Teachers' assessment practices and students' perceptions of the classroom assessment environment

Hussain Alkharusi [a] *

[a] Department of Psychology, Sultan Qaboos University, Oman

Abstract

The purpose of this study was to examine the relationship between teachers' assessment practices and students' perceptions of the classroom assessment environment. A total of 1,636 students and 83 science teachers in the ninth grade participated in the study. Results from hierarchical linear modeling techniques showed that students' perceptions of the assessment environment were shaped by student characteristics such as self-efficacy, class contextual features such as aggregate perceived assessment environment and self-efficacy levels of the class, and teacher's teaching experience and assessment practices. These results point to a conclusion that each class has an assessment environment that originates from the teacher's assessment practices, and that students develop their perceptions of the classroom assessment environment based in part on their group experiences. Therefore, researchers studying classroom environment may need to consider not only the individual student perception of the assessment environment, but also the aggregate perceptions of students in a class about their classroom assessment practices.

Keywords: Assessment environment; assessment practices; students' perceptions; classroom environment; classroom assessment.

---

* Hussain Alkharusi. Tel.: +968-96222535; fax: +968-24413817
  E-mail address: hussein5@squ.edu.om

1. INTRODUCTION

Classroom assessment involves a wide range of activities from designing paper-pencil tests and performance measures to grading, communicating assessment results, and using them in decision-making (Zhang & Burry-Stock, 2003). Although there is a great deal of research on teachers' assessment practices, few empirical research attempts have been made to link these practices to students' perceptions of classroom assessment environment. Theoretically, perceived classroom assessment environment refers to the overall meaning that students make out of the various teachers' assessment practices in the classroom (Brookhart & DeVoge, 1999). Operationally, it has been represented by two dimensions: learning- and harsh-oriented assessment environments (Alkharusi, 2008). The learning-oriented environment refers to the extent to which students perceive assessment tasks as moderately challenging, assessment standards and criteria are clear, assessment feedback is informative, and that they have chances to improve their performance. The harsh-oriented environment refers to the extent to which students perceive that the assessment tasks and grading are difficult, greater importance is given to the grades than learning, and that the evaluation and recognition practices are public highlighting social comparison. Given that a substantial proportion of classroom time is devoted to the assessment of student learning and that students' perceptions of the meaning of the classroom activities play a critical role in the learning process, it seems reasonable to argue that the impact of teachers' assessment practices on students' perceptions of the classroom assessment environment deserve recognition and investigation.

1.1. Purpose of The Study

Having identified the dimensions of student perceived assessment environment, this study sought to find out what student and class characteristics are related to students' perceptions of the classroom assessment environment? The objects of interest and measurement in this study were students and teachers and as such hierarchical linear modeling (HLM) (Raudenbush & Bryk, 2002) is the appropriate analytic technique to handle the nested nature of the data. This kind of investigation may contribute to the understanding of both theory and practice of classroom assessment, and could provide a road sign for improving the learning and assessment climate in the classroom.

1.2. Academic Self-Efficacy and Gender

Academic self-efficacy and gender have been detected as potential correlates of students' perceptions of the classroom assessment environment. Academic self-efficacy refers to students' judgments of their capabilities to successfully perform specific tasks (Bandura, 1997). Research has shown that students with a high sense of self-efficacy tended to hold more positive perceptions of their classroom environment than students with a low sense of self-efficacy (Greene, Miller, Crowson, Duke, & Akey, 2004). Also, females have been found to report more positive perceptions of their classroom environment (Meece, Herman, & McCombs, 2003) and tended to hold stronger self-efficacy beliefs than males (Britner & Pajares, 2006). It should be noted that gender in this study

varied across classes because in Oman, students within the same class and their teacher are of the same gender, either all of them are males or all of them are females. Therefore, unlike past research studies, gender in the present study would more appropriately to be treated as an independent variable at the class-level. Thus, the hierarchical modeling approach to account for the gender effects would be applied in this study by assuming that student self-efficacy is a confounding variable in the within-class model and that the composition of students in each class (i.e., class average for self-efficacy) and its interaction with class gender are confounding variables in the between-class model.

1.3. Aggregate Perceived Classroom Assessment Environment

Students within a classroom share common characteristics of the teacher and his or her assessment practices, and as such even though students respond differently to the same classroom assessment process, their responses may have commonality. Yet, research on classroom environment research has used individual student scores as the unit of analysis rather than the average score of students at the classroom level (e.g., Church, Elliot, & Gable, 2001; Kaplan, Middleton, Urdan, Midgley, 2002). Proponents of this approach argue that students within the same classroom differ in how they interpret and perceive the various practices in the classroom as a result of differential treatment and their different prior experiences brought to the classroom (Ames, 1992; Kaplan et al., 2002). Since students in the classroom "are not social isolates of the influence of those around them" (Bandura, 1997, p. 469), it seems reasonable to argue that the aggregate perceived classroom assessment environment might act as a cogent attribute to characterize the social influence of the classroom. Therefore, the present study would attempt to address the question of how is the aggregate perceived assessment environment as an emergent attribute of the classroom related to differences in individual perceived assessment environment as a student-level attribute.

1.4. Teachers' Assessment Practices

Teachers' uses of different forms of assessment methods (traditional vs. alternative assessments) represent one facet of the classroom assessment environment (Brookhart, 1997). Alternative assessments (e.g., portfolio and performance assessments) have been thought to be a more authentic way to assess student learning because they are based on what might students be called to do the real world (Darling-Hammond, 1994). In comparison, traditional assessments (e.g., multiple-choice test items) have been criticized for being focusing on the product of learning rather than on the process of learning (Henning-Stout, 1994). Findings from survey studies of teachers' assessment practices have suggested that teacher's frequent use of assessment methods might depend on teacher's teaching experience (Bol, Stephenson, O'Connell, & Nunnery, 1998; Mertler, 1998). However, little empirical research (e.g., Maslovaty & Kuzi, 2002) exists about the effects of teacher's use of a particular form of assessment on the classroom assessment environment as perceived by students. Furthermore, although teachers are expected to conduct classroom assessment practices that are in agreement with those recommended by experts of educational

assessment (American Federation of Teachers, National Council on Measurement in Education, & National Educational Association, 1990), considerable amount of research (e.g., McMillan, Myran, & Workman, 2002; Mertler, 1999) have shown that teachers' assessment practices are often not consistent with the recommended practices. However, empirical studies in the classroom assessment literature investigating how the use of recommended classroom assessment practices affects students' perceptions of the classroom assessment environment are extremely limited. The present study would attempt to address this question in relation to teachers' teaching experience.

2. METHODS

2.1. Participants and Procedures

A total of 1,636 students and their corresponding 83 science teachers in the ninth grade participated in the study. The number of participating students in each class ranged from 14 to 21 with an average of 20 students. Of all participating students, 735 were males and 901 were females. Of all participating teachers, 37 were males and 46 were females. The teaching experience of the teachers ranged from 1 to 13.5 years with an average of 5.20 and a standard deviation of 2.64. Permission for the study was granted by the Ministry of Education in Oman. The data collection process took place during a regular scheduled class meeting. The participants were informed that they were not obligated to participate in the study, and if they wished to participate in the study, their responses would remain anonymous and confidential.

2.2. Instrumentation

Two questionnaires were used, one for students and one for teachers. The students' questionnaire included items asking students to indicate their perceptions of the assessment environment and self-efficacy using a 4-point Likert scale ranging from 1 (*completely not true*) to 4 (*completely true*). The teachers' questionnaire included items asking teachers to rate the frequency of use of traditional assessments, alternative assessments, and various assessment practices recommended by experts of educational assessment on a 5-point Likert scale ranging from 1 (*never*) to 5 (*always*).

2.2.1. Classroom Assessment Environment
This section included items from the two dimensions of the perceived classroom assessment environment scale developed by Alkharusi (2008). The first dimension consisted of five items (α = .63) measuring the perceived *learning* assessment environment (e.g., *in this class, students are given a chance to correct their mistakes*). The second dimension consisted of six items (α = .60) measuring the perceived *harsh* assessment environment (e.g., *the science tests in this class are difficult to students*). In this study, the perceived assessment environment construct was measured at two levels: student and class. At the student-level, the individual (student) perceived assessment environment score on each dimension was constructed for each student as the sum of the student's

responses to all items defining that dimension of the classroom assessment environment scale. At the class-level, the aggregate perceived assessment environment score on each dimension was constructed for each class as the average levels of individual students' perceptions within the class on that dimension of the perceived classroom assessment environment scale. These last values were added to the class-level data set with a corresponding class's identification number as the selection variable.

2.2.2. Academic Self-Efficacy

The measure of student self-efficacy was adapted from Greene, Miller, Crowson, Duke, and Akey (2004) and Midgley et al. (2000). This measure contained six positively worded items measuring students' perceptions of their competence to do their science class work in the current semester (e.g., *I'm certain I can master the skills taught in science class this semester*). Internal consistency reliability was found to be .74 as indicated by Cronbach's alpha. In this study, the self-efficacy construct was measured at two levels: student and class. At the student-level, the student self-efficacy was reflected by a total rating score across all the six items of the scale. At the class-level, the class self-efficacy was reflected by the average levels of individual students' self-efficacy within the class. These last values were added to the class-level data set with a corresponding class's identification number as the selection variable. Higher scores represented higher levels of self-efficacy.

2.2.3. Type of Assessment

This section contained 13 items drawn from a questionnaire developed by Alsarimi (2000) to measure frequent use of various types of traditional and alternative assessments by the ninth grade science teachers in Oman. Traditional assessments included seven elements: oral exams, true-false, multiple-choice, matching, completion, short-answer, and extended short-answer test items. Alternative assessments included six elements: essay items, research papers, portfolios, models, and structured and unstructured performance assessments. Two scale scores, one for traditional assessment and one for alternative assessment, were derived to indicate teachers' frequent use of a particular type of assessment. Scale scores were calculated as the total rating score obtained across the items comprising each scale. Higher scores represented a more frequent use of that type of assessment by the teacher. Internal consistency reliability was established for each scale's scores through Cronbach's alpha. The reliability coefficient for scores representing traditional assessments was .69, and the reliability coefficient for scores representing alternative assessments was .61.

2.2.4. Recommended Assessment Practices

This section was developed by the author to measure teachers' frequent use of classroom assessment practices recommended by experts of educational assessment. The section consisted of 30 items divided into five areas representing various aspects of classroom assessment that were identified from the literature. The items were drawn and adapted from previous similar questionnaires and studies in the literature of classroom assessment (e.g., Alsarimi, 2000; Ames, 1992; Church et al., 2001; Stiggins, Frisbie, & Griswold, 1989; Zhang & Burry-Stock, 2003). The areas were revision of assessment (6 items; e.g., *using a table of specifications to plan assessments*), communicating assessment (9 items; e.g., *informing students about the purpose of assessment prior*

to its administration), assessment standards and criteria (5 items; e.g., *defining a rating scale for performance criteria in advance*), student-involved assessment (4 items; e.g., *providing students opportunities to write test questions based on their understanding of the instructional objectives*), and nonachievement-based grading factors (6 items; e.g., *incorporating student' class attendance in the calculation of grades*). Scores for items reflecting un-recommended assessment practices were reversed so that higher scores represent greater alignment or agreement with the recommended assessment practices. A teacher's frequent use of the recommended assessment practices was reflected by a total rating score across all the items. The reliability coefficient was found to be .65 as indicated by Cronbach's alpha.

2.3. Data Analysis

The data in this study were hierarchically structured, in that students were nested within classes. Therefore, two HLM analyses were conducted, one for each dependent variable: (a) perceived-learning assessment environment, (b) perceived-harsh assessment environment. In order to facilitate interpretation of the HLM results, all variables, except for class's gender which was a dummy variable (1 = female classes and -1 = male classes), were standardized to a mean of zero and a standard deviation of one. The student-level independent variable (i.e., student self-efficacy) was group-mean centered. The modeling process for each dependent variable began with a fully unconditional model. The next step involved posing a random-coefficient regression model to examine the relationship of the student-level independent variable to each dimension of the assessment environment, and whether this relationship varied significantly across classes. The analyses proceeded with intercepts-and-slopes-as-outcomes regression models to explain variability in the intercepts and slopes using class-level variables. Following Raudenbush and Bryk's (2002, p. 267) suggestion, the class-level variables were divided into three sets. The first set represented the contextual effect of academic self-efficacy on perceived classroom assessment environment along with its differential contextual effect by class gender. The second set represented the joint effects of class gender and class average for perceived classroom assessment environment. The third set represented the joint effects of class gender, teacher's teaching experience, and teacher's assessment practices. Then, three submodels of the intercepts-and-slopes-as-outcomes regression model were fitted, one for each of the three sets of the class-level variables. Next, the analyses involved combining together statistically significant class-level variables detected in the early steps to produce a parsimonious overall intercepts-and-slopes-as-outcomes regression model of each dimension of the perceived assessment environment. The validity of inferences based on the final models was assessed by verifying the tenability of two-level HLM assumptions (Raudenbush & Bryk, p. 255). Prior to the HLM analyses, the data were screened at both levels of the analyses, student and class. The data screening process showed no concern about normality, outliers, and collinearity. Also, the bivariate correlations at each level were in the expected directions. Readers are invited to contact the authors for details about the analysis.

3. RESULTS

3.1. Modeling Perceived Learning Assessment Environment

3.1.1. A Fully Unconditional Model

Based on this model, a statistically significant variation was found among class means on perceived learning assessment environment; $\hat{\tau}_{00} = .1107$, $\chi^2(82) = 283.75$, *p* < .001. The estimated within-class variance ($\hat{\sigma}^2$) was .8901. Hence, the intraclass correlation ($\hat{p}$) was estimated as .1106, indicating that approximately 11% of the variance in perceived learning assessment environment was between classrooms.

3.1.2. A Random-Coefficient Regression Model

After taking student self-efficacy into account, the estimated within-class variance ($\hat{\sigma}^2$) was reduced from .8901 in the random-effects ANOVA model to .8077. Hence, student self-efficacy accounted for about 9% of the within-class variance in perceived learning assessment environment.

3.1.3. A Final Explanatory Model of Perceived Learning Assessment Environment.

Table 1 presents results of the final reduced composite model of perceived learning assessment environment. As shown in Table 1, the average self-efficacy of students was positively related to class mean perceived learning assessment environment. Also, the effect of class gender on class average perceived learning assessment environment depended significantly on teacher's frequent use of alternative assessments. Specifically, in classes using alternative assessments more frequently, the average perceived learning assessment environment was higher in female classrooms. The opposite was true in classes using alternative assessments less frequently, in that, the average perceived learning assessment environment was higher in male classrooms. Using the random-coefficient regression model as the base model, approximately 69% of the variance among classrooms in average perceived learning assessment environment was explained once class gender, class average self-efficacy, class average perceived harsh assessment environment, frequent use of alternative assessments, and interaction of class gender-by- frequent use of alternative assessments were taken into account. With regard to self-efficacy slope, perceptions of high efficacious students about their classroom assessment environment as being learning-oriented were on average not only higher; $\hat{\gamma}_{10}$ = .265, *t*(1629) = 9.430, *p* < .001; but also less variable; $\hat{\alpha}_1$ = -.149, *z* = 3.873, *p* < .001; than those for less efficacious students.

Table 1 Final Reduced Composite Model of Perceived Learning Assessment Environment With Heterogeneous Level-1 Variance

Fixed effect	Coefficient	SE	t-value
Class PLAE mean, β_{0j}			
Base, γ_{00}	.075	.037	2.022*
GNDR, γ_{01}	-.677	.166	4.080***
ALTR, γ_{02}	.011	.027	.394
CEFC, γ_{03}	.129	.036	3.554**
CPHAE, γ_{04}	-.163	.031	5.208***
GNDR × ALTR, γ_{05}	.705	.168	4.200***
SEFC slope, β_{1j}			
Base, γ_{10}	.265	.028	9.430***

Random effect	Variance component	df	χ^2
PLAE mean, u_{0j}	.0359	77	155.901***

Model for level-1 variance Parameter	Coefficient	SE	z-value
Intercept, α_0	-.206	.036	5.738***
SEFC, α_1	-.149	.038	3.873***

Note. PLAE = perceived learning assessment environment. GNDR = class gender (1 = female and -1 = male). ALTR = teacher's frequent use of alternative assessments. CEFC = class average self-efficacy. CPHAE = class average perceived harsh assessment environment. SEFC = student self-efficacy.
*$p < .05$. **$p < .01$. ***$p < .001$.

3.2. Modeling Perceived Harsh Assessment Environment

3.2.1. A Fully Unconditional Model

Based on this model, a statistically significant variation was found among class means on perceived harsh assessment environment; $\hat{\tau}_{00} = .1829$, $\chi^2(82) = 445.945$, $p < .001$. The estimated within-class variance ($\hat{\sigma}^2$) was .8179. Hence, the intraclass correlation ($\hat{p}$) was estimated as .1828, indicating that approximately 18% of the variance in perceived harsh assessment environment was between classrooms.

3.2.2. A Random-coefficient Regression Model

After taking student self-efficacy into account, the estimated within-class variance ($\hat{\sigma}^2$) was reduced from .8179 in the random-effects ANOVA model to .7761. Hence, student self-efficacy accounted for about 5% of the within-class variance in perceived harsh assessment environment.

3.2.3. A Final Explanatory Model of Perceived Harsh Assessment Environment

Table 2 presents results of the final fitted explanatory model of perceived harsh assessment environment. As shown in Table 2, the average perceived learning assessment environment was negatively related to class mean perceived harsh assessment environment. Using the random-coefficient regression model as the base model, approximately 37% of the variance among classrooms in average perceived harsh assessment environment was explained by class average perceived learning assessment environment. With regard to the self-efficacy slope, on average, student self-efficacy was negatively related to perceived harsh assessment environment within classrooms. Also, there was a statistically significant contextual effect in the data for the relationship between student self-efficacy and perceived harsh assessment environment. Specifically, the relationship between student self-efficacy and perceived harsh assessment environment tended to be stronger in classes with a low average self-efficacy than in classes with a high average self-efficacy. Further, student's self-efficacy tended to have a weaker effect on perceived harsh assessment environment in classes with a high adherence to the recommended assessment practices than in classes with a low adherence to the recommended assessment practices. In addition, the differentiating effect of student self-efficacy on perceived harsh assessment environment within a classroom depended jointly on class gender, teacher's teaching experience, and teacher's frequent use of alternative assessments. Specifically, in both classes having a high experienced teacher using alternative assessments more frequently and classes having a low experienced teacher using alternative assessments less frequently, female classrooms were less differentiating with regard to student self-efficacy than were male classrooms, holding other factors constant. The opposite was true in both classes having a high experienced teacher using alternative assessments less frequently and classes having a low experienced teacher using alternative assessments more frequently, in that, female classrooms were more differentiating with regard to student self-efficacy than were male classrooms, holding other factors constant. After controlling for class gender, class average self-efficacy, teaching experience, frequent use of recommended assessment practices, frequent use of alternative assessments, and interaction of class gender-by-teaching experience-by-frequent use of alterative assessments; no significant variation remained unexplained in the relationship between student self-efficacy and perceived harsh assessment environment.

Table 2 Final Reduced Composite Model of Perceived Harsh Assessment Environment

Fixed effect	Coefficient	SE	t-value
Class PHAE mean, β_{0j}			
Base, γ_{00}	-.001	.043	.024
CPLAE, γ_{01}	-.264	.051	5.151***
SEFC slope, β_{1j}			
Base, γ_{10}	-.185	.025	7.481***
GNDR, γ_{11}	-.062	.051	1.203
CEFC, γ_{12}	-.053	.022	2.415*
RECOM, γ_{13}	-.072	.029	2.472*
ALTR, γ_{14}	.032	.031	1.028
TEXP, γ_{15}	.018	.026	.689
GNDR × TEXP, γ_{16}	.529	.148	3.572**
GNDR × TEXP × ALTR, γ_{17}	-.452	.155	2.907**

Random effect	Variance component	df	χ^2
PHAE mean, u_{0j}	.1165	81	317.604***
Level-1 effect, r_{ij}	.7829		

Note. PHAE = perceived harsh assessment environment. CPLAE = class average perceived learning assessment environment. SEFC = student self-efficacy. GNDR = class gender (1 = female and -1 = male). CEFC = class average self-efficacy. RECOM = teacher's frequent use of recommended assessment practices. ALTR = teacher's frequent use of alternative assessments. TEXP = teacher's teaching experience.

*p < .05. **p < .01. ***p < .001.

4. DISCUSSION AND CONCLUSION

This study utilized hierarchical linear modeling techniques to examine the effects of teachers' assessment practices on ninth grade students' perceptions of the classroom assessment environment. The results showed that students' perceptions of the assessment environment were shaped by student characteristics such as self-efficacy, class contextual features such as aggregate perceived assessment environment and self-efficacy levels of the class, and teacher's teaching experience and assessment practices. These results lend support to the assertion that "classes have an assessment 'character' or environment" that originates from the teacher's classroom assessment practices, and that "students construct their own meaning [of the classroom assessment

environment] based in part on their group experiences" (Brookhart, 2004, pp. 444 – 445). Given that the classes sampled in this study were independent in the sense that each teacher taught only one class, the findings of the study indicated that students' perceptions of the classroom assessment environment did vary systematically among classrooms. On one hand, this finding tends to confirm McMillan and Workman's (1998, p. 29) conclusion that "Assessment and grading continue to be a private activity, with considerable variation among teachers." Such results may seem to contradict the assumption about the assessment context in Oman, in that the educational system in Oman is centralized and regulated to all science teachers by the Ministry of Education, and as such it was assumed that Omani science teachers follow the Ministry's classroom assessment policy with no variation among teachers' practices and their possible effects on students (Alsarimi, 2000). On the other hand, the finding of the present study highlights the shared common experience and thus perception of students within the same class about their teacher's classroom assessment practices. Therefore, the implication of this study finding for future research studying classroom environment is that researchers may need to consider not only the individual student perception of the classroom assessment environment which is referred to by Maehr and Midgley (1991, p. 405) as the "psychological environment", but also the aggregate perceptions of students in a class about their classroom assessment practices which is referred to as the "objective environment" (Church et al., 2001, p. 44).

The findings that self-efficacy was positively related to perceived learning assessment environment and negatively related to perceived harsh assessment environment seem plausible and agree with previous research findings (e.g., Brookhart & Bronowicz, 2003; Brookhart & DeVoge, 1999). According to the social cognitive theory (Bandura, 1986; Deemer, 2004), high efficacious students tend to persist in the face of difficulty, seek moderately challenging learning situations, and view failures as learning opportunities. All these aspects are theoretically consistent with the perceived learning assessment environment specified in this study. In contrast, low efficacious students tend to show little persistence with difficult tasks, try to avoid challenging achievement experiences, and may view failures as lack of ability. All these aspects are theoretically consistent with the perceived harsh assessment environment specified in this study. Therefore, the implication of these findings for practice is that one way to positively impact students' perceptions of the classroom assessment environment as being learning oriented is through self-efficacy. For example, teaching students to set short term goals for themselves when assigned a task, accompanied with clearly-defined assessment standards and criteria as well as frequent informative feedback may convey a positive sense of efficacy and self-improvements (Stipek, 2002). Although previous research has supported the influence of perceived classroom environment on student self-efficacy (e.g., Anderman & Midgley, 1997); it may be insensible to make causal inferences from the correlational nature of these results. Reciprocal causation between these two constructs may occur over time. As such, future research should be conducted to further examine the relationship between perceptions of the classroom assessment environment and self-efficacy.

Further, the present study extends previous research findings by suggesting that class gender, teacher's teaching experience, and assessment practices may be possible explanations for the relationship between self-efficacy and perceived assessment environment. First, the findings of this

study showed that the negative effect of self-efficacy on perceived assessment environment tended to be strengthened in classes with a low adherence to the recommended assessment practices, thereby demonstrating how important assessment practices recommended by educational assessment experts are for desirable classroom learning environment. This finding demonstrates how important classroom assessment practices recommended by educational assessment experts are for desirable classroom learning environment. However, the present study was based on the assumption that aspects of the classroom assessment are interdependent, operating in a multiplicative manner, and as such the study followed an integrative approach to the investigation of the effects of the recommended classroom assessments on student perceived classroom assessment environment. Stated differently, one limitation of this study was that it construed teacher's frequent use of the recommended classroom assessment practices as an omnibus combination of variables including revision of assessment, communication of assessment, student-involved assessment, and grading factors. This might have made it difficult to know which aspects of the recommended classroom assessment practices could be considered responsible for the observed effects detected in the study.

Second, the findings of this study showed that the negative effect of self-efficacy on perceived assessment environment tended to be weakened in male classes having a high experienced teacher than in male classes having a low experienced teacher. The opposite was found true in female classes. In light of teaching efficacy research (e.g., Daugherty, 2005; Deemer, 2004), the sampled high experienced male teachers and low experienced female teachers in this study might have high levels of teaching efficacy defined as strong beliefs about their capabilities to help students learn (Stipek, 2002), which in turn might be responsible for weakening the negative effect of student self-efficacy on perceived assessment environment in their classes. The implications of these findings for theory and practice are that given that the classroom assessment environment is often structured by the teacher (Brookhart, 1997) in the sense that classroom assessment is, to a large extent, a teacher-centered activity, teacher's psychological belief systems may influence their views and practice about certain aspects of the classroom assessment. Therefore, changing classroom assessment environment as perceived by students may require considering not only students' self-efficacy, but also teachers' teaching experience, self-beliefs and goals for students' learning, and their adherence to the recommended assessment practices.

It has been reported that females generally tend to hold higher positive perceptions of their classroom learning environment (e.g., Meece et al., 2003) and that students in alternative assessment classes generally tend to express higher positive perceptions of their classroom environment (Maslovaty & Kuzi, 2002). The current study findings offer additional clarifications for this prior work. Specifically, the findings of this study indicated that the higher levels of perceptions about classroom assessment environment as being learning-oriented were salient for female students in classes using alternative assessment more frequently and for male students in classes using alternative assessments less frequently. On one hand, these findings suggest that the public nature of alternative assessments may be responsible for promoting low levels of perceived learning classroom assessment environment in male classes, in that students are required to publicly demonstrate their knowledge (Darling-Hammond, Ancess, & Falk, 1995). On the other hand, the

findings also suggest that the challenging and contextual nature of alternative assessments may be responsible for promoting high levels of perceived learning classroom assessment environment in female classes, in that students are provided opportunities to develop higher order skills through challenging and authentic forms of assessment tasks linked to real life experiences (Darling-Hammond et al., 1995). To sum, the findings from this study imply that alternative assessments may be more advantageous for female students than for male students in depicting their classroom assessment environment. Finally, the generalizability of this study's findings may be limited by the use of self-report questionnaires and by the particular participating sample of students and teachers. Future research should be conducted to testify the findings from this study in various subject areas and grade levels using mixed-methods research design.

REFERENCES

Alkharusi, H. (2008). Effects of classroom assessment practices on student's achievement goals. *Educational Assessment, 13*, 243-266.

Alsarimi, A. M. (2000). *Classroom assessment and grading practices in the Sultanate of Oman*. Unpublished dissertation, University of Pittsburgh, Pennsylvania.

American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). Standards for teacher competence in educational assessment of students. *Educational Measurement: Issues and Practice, 9*, 30 – 32.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*, 261 – 271.

Anderman, E. M., & Midgely, C. (1997). Changes in achievement goal orientations, perceived academic competence, and grades across the transition to middle-level schools. *Contemporary Educational Psychology, 22*, 269 – 298.

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman and Company.

Bol, L., Stephenson, P. L., O'Connell, A. A., & Nunnery, J. A. (1998). Influence of experience, grade level, and subject area on teachers' assessment practices. *The Journal of Educational Research, 91*, 323 – 330.

Britner, S. L., & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school student. *Journal of Research in Science Teaching, 43*, 485 – 499.

Brookhart, S. M. (1997). A theoretical framework for the role of classroom assessment in motivating student effort and achievement. *Applied Measurement in Education, 10*, 161 – 180.

Brookhart, S. M., & Bronowicz, D. L. (2003). 'I don't like writing. It makes my fingers hurt': Students talk about their classroom assessment. *Assessment in Education: Principles, Policy & Practice, 10,* 221 – 242.

Brookhart, S. M. (2004). Classroom assessment: Tensions and intersections in theory and practice. *Teachers College Record, 106*, 429 – 458.

Brookhart, S. M., & DeVoge, J. G. (1999). Testing a theory about the role of classroom assessment in student motivation and achievement. *Applied Measurement in Education, 12*, 409 – 425.

Church, M. A., Elliot, A. J., & Gable, S. L. (2001). Perceptions of classroom environment, achievement goals, and achievement outcomes. *Journal of Educational Psychology, 93*, 43 – 54.

Darling-Hammond, L. (1994). Setting standards for students: The case for authentic assessment. *The Educational Forum, 59*, 14 – 21.

Darling-Hammond, L., Ancess, J., & Falk, B. (1995). *Authentic assessment in action: Studies of schools and students at work.* New York: Teachers College Press.

Daugherty, S. G. (2005). Teacher efficacy and its relation to teachers' behaviors in the classroom (Doctoral dissertation, University of Houston, 2005). *Dissertation Abstracts International, 66*, 553.

Deemer, S. A. (2004). Classroom goal orientation in high school classrooms: Revealing links between teacher beliefs and classroom environments. *Educational Research, 46*, 74 – 90.

Greene, B. A., Miller, R. B., Crowson, H. M., Duke, B. L., & Akey, K. L. (2004). Predicting high school students' cognitive engagement and achievement: Contributions of classroom perceptions and motivation. *Contemporary Educational Psychology, 29*, 462 – 482.

Henning-Stout, M. (1994). Responsive assessment: A new way of thinking about learning. San Francisco: Jossey-Bass.

Kaplan, A., Middleton, M. J., Urdan, T., & Midgley, C. (2002). Achievement goals and goal structures. In C. Midgley (Ed.), *Goals, goal structures, and patterns of adaptive learning* (pp. 21 – 53). Mahwah, NJ: Lawrence Erlbaum.

Maehr, M. L., & Midgley, C. (1991). Enhancing student motivation: A school-wide approach. *Educational Psychologist, 26*, 399 – 427.

Maslovaty, N., & Kuzi, E. (2002). Promoting motivational goals through alternative or traditional assessment. *Studies in Educational Evaluation, 28*, 199 – 222.

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., & Workman, D. J. (1998). *Classroom assessment and grading practices: A review of the literature.* (ERIC Document Reproduction Service No. ED453263)

Meece, J. L., Herman, P., & McCombs, B. L. (2003). Relations of learner-centered teaching practices to adolescents' achievement goals. *International Journal of Educational Research, 39,* 457 – 475.

Mertler, C. A. (1998, October). *Classroom assessment practices of Ohio teachers*. Paper presented at the meeting of the Mid-Western Educational Research Association, Chicago.

Mertler, C. A. (1999, October). *Teachers' (Mis)conceptions of classroom test validity and reliability.* Paper presented at the meeting of the Mid-Western Educational Research Association, Chicago.

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., Gheen, M., Kaplan, A., Kumar, R., Middleton, M. J., Nelson, J., Roeser, R., & Urdan, T. (2000). *Manual for the patterns of adaptive learning scales*. Retrieved August 5, 2005, from http://www.umich.edu/~pals/PALS%202000_V13Word97.pdf

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

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

Stipek, D. (2002). *Motivation to learn: Integrating theory and practice* (4th ed.). Boston: Allyn and Bacon.

Zhang, Z., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers' self-perceived assessment skills. *Applied Measurement in Education, 16*, 323 -342.