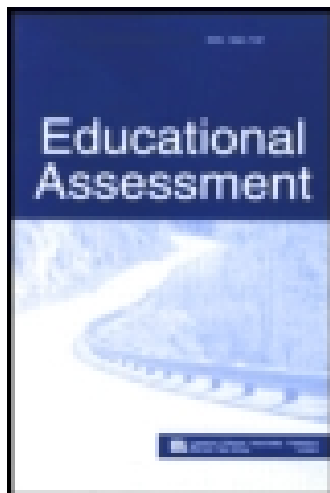


This article was downloaded by: [University of Bath]

On: 15 April 2015, At: 04:31

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Educational Assessment

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/heda20>

Effects of Classroom Assessment Practices on Students' Achievement Goals

Hussain Alkharusi^a

^a Sultan Qaboos University, Oman

Published online: 16 Dec 2008.

To cite this article: Hussain Alkharusi (2008) Effects of Classroom Assessment Practices on Students' Achievement Goals, Educational Assessment, 13:4, 243-266, DOI: [10.1080/10627190802602509](https://doi.org/10.1080/10627190802602509)

To link to this article: <http://dx.doi.org/10.1080/10627190802602509>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Effects of Classroom Assessment Practices on Students' Achievement Goals

Hussain Alkharusi
Sultan Qaboos University, Oman

This study examined the effects of classroom assessment practices on students' achievement goals. The study included 1,636 ninth grade students and 83 science teachers from Muscat public schools in Oman. Results from hierarchical linear modeling techniques showed that class contextual features and teachers' teaching experiences and assessment practices interacted significantly with students' characteristics in influencing students' achievement goals. Recommendations, implications, and suggestions for practice and future research are discussed.

The classroom assessment environment as perceived by students is of increasing interest to the educational assessment community. Many educators have claimed that assessment-related activities used in the classroom convey important information about what is valued there, and hence have an influence on students' achievement goals (Ames, 1992; Harlen & Crick, 2003). Students' achievement goals have been regarded as important achievement-related outcomes to be promoted in the classroom learning process (Harlen & Crick, 2003). Since a large amount of the classroom time is assigned to assessment-related activities (Mertler, 2003), identifying assessment practices conducive to the adoption of desirable achievement goals becomes crucial. The present study attempts to address this issue by utilizing achievement goal theory and classroom assessment literature to examine the effects of certain student characteristics (e.g., self-efficacy) and classroom characteristics (e.g., assessment practices) on achievement goals for ninth-grade students in Muscat science classrooms in Oman.

Correspondence should be addressed to Hussain Alkharusi, Sultan Qaboos University, P.O. Box 32 Al-Khoud P.C. 123 Oman. E-mail: hussein5@squ.edu.om

ACHIEVEMENT GOAL THEORY

Achievement goal theorists have traditionally identified two types of achievement goals: mastery and performance goals (Ames, 1992; Dweck, 1986; Nicholls, 1984). Mastery goals center on the development of competence whereas performance goals center on the outward showing of competence (Ames, 1992; Dweck, 1986; Nicholls, 1984). Students who adopt mastery goals are expected to persist in the face of difficulty, seek challenging tasks, and have high intrinsic motivation (Ames, 1992; Dweck, 1986; Nicholls, 1984). In contrast, students who adopt performance goals are expected to minimally persist in the face of difficulty, avoid challenging tasks, and have low intrinsic motivation (Ames, 1992; Dweck, 1986; Nicholls, 1984). Since its origin in the late 1970s and early 1980s, achievement goal theory has undergone a number of theoretical advances. For example, Elliot and Church (1997) and Middleton and Midgley (1997) have developed a trichotomous framework of achievement goals that further differentiates performance goals into approach and avoidance goals. In particular, three achievement goals have been identified: (a) mastery goals that focus on improving competence, (b) performance-approach goals that focus on displaying competence, and (c) performance-avoidance goals that focus on avoiding a display of incompetence (Elliot & Church, 1997; Middleton & Midgley, 1997). This framework has been validated with college (Elliot & Church, 1997) and middle school students (Middleton & Midgley, 1997). The present study focuses on middle school students and, as such, its conceptualization of achievement goals is based on the approach suggested by Middleton and Midgley.

CLASSROOM ASSESSMENT ENVIRONMENT

To illustrate the role of the assessment environment on student motivation and achievement, Brookhart (1997) has developed a theoretical model based on a synthesis of classroom assessment literature and social cognitive theories of learning and motivation. In this model, the classroom assessment environment is conceived as a classroom context experienced by students as the teacher establishes assessment purposes, assigns assessment tasks, sets performance criteria and standards, gives feedback, and monitors outcomes. Based on this model, students' perceptions of the classroom assessment environment have been thought to influence students' motivational beliefs and achievement-related outcomes. This postulation has been examined for elementary, middle, and high school students as well as for college-level students (e.g., Brookhart & Bronowicz, 2003; Church, Elliot, & Gable, 2001; Rodriguez, 2004; Wang, 2004).

However, most research on the assessment environment has used individual student scores as the unit of analysis rather than the average score of students

at the classroom level (e.g., Church et al., 2001; Wang, 2004). Proponents of this approach argued 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). Nevertheless, Church et al. asserted that “composite measures of perceived classroom [assessment environment] have been shown to be internally consistent, and composite indicators yield a more comprehensive assessment of the perceived classroom environment than do individual indicators” (p. 51). Likewise, Brookhart (2004, p. 444) maintained that “classes have an assessment ‘character’ or environment” that originates from the teacher’s assessment practices, and that “students construct their own meaning [of the classroom assessment environment] based in part on their group experiences” (p. 445). These assertions have the following implication for the present study.

There seem to be two conceptually related types of students’ perceptions about the classroom assessment environment: individual and aggregate. The individual perception of the classroom assessment environment refers to the self-perception of an individual student about the various aspects of the classroom assessment activities as the teacher establishes assessment purposes, assigns assessment tasks, sets performance criteria and standards, provides feedback, and monitors outcomes (Brookhart, 1997). The aggregate (i.e., class) perception of the assessment environment refers to the overall shared perception of students in a classroom about the various aspects of the classroom assessment activities done by the teacher; and it can be reflected for each class by the average levels of individual students’ perceptions within the classroom (i.e., class average perceived assessment environment). The vital distinction between individual and aggregate perceptions of the assessment environment entails the object of interest: self or group perceptions of the assessment environment, respectively.

One may argue that the aggregate perception of the assessment environment is a cogent attribute for describing the social influence of the classroom. From the perspective of social theory, norms are developed to give class members some power to regulate their actions over others when those actions have effects for the class (Coleman, 1987, 1990, as cited in Goddard & Goddard, 2001, p. 810). When the individual student’s perceptions do not agree with the shared perceptions of the class, the student’s perceptions can be weakened by the class members (Coleman, 1987, 1990, as cited in Goddard & Goddard, 2001, p. 810). Yet, except for Meece, Herman, and McCombs’ (2003) study, no recent study to my knowledge has considered how perceptions of the assessment environment interact across levels, student, and class in influencing students’ achievement goals. Contrary to Meece et al.’s study that combined participants from various subject areas and grade levels, the present study will control for the effects of academic subject and grade level by focusing on ninth-grade science classrooms.

In addition, instead of general classroom teaching practices, the present study will focus on classroom assessment practices.

SELF-EFFICACY

Findings from previous studies (e.g., Brookhart & Bronowicz, 2003) suggest that students' academic self-efficacy may need to be considered when investigating the impact of classroom assessment practices on students' achievement goals. Self-efficacy pertains to students' judgments of their performance capability on a particular type of assessment task (Bandura, 1986). According to Bandura's social cognitive theory, performances in previous assessment tasks influence self-efficacy judgments for tasks of the same assessment type. If students have experienced success in earlier assessment tasks, they are more likely to feel capable to succeed in future tasks of the same type of assessment, which in turn may lead to adoption of performance-approach and/or mastery goals (Schunk, 1996). As such, one could argue that there may be a cross-level interaction effect for teachers' use of a particular assessment type and students' self-efficacy on students' achievement goals. The present study will attempt to shed light on the relationship between students' self-efficacy and achievement goals as a function of teachers' assessment practices in ninth-grade Muscat science classrooms in Oman.

GENDER

Previous research findings (e.g., Wang, 2004) suggest that gender may also need to be considered when investigating the impact of classroom assessment on students' achievement goals. Specifically, in Wang's study, performance-approach goals were found to be positively related to both perceptions of the classroom assessment environment as being learning-oriented and test-oriented for male students, but not for female students. In addition, mastery goals were found to be positively related to perceptions of the classroom assessment environment as being learning-oriented for male students but not for female students. Having concluded that the relationship between the perceived classroom assessment environment and achievement goals differ as a function of gender, it seems necessary to account for these differences.

It should be noted, however, that gender in this study varies across classes because in Oman, students within the same class and their science teacher are of the same gender: either all of them are male or all of them are female. Therefore, unlike previous studies, gender in this study will be more appropriately treated as a class-level independent variable. Hierarchical linear modeling

(HLM; Raudenbush & Bryk, 2002), which is the appropriate analytic technique for this study to handle the nested nature of the data in that students are nested within classrooms, provides two ways for accounting for the gender effects (a) controlling for the student-level differences and (b) controlling for the class-level differences. Among middle school students, research has shown that girls tend to hold stronger science self-efficacy beliefs than boys (Britner & Pajares, 2006). Thus, the HLM approach to accounting for the gender effects will be applied in this study by assuming that student's self-efficacy is a confounding variable in the within-class models 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 models.

TRADITIONAL AND ALTERNATIVE ASSESSMENTS

According to Brookhart's (1997) model, teachers' use of distinct types of assessments is one aspect of the classroom assessment environment. For the most part, assessment methods can be classified as traditional or alternative based on the realism and complexity of the assessment tasks and the amount of time needed for the assessment (Gronlund, 2006). Traditional assessments such as multiple choice, true-false, and matching items are often lower in realism and complexities of the tasks assessed but require little time to administer and score (Gronlund, 2006). Alternative assessments such as portfolios, observations, and other performance-based assessments are higher in both realism and complexity of the tasks assessed but require more time to use and score than traditional assessments (Gronlund, 2006). There has been a movement toward the use of more alternative assessments than traditional assessments. The arguments in favor of alternative assessments over traditional ones are based on the notion that alternative assessments are more intrinsically motivating than traditional assessments (Shepard, 2000). However, little empirical research exists about the effects of assessment type on achievement goals. When this topic was considered (e.g., Stefanou and Parkes, 2003), not enough attention was given to the nested structure of the data (students nested within classrooms), thereby ignoring the consequences that clustering of the data might have for the relationships between assessment practices and achievement goals. This study will attempt to overcome this issue by not only considering both levels of the data but also by estimating the relations that might cross the student-level and the class-level characteristics.

In addition, classroom assessment research has reported significant differences based on teaching experience and gender regarding teachers' frequent use of alternative assessments but not with traditional assessments (e.g., Bol, Stephenson, O'Connell, & Nunnery, 1998; Mertler, 1998). Therefore, in this study, the effects of teachers' frequent use of alternative assessments on stu-

dents' achievement goals will be studied in interaction with class gender and teaching experience, whereas only main and cross-level interaction effects will be specified and tested for teachers' frequent use of traditional assessments.

RECOMMENDED ASSESSMENT PRACTICES

Teachers are required to develop classroom assessment that aligns with practices recommended by experts of educational assessment. For example, it is recommended that teachers use multiple assessment methods to gather evidence about student learning (Stiggins, Frisbie, & Griswold, 1989); students should clearly be informed about the grading procedure in advance; nonachievement factors such as neatness of work and class attendance should not be incorporated into grading; and students should be given continuous and informative assessment feedback rather than judgmental feedback about their academic performance (Brookhart, 1994). Unfortunately, findings from classroom assessment research has revealed a gap between the recommended and the actual assessment practices regardless of teachers' gender and teaching experience (e.g., Alsarimi, 2000). Brookhart stated that "grading theory and practices will be better connected once the role of classroom assessment and grading practices in student achievement motivation . . . is understood" (p. 279) and, as such, it seems reasonable to argue that to be able to understand and make sense out of the gap between assessment experts' recommendations and teachers' assessment practices, it is important to find out the possible effects of these practices on students' achievement goals as one aspect of student motivation for learning. The present study will attempt to shed light on this issue by testing the main and cross-level interaction effects of teachers' frequent use of the recommended assessment practices on students' achievement goals.

METHODS

Participants and Procedures

The study included 1,636 ninth-grade students and their 83 science teachers from Muscat public schools in Oman. Each class consisted of 14 to 21 students with an average of 20. Of all students, 735 were males and 901 were females. Of all 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 regularly scheduled class meet-

ing. The participants were told that a study was being conducted to investigate the relationship of teachers' assessment practices to students' achievement goals in the ninth-grade science classrooms. They were informed that they were not obligated to participate, and if they wished, their responses would remain anonymous and confidential.

Instrumentation

Two questionnaires were used, one for students and one for teachers. The students' questionnaire included items that asked students to indicate their perceptions of the assessment environment, achievement goals, and self-efficacy using a 4-point Likert scale ranging from 1 (*completely not true*) to 4 (*completely true*). The teachers' questionnaire included items that asked 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*). All questionnaires were administered in Arabic, which is the medium language of instruction.

Students' Questionnaire

Classroom assessment environment. This section included items from the three dimensions of the perceived classroom assessment environment scale developed by Alkharusi (2007). The first dimension consisted of five items ($\alpha = .63$) that measured the perceived *learning* assessment environment (e.g., *in this class, students are given a chance to correct their mistakes*). The second dimension consisted of three items ($\alpha = .52$) that measured the perceived *harsh* assessment environment (e.g., *the science tests in this class are difficult to students*). The third dimension consisted of five items ($\alpha = .45$) that measured the perceived *public* assessment environment (e.g., *in this class, students who do well are praised in front of the whole class*). 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.

Achievement goals. This section contained 14 items from the Patterns of Adaptive Learning Scales (PALS) developed by Midgley et al. (2000). Alkharusi (2007) validated the items for use with ninth-grade students in Oman. The items measure students' mastery (5 items; e.g., *One of my goals is to master a lot of new science skills this semester*), performance-approach (5 items; e.g., *One of my goals is to show others that I'm good at science class work*), and performance-avoidance goals (4 items; e.g., *It is important to me that I don't look stupid in science class*). Three scales were created by averaging scores on the items of each achievement goal. Higher scores on each scale represented a higher adoption of the achievement goal defined in that scale. Internal consistency reliability coefficients were .66, .75, and .54 for scores representing mastery, performance-approach, and performance-avoidance goals, respectively.

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*). A principal components analysis was conducted on the six items to determine whether they represent a single construct. This analysis yielded a single factor with an eigenvalue of 2.625, accounting for 43.74% of the total variance. All items loaded higher than .60 on the factor. 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.

Teachers' Questionnaire

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 .29 with a standard error of measurement (SEM) of 2.24. The reliability coefficient for scores representing alternative assessments was .41 with an SEM of 2.60. As might be noted, the reliability coefficients were low. Yet, their corresponding SEMs were also low, suggesting small discrepancies between an individual teacher's true score and the observed scores over repeated administrations of the scales (Crocker & Algina, 1986). As suggested by Roberts, Onwuegbuzie, and Eby (2001, p. 15), "the data should still be used in the analysis ... because the low reliability estimate is due to individual homogeneity and thus appears acceptable considering the context of the study." Therefore, the homogeneous nature of the present study's sample of teachers needs to be considered when interpreting the results pertaining to teachers' frequent uses of traditional and alternative assessments.

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 that represented 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 et al., 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 .56 as indicated by Cronbach's alpha.

DATA ANALYSIS

The data in this study were hierarchically structured in that students were nested within classes. Therefore, three HLM analyses were conducted, one for each dependent variable: (a) mastery, (b) performance-approach, and (c) performance-avoidance goals. 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 variables were 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 relationships of student-level independent variables to each achievement goal and whether these relationships varied significantly across classes. The student-level independent variables were student self-efficacy, perceived learning assessment environment, perceived harsh assessment environment, and perceived public assessment environment. 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 two sets. The first set represented the contextual-effects of self-efficacy and perceived assessment environment along with their differential contextual effects by class gender. The second set represented the joint effects of class gender, teacher's teaching experience, and teacher's assessment practices. Then, two submodels of the intercepts-and-slopes-as-outcomes regression model were fitted, one for each of the two 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 achievement goal. The validity of inferences based on the final models was assessed by verifying the tenability of two-level HLM assumptions (Raudenbush & Bryk, 2002, 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. In addition, the bivariate correlations at each level were in the expected directions. Readers are invited to contact the authors for details about the analysis.

RESULTS

Modeling Mastery Goal

A fully unconditional model. Based on this model, a statistically significant variation was found among class means on mastery goal; $\hat{\tau}_{00} = 0.0851$,

$\chi^2(82) = 231.861$, $p < .001$. The estimated within-class variance ($\hat{\sigma}^2$) was .9161. Hence, the intraclass correlation ($\hat{\rho}$) was estimated as .0850, indicating that 8.5% of the variance in mastery goal was between classrooms.

A random-coefficient regression model. After taking student self-efficacy and perceived assessment environment into account, the estimated within-class variance ($\hat{\sigma}^2$) was reduced from .9161 in the fully unconditional model to .6479. Hence, student self-efficacy and perceived assessment environment accounted for about 29% of the within-class variance in mastery goal.

A final explanatory model of mastery goal. Based on the early steps of the analyses, Table 1 presents results of the final fitted explanatory model of mastery goal.

With regard to class mean mastery goal, as shown in Table 1, female classrooms had significantly higher average mastery goals than did male classrooms; $\hat{\gamma}_{01} = .139$, $t(80) = 5.262$, $p < .001$; controlling for class average self-efficacy. In addition, holding other factors constant, average self-efficacy of students was positively related to class mean mastery goal; $\hat{\gamma}_{02} = .168$, $t(80) = 5.664$, $p < .001$. The estimated between-class variance in average mastery goal ($\hat{\tau}_{00}$) was reduced from .0989 in the random coefficient regression model to .0308, suggesting that about 69% of the variance among classrooms in average mastery goal was explained by class gender and class average self-efficacy.

With regard to self-efficacy slope, as shown in Table 1, mastery goal levels of high efficacious students were on average not only higher, $\hat{\gamma}_{10} = .334$, $t(81) = 10.892$, $p < .001$, but also less variable than those for less efficacious students; $\hat{\alpha}_1 = -.259$, $z = -5.927$, $p < .001$. In addition, the positive relationship between student self-efficacy and mastery goal tended on average to be stronger in male classrooms than in female classrooms; $\hat{\gamma}_{11} = -.071$, $t(81) = -2.292$, $p < .05$. Using the random-coefficient regression model of mastery goals as the base model, class gender explained approximately 2% of the variance in the relationship between student self-efficacy and mastery goal.

Regarding the perceived learning assessment environment slope, as shown in Table 1, the differentiating effect of the perceived learning assessment environment within a classroom depended jointly on the class average perceived learning assessment environment and class gender; $\hat{\gamma}_{23} = -1.572$, $t(80) = -7.082$, $p < .001$. Specifically, in male classrooms, classes with a high average perceived learning assessment environment were more differentiating with regard to student perceived learning assessment environment than were classes with a low average perceived learning assessment environment; $\hat{\gamma}_{22} + (GNDR)_j \hat{\gamma}_{23} = .059 + (-1)(-1.571) = 1.631$. The opposite was true in female classrooms. In addition, mastery goal levels of students with a high perceived learning assessment environment were less variable than those for students with low

TABLE 1
Final Fitted Composite Model of Mastery Goals with
Heterogeneous Level-1 Variance

<i>Fixed Effect</i>	<i>Coefficient</i>	<i>SE</i>	<i>t-value</i>
Class MASG mean, β_{0j}			
Base, γ_{00}	-.006	.029	-.216
GNDR, γ_{01}	.139	.026	5.262***
CEFC, γ_{02}	.168	.030	5.664***
SEFC slope, β_{1j}			
Base, γ_{10}	.334	.031	10.743***
GNDR, γ_{11}	-.071	.031	-2.292*
SLAE slope, β_{2j}			
GNDR, γ_{21}	1.557	.216	7.214***
CLAE, γ_{22}	.059	.022	2.747**
GNDR \times CLAE, γ_{23}	-1.572	.222	-7.082***
SHAE slope, β_{3j}			
Base, γ_{30}	-.090	.024	-3.720***
<i>Random Effect</i>	<i>Variance Component</i>	<i>df</i>	χ^2
MASG mean, u_{0j}	.0308	80	157.817***
SEFC slope, u_{1j}	.0311	81	138.965***
SLAE slope, u_{2j}	.0198	80	120.549**
<i>Parameter</i>	<i>Coefficient</i>	<i>SE</i>	<i>z-value</i>
Model for Level-1 Variance			
Intercept, α_0	.504	.038	13.389***
SEFC, α_1	-.259	.044	-5.927***
SLAE, α_2	-.181	.045	-4.057***
SHAE, α_3	.177	.042	4.172***

Note. MASG = mastery goal. GNDR = class gender (1 = female and -1 = male). CEFC = class average self-efficacy. SEFC = student self-efficacy. SLAE = student perceived learning assessment environment. CLAE = class average perceived learning assessment environment. SHAE = student perceived harsh assessment environment.

* $p < .05$; ** $p < .01$; *** $p < .001$.

perceived learning assessment environment levels; $\hat{\alpha}_2 = -.181$, $z = -4.057$, $p < .001$. Using the random-coefficient regression model of mastery goals as the base model, approximately 46% of the variance among classrooms in the perceived learning assessment environment differentiation effect was explained by class gender, class average perceived learning assessment environment, and interaction of class gender-by-class average perceived learning assessment environment. With regard to a perceived harsh assessment environment slope, as shown in Table 1, mastery goal levels of students with a high perceived harsh assessment

environment were on average lower; $\hat{\gamma}_{30} = -.090$, $t(1627) = -3.720$, $p < .001$; and more variable than those for students with a low perceived harsh assessment environment levels; $\hat{\alpha}_3 = .177$, $z = 4.172$, $p < .001$.

Modeling Performance-Approach Goal

A fully unconditional model. Based on this model, a statistically significant variation was found among class means on performance-approach goals; $\hat{\tau}_{00} = .0571$, $\chi^2(82) = 179.7803$, $p < .001$. The estimated within-class variance ($\hat{\sigma}^2$) was .9436. Hence, the intraclass correlation ($\hat{\rho}$) was estimated as .0571, indicating that about 6% of the variance in performance-approach goal was between classrooms.

A random-coefficient regression model. After taking student self-efficacy and perceived public assessment environment into account, the estimated within-class variance ($\hat{\sigma}^2$) was reduced from .9436 in the fully unconditional model to .8052. Hence, student self-efficacy and perceived public assessment environment accounted for about 15% of the within-class variance in performance-approach goal.

A final explanatory model of performance-approach goal. Based on the early steps of the analyses, Table 2 presents results of the final composite model of performance-approach goal.

With regard to class mean performance-approach goal, as shown in Table 2, holding other factors constant, female classrooms had a significantly higher average performance-approach goal than did male classrooms; $\hat{\gamma}_{01} = .069$, $t(78) = 2.049$, $p < .05$. In addition, holding other factors constant, although not statistically significant, there was a trend for classes with a high frequent use of alternative assessments to have a smaller average performance-approach goal than did classes with a low frequent use of alternative assessments; $\hat{\gamma}_{02} = -.068$, $t(78) = -1.895$, $p = .061$. Further, holding other factors constant, the average self-efficacy of students was positively related to class mean performance-approach goal; $\hat{\gamma}_{03} = .114$, $t(78) = 3.914$, $p < .001$. Moreover, holding other factors constant, the average perceived public assessment environment was positively related to the class mean performance-approach goal; $\hat{\gamma}_{04} = .075$, $t(78) = 2.697$, $p < .01$. The estimated between-class variance in average performance-approach goal was reduced from .0644 in the random coefficient regression model to .0372, suggesting that about 42% of the variance among classrooms in average performance-approach goal was explained after class gender, teacher's frequent use of alternative assessments, class average self-efficacy, and the class average perceived public assessment environment were taken into account.

TABLE 2
Final Fitted Composite Model of Performance-Approach Goal Orientation
with Heterogeneous Level-1 Variance

<i>Fixed Effect</i>	<i>Coefficient</i>	<i>SE</i>	<i>t-value</i>
Class PAPG mean, β_{0j}			
Base, γ_{00}	-.006	.031	-.190
GNDR, γ_{01}	.069	.034	2.049*
ALTR, γ_{02}	-.068	.036	-1.895
CEFC, γ_{03}	.114	.029	3.914***
CPAE, γ_{04}	.075	.028	2.697**
SEFC slope, β_{1j}			
Base, γ_{10}	.382	.045	8.476***
GNDR, γ_{11}	-.902	.321	-2.806**
ALTR, γ_{12}	.005	.026	.203
TEXP, γ_{13}	.008	.021	.405
GNDR \times TEXP, γ_{14}	.617	.263	2.342*
GNDR \times ALTR, γ_{15}	.955	.341	2.800**
GNDR \times TEXP \times ALTR, γ_{16}	-.715	.286	-2.502*
SPAPE slope, β_{2j}			
Base, γ_{20}	.144	.028	5.079***
GNDR, γ_{21}	.064	.028	2.269*
<i>Random Effect</i>	<i>Variance Component</i>	<i>df</i>	χ^2
PAPG mean, u_{0j}	.0372	78	160.706***
SEFC slope, u_{1j}	.0155	76	117.481**
<i>Parameter</i>	<i>Coefficient</i>	<i>SE</i>	<i>z-value</i>
Model for level-1 variance			
Intercept, α_0	.213	.037	5.802***
SPAPE, α_1	-.196	.041	-4.789***

Note. PAPG = performance-approach goal. GNDR = class gender (1 = female and -1 = male). ALTR = teacher's frequent use of alternative assessments. CEFC = class average self-efficacy. CPAE = class average perceived public assessment environment. SEFC = student self-efficacy. TEXP = teacher's teaching experience. SPAPE = student perceived public assessment environment.

* $p < .05$; ** $p < .01$; *** $p < .001$.

Regarding self-efficacy slope, as shown in Table 2, holding other factors constant, on average, student self-efficacy was positively related to performance-approach goal within classrooms; $\hat{\gamma}_{10} = .382$, $t(76) = 8.476$, $p < .001$. In addition, holding other factors constant, the differentiating effect of self-efficacy within a classroom depended jointly on class gender, teacher's teaching experience, and teacher's frequent use of alternative assessments; $\hat{\gamma}_{16} = -.715$, $t(76) = -2.502$, $p < .05$. This can be seen by computing the differentiating effect of

self-efficacy separately for male and female classrooms having high ($TEXP = 1$) and low ($TEXP = -1$) experienced teachers using alternative assessments more ($ALTR = 1$) or less ($ALTR = -1$) frequently based on the following equation:

$$\begin{aligned} &\hat{\gamma}_{10} + \hat{\gamma}_{11}(GNDR)_j + \hat{\gamma}_{12}(ALTR)_j + \hat{\gamma}_{13}(TEXP)_j + \hat{\gamma}_{14}(GNDR \times TEXP)_j \\ &+ \hat{\gamma}_{15}(GNDR \times ALTR)_j + \hat{\gamma}_{16}(GNDR \times TEXP \times ALTR)_j \end{aligned}$$

Accordingly, for male classrooms using alternative assessments less frequently, classes having a highly experienced teacher were less differentiating with regard to student self-efficacy than were classes having a less experienced teacher. The opposite was true in male classrooms that used alternative assessments more frequently. For female classrooms that used alternative assessments less frequently, classes having a highly experienced teacher were more differentiating with regard to student self-efficacy than were classes having a less experienced teacher. The opposite was true in female classrooms that used alternative assessments more frequently.

Using the random-coefficient regression model of performance-approach goal as the base model, approximately 32% of the variance among classrooms in self-efficacy differentiating effect was explained by class gender, frequent use of alternative assessments, teaching experience, interaction of class gender-by-teaching experience, interaction of class gender-by-frequent use of alternative assessments, and interaction of class gender-by-teaching experience-by-frequent use of alternative assessments. As shown in Table 2, performance-approach goal levels of students with a high perceived public assessment environment were higher, $\hat{\gamma}_{20} = 144$, $t(1622) = 5.079$, $p < .001$, and less variable, $\hat{\alpha}_1 = -.196$, $z = -4.789$, $p < .001$, than those for students with low levels of perceived public assessment environment. In addition, the positive relationship between the perceived public assessment environment and performance-approach goal tended to be stronger in female classrooms than in male classrooms; $\hat{\gamma}_{21} = .064$, $t(1622) = 2.269$, $p < .05$.

Modeling Performance-Avoidance Goal

A fully unconditional model. Based on this model, a statistically significant variation was found among class means on performance-avoidance goal; $\hat{\tau}_{00} = .0536$, $\chi^2(82) = 173.270$, $p < .001$. The estimated within-class variance ($\hat{\sigma}^2$) was .9474. Hence, the intraclass correlation ($\hat{\rho}$) was estimated as .0535, indicating that about 5% of the variance in performance-avoidance goal was between classrooms.

A random-coefficient regression model. After taking student self-efficacy and perceived public assessment environment into account, the estimated within-class variance ($\hat{\sigma}^2$) was reduced from .9474 in the fully unconditional model to .8846. Hence, student self-efficacy and perceived public assessment environment accounted for about 7% of the within-class variance in performance-avoidance goal.

A final explanatory model of performance-avoidance goal. Based on the early steps of the analyses, Table 3 presents results of the final fitted explanatory model of performance-avoidance goal.

Based on Table 3, holding other factors constant, female classrooms had higher average performance-avoidance goals than did male classrooms; $\hat{\gamma}_{01} = .086$, $t(80) = 2.512$, $p < .05$. In addition, holding other factors constant, the average self-efficacy of students was positively related to class mean performance-avoidance goal; $\hat{\gamma}_{02} = .093$, $t(80) = 2.789$, $p < .01$. The estimated between-

TABLE 3
Final Fitted Composite Model of Performance-Avoidance Goal Orientation

<i>Fixed Effect</i>	<i>Coefficient</i>	<i>SE</i>	<i>t-value</i>
Class PAVG mean, β_{0j}			
Base, γ_{00}	-.010	.031	-.317
GNDR, γ_{01}	.086	.034	2.512*
CEFC, γ_{02}	.093	.033	2.789**
SEFC slope, β_{1j}			
Base, γ_{10}	.196	.029	6.877***
GNDR, γ_{11}	.083	.065	1.293
TRAD, γ_{12}	.054	.027	2.031*
TEXP, γ_{13}	-.001	.023	-.049
GNDR \times TEXP, γ_{14}	-.135	.054	-2.529*
SPAE slope, β_{2j}			
Base, γ_{20}	.071	.029	2.394*
<i>Random Effect</i>	<i>Variance Component</i>	<i>df</i>	χ^2
PAVG mean, u_{0j}	.0369	80	146.244***
SEFC slope, u_{1j}	.0165	78	109.237*
Level-1 effect, r_{ij}	.8850		

Note. PAVG = performance-avoidance goal. GNDR = class gender (1 = female and -1 = male). CEFC = class average self-efficacy. SEFC = student self-efficacy. TRAD = teacher's frequent use of traditional assessments. TEXP = teacher's teaching experience. SPAE = student perceived public assessment environment.

* $p < .05$; ** $p < .01$; *** $p < .001$.

class variance in average performance-avoidance goal was reduced from .0569 in the random coefficient regression model to .0369, suggesting that about 35% of the variance among classrooms in average performance-avoidance goal was explained by class gender and class average self-efficacy.

According to Table 3, holding other factors constant, on average, student self-efficacy was positively related to performance-avoidance goal within classrooms; $\hat{\gamma}_{10} = .196$, $t(78) = 6.877$, $p < .001$. In addition, holding other factors constant, classes with a high frequent use of traditional assessments had a stronger positive effect of student self-efficacy on performance-avoidance goal than did classes with a low frequent use of traditional assessments; $\hat{\gamma}_{12} = .054$, $t(78) = 2.031$, $p < .05$.

Further, holding other factors constant, the differentiating effect of self-efficacy within a classroom depended jointly on class gender and teacher's teaching experience; $\hat{\gamma}_{14} = -.135$, $t(78) = -2.529$, $p < .05$. In female classrooms, classes having a highly experienced teacher were less differentiating with regard to student self-efficacy than were classes having a less experienced teacher; $\hat{\gamma}_{13} + (GNDR)_j \hat{\gamma}_{14} = -.001 + (1)(-.135) = -.136$; holding other factors constant. In male classrooms, the opposite was true. Using the random-coefficient regression model of performance-avoidance goal as the base model, about 23% of the variance in the self-efficacy differentiating effect was accounted for, once class gender, frequent use of traditional assessments, teaching experience, and interaction of class gender-by-teaching experience were taken into account. Finally, although the student perceived public assessment environment was on average positively related to performance-avoidance goal, $\hat{\gamma}_{20} = .071$, $t(1627) = 2.394$, $p < .05$, this relationship was invariant across classrooms.

DISCUSSION

Summary of Findings

Classroom assessment is a continual activity for teachers to improve the quality of instruction and motivate students to learn (Gronlund, 2006). Although there is a great deal of research on teachers' assessment practices, few empirical research attempts have been made to link assessment practices to achievement goals. When this issue was considered, little attention was devoted to the hierarchical structure of the data (e.g., Wang, 2004). This study attempted to fill this gap by investigating the possible effects of teachers' assessment practices on ninth-grade students' achievement goals in Muscat science classrooms in Oman.

Of primary importance in this study were the final hierarchical linear models that explained the variation in each achievement goal. The findings suggested that student characteristics such as self-efficacy and a perceived harsh assessment

environment as well as class characteristics such as class gender, class average self-efficacy, and class average perceived learning assessment environment provided the best model for explaining differences in mastery goals. In addition, the findings indicated that student characteristics such as self-efficacy and perceived public assessment environment as well as class characteristics such as class gender, class average self-efficacy, class average perceived public assessment environment, teaching experience, and frequent use of alternative assessments provided the best model for explaining differences in performance-approach goals. Furthermore, the findings revealed that student characteristics such as self-efficacy and perceived public assessment environment as well as class characteristics such as class gender, class average self-efficacy, teaching experience, and frequent use of traditional assessments provided the best model for explaining differences in performance-avoidance goals. Taking together, the present study's results found in Oman support many perspectives reported in the United States regarding classroom assessment and student achievement motivation (e.g., Ames, 1992; Brookhart, 1997). This implies that the general principles of classroom assessment and achievement goal theory tend to apply equally in both cultures.

Self-efficacy and Achievement Goals

Given that both mastery and performance-approach goals represent forms of approach achievement motivation (Elliot, 1999), it was not surprising that self-efficacy in this study was positively related with these forms of goals both at the student-level and class-level. What was surprising in this study, however, was the positive relationship between self-efficacy and performance-avoidance goal at both levels of the hierarchy: student and class. Many studies reported negative relationships between these two constructs (e.g., Middleton & Midgley, 1997). Given that performance-avoidance goal is a form of avoidance motivation and that low academic competence perceptions lead to avoidance motivation (Elliot, 1999), why then high efficacious students in this study strive to avoid the demonstration of lack of academic competence? The findings showed that teacher's frequent use of traditional assessments tended to moderate the effect of student self-efficacy on performance-avoidance goal. Specifically, there was a trend in this study for classes with a high frequent use of traditional assessments to have a stronger positive effect of student self-efficacy on performance-avoidance goal than did classes with a low frequent use of traditional assessments. Theoretically, traditional assessments featuring objective close-ended test items fit closely with the associationist learning theories where motivation is extrinsic (Shepard, 2000). As such, it may not be surprising that traditional assessments tend to convey to students that effort in class work should be in response to avoid the demonstration of incompetence to others. Therefore, results of this study suggest that educators may need to become aware that students with high levels of

efficacy may be vulnerable to the negative consequences (e.g., reduced intrinsic motivation) of pursuing performance-avoidance goals in classes where traditional assessments are more prevalent.

Gender Differences on Achievement Goals

Previous research findings have indicated that male students generally hold higher levels of performance goals and lower levels of mastery goals than female students (e.g., Middleton & Midgley, 1997) and that teacher's teaching experience is negatively related to student's adoption of goals that stress avoiding engagement in academic skills (e.g., Herman, 2001). While these previous research findings were based on single-level analyses that devote little attention to the nested nature of the data, the present study offered extensions to the previous research by using hierarchical linear modeling analyses. First, the present study revealed a similar pattern for gender differences in mastery goals, but at the classroom-level, in that female classrooms tended to have higher levels of mastery goals than male classrooms. The present study, however, differed from previous studies regarding gender differences in performance goals in that female classrooms tended to be more performance-oriented than male classrooms. The difference between this study and previous studies (e.g., Middleton & Midgley, 1997) was that ninth-grade male and female students in Omani public schools are segregated in that male teachers only teach in male students' schools and female teachers only teach in female students' schools and, as such, gender in this study was a class-level variable, whereas in the previous studies it was a student-level variable. Hence, these findings tend to suggest that when classes are segregated by gender, females are likely to not only seek achieving mastery but also to focus on competitive success. Classroom observations and interviews might shed more light on gender differences in achievement goals between segregated and desegregated classrooms.

Traditional and Alternative Assessments

The current study findings reveal that there was a trend for teacher's frequent use of alternative assessments to be correlated negatively with performance-approach goals, suggesting that a strong emphasis on alternative assessments may be less likely to orient students toward the adoption of performance-approach goals. In light of classroom assessment literature, alternative assessments tend to be more authentic, engaging, challenging, and emphasizing deep rather than superficial learning (Darling-Hammond, Ancess, & Falk, 1995; Hargreaves, Early, & Schmidt, 2002). In light of achievement goal research, adoption of performance-approach goals might lead to maladaptive patterns of achievement-related behaviors (Linnenbrink, 2005). Therefore, armed with the classroom assessment

literature regarding the advantages of alternative assessments as well as with the achievement goal research regarding the potential negative consequences of adopting performance-approach goals, the present study findings tend to support the movement toward the use of more alternative assessments.

To perplex matters more, the differentiating effect of self-efficacy on performance-approach goal within a classroom was found in this study to be dependent jointly on class gender, teaching experience, and frequent use of alternative assessments. For example, in female classrooms where alternative assessments were more prevalent, a highly experienced teacher tended to narrow the gap in performance-approach goals between high and low efficacy students. The opposite was found true in male classrooms where alternative assessments were more prevalent. There was evidence in the literature based on studies of single-level analyses to suggest that teaching experience might affect student motivation-related outcomes (Herman, 2001). Accordingly, the present multilevel study findings suggest, at least for the sampled female classrooms, that the most experienced teachers not only might have learned through experience how to enhance student motivation to learn but also might have observed through experience the hypothesized positive effects of using alternative assessments on student motivation to learn (Darling-Hammond et al., 1995; Hargreaves et al., 2002; Shepard, 2000), and hence they use them more frequently than the least experienced teachers. Future research may need to be conducted to further examine the role of teaching experience and frequent use of certain assessment practices in student motivation to learn.

Assessment Environment

According to Brookhart's (1997) model, students' perceptions of the classroom assessment environment influence their achievement motivational beliefs. The current study findings show that the student perceived public-oriented assessment environment was positively related to both types of performance goals. These findings are consistent with the findings from previous research studies investigating relationships between the classroom environment and achievement goals (e.g., Church et al., 2001; Wang, 2004). Ames (1992) stated that in classrooms characterized by public evaluation and recognition practices, "students become focused on their ability and the distribution of ability in the classroom" (p. 264), which in turn may orient them toward the adoption of performance goals. As mentioned previously, the adoption of performance goals has been associated with maladaptive patterns of achievement-related behaviors (Linnenbrink, 2005). Therefore, results of the present study suggest that educators may need to be aware of the detrimental effects of classroom assessment that emphasize the importance of grades rather than learning and focusing on public rather than private evaluation and recognition practices in student achievement motivation.

Moreover, the present study findings suggest that the collective (i.e., aggregate) perceived assessment environment is operative in the assessment settings and, as such, it may be important to begin attending to this construct in future empirical work on classroom assessment. From a sociocognitive perspective, students “are not social isolates of the influence of those around them” (Bandura, 1997, p. 469). In this particular regard, the findings showed that the shared perceptions of class members about the assessment environment might influence students’ adoption of achievement goals in ways that are consistent with the class shared perception. For example, as shown in this study, when the class members collectively perceive their assessment environment as being public-oriented, an individual student within that class is likely to pursue performance-approach goals by trying to demonstrate his or her competence to others, such as approaching a class assignment with a goal of doing better than others rather than learning as much as possible from the assignment. Practically, these findings imply that group experience could be instrumental in maintaining a positive assessment climate in the class, which in turn is conducive for desirable patterns of student achievement motivation. Stiggins (2005) stated that “the greatest value of classroom assessment is realized when we open the process up and welcome students in as full partners” (p. 29). Therefore, one practical suggestion from the present study is that teachers might conduct informal interviews throughout the semester with the students, one-on-one or in groups, to discuss how well the classroom assessment process is going for the students as individuals and as a group (Stiggins, 2005). This kind of discussion might help teachers make classroom assessments that increase students’ motivation to learn.

Recommended Assessment Practices

In this study, there was no evidence of teachers’ frequent use of recommended assessment practice effects on achievement goals. This could be attributed to the assumption underlying the current investigation. Specifically, the present study was based on the assumption that aspects of classroom assessment are interdependent, operating in an additive manner and, as such, the study followed an integrative approach (Ames, 1992) in the investigation of the effects of the recommended assessment practices on students’ achievement goals. In other words, one limitation of this study was that it construed teachers’ frequent use of the recommended assessments as an omnibus combination of several factors including revision of assessment, communication of assessment, student-involved assessment, and nonachievement-based grading. This might have made it difficult to know which aspect of the recommended assessment practices could have effects on students’ achievement goals. Future research may need to independently consider each aspect of the recommended assessment practices.

Limitations of the Study Findings

The present study findings might have been affected by the low reliability estimates from the self-report questionnaires for a number of reasons. First, the questionnaire items were based on theory and research developed in the United States and published in English-language journals and manuals. Even though the translation of the items was verified by bilingual professors and the translated items were subjected to a content validation process, some of the items and/or instructions might not have been clear, causing participants to misinterpret the items and respond on some basis unrelated to the content being measured, thereby lowering the internal consistency of the responses. Second, the use of fewer response options (i.e., four points) in the students' questionnaires might have reduced the variance and as a result might have contributed to the lower reliability estimates. Third, the present study sampled a much narrower population of teachers than past studies (e.g., Alsarimi, 2000) and, as such, a restriction of the scores' range might have occurred and reduced the reliability estimates. Future research may need to use multiple data collection methods including teachers' lesson and assessment plans, classroom observations, and interviews with students and teachers to validate the self-report questionnaires. In addition, the questionnaires may need to be administered to a more representative sample selected from different geographic regions across the country. Future research should also be conducted to testify the findings from this study in various subject areas and grade levels.

Summary of the Implications of the Findings for Classroom Practice

The present study attempted to link teachers' assessment practices to students' achievement goals. The findings revealed a number of implications for classroom practice. First, classes with a high emphasis on traditional assessments featuring objective close-ended test items might encourage students to pursue performance-avoidance goals. Second, classes with a high emphasis on alternative assessments featuring performance-based assessments tend to discourage students from pursuing performance-approach goals. Third, classroom assessment emphasizing the importance of grades rather than learning and focusing on public rather than private evaluation practices might orient students toward the adoption of performance goals. Fourth, the collective perception of the class members about their classroom assessment environment as being public-oriented tend to orient an individual student within that class toward the adoption of performance-approach goals. Hence, the collective experience of the class members about the classroom assessment climate could be instrumental in having desirable patterns of student achievement motivation.

REFERENCES

- Alkharusi, H. (2007). *Effects of Teachers' Assessment Practices on Ninth Grade Students' Perceptions of Classroom Assessment Environment and Achievement Goal Orientations in Muscat Science Classrooms in the Sultanate of Oman*. Unpublished dissertation, Kent State University, Kent, Ohio.
- Alsarimi, A. M. (2000). *Classroom assessment and grading practices in the Sultanate of Oman*. Unpublished dissertation, University of Pittsburgh, Pennsylvania.
- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*, 261–271.
- 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. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education, 7*, 279–301.
- 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. (2004). Classroom assessment: Tensions and intersections in theory and practice. *Teachers College Record, 106*, 429–458.
- 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.
- 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.
- Coleman, J. S. (1987). Norms as social capital. In G. Radnitzky & P. Bernholz (Eds.), *Economic imperialism: The economic approach applied outside the field of economics*. New York: Paragon House Publishers.
- Coleman, J. S. (1990). *Foundations of social theory*. Cambridge, MA: Harvard University Press.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Belmont, CA: Wadsworth Group/Thomson Learning.
- Darling-Hammond, L., Ancess, J., & Falk, B. (1995). *Authentic assessment in action: Studies of schools and students at work*. New York: Teachers College Press.
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist, 41*, 1040–1048.
- Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist, 34*, 169–189.
- Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 72*, 218–232.
- Goddard, R. D., & Goddard, Y. L. (2001). A multilevel analysis of the relationship between teacher and collective efficacy in urban schools. *Teaching and Teacher Education, 17*, 807–818.
- 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.
- Gronlund, N. E. (2006). *Assessment of student achievement* (8th ed.). Boston: Pearson.

- Hargreaves, A., Early, L., & Schmidt, M. (2002). Perspectives on alternative assessment reform. *American Educational Research Journal*, 39, 69–95.
- Harlen, W., & Crick, R. D. (2003). Testing and motivation for learning. *Assessment in Education: Principles, Policy & Practice*, 10, 169–207.
- Herman, P. (2001). Teacher experience and teacher efficacy: Relations to motivation and achievement (Doctoral dissertation, University of North Carolina at Chapel Hill, 2001). *Dissertation Abstracts International*, 61, 2592.
- Linnenbrink, E. A. (2005). The dilemma of performance-approach goals: The use of multiple goal contexts to promote students' motivation and learning. *Journal of Educational Psychology*, 97, 197–213.
- 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. (2003, October). *Preservice versus inservice teachers' assessment literacy: Does classroom experience make a difference?* Paper presented at the meeting of the Mid-Western Educational Research Association, Columbus, OH.
- Middleton, M., & Midgley, C. (1997). Avoiding the demonstration of lack of ability: An under-explored aspect of goal theory. *Journal of Educational Psychology*, 89, 710–718.
- Midgley, C., Maehr, M. L., Huda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., et al. (2000). *Manual for the patterns of adaptive learning scales*. Retrieved August 5, 2005, from http://www.umich.edu/~pals/PALS%202000_V13Word97.pdf
- Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, 91, 328–346.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Roberts, J. K., Onwuegbuzie, A. J., & Eby, J. R. (2001, April). *The introduction of a measure of instrument homogeneity for interpreting low reliability coefficients*. Paper presented at the meeting of the American Educational Research Association, Seattle, WA.
- Rodriguez, M. C. (2004). The role of classroom assessment in student performance on TIMSS. *Applied Measurement in Education*, 17, 1–24.
- Schunk, D. H. (1996). Goal and self-evaluative influences during children's cognitive skill learning. *American Educational Research Journal*, 33, 359–382.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29, 4–14.
- Stefanou, C., & Parkes, J. (2003). Effects of classroom assessment on student motivation in fifth-grade science. *The Journal of Educational Research*, 96, 152–162.
- Stiggins, R. J. (2005). *Student-involved assessment FOR learning* (4th ed.). Columbus, OH: Merrill Prentice Hall.
- 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.
- Wang, X. (2004). *Chinese EFL students' perceptions of classroom assessment environment and their goal orientations in the college English course*. Unpublished master's thesis, Queen's University, Kingston, Ontario, Canada.
- Zhang, Z., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers' self-perceived assessment skills. *Applied Measurement in Education*, 16, 323–342.