008; McMillan & Lawson, 2001; McMillan, 2001) indicate that few studies concentrate on grading practices among basic school teachers. This study was focused on basic school teachers, because, substantial evidence shows that teachers differ in their scoring/grading practices (Kubiszyn & Borich, 2013; Guskey, 2006; Wormeli, 2006). Several studies have reported no significant difference in factors teachers considered when grading (e.g., McMillan & Lawson, 2001; Liu, 2008). Results from these researches also show that teachers do not differ in factors considered when scoring and grading students' responses to tasks.

McMillan and Lawson (2001) explored the grade level influence on the variables considered in grading practices in USA. The investigators established that while there was significant difference in the factors secondary science teachers considered when assigning grades to students' work, there was no substantial difference between teachers at different grade levels. Liu (2008) postulated that there was no significant difference between middle and high school teachers' views and practices of scoring and rating students' work. In fact, the results from their study showed that middle school and high school teachers were not significantly different on the variables they took into account in formal achievement measures (e.g., tests/quizzes). They considered attributes such as student effort, student ability, and student attendance or participation (Liu, 2008; McMillan & Lawson, 2001). Further findings by Liu showed no significant difference in perceived self-efficacy between middle school and high school teachers. Similarly, McMillan (2001) suggested, based on his research findings that teachers have implemented some of the expectations for assessment reform based on measurement experts, while maintaining some traditional practices in others.

A number of studies conducted in the area of teacher classroom assessments have focused on the general outlook of assessment of students' performance and the practice of continuous assessment in Ghanaian schools. Available evidence suggests that notwithstanding the fact that scoring and grading serves as an important responsibility, many teachers in the basic school system still challenging find the process of determining which academic and non-academic factors correctly represent student achievement. There is the need to present a deeper analysis of what teachers do when it comes to scoring of students' responses and the grading practices adopted to reflect students' behaviour and achievement. Again, there seems to be little ecological literature on teacher scoring and grading practices in a context like Ghana. The present study, therefore, sought to explore teachers' grading practices in Ghanaian basic school classrooms. In summary, this study was guided by the following three research questions:

1. What factors do teachers consider when assigning grades to students' responses to tasks?
2. What differences exist between primary and junior high school teachers with respect to students' effort considered when grading students' responses to tasks?
3. What differences exist between primary and junior high school teachers with respect to students' classroom behaviour considered when grading students' responses to tasks?

2. Methodology

2.1 Design and Participants

This study used a descriptive survey research design. The choice of this research approach was necessitated by the fact that the researchers wanted to quantitatively describe grading practices among basic school teachers in Ghana. The researchers examined the assessment practices of teachers without manipulating any variables such as assessment formats or the scoring processes or teachers' academic qualifications. The study was to establish how and what they consider when

they are scoring students' responses to tasks. The stratified sampling procedure was used to select 278 respondents, 139 each from primary and junior high schools, from whom data were collected for the study. The sample were drawn from 50 basic schools in Obuasi Metropolis of Ghana.

2.2 Measures

In order to gather valid data from the respondents (primary and junior high school teachers), a close-ended questionnaire designed in a 4-point Likert-type scale (strongly agree = 4, agree = 3, disagree = 2 and strongly disagree = 1) was used by the researchers. The questionnaire was in two parts. The first part collected demographic information of the respondents. The second part collected information about the factors respondents consider when scoring and grading students' responses to tasks. In view of the moderately large sample size, and the distribution of the schools used for the study, the questionnaires were administered to the participants by the researchers with two research assistants who were given a one-day training for the data collection. The data collectors visited the participants in their schools. Two weeks was used to collect the information. Data collected from the practising teachers were organised and analysed with descriptive statistics and independent samples t-test. Data were analysed using SPSS v21.0. The internal consistency of the instrument was high, with an alpha of 0.77. Of the 278 respondents from primary and junior high schools, 241 representing about 87% indicated they have had a formal training in test grading and score reporting. All information was kept private and only investigators in this study could have access to the data.

3. Results

3.1 Factors teachers considered when scoring and grading students' responses

The first question of the study was to find out the factors that teachers consider when scoring and grading students' responses to academic tasks. These factors include students' behaviour, work habits, students' relationships with teachers, and students' general conduct. The researchers computed descriptive statistics, and a mean of ≥ 2.45 was used as the criterion score of affirmation or agreement to the statements which we used to assess teachers' scoring and grading practices for the interpretation of the results. The sample's mean scores of the subscales of the scoring and grading practices surveyed are presented in Tables 1 and 2 in descending order.

Table 1: Means and Standard Deviations for Students' Effort Teachers consider when scoring and grading

Behaviour	Mean	SD
Scores and grades are based on:		
1. student's participation in class activities	2.71	0.82
2. items on the test attempted by student	2.67	0.80
3. how often a student asks questions in class	2.56	0.79
4. student's completion of homework	2.42	0.91
5. how often students submit assignment	2.36	0.77
Mean of means	2.54	0.82

n = 278

In general, the teachers in the study represented themselves as professionals who consider students' effort when scoring and grading students' responses to tasks. As shown in Table 1, teachers agreed with the statement that they considered a student's participation in class when assessing student's responses to academic tasks ($M = 2.71$, $SD = 0.82$). Similarly, most teachers agreed with the statement that they reflected the number of items attempted by a student on a task when assigning scores and grades ($M = 2.56$, $SD = 0.79$). With a mean of 2.56, the results show that teachers considered how often a student asks questions in class when assessing student learning (thus, scoring and grading the student's responses to tasks). However, it is worthy of note that teachers in the study appeared not to consider, to some extent, how often a student submitted an assignment.

Table 2: Means and Standard Deviations for Students' Behaviour teachers consider when scoring and grading

Practices	Mean	SD
I award marks for students' punctuality in class	2.67	0.81
I consider a student's interest in subject taught by making reference to previous performance	2.56	0.78
I award marks for neatness of work presented by students	2.45	0.85
I often use grade to punish truant students	1.91	0.89
I consider teacher-student relationship	1.83	0.82
I grade undisciplined students low	1.62	0.73
Overall mean	2.05	0.81

n = 278

As indicated in Table 2, most teachers in the study agreed with the statement that they award marks for students' punctuality in class ($M = 2.67$, $SD = 0.81$). In the same vein, teachers agreed with the statement that they considered students' interest in the subject they taught when scoring or assigning grades ($M = 2.56$, $SD = 0.78$). The results from the analysis of data also shows that most teachers in the study awarded marks for neatness of work presented by students when assessing their responses to tasks for grading purposes ($M = 2.45$, $SD = 0.85$). However, it is worthy to note that the teachers in the study indicated they did not often use grades as means of punishment in checking truant students ($M = 1.91$, $SD = 0.89$), and undisciplined students ($M = 1.62$, $SD = 0.73$).

3.2 Differences in factors teachers consider in scoring and grading across levels

The second research question sought to examine whether teachers differed on their scoring and grading practices with respect to the level they teach. An independent samples t-test was used to test the difference at 0.05 alpha level. A preliminary analysis to test the assumption of Homogeneity of Variance (HOV) was conducted using the Levene's test and the results are as presented in Table 3.

Table 3: Test for Homogeneity of Variance

	F	p	t	df	Sig(2-tail)
Equal variance assumed	2.009	.157	1.858	276	0.064
Equal variance not assumed				270.202	0.064

$\alpha = 0.05$

As indicated in Table 3, the p -value for Levene's test is 0.157 which is greater than the alpha (critical value) of 0.05. This implies that the assumption of homogeneity has been met for this sample, therefore, equal variance assumed. Table 4 then displays the output for the independent samples t-test for mean scores of teachers in respect to students' effort considered by teachers when grading.

Table 4: A Comparison of Teachers' Consideration of Students' Effort when Grading

Academic level	n	Mean	SD	t	df	p
Primary teachers	139	13.04	2.59	1.858	276	0.064
Junior High teachers	139	12.41	3.01			

$\alpha = 0.05$

As indicated in Table 4, that primary level teachers considered students' effort ($M = 13.04$, $SD = 2.59$) more than junior high level teachers ($M = 12.41$, $SD = 3.01$) when grading. The independent samples t-test statistics revealed that there was no statistically significant difference in teachers' consideration of effort by students during scoring and grading, $t(276) = 1.85$, $p > .05$. Therefore, primary and junior high school teachers were not significantly different in their attitude towards students' effort in class when assigning grades to responses to tasks.

3.3 Differences in factors teachers consider when grading students' responses to tasks with respect to students' classroom behaviour

To answer the third research question, both primary and junior high school teachers were asked to respond to behavioural items regarding factors they considered when scoring and grading students' responses to task. The factors included classroom behaviour of students such as attendance and participation in class activities. An independent samples t-test was used at 0.05 alpha level of significance. The Levene's test of Homogeneity of Variance (HOV) was conducted and the results are as shown in Table 5.

Table 5: Levene's Test for Homogeneity of Variance

	<i>F</i>	<i>Sig.</i>	<i>t</i>	<i>df</i>	<i>Sig(2-tail)</i>
Equal variance assumed	1.671	.197	1.614	276	0.108
Equal variance not assumed				273.559	0.108

The Levene's test revealed that the variation in the students' classroom behaviour teachers considered were homogeneous across school level, $F(2, 276) = 1.67, p > 0.05$. Table 6, therefore, presents the output of results for the independent samples t-test for students' classroom behaviour teachers considered when assigning grades to their responses.

Table 6: A Comparison of Teachers' Consideration of Students' Behaviour when Assigning Grades

Academic level	<i>N</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>
Primary teachers	139	12.65	3.18	1.614	276	0.108
Junior High teachers	139	12.01	3.49			

An independent samples t-test was used to examine the differences in mean scores of teachers at the primary and junior high school levels. No statistically significant differences in mean scores was found between primary and junior school teachers concerning students' classroom behaviour (e.g., punctuality, interest in subject taught or previous performance). That is, teachers did not differ on students' behaviour they considered when scoring and grading their responses to task ($t(276) = 1.614; p > .05$). This suggests that the classroom behavioural factors considered by primary school teachers ($M = 12.65, SD = 3.18$) were not significantly different from that of junior high school teachers ($M = 12.01, SD = 3.49$).

4. Discussion

In this study, we looked at the scoring and grading practices classroom teachers adopt in Ghanaian basic schools. Descriptive statistics was used to examine the effort and classroom behaviour variables teachers considered when scoring and assigning grades to students' responses to a task. Independent samples t-tests were conducted to examine whether teachers in primary school and junior high school differed in the students' effort and classroom behaviour they considered when scoring and assigning grades. Findings from the present study indicate that teachers considered students' participation in class when assigning grades to students' responses to academic tasks. This finding supports Dueck's (2014) conclusion that students are likely to benefit in their performance when they participate in class activities. The results also showed that teachers awarded marks for the number of items attempted by a student on a task, and this is consistent with Munk and Bursuck's (2004) assertion that teachers award marks to motivate students to do better. Teachers considering such factors might stem from the belief that students who frequently ask questions give feedback that illuminate how well a student understands learning objectives. What is more, it might be that continual questioning from students helped teachers in identifying misconceptions or misunderstanding of steps in a process on instructional material taught. This could possibly help teachers to determine the extent to which instructional objectives have been achieved.

The findings also revealed that teachers awarded marks for extraneous factors, such as students' punctuality to class. This finding is consistent with Dueck's (2014) finding that teachers often use grades to compel students to put up good behaviour. It is possible that in this study, teachers considered class attendance to be related to high academic achievement; hence, awarded marks for punctuality. This notwithstanding, the converse could be true that most teachers have scanty knowledge on issues relating to grading practices (Brookhart & Nitko, 2008; Stiggins, 2008), hence, replicating such unprofessional grading practices they might have experienced as students (Guskey & Bailey, 2001).

Furthermore, results obtained from the independent samples t-tests indicated that there are no statistically significant differences between primary and junior high school teachers' scoring and grading practices with respect to students' efforts and classroom behaviour. These findings corroborate the findings of earlier studies by McMillan and Lawson (2001) and Liu (2008). McMillan and Lawson (2001) discovered in their study of K-12 school settings in the USA that teachers at different grade levels do not significantly differ with respect to students' effort considered when grading. Similarly, the finding from the study is consistent with the previous study conducted by Liu (2008), who found no significant differences between middle school and high school teachers with regard to the consideration of students' effort when assigning final grades.

Concerning students' classroom behaviour teachers considered in their grading practices, the results support Liu's (2008) finding that middle school and high school teachers do not differ on students' classroom behaviour they considered when assigning final grades. Similarly, the finding from this study supports McMillan and Lawson's (2001) conclusion that teachers at different grade levels do not differ with respect to non-achievement factors considered when grading. These findings could be due to the fact that assessment practices at the basic schools are guided by the School-Based Assessment Policy. However, these common grading practices are interesting as primary and junior high school teachers may need to engage in somewhat different forms of grading practices in consonance with developmentally appropriate assessment principles. There is, therefore, the need for further investigation.

5. Conclusions and Implications for Practice

In order to gain a complete understanding of how teachers in basic schools score and grade their students, it is necessary to explore their practices in the classroom. This includes, but not limited to the behaviour and other factors teachers take into consideration when grading students' responses to tasks in the basic schools in Ghana. Based on the findings of the study, the researchers conclude that teachers in the present study agreed to have engaged some unprofessional grading practices. Teachers were influenced by extraneous factors such as a student's ability to 'pool', neatness of work presented and punctuality to class when scoring and grading students' responses to academic tasks.

Considering the fact that the effectiveness of grading practices is an important research area in psychometrics, this study contributes circumstantial data to the field of assessment practices in schools by exploring basic school teachers' grading practices in Ghanaian schools. Grading is one of the primary responsibilities of classroom teachers, and when extraneous factors continue to influence teachers' grading practices, grades assigned would not adequately reflect students' abilities or academic achievements. In addition, when a grade is a representation of a hodgepodge of factors such as student's effort and behaviour, it would make interpretation or comparison of such grade a difficult task to undertake. The results of this study would help to elucidate some misconceptions of the factors that influence teachers' decision in scoring and grading students' responses to tasks.

6. Recommendation

For teachers' inclusion of non-academic factors such as students' participation in class and the number of items attempted, we recommend that teacher training should provide measures on how to assess these non-academic and academic factors in scoring and grading. Again, regular in-service training should be given to teachers on how the reliability and validity of grades are

influenced by the inclusion of such non-academic factors. Finally, we recommend that, for student teachers, it would be supportive to change or review their core curriculum to include skills and dispositions that would enhance their ability to fuse assessment(s) into teaching to improve students' learning. Teachers should be trained to build their capacity on assessment and grading practices through teacher preparation programmes, and be given opportunities to practice and refine their skills through practicum sessions.

References

- Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. (2013). Observations of effective teacher–student interactions in secondary school classrooms: Predicting student achievement with the classroom assessment scoring system - secondary. *School Psychology Review*, 42(1), 76.
- Anane, E. (2011). Predictive validity of continuous assessment scores of students' performance in end-of-semester examinations in colleges of education in Ghana. *Journal of Business and Educational Policies*, 7(1), 95-104.
- Black, P., Harrison, C., Lee, C., Marshall, B., & William, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86(1), 8-21.
- Bonesrønning, H. (2004). Do the teachers' grading practices affect student achievement? *Education Economics*, 12(2), 151-167.
- Brookhart, S. M. (2004). Classroom assessment: Tension and intersections in theory and practice. *Teachers College Record* 106(3), 429-458.
- Dueck, M. (2014). *Grading smarter not harder: Assessment strategies that motivate kids and help them learn*. USA: Alexandria, VA.
- Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement* (3rd ed.). Englewood: Prentice Hall Inc.
- Emmanuel, E. E. (1990). *Evaluation of continuous assessment practice in Cross River State*. Unpublished M.Ed. dissertation, University of Ibadan.
- Green, K., & Emerson, W. A. (2007). *A new framework for grading*. Retrieved on September 20, 2016 from Fisher Digital Publications: http://fisherpub.sjfc.edu/math_facpub/3/
- Guskey, T. R. (2006, April). It wasn't fair! Educators' recollections of their experiences as students with grading. In *annual meeting of the American Educational Research Association, San Francisco*.
- Guskey, T. R., & Bailey, J. M. (2001). *Developing grading and reporting systems for student learning*. Thousand Oaks, CA: Corwin.
- Kubiszyn, T., & Borich, G. (2013). *Educational testing and measurement: Classroom application and practice* (10th ed.). Hoboken, NJ: John Wiley & Sons, Inc.
- McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice*, 20(1), 20-32.
- 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*. Metropolitan Education Research Consortium, Richmond, VA.
- Munk, D. D., & Bursuck, W. D. (2004). Personalized grading plans: A systematic approach to making the grades of included students more accurate and meaningful. *Focus on Exceptional Children* 36(9), 1-11.
- Stiggins, R. J. (2008). *Report cards: Assessments for learning*. (5th ed.). Upper Saddle River, NJ: Merrill/Prentice.
- WAEC. (1990). *A research report on the relationship between school estimates and ordinary level examination results: Science and mathematics*. WAEC publication. Accra department of WAEC research division.
- WAEC. (1993). *A research report on Ghana basic education certificate examination (BECE): Relationship between Internal and external assessments*. Accra: WAEC publications. Accra department of WAEC research division.
- Winger, G. T. (2005). Grading to communicate. *Educational Leadership*, 11, 61-65.
- Wormeli, R. (2006). Accountability: Teaching through assessment and feedback, not grading. *American Secondary Education*, 34(3), 14-27.
- Liu, X., (2008, October). *Measuring teachers' perceptions of grading practices: Does school level make a difference?* Paper presented at the 2008 Annual Conference of the Northeastern Educational Research Association (NERA), Rocky Hill, CT.