

10-23-2008

Measuring Teachers' Perceptions of Grading Practices: Does School Level Make a Difference?

Xing Liu

Eastern Connecticut State University, liux@easternct.edu

Follow this and additional works at: https://opencommons.uconn.edu/nera_2008

 Part of the [Educational Assessment, Evaluation, and Research Commons](#)

Recommended Citation

Liu, Xing, "Measuring Teachers' Perceptions of Grading Practices: Does School Level Make a Difference?" (2008). *NERA Conference Proceedings* 2008. 4.

https://opencommons.uconn.edu/nera_2008/4

Running head: Measuring Teachers' Perceptions

Measuring Teachers' Perceptions of Grading Practices: Does School Level Make a Difference?

Xing Liu

Eastern Connecticut State University^{1,2}

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.

¹ This study was supported by CSU-AAUP Research Grant. The opinions reflect those of the author and do not necessarily reflect those of the granting agency.

² The author is grateful for the data collection assistance of Dr. Scott W. Brown and the Teachers for a New Era Project at the University of Connecticut.

Measuring Teachers' Perceptions of Grading Practices: Does School Level Make a Difference?

The effectiveness of classroom assessment and grading practices has become an increasingly important research topic in education (Bonesronning, 1998, 2004; Brookhart, 1993, 1994; McMillan & Lawson, 2001; McMillan & Nash, 2000; McMillan, Myran, & Workman, 2002). For policy makers, grading reflects school and student accountability, which influences instruction and curriculum. For instructors, grading assesses teaching effectiveness, and helps educators to make informed decisions regarding students' progress and their own teaching. Grades can also help students to understand their strengths and weaknesses (Kubiszyn & Borich, 1990; Thorndike, 1997).

Many studies have been conducted regarding factors affecting teachers' grading practices in K-12 school settings (Brookhart, 1993, 1994; McMillan & Lawson, 2001; McMillan, Myran, & Workman, 2002; McMunn, Schenck, & McColskey, 2003; Stiggins, Frisbie, & Griswold, 1989). These studies found that teachers utilized many factors in assigning their grades to their students, and students' academic achievement was a major factor influencing teachers' grading practices. However, results were contradictory in terms of whether it was necessary for teachers to consider factors other than student academic achievement, such as ability, tardiness, or behavior, in their grading practices. Some studies (Brookhart, 1993, 1994; Stiggins, Frisbie, & Griswold, 1989) contended that teachers tried to be fair to students and were concerned with the motivation, self-esteem and social consequences of students; thus they found that teachers also included other factors such as student effort and ability in grading practices. On the other hand, a recent empirical study (McMunn, Schenck, & McColskey, 2003) revealed that it became problematic when teachers included non-achievement factors in their grades that might not

reflect student academic achievement, or when teachers did not follow established procedures and policies of grading.

Previous research also had no consensus on whether school level had an effect on teachers' grading practices. O'Donnell and Woolfolk's study (1991) found a school type effect on testing and grading. Two hundred twenty six elementary school teachers and 117 secondary school teachers participated in the study. They were asked to provide information about their views on the nature and dimensions of intelligence, educational goals, and beliefs about testing and grading by completing questionnaires. The study found elementary teachers were distinguished from secondary school teachers in using tests and subjective assessment.

Cizek, Fitzgerald, and Rachor (1996) surveyed 143 mid-western elementary and secondary school teachers regarding their assessment practices, with a focus on grading practices. The results of a logistic regression analysis suggested that the teacher characteristics (gender, grade level, and teaching experience) had no significant effects on whether the teacher reported having knowledge of a district policy on grading.

McMillan and Lawson (2001) investigated the grade level effect on the factors considered in grading practices using Univariate Analysis of Variance (ANOVA). A sample size of 213 teachers from 58 secondary schools (Grade 6-12) completed and returned surveys. The researchers concluded that although there was considerable variation in the factors secondary science teachers included in assigning grades, there was no significant difference between teachers at different grade levels.

Given these conflicting results relative to teachers' grading practices and whether elementary and secondary school teachers differ on grading practices, this study seeks to clarify

and understand how teachers perceive various factors associated with grading practices, and to investigate whether there is a school level effect on teachers' perceptions of grading practices.

Purpose of the Study

The purpose of this study was to examine whether middle and high school teachers differ in regard to ratings of the importance and usefulness of grading practices, teachers' perceived self-efficacy of the grading process, and the degree to which factors such as student effort and ability and teachers' personal grading habits affect their grading decisions. The research question mainly focused on: Were there any significantly different perceptions of grading between middle school and high school teachers?

Method

Instrumentation and Data Collection

An instrument, the Teachers' Perceptions of Grading Practices (TPGP), has been developed and validated to assess teachers' perceptions using both exploratory factor analysis and confirmatory factor analysis (Liu, 2004; Liu, O'Connell, & McCoach 2006, 2008). This instrument measuring teachers' perceptions of grading practices has six sections. To complete the survey, participants were asked to circle or click on their answer to each item with responses ranging from strongly disagree to strongly agree based on 5-point Likert rating scale (1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, and 5 = Strongly Agree). This instrument also includes four behavioral questions regarding factors teachers consider in assigning final grades. Respondents were asked to indicate whether they included formal achievement measures (e.g., tests/quizzes), student effort, student ability, classroom behavior

(e.g., laudatory or disruptive behavior), and attendance or participation in grading (1= yes, 0 = no). The data were collected from a state in the northeast. Self-report web-based surveys were used to gather the data. Participants were asked to respond to the survey items by following the directions online. Responses were anonymous; respondents were not required to provide names that could be linked to their responses. To increase the response rate of the online surveys, two iPods were used as incentives for a raffle, since previous research found that using an incentive could increase response rates to the online survey (Dommeyer, Baum, & Hanna, 2004). The participants who received emails were asked to enter their email addresses at the end of the online survey if they were willing to join in the raffle. Those participants who received requests through the regular mail were asked to enter codes (numbers) which were assigned and mailed to each of them in the letter, or enter their email addresses. These codes (numbers) were used for raffle purposes only, since the raffle needed to link the name with the corresponding code.

Ensuring confidentiality has been found to increase response rates in survey research (Asch, Jedrziwski, & Christakis, 1997). To ensure confidentiality, two separate files were programmed into the on-line survey, one to collect the email addresses or codes so we knew who completed the survey to enter the raffle, and one to collect the actual data, which was not linked back to the email address or codes. In the cleaned final data set, no information on the teachers' name, email addresses and codes are identified. All survey data were entered into a secure, restricted database. All information was kept confidential and only researchers could have access to the data, which were kept in a locked cabinet in the researcher's office. When reporting the results, no individual names or school district could be identified. The results are reported only on the group level. The purpose of using the email addresses or codes and these procedures to ensure confidentiality were fully explained to the participants in the emails and letters.

A total of 107 teachers including 52 middle school teacher and 57 high school teachers participated in this study.

Data Analysis

A Multivariate Analysis of Variance (MANOVA) procedure was conducted to test the between-group (school level) differences in teachers' perceptions of grading practices since this procedure allows the dependent variables to be correlated and is more powerful to detect group differences than the ANOVA procedure (Tabachnick & Fidell, 2007; Weinfurt, 1995). The One-Way MANOVA was conducted on six dependent variables which corresponded to six factors of the TPGP instrument: importance, usefulness, student effort (effort), student ability (ability), grading habits (habit), and perceived self-efficacy of the grading process (efficacy). Composite scores of these six dependent variables were created by summing up the items' scores and dividing the sum by the total number of items contributed to that factor. The independent variable in the MANOVA was school level.

Chi-square analyses were conducted to examine whether middle and high school teachers differed on the factors of formal achievement measures (e.g., tests/quizzes), student effort or hard work, student ability, and classroom behavior (e.g., laudatory or disruptive behavior) they considered when assigning final grades. Data was analyzed using SPSS 15.0.

Results

The results of the One-way MANOVA analysis were presented and interpreted in the paper. Pillai's statistics were reported together with Wilk's lambda, since the former statistic was more robust to the violation of the Homogeneity of Covariance (HOC) assumption,

compared to the latter (Tabachnick & Fidell, 2007). Table 1 presents means and standard deviations of these six DVs across school level.

Table 1

Means and Standard Deviations for Factor Variables by School Level

	Middle School (N = 52)		High School (N = 55)	
	Mean	S.D.	Mean	S.D.
<i>Importance</i>	3.728	.586	3.933	.519
<i>Usefulness</i>	3.739	.649	3.760	.473
<i>Student effort</i>	3.543	.618	3.524	.547
<i>Student ability</i>	3.735	.554	3.795	.359
<i>Teachers' grading habits</i>	3.510	.776	3.109	1.026
<i>Perceived control over grading process</i>	2.128	.686	2.194	.733

Three assumptions including normality, linearity and HOC were examined first. The descriptive statistics indicated that the normality was not violated for all six dependent variables across school level. A multiple regression analysis with case number as the outcome variable was used to check whether there were multivariate outliers within each of the two groups. A chi-square test was used to examine the values of Mahalanobis distance at the .001 level. No multivariate outliers were identified. The Box-M test for Homogeneity of Covariance (HOC) was statistically significant at the 0.001 level, Box-M = 51.506, $F(21, 40275) = 2.303$, $p = .001$. Therefore, the assumption of HOC was violated. This violation might be due to the heterogeneity of variance on some dependent variables across groups. Further examinations of univariate

homogeneity of variance (HOV) tests found that variances of ability and habit were not homogeneous across school level.

The multivariate statistics revealed that there was no significant main effect for school level: Pillais statistic = .093, Wilk's Λ = .907, $F(6, 100) = 1.700$, $p = .129$, partial eta-square = .093, which indicated that 9.3% of the generalized variance in teachers' perceptions of grading practices was accounted for by school level, a medium effect size by Cohen's standards. Therefore, middle and high school teachers were not significantly different in their perceptions of the importance and usefulness of grading practices, student effort, student ability, grading habits, and perceived self-efficacy of the grading process.

Frequency differences in factors teachers consider in grading across school level

Both middle and high school teachers were asked to respond to five behavioral items regarding factors they considered when they assigned final grades for a marking period or a semester. These factors included formal achievement measures (e.g., tests/quizzes), student effort, student ability, classroom behavior (e.g., laudatory or disruptive behavior), and attendance or participation. Chi-square analyses were conducted to examine whether teachers in middle school and high school differed on the factors they considered when assigning final grades. Table 2 through Table 6 present the cross-tabulated data and chi-square results for each analysis. The results indicated that middle school and high school teachers were not significantly different on the factors of formal achievement measures (e.g., tests/quizzes), student effort, student ability, and student attendance or participation they considered when assigning final grades. However, middle school and high school teachers were significantly different on whether they considered classroom behavior (e.g., laudatory or disruptive behavior) as a factor in grading. The result of

the analysis revealed that high school teachers were more likely to consider the factor of student attendance or participation when they assigned grades.

Table 2

Chi-square Analysis of Formal Achievement Measures by School Level

Variable	n	Formal Achievement Measures		χ^2	p
		yes	no		
School level				.403	.525
Middle school	52	50 96.2%	2 3.8%		
High school	55	54 98.2%	1 1.8%		
Total	107	104	3		

Table 3

Chi-square Analysis of Student Effort by School Level

Variable	n	Student Effort		χ^2	p
		yes	no		
School level				.007	.61
Middle school	52	48 92.3%	4 7.7%		
High school	55	51 92.7%	4 7.3%		

Total	107	99	8
-------	-----	----	---

Table 4

Chi-square Analysis of Student Ability by School Level

Variable	Student Ability			χ^2	p
	n	yes	no		
School level				.036	.850
Middle school	52	34	18		
	100%	65.4%	34.6%		
High school	55	35	20		
	100%	63.6%	36.4%		
Total	107	69	38		

Table 5

Chi-square Analysis of Classroom Behavior by School Level

Variable	Classroom Behavior			χ^2	p
	n	yes	no		
School level				4.867	.027
Middle school	52	23	29		
	100%	44.2%	55.8%		
High school	55	36	19		
	100%	65.5%	34.5%		
Total	107	59	48		

Table 6

Chi-square Analysis of Attendance/participation by School Level

Variable	Attendance/participation			χ^2	p
	n	yes	no		
School level				2.302	.129
Middle school	52	36	16		
	100%	69.2%	30.8%		
High school	55	45	10		
	100%	81.8%	18.2%		
Total	107	81	26		

Discussion

In this study, the One-way MANOVA was used to investigate the school level difference in teachers' perceptions of grading practices, and Chi-square tests were conducted to examine whether teachers in middle school and high school differed on the factors they considered when assigning final grades. Findings from the present study indicated that there were no significant differences between middle and high school teachers' perceptions of grading practices. The results also indicated that middle school and high school teachers did not differ significantly on the factors of formal achievement measures (e.g., tests/quizzes), student effort, student ability, and student attendance or participation they considered when assigning final grades. This finding support prior research indicating there was no grade level effect on factors secondary teachers

considered in grading practices (McMillan & Lawson, 2001). This is good news for teacher training in grading practices and curriculum development in teacher preparation program: there is no need to train middle and high school teachers differently in grading, and pre-service teachers in the secondary education program can take the same course related to grading, no matter which school level (middle or high school) that they plan to teach. However, middle and high school teachers disagreed on whether they considered classroom behavior (e.g., laudatory or disruptive behavior) as a factor in grading, and more high school teachers than middle school teachers considered student classroom behavior as a factor in grading. This is an interesting finding and needs further investigation.

Educational Implications

This study would help educators, administrators, and researchers evaluate and understand teachers' perceptions and practices of grading in secondary schools. Considering the effectiveness of grading has been an important research topic in education, this study would contribute to the field by examining the current status of teachers' perceptions of grading practices. Secondly, this study would provide empirical evidence for decision making regarding whether it is appropriate for teachers to utilize many other factors such as student effort and student ability in assigning grades in addition to student achievement measures. The result of this study would help to clarify some misunderstandings of factors influencing teachers' decision making in grading. Since both middle and high school teachers considered non-achievement factors in their grading, this study supported that teacher training should provide procedures on how to measure these non-achievement factors in grading (Stiggins, Frisbie, & Griswold, 1989). Finally, this study might provide direction for school administrators and policy makers to take

actions to enhance teachers' self-efficacy in grading since both middle and high school teachers had low self-efficacy in grading. For instance, for pre-service teachers, it might be helpful to adjust the curriculum and train them how to build their capability of grading and conducting classroom assessment through teacher preparation program; for beginning teachers, systematic support systems should be built to help them improve grading skills; and for other teachers, they should have opportunities to participate in periodic training programs related to grading practices and classroom assessment.

References

- Asch, D., Jedrziwski, M., & Christakis, N. (1997). Response rates to mail survey published in medical journal. *Journal of Clinical Epidemiology*, 50 (10), 1129-1136.
- Bonesronning, H. (1998). The variation in teachers' grading practices: causes and consequences. *Economics of Education Review*, 18, 89-105.
- Bonesronning, H. (2004). Do the teachers' grading practices affect student achievement? *Education Economics*, 12, 151-167.
- Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement*, 30, 123-142.
- Brookhart, S. M. (1994). Teacher's grading: Practice and theory. *Applied Measurement in Education*, 7, 279-301.
- Cizek, G. J., Fitzgerald, S. M., & Rachor, R. E. (1996). Teachers' assessment practices: Preparation, isolation, and the kitten sink. *Educational Assessment*, 3, 159-179.
- Dommeier, C. J., Baum, P., & Hanna, R. W. (2004). Gathering faculty teaching evaluations by in-class and online surveys: Their effects on response rates and evaluations. *Assessment & Evaluation in High Education*. 29 (5) 611-623.
- Kubiszyn, T., & Borich, G. (1990). *Educational testing and measurement: Classroom application and practice* (3rd ed.). Glenview, IL: Scott, Foresman and Company.
- Liu, X. (2004, October). *The initial validation of teacher's perception of grading practices*. Paper presented at the 2004 Northeastern Educational Research Association annual meeting, Kerhonkson, NY.

Liu, X., O'Connell, A.A., & McCoach, D.B. (2006, April). *The initial validation of teachers' perceptions of grading practices*. Paper presented at the 2006 Annual Meeting of American Educational Research Association (AERA), San Francisco, CA.

Liu, X., O'Connell, A.A., & McCoach, D.B. (2008, April). *Teachers' perceptions of grading practices: A cross-cultural study in the U.S. and China*. Paper accepted for presentation at the 2007 Annual Meeting of American Educational Research Association (AERA), New York, NY.

McMillan, J. H., & Lawson, S. R. (2001). *Secondary science teachers' classroom assessment and grading practices*. Metropolitan Education Research Consortium, Richmond, VA. (ERIC Document Reproduction Service N. Ed 450 158)

McMillan, J. H., Myran S., & Workman D. (2002). Elementary teachers' classroom assessment and grading practices. *The Journal of Educational Research*, 95, 203-213.

McMillan, J. H., & Nash, S. (2000, April). *Teacher classroom assessment and grading practice decision making*. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA.

McMunn, N., Schenck, P., McColskey, W. (2003 April). *Standards-based assessment, grading, and reporting in classroom: Can district training and support change teacher practice?* Paper presented at the Annual Meeting of American Educational Research Association, Chicago, IL.

O'Donnell, A., & Woolfolk, A. E. (1991). *Elementary and secondary teachers' beliefs about testing and grading*. Paper presented at the Annual Meeting of American Psychological Association, San Francisco, CA.

Stiggins, R. J., Frisbie, D. A., & Griswold, P. A. (1989). Inside high school: Building a research agenda. *Educational Measurement: Issues and Practice*, 8, 5-14.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Allyn & Bacon.

Thorndike, R. M. (1997). *Measurement and evaluation in psychology and education* (6th ed.). Upper Saddle River, NJ: Prentice-Hall.

Weinfurt, K. P. (1995). Multivariate analysis of variance. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp.245-276). Washington DC: American Psychological Association.