Classroom assessment is a continual activity for teachers to improve the quality of instruction and motivate students to learn (Brookhart, 1999; Gronlund, 2006). Although there is a great deal of research on teachers’ classroom assessment practices, few empirical research attempts have been made to link these practices to students’ perceptions of classroom assessment environment and motivation defined in terms of achievement goal orientations. This study examined teachers’ assessment practices within the framework of classroom assessment literature and achievement goal theory. More specifically, the purposes of this study were to identify the underlying dimensions of students’ perceptions of classroom assessment environment and achievement goal orientations and to investigate the possible effects of certain student-level and class-level characteristics on perceived classroom assessment environment and achievement goal orientations.

The participants were 1,636 ninth grade students and their corresponding 83 science teachers enrolled in public schools within Muscat educational region in Oman.
during the spring semester 2007. Two questionnaires were developed and used, one for students and one for teachers. The student’s questionnaire focused on students’ perceived classroom assessment environment, achievement goal orientations, and academic self-efficacy. The teacher’s questionnaire focused on teachers’ frequent uses of traditional assessments, alternative assessments, and classroom assessment practices recommended by experts of educational measurement and assessment.

Principal components/exploratory factor analyses (PCA/EFA) were conducted to identify the underlying dimensions of students’ perceptions of classroom assessment environment and achievement goal orientations. Hierarchical linear modeling (HLM) analyses were employed to examine the effects of certain student-level and class-level characteristics on students’ perceptions of classroom assessment environment and achievement goal orientations. Results of the PCA/EFA revealed three dimensions of perceived classroom assessment environment: learning-, harsh-, and public-oriented assessment environments; and three dimensions of achievement goal orientations: mastery, performance-approach, and performance-avoidance goals. Results of the HLM showed that class contextual features and teacher’s teaching experience and assessment practices interacted significantly with student characteristics in influencing students’ perceptions of classroom assessment environment and achievement goal orientations. The findings were compared with findings from previous studies related to classroom assessment environment and achievement goal orientations. Recommendations, implications, and suggestions for future research were discussed.
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

A dissertation submitted to the
Kent State University College and Graduate School
of Education, Health, and Human Services
in partial fulfillment of the requirements
for the degree Doctor of Philosophy

by

Hussain A. Al Kharusi

August, 2007
A dissertation written by

Hussain A. Al Kharusi

B.E., Sultan Qaboos University, 1999

M.A., Kent State University, 2002

Ph.D., Kent State University, 2007

Approved by

_________________________  Co-director, Doctoral Dissertation Committee
Rafa Kasim

_________________________  Co-director, Doctoral Dissertation Committee
Tricia Niesz

_________________________  Member, Doctoral Dissertation Committee
Christopher Was

Accepted by

_________________________  Chairperson, Department of Educational Foundations and Special Services
Awilda Hamilton

_________________________  Dean, College and Graduate School of Education, Health, and Human Services
David England
ACKNOWLEDGMENTS

The accomplishment of this dissertation was a blessing from Allah who is blessings can never be counted. Thanks to Him from everlasting to everlasting.

I would like to especially thank my academic advisor and dissertation’s director, Dr. Rafa Kasim, for his insight and critique of my work. He is the one who supported and guided me through my entire doctoral program. I would also like to express my appreciation to Dr. Tricia Niesz and Dr. Christopher Was for their valuable time contributed to the dissertation. Having them in my dissertation committee helped me learn a lot from them. I would also like to thank Dr. Melody Tankersley for agreeing to serve on the committee in the role of graduate faculty representative. My thanks also go to my professors, fellow students, and friends at Kent State University who have assisted and motivated me throughout my coursework.

I would also like to thank Prof. Abdulgawi Al-Zubaidi, Dr. Ali Kazem, Dr. Amean Sulaiman, Dr. Said Al-Dhafri, and Dr. Skreen Al-Mashadani from Department of Psychology at Sultan Qaboos University in the Sultanate of Oman for their helpful comments and suggestion in reviewing the questionnaires used in the dissertation. Thanks to them and to all members of the Department of Psychology at Sultan Qaboos University for their support and encouragement. I extend my thanks to my country Sultanate of Oman’s government represented by Sultan Qaboos University for granting me a full scholarship to get this degree of doctor of philosophy in the field of Evaluation and Measurement from Kent State University.
I would also like to express my appreciation to the Ministry of Education in the Sultanate of Oman for granting me permission to collect data from Muscat educational region’s public schools. I am especially thankful to all principals, teachers, and students of the sampled schools for their cooperation during the data collection.

My great appreciation and thanks go to my mother and father, to my brothers and sisters, and to all members of my family for their endless love, support, and encouragement throughout my study in the United States.
TABLE OF CONTENTS

ACKNOWLEDGMENTS ........................................................................................................ iv

LIST OF TABLES .............................................................................................................. viii

CHAPTER

I INTRODUCTION ........................................................................................................ 1
  Background for the Study .................................................................................. 1
  Statement of the Problem ............................................................................ 17
  Rationale for the Study ................................................................................. 18
  Significance of the Study .......................................................................... 25
  Purposes of the Study .................................................................................... 27
  Research Questions ..................................................................................... 27
  Limitations of the Study ........................................................................... 28
  Definition of Terms ..................................................................................... 29
  Summary ...................................................................................................... 39

II REVIEW OF THE LITERATURE ........................................................................ 40
  Achievement Goal Orientations. ............................................................... 40
  Academic Self-Efficacy .............................................................................. 50
  Traditional and Alternative Assessments ................................................. 54
  Recommended Classroom Assessment Practices .................................... 60
  Classroom Assessment Environment ....................................................... 67
  Summary ...................................................................................................... 82
  Synthesis of the Literature and Research Questions .................................. 83

III METHODS ......................................................................................................... 103
  Participants ................................................................................................ 103
  Setting ....................................................................................................... 105
  Instrumentation .......................................................................................... 106
  Procedures ................................................................................................ 115
  Research Design ....................................................................................... 118
  Analysis ..................................................................................................... 118
  Summary .................................................................................................. 148
TABLE OF CONTENTS (Continued)

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV</td>
<td>RESULTS .......................................................... 149</td>
</tr>
<tr>
<td></td>
<td>Research Question One ........................................ 150</td>
</tr>
<tr>
<td></td>
<td>Research Question Two ......................................... 159</td>
</tr>
<tr>
<td></td>
<td>Research Questions Three, Four, Five, and Six .......... 167</td>
</tr>
<tr>
<td></td>
<td>Research Questions Seven, Eight, Nine, and Ten .......... 225</td>
</tr>
<tr>
<td>V</td>
<td>DISCUSSION ...................................................... 288</td>
</tr>
<tr>
<td></td>
<td>Classroom Assessment Environment .......................... 289</td>
</tr>
<tr>
<td></td>
<td>Achievement Goal Orientations .............................. 300</td>
</tr>
<tr>
<td></td>
<td>Conclusion ....................................................... 310</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>A STUDENT’S QUESTIONNAIRE ................................. 319</td>
</tr>
<tr>
<td></td>
<td>B TEACHER’S QUESTIONNAIRE .................................. 329</td>
</tr>
<tr>
<td></td>
<td>C OMANI MINISTRY OF EDUCATION’S PERMISSION LETTER FOR DATA COLLECTION .......................... 338</td>
</tr>
<tr>
<td></td>
<td>D CONSENT FORMS ................................................ 340</td>
</tr>
<tr>
<td></td>
<td>E COVER LETTERS ................................................ 347</td>
</tr>
<tr>
<td></td>
<td>REFERENCES ........................................................ 352</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Assigned Raw Scores to the Outlying Cases on the Offending Variables (Items) of Perceived Classroom Assessment Environment</td>
</tr>
<tr>
<td>2</td>
<td>Summary of Factor Loadings by Principal Components Analysis for the Orthogonal Three-Factor Model of Perceived Classroom Assessment Environment</td>
</tr>
<tr>
<td>3</td>
<td>Descriptive Statistics and Reliabilities for, and Zero-Order Correlations among the Three Scales of Perceived Classroom Assessment Environment</td>
</tr>
<tr>
<td>4</td>
<td>Assigned Raw Scores to the Outlying Cases on the Offending Variables (Items) of Achievement Goal Orientations</td>
</tr>
<tr>
<td>5</td>
<td>Summary of Factor Loadings by Principal Components Analysis for the Oblique Three-Factor Model of Achievement Goal Orientations</td>
</tr>
<tr>
<td>6</td>
<td>Descriptive Statistics and Reliabilities for, and Zero-Order Correlations among the Three Scales of Achievement Goal Orientations</td>
</tr>
<tr>
<td>7</td>
<td>Assigned Raw Scores to the Outlying Cases on the Offending Student-Level Variables</td>
</tr>
<tr>
<td>8</td>
<td>Descriptive Statistics of Student-Level Variables for Overall Sample of and for Male and Female Students</td>
</tr>
<tr>
<td>9</td>
<td>Zero-Order Correlations among Student-Level Variables</td>
</tr>
<tr>
<td>10</td>
<td>Assigned Raw Scores to the Outlying Cases on the Offending Class-Level Variables</td>
</tr>
<tr>
<td>11</td>
<td>Descriptive Statistics of Class-Level Variables for Overall Sample of and for Male and Female Classes</td>
</tr>
<tr>
<td>12</td>
<td>Zero-Order Correlations among Class-Level Variables</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
</tr>
<tr>
<td>13</td>
<td>Random-Coefficient Regression Model of Perceived Learning Classroom Assessment Environment</td>
</tr>
<tr>
<td>14</td>
<td>Reduced Contextual Effect Model of Academic Self-Efficacy on Perceived Learning Classroom Assessment Environment</td>
</tr>
<tr>
<td>15</td>
<td>Reduced Joint Effects Model of Class Gender and Class Averages for Perceived Harsh and Public Classroom Assessment Environments on Perceived Learning Classroom Assessment Environment</td>
</tr>
<tr>
<td>16</td>
<td>Reduced Joint Effects Model of Class Gender and Teacher’s Assessment Practices on Perceived Learning Classroom Assessment Environment</td>
</tr>
<tr>
<td>17</td>
<td>Final Fitted Composite Model of Perceived Learning Classroom Assessment Environment with Heterogeneous Level-1 Variance</td>
</tr>
<tr>
<td>18</td>
<td>Random-Coefficient Regression Model of Perceived Harsh Classroom Assessment Environment</td>
</tr>
<tr>
<td>19</td>
<td>Reduced Contextual Effect Model of Academic Self-Efficacy on Perceived Harsh Classroom Assessment Environment</td>
</tr>
<tr>
<td>20</td>
<td>Reduced Joint Effects Model of Class Averages for Perceived Learning and Public Classroom Assessment Environments on Perceived Harsh Classroom Assessment Environment</td>
</tr>
<tr>
<td>21</td>
<td>Reduced Joint Effects Model of Class Gender, Teaching Experience, and Teacher’s Assessment Practices on Perceived Harsh Classroom Assessment Environment</td>
</tr>
<tr>
<td>22</td>
<td>Final Fitted Composite Model of Perceived Harsh Classroom Assessment Environment</td>
</tr>
<tr>
<td>23</td>
<td>Final Fitted Composite Model of Perceived Public Classroom Assessment Environment</td>
</tr>
</tbody>
</table>
### LIST OF TABLES (Continued)

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>Reduced Random-Coefficient Regression Model of Mastery Goal Orientation.</td>
</tr>
<tr>
<td>25</td>
<td>Reduced Contextual Effects Model of Academic Self-Efficacy and Perceived Classroom Assessment Environment along with the Differential Contextual Effects by Class Gender on Mastery Goal Orientation</td>
</tr>
<tr>
<td>26</td>
<td>Final Fitted Composite Model of Mastery Goal Orientation with Heterogeneous Level-1 Variance</td>
</tr>
<tr>
<td>27</td>
<td>Reduced Random-Coefficient Regression Model of Performance-Approach Goal Orientation</td>
</tr>
<tr>
<td>28</td>
<td>Reduced Contextual Effects Model of Academic Self-Efficacy and Perceived Classroom Assessment Environment on Performance-Approach Goal Orientation</td>
</tr>
<tr>
<td>29</td>
<td>Reduced Joint Effects Model of Class Gender, Teaching Experience, and Teacher’s Assessment Practices on Performance-Approach Goal Orientation</td>
</tr>
<tr>
<td>30</td>
<td>Final Fitted Composite Model of Performance-Approach Goal Orientation with Heterogeneous Level-1 Variance</td>
</tr>
<tr>
<td>31</td>
<td>Reduced Random-Coefficient Regression Model of Performance-Avoidance Goal Orientation</td>
</tr>
<tr>
<td>32</td>
<td>Reduced Contextual Effect Model of Academic Self-Efficacy on Performance-Avoidance Goal Orientation</td>
</tr>
<tr>
<td>33</td>
<td>Reduced Joint Effects Model of Class Gender and Teacher’s Assessment Practices on Performance-Avoidance Goal Orientation</td>
</tr>
<tr>
<td>34</td>
<td>Final Fitted Composite Model of Performance-Avoidance Goal Orientation</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

This dissertation is organized and presented in five chapters. The first chapter serves as an introduction to the study. It provides a background for and statement of the research problem, a rationale for conducting the study, significance of the study, purposes of the study, research questions, limitations of the study, and definitions of terms used in the study. The ideas presented in the first chapter are further elaborated in the second chapter, where a detailed review of the literature is discussed. The third chapter includes the research methods that were followed to conduct the study. The fourth chapter presents results of each research question. The fifth chapter includes a discussion of the results along with their implications for theory, research, and practice.

Background for the Study

Classroom assessment involves a wide range of activities from designing assessment tasks to grading, communicating assessment results, and using them in decision-making (Zhang & Burry-Stock, 2003). Most of these activities are created by the classroom teacher to assess and motivate students to learn (Brookhart, 1997a, 1999; Gronlund, 1998, 2006; Stiggins, 1999). The overall sense or meaning that students make out of the various aspects of classroom assessment activities comprises the classroom
assessment environment (Brookhart & DeVoge, 1999). Since a substantial proportion of
classroom time is devoted to the assessment for and of student learning (Mertler, 2003)
and that students’ perceptions of classroom activities and achievement goals, their
purposes of task engagement (Maehr, 1989), play a critical role in the learning process
(Ames, 1992b; Brookhart, 1997a; Brophy, 1999), it seems reasonable to argue that
careful consideration of the impact of teachers’ classroom assessment practices on
students’ perceptions of the classroom assessment environment and achievement goal
orientations is certainly warranted.

Achievement Goal Orientations

Achievement goal theorists have traditionally identified two types of achievement
goal orientations (also called a dichotomous framework of achievement goals): mastery
goals and performance goals (Ames, 1992b; Dweck, 1986; Nicholls, 1984). Mastery
goals center on the development of competence, whereas performance goals center on the
outward showing of competence (Ames, 1992b; 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, 1992b; 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,
1992b; Dweck, 1986; Nicholls, 1984). Hence, students’ achievement goal orientations
should deserve recognition and investigation as a valuable achievement-related outcome
to be promoted in the classroom.
Since its origin in the late 1970s and early 1980s (Kaplan, Middleton, Urdan, & Midgley, 2002b), achievement goal theory has undergone a number of theoretical advances, which are discussed more fully in Chapter II of the dissertation. For example, Elliot and his colleagues have developed a trichotomous framework of achievement goals that further differentiates performance goals into performance-approach and performance-avoidance goals (Elliot & Church, 1997; Elliot & Harackiewicz, 1996). In particular, in this trichotomous framework of achievement goal theory, 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; Elliot & Harackiewicz, 1996). This framework has been validated with college and middle school students.

Although factor analytic procedures from studies conducted with college students yielded the three anticipated achievement goals (mastery, performance-approach, and performance-avoidance) with a low correlation between the two types of performance goals (Elliot & Chuch, 1997), results from studies conducted with middle school students tended to reveal some overlap between these two goals. For example, in a study of 703 sixth grade students, Middleton and Midgley (1997) found that the scales measuring performance-approach and performance-avoidance goals were moderately correlated ($r = .56$), suggesting that these two goals may not be descriptively different. These somewhat contradictory findings might be due to the lack of consistency in the measurement of achievement goal orientations, which will further be elaborated in Chapter II of the
dissertation. The present study sought to contribute to the clarification of the relationships among achievement goals by identifying the underlying dimensions of achievement goal orientations for ninth grade students in Muscat science classrooms in the Sultanate of Oman, hereinafter referred to as Oman, utilizing measures that have been validated for use with middle school students. The study adopted the trichotomous framework of achievement goal theory because it was assumed to be the most prevalent goal framework in achievement settings (Elliot, 1999; Elliot & Church, 1997; Elliot & Thrash, 2001). Thus, in accord with the trichotomous framework of achievement goal theory, it was expected that factor analytic procedures in this study would yield the three dimensions of achievement goals (mastery, performance-approach, and performance-avoidance) with (a) a positive correlation between the two types of performance goals and (b) no correlations between mastery goals and each of the two types of performance goals.

Classroom Assessment Environment

An emergent movement in classroom assessment research has been toward the role of classroom assessment on student motivation. Brookhart (2004, p. 444) contends that each classroom has “an assessment ‘character’ or environment” that largely springs from the teacher’s classroom assessment practices. As such, Brookhart (1997a) has developed a theoretical model based on a synthesis of classroom assessment literature and social cognitive theories of learning and motivation. In this model, classroom assessment environment is construed 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 (Brookhart, 1997a).
Brookhart (1997a) has postulated that students’ perceptions of the classroom assessment
environment influence students’ motivational beliefs and achievement-related outcomes.
This suggests that students’ perceptions of their classroom assessment environment
should deserve recognition and investigation as a valuable instructional factor in the
learning process.

Brookhart’s (1997a) model has been examined for elementary, middle, and high
school students as well as for college level students. For example, Brookhart and DeVoge
(1999) designed a study to investigate part of the Brookhart’s (1997a) theoretical model
of classroom assessment and student motivation. They collected data from four classroom
assessment events in each of two third-grade language arts classes using survey,
observation, and interview techniques. Findings revealed positive relationships among
students’ perceptions of the classroom assessment tasks (as one aspect of the classroom
assessment environment) and their ability to do the tasks (known as self-efficacy).
Students expressed importance of the task in accordance with their goal orientations.
Specifically, some students expressed the importance of the task in terms of its value for
learning, which is consistent with the mastery goal orientation. Others expressed it in
terms of getting good grades, which is consistent with the performance goal orientation.
At the conclusion of the study, Brookhart and DeVoge (1999) suggested that student’s
goal orientations should be considered when investigating the impact of classroom
assessment environment on student motivation.
Although Brookhart and DeVoge’s (1999) mixed methods research design had the strength of convergence and corroboration of findings (Johnson & Onwuegbuzie, 2004), the construct of perceived classroom assessment environment in their study was not clearly defined in operational terms thereby making it difficult to replicate their study findings. Therefore, the present study sought to identify the underlying dimensions of perceived classroom assessment environment for ninth grade students in Muscat science classrooms in Oman utilizing factor analytic techniques. In light of achievement goal theory (Ames, 1992a, 1992b) and classroom assessment literature (McMillan & Workman, 1998) that will both be discussed more fully in Chapter II of the dissertation, it was expected that factor analytic procedures in this study would yield two independent dimensions of perceived classroom assessment environment (a) a learning-oriented and (b) a normative-oriented classroom assessment environment. Further, in light of achievement goal theory (Ames, 1992a, 1992b) and classroom assessment literature (McMillan & Workman, 1998) and consistent with Brookhart and DeVoge’s (1999) study findings, it was expected in this study that student’s perceived learning-oriented classroom assessment environment would on average be positively related to adoptions of mastery goals and performance-approach goals, but negatively related to the adoption of performance-avoidance goals within classes. In addition, it was expected in this study that student’s perceived normative-oriented classroom assessment environment would on average be strongly related to adoptions of performance-approach goals and performance-avoidance goals, but weakly related to the adoption of mastery goals within classes.
In response to Brookhart and DeVoge’s (1999) suggestion with regard to the consideration of student’s goal orientations when examining the impact of classroom assessment environment on student motivation, Brookhart and Bronowicz (2003) conducted a multiple case study in which a total of 161 elementary, middle, and high school students were interviewed about their perceptions of classroom assessment environment, self-efficacy, and achievement goal orientations. Findings indicated that students’ perceptions of the classroom assessment tasks revolved around their self-efficacy for accomplishing the tasks and the goal orientations behind their efforts to do the tasks. Although Brookhart and Bronowicz (2003) concluded that students’ perceptions tended to hold across the participating classrooms, the lack of objective measurement of the variables considered in their study and the lack of focus on a specific instructional setting by combining participants from various subject areas and grade levels make it difficult to draw any meaningful generalizations. The subject area and grade level in Brookhart and Bronowicz’s (2003) study might be confounding variables that need to be held constant across the observational units. Therefore, the present study attempted to overcome these issues by exploring the possible effects of classroom assessment environment as perceived only by ninth grade students in Muscat science classrooms in Oman on achievement goal orientations utilizing objective measures.

Although the focus was at the college level, Wang (2004) explored the underlying dimensions of classroom assessment environment and achievement goal orientations for 503 first year non-English major undergraduates from one university in China. Three characteristics of classroom assessment environment and four types of goal orientations
were identified for the participating students. Learning-oriented classroom assessment environment positively predicted adoptions of mastery in-class goals, mastery-external goals, and performance-approach goals. Classroom assessment environment that is test-oriented positively predicted adoption of performance-avoidance goals and negatively predicted adoption of mastery-external goals. Praise-oriented classroom assessment environment positively predicted adoption of performance-approach goals. Student’s gender mediated the relationships of perceived classroom assessment environment to performance-approach and mastery-in class goals. In contrast to Wang’s (2004) study that focused on college level students, the current study aimed at identifying the underlying dimensions of classroom assessment environment as perceived by ninth grade students in Muscat science classrooms in Oman utilizing factor analytic techniques.

Despite the hierarchical structure of Wang’s (2004) study data, in that, students were nested within classrooms, the results were based on single-level linear-model analyses. As might be expected, ignoring the nested nature of the data may lead to invalid inferences about the relationship between perceived classroom assessment environment and adoption of achievement goals. Therefore, the present study attempted to resolve this issue by applying a hierarchical linear model (HLM) analysis, which not only facilitates a decomposition of the relationship between the variables into separate student-level and class-level components, but also recognizes the dependence among the outcomes of students within the same classroom (Raudenbush & Bryk, 2002). This dependence may arise as a result of shared class experiences with regard to the teacher’s characteristics and his or her assessment practices.
Brookhart and Bronowicz (2003, p. 222) stated that “classroom assessments are usually group experiences, dependent upon their context. Assessments differ from class to class…because of different instructional experiences.” Thus, one could argue that failure to consider such variation between classes may lead to invalid inferences about the effects of classroom assessment practices on students’ perceptions of classroom assessment environment and achievement goal orientations. Accordingly, although Wang’s (2004) study pointed to the conclusion that teachers’ classroom assessment practices may affect the way students perceive and approach assessment tasks, its generalizability was limited by ignoring the classroom as a key sociological unit of analysis (Black & Wiliam, 1998a, 1998b). The present study attempted to consider the group nature of the classroom assessment environment by not only taking the nested structure of the data (students nested within classrooms) into account, but also estimating the relations that might cross the class-level and the student-level characteristics.

*Academic Self-Efficacy*

As might be noted, findings from the aforementioned studies conducted by Brookhart and her colleagues (Brookhart & Bronowicz, 2003; Brookhart & DeVoge, 1999) suggested that students’ academic self-efficacy might need to be considered when investigating the impact of classroom assessment practices on students’ achievement goal orientations. Self-efficacy pertains to student’s judgments of his or her performance capability on a particular type of assessment tasks (Bandura, 1986). According to Bandura’s (1986) 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 goal orientation (Schunk, 1991, 1996). As such, one could argue that there may be a cross-level interaction effect for teacher’s use of a particular assessment type and student’s self-efficacy on student’s achievement goal orientation. The present study attempted to shed light on the relationships of students’ academic self-efficacy to perceived classroom assessment environment and achievement goal orientations as a function of teachers’ classroom assessment practices in ninth grade Muscat science classrooms in Oman.

Gender

As might also be noted, findings from the aforementioned Wang’s (2004) study suggested that students’ gender might need to be considered when investigating the impact of teacher’s classroom assessment practices on students’ perceptions of their classroom assessment environment and achievement goal orientations. Specifically, in Wang’s (2004) 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. Also, mastery in-class 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. It should be noted, however, that gender in this study varied across classes because in
Oman, students within the same class and their science teacher are of the same gender, either all of them are males or all of them are females. Therefore, unlike previous studies of classroom assessment environment and achievement goals, gender in this study was more appropriately to be treated as an independent variable at the class-level.

The effect of gender on perceived classroom assessment environment and achievement goal orientations has been detected in several studies of middle and high school students. Specifically, middle school girls tend to report more positive perceptions of their classroom environment than boys (Anderman & Midgley, 1997; Meece, Herman, & McCombs, 2003). Likewise, female adolescents tend to be more oriented to mastery goals and less oriented to performance goals than male adolescents (Ablard & Lipschultz, 1998; Anderman & Midgley, 1997; Meece et al., 2003; Pajares, Britner, & Valiante, 2000). Having concluded from these studies that perceived classroom assessment environment and achievement goal orientations differ as function of gender, it seemed necessary to account for these differences.

As noted above, gender in this study was an independent variable at the class-level. Therefore, hierarchical linear modeling (HLM), which was the appropriate analytic technique for this study to handle the nested nature of the data in that students were nested within classrooms, provides two ways for accounting the gender effects (a) controlling for the student-level differences and (b) controlling for the class-level differences (Raudenbush & Bryk, 1986). Among middle school students, research has shown that girls tend to hold stronger science self-efficacy beliefs than boys (Britner & Pajares, 2001, 2006; Pajares et al., 2000). Thus, the HLM approach to accounting for the
gender effects would be applied in this study by assuming that student’s academic 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 academic self-efficacy) and its interaction with class’s gender are confounding variables in the between-class model. The hypothesized directional nature of these contextual effects along with their theoretical and empirical support will further be discussed at the end of Chapter II of the dissertation.

**Traditional and Alternative Assessments**

According to Brookhart’s (1997a) theoretical model, teacher’s use of different forms of assessment is one component of the classroom assessment environment. There are two distinct forms of classroom assessment methods: (a) traditional assessments that feature close-ended test items and (b) alternative assessments that feature performance-based assessments (Bol, Stephenson, O’Connell, & Nunnery, 1998). Many educators have promoted a move away from traditional to alternative forms of assessment (e.g., Darling-Hammond, 1994; Shepard, 2000). The arguments in favor of alternative assessments over traditional assessments rest on motivational aspects of the alternative assessments for being more authentic and challenging (McMillan & Workman, 1998), which in turn may orient students to mastery-related outcome beliefs (Ames, 1992b). However, these arguments were based on theory and little empirical research existed about the effects of assessment type on students’ perceptions of the classroom assessment environment and achievement goal orientations. The few studies that did address these issues empirically tended to focus on elementary school level.
For example, Stefanou and Parkes (2003) studied the effects of classroom assessment type on student motivational goals. Seventy nine students in three fifth grade science classes, taught by the same teacher, were exposed to three different classroom assessment conditions: (a) traditional paper-pencil tests, (b) a laboratory assessment task format, and (c) a performance assessment. Results indicated a statistically significant effect for assessment type on goal orientation. Traditional paper-pencil tests and performance assessment fostered more a task goal orientation than the laboratory assessment. Interviews with students revealed that traditional paper-pencil tests seemed to foster more a performance-approach goal orientation. Unlike Stefanou and Parkes’ (2003) study that focused on the fifth grade science classrooms, the current study focused on the ninth grade students in Muscat science classrooms in Oman. Also, although Stefanou and Parkes’ (2003) experimental research design had the strength of having the same teacher across the groups, the findings had a limited ecological validity because the study was not part of the normal classroom work (Black & Wiliam, 1998a; Crooks, 1988). Therefore, the present study examined the effects of teachers’ frequent uses of traditional and alternative assessments on students’ perceptions of classroom assessment environment and achievement goals using a survey research method within a correlational research design.

In light of classroom assessment literature and achievement goal theory (Ames, 1992a, 1992b; McMillan & Workman, 1998), it was expected in this study that classes with a high emphasis on traditional assessments would be associated with performance-related outcomes more than mastery-related outcomes. The opposite was expected to be
true for classes with a high emphasis on alternative assessments. In addition, since laboratory assessment tasks are one kind of alternative assessments, their effects on student motivational goals were not clear in Stefanou and Parkes’ (2003) study, thereby casting doubt about the effect of alternative assessments on achievement goals. Therefore, the current study sought to clarify this relationship by including performance assessments, both structured and unstructured, within alternative assessments used by the teacher following Gronlund’s (1998) definitions of performance and alternative assessments that will be provided later in this Chapter.

Moreover, survey research studies investigating teachers’ frequent uses of traditional and alternative assessments have suggested the need to consider teacher’s gender and years of teaching experience. With regard to the frequent use of traditional assessments, no statistically significant differences were found between teachers based on their gender and years of teaching experience (Alsarimi, 2000; Bol et al., 1998; Mertler, 1998). With regard to the frequent use of alternative assessments, however, the findings were not conclusive. For example, although Alsarimi (2000) found no statistically significant gender differences for the frequent use of alternative assessments for 246 third preparatory science teachers in Oman, Mertler (1998) found that females reported using alternative assessments more frequently than males for 625 K-12 teachers in Ohio. Also, although Bol et al. (1998) found that the most experienced teachers (20 years or more) indicated using alternative assessments more frequently than the least experienced teachers (6 years or less) for 893 teachers at various grade levels in a large mid-southern urban school district, Mertler (1998) found that the least experienced teachers (5 years or
less) reported using alternative assessments more frequently than the most experienced teachers (31 to 35 years). In contrast, Alsarimi (2000) found no statistically significant differences based on years of teaching experience on the frequent use of alternative assessments between the participating teachers in his study.

In addition to the context where the study was conducted such as Ohio versus Oman, the aforementioned contradictory findings from survey research studies on teacher’s frequent uses of traditional and alternative assessments could be attributed to the lack of consistency in operationalizing the variable years of teaching experience and to the lack of focus on a specific instructional setting by combining participants from various geographic regions, subject areas, and grade levels. The present study focused only on ninth grade science teachers in Muscat science classrooms in Oman. Nevertheless, in light of classroom assessment literature and achievement goal theory, the implication of these findings for the present study was that the effect of teacher’s frequent use of alternative assessments on student’s perceived classroom assessment environment and achievement goals might depend on class’s gender and teacher’s levels of teaching experience. Specifically, in this study, two-and three-way interaction effects for class’s gender, teacher’s years of teaching experience, and teacher’s frequent use of alternative assessments on perceived classroom assessment environment and achievement goals would be specified and tested within the class-level model in addition to their main and cross-level interactions effects. In contrast, since teacher’s frequent use of traditional assessments has consistently not been found to vary as a function of teacher’s gender and years of teaching experience (Alsarimi, 2000; Bol et al., 1998; Mertler, 1998), only main
and cross-level interaction effects for teacher’s frequent use of traditional assessments on perceived classroom assessment environment and achievement goals would be specified and tested in this study. The hypothesized directional nature of these effects along with their theoretical and empirical support will further be discussed at the end of Chapter II of the dissertation.

**Recommended Classroom Assessment Practices**

Furthermore, both qualitative and quantitative research studies in classroom assessment have repeatedly documented that teachers do not often follow the assessment practices recommended by experts of educational measurement and assessment. For example, some teachers have been found assigning grades based on nonachievement factors such as class attendance, using assessment tasks that are technically inadequate, administering assessment tasks in ways that deviate from established standardized administration procedures, and not to conduct statistical analyses of quantitative assessment results (e.g., Hills, 1991; McMillan, Myran, & Workman, 2002; Mertler, 1998, 1999). 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 and what aspects of these assessment practices influence the adoption of each achievement goal are extremely limited. These questions are worth recognition and investigation because they may contribute to the understanding of both theory and practice of classroom assessment and achievement goals. The present study empirically addressed these questions to fill the
void in the knowledge base. Since teacher’s frequent use of recommended assessment practices has consistently not been found to vary as a function of teacher’s gender, years of teaching experience, grade level, subject area, and geographic region (Alsarimi, 2000; Cizek, Rachor, & Fitzgerald, 1995; Feldman, Alibrandi, & Kropf, 1998; Mertler, 1998, 1999), only main and cross-level interaction effects for teacher’s frequent use of recommended assessment practices on perceived classroom assessment environment and achievement goals would be specified and tested in this study. The hypothesized directional nature of these effects along with their theoretical and empirical support will further be discussed in Chapter II of the dissertation.

Statement of the Problem

Many studies have investigated teachers’ classroom assessment practices (e.g., Alsarimi, 2000; Mertler, 1998). However, few have considered their effects on student perceived classroom assessment environment and achievement goal orientations (e.g., Stefanou & Parkes, 2003; Wang, 2004). When this has been investigated, methodological nuisances have plagued the research. The ramification of this was the development of inadequate models in explaining student perceived classroom assessment environment and achievement goal orientations. Therefore, the problem addressed in this study, within the framework of classroom assessment literature and achievement goal theory, was the development of models that might explain the variation in student perceived classroom assessment environment and achievement goal orientations for ninth grade students in Muscat science classrooms in Oman.
Rationale for the Study

There were a number of reasons for conducting this study. First, although several studies have demonstrated that classroom assessment environment might influence students’ achievement goals, most of these studies tended to overlook the hierarchical structure of the data, in that students were nested within classrooms. As might be expected, ignoring the nested nature of the data might lead to statistical and conceptual problems such as unit of analysis, violation of independence assumption, and loss of information (Kasim, 1994; Raudenbush & Bryk, 2002). These problems are further explained in Chapter II of the dissertation. For now, the point is that students within a classroom may 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 (Kasim, 1994).

There are empirical evidences that support this argument. For example, in a study of 516 sixth grade students situated in 63 math classrooms, Ryan, Gheen, and Midgley (1998) found a small within-class variance in students’ perceptions of the purposes of engaging in academic work in the classroom. This finding suggested that there was some degree of consensus among students’ perceptions within classrooms, and hence it seemed reasonable to be identified as classroom-level variables (Ryan et al., 1998). The implication of Ryan et al.’s (1998) study findings for the present study was that classroom assessment research might need to take advantage of the hierarchical linear modeling (HLM) to avoid the conceptual and empirical problems associated with conventional methods of single-level analyses when examining the impact of teachers’
assessment practices on students’ perceptions of classroom assessment environment and achievement goals. Hence, the current study was undertaken to address this need.

Second, research on classroom assessment environment has used individual student scores as the unit of analysis rather than the average score of students at the classroom level (Ames, 1992b; Ames & Archer, 1988; Church, Elliot, & Gable, 2001; Kaplan et al., 2002b; 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, 1992b; Kaplan et al., 2002b). Nevertheless, Church et al. (2001) 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 classroom 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 had the following implication for the present study.

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

One may argue that the aggregate perception of classroom 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, 1985, 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, 1985, 1987, 1990, as cited in Goddard & Goddard, 2001, p. 810). Yet, except for Meece and her colleagues’ (2003) study, no recent study to my knowledge has considered how perceptions of classroom assessment environment interact across levels: student and class. In particular, Meece et al. (2003) examined the effects of students’ perceptions of
teaching practices on achievement goals for 4615 middle and high school students nested within 256 middle and high school teachers. Results indicated that individual students’ perceptions of teaching practices positively predicted mastery goals at the student-level. At the teacher-level, average class perceptions of teaching practices increased the relation between individual perceptions of teaching practices and mastery goals. Although individual perceptions of teaching practices positively predicted students’ performance goals, this effect did not vary significantly across classrooms.

Consistent with Meece et al.’s (2003) study findings, it was expected in this study that individual student’s perceived classroom assessment environment would vary significantly across classes. Contrary to Meece et al.’s (2003) study that combined middle and high school participants across various subject areas, the present study, however, controlled for the effects of academic subject and grade level by focusing on ninth grade science classrooms. In other words, academic subject and grade level were not variables in this study because they did not represent non-uniform characteristics of the observational units (i.e., students and teachers). Also, instead of general classroom teaching practices, the present study focused on classroom assessment practices. In addition, important research questions have not yet been answered in the research literature about whether an aggregate perceived classroom assessment environment (an emergent attribute of the classroom) is related to differences in individual perceived classroom assessment environment and achievement goals (student-level attributes). Hence, one of the purposes of this study was to empirically address these questions for ninth grade students in Muscat science classrooms in Oman.
Third, although teachers are expected to conduct classroom assessment practices that are in agreement with those recommended by experts of educational measurement and assessment (American Federation of Teachers [AFT], National Council on Measurement in Education [NCME], & National Educational Association [NEA], 1990; Gronlund, 1998), considerable amount of research have shown that teachers’ assessment practices are often not consistent with the recommended practices (e.g., Cizek et al., 1995; Cross & Frary, 1999; Frary, Cross, Weber, 1993; Mertler, 2003; Stiggins, Frisbie, & Griswold, 1989). Yet, many unanswered questions remain regarding the effects of teachers’ frequent use of the recommended classroom assessment practices on students’ perceptions of classroom assessment environment and motivation defined in terms of achievement goals; and hence the need for this study.

Motivating students to learn is one of the classroom assessment purposes (Gronlund, 2006; Harlen & Crick, 2003; Stipek, 2002). One then could argue that the consideration of nonachievement factors such as student’s class participation when determining academic grades might motivate students to exert more effort for learning (McMillan & Nash, 2000). However, some experts of educational measurement and assessment recommend that nonachievement factors not be used in grading because they are difficult to be objectively defined and measured (Frary et al., 1993; Stiggins et al., 1989). Taking together, it seems that these purposes of motivating students to learn and having a psychometrically sound assessment may be at odds with one another. As alluded to by Stiggins et al. (1989), how then can these purposes both be simultaneously achieved and neither is eroded? The present study attempted to shed light on this issue.
In a review of research on teachers’ grading practices, Brookhart (1994) argued that “grading theory and practice will be better connected once the role of classroom assessment and grading practices in student achievement motivation… is understood” (p. 279). Consequently, there have been repeated calls for systematic investigations of the impact of classroom assessment on student motivation (Brookhart, 1994, 1997a, 1997b, 2004; Harlen & Crick, 2003; McMillan & Lawson, 2001; McMillan & Workman, 1998; Stiggins, 1999; Stiggins & Chappuis, 2005). This study was undertaken to respond for more knowledge, understanding, and research in this area. It may be argued that to be able to understand and make sense out of the gap between the recommended and the actual classroom assessment practices, it is important to find out the possible effects of these practices on students’ achievement goals. This is because students’ achievement goals have been thought to be affected by teacher’s classroom practices (Black & Wiliam, 1998a; Deeevers, 2005; Harlen & Crick, 2003; Nolen & Haladyna, 1990; Stipek, 2002).

Fourth, of increasing interest to educational assessment community is classroom assessment environment as perceived by students. 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, 1992a, 1992b; Harlen & Crick, 2003; Hidi & Harackiewicz, 2000; Linnenbrink & Pintrich, 2001). Students’ achievement goals have been recognized as an important part to be promoted in the learning process (Harlen & Crick, 2003; Maslovaty & Kuzi, 2002).
Therefore, identifying characteristics of classroom assessment environment conducive to adoption of particular achievement goals becomes crucial.

In this study, the focus was on ninth grade Muscat science classrooms in Oman. The ninth grade science was chosen because of the importance of the transition period from middle school science curriculum to the greatest differentiation of the science curriculum at the secondary school in Oman as compared to other academic subjects. In addition, science is a required subject for all ninth grade students in Omani middle schools. It has a variety of assessment components both traditional and alternative. As such, it seems an intriguing research context for studying variability in both students’ perceptions of classroom assessment environment and achievement goals. Moreover, the ninth grade is the exit grade for middle schools in Oman. As such, it represents the culmination of the middle school years in Oman. This adds an interesting twist to the effects of classroom assessment practices on motivational beliefs, in that, fostering mastery-oriented beliefs in this particular grade level may facilitate adoption of these beliefs as students move to a higher level of schooling. This argument is supported by a number of empirical evidences. For example, as part of a longitudinal study examining student motivation as they made the transitions from elementary to middle school and middle to high school, Middleton, Kaplan, and Midgley (2004) found that student mastery goal orientation tended to be stable over time and that “being oriented to demonstrating ability may be problematic over time” (p. 307). Likewise, in their study of student motivational beliefs in high school science classes, Nolen and Haladyna (1990) reported that “the motivational orientation that a student brings to [the high] school
Significance of the Study

The current study was expected to reveal both theoretical and practical implications. From a theoretical point of view, except for Wang’s (2004) study, the construct of perceived classroom assessment environment has not been clearly operationalized in the classroom assessment literature. Also, as will further be elaborated in Chapter II of the dissertation, there has been a debate regarding the independence of achievement goal constructs. As such, exploring the underlying dimensions of perceived classroom assessment environment and achievement goal orientations may further research agenda in both classroom assessment literature and achievement goal theories.

Moreover, the study would provide tests of the hypothesized relationships between some of the classroom assessment and motivation variables postulated by Ames (1992a, 1992b), Brookhart (1997a), and McMillan and Workman (1998), which will be discussed more fully in Chapter II of the dissertation. Although previous studies have reported some relationships among these variables, it is quite difficult to draw meaningful definitive generalizations from their findings. Most of these studies suffered from insufficient sample size; little devotement to the nested nature of the data; and lack of focus on a specific instructional setting by combining participants from various geographic regions, subject areas, and grade levels. Also, of those variables considered in
the previous studies, there was little consistency in instrumentation, and, as such, the somewhat contradictory findings are not surprising. The present study attempted to take these concerns into account.

In addition, despite its promising role that has been alluded to, the aggregate perception (i.e., class’s average perception) of classroom assessment environment seems to be a neglected construct in the past studies of classroom assessment and student motivation. Studying this construct may offer a new possibility for improving student motivational outcomes and perhaps at least diminishing the documented patterns of problematic changes in motivation experienced by students during the transition to middle school and within middle school grade levels (see Anderman & Midgley, 1997; Middleton et al., 2004; Urdan & Midgley, 2003). Such a study would also have an import regarding the manner in which students influence each other and are influenced by their teacher, and consequently could provide a road sign for the teacher to improve the learning climate within the classroom. Although there is a great deal of research on classroom assessment practices of teachers, few attempts have been made to link these practices to students’ perceptions of classroom assessment environment and achievement goals. Finally, the study was expected to provide Muscat ninth grade science teachers with insights to implement classroom assessment in ways that encourage students to master and understand the content materials and demonstrate a willingness to actively engage in the learning process.
Purposes of the Study

The purposes of this study were fourfold:

1. To identify the underlying dimensions of ninth grade students’ perceptions of classroom assessment environment in Muscat science classrooms in Oman.

2. To identify the underlying dimensions of ninth grade students’ achievement goal orientations in Muscat science classrooms in Oman.

3. To investigate the possible effects of certain student characteristics (e.g., academic self-efficacy) and classroom characteristics (e.g., class gender) on perceived classroom assessment environment for ninth grade students in Muscat science classrooms in Oman.

4. To investigate the possible effects of certain student characteristics (e.g., academic self-efficacy) and classroom characteristics (e.g., class gender) on achievement goal orientations for ninth grade students in Muscat science classrooms in Oman.

Research Questions

The current research study investigated the effects of teachers’ assessment practices on ninth grade students’ perceptions of classroom assessment environment and achievement goal orientations in Muscat science classrooms in Oman. This investigation was shaped by ten general research questions that were formulated as follows:

1. What are the underlying dimensions of ninth grade students’ perceptions of classroom assessment environment in Muscat science classrooms in Oman?

2. What are the underlying dimensions of ninth grade students’ achievement goal orientations in Muscat science classrooms in Oman?

3. Do ninth grade Muscat science classrooms in Oman vary in perceived classroom assessment environment (PCAE)?

4. What are some of the student characteristics (e.g., academic self-efficacy) that might have effects on PCAE?
5. Do the effects of student characteristics (e.g., academic self-efficacy) on PCAE vary across classrooms?

6. What are some of the classroom characteristics (e.g., class gender) that might help explain the variability in the PCAE and in the effects of student characteristics (e.g., academic self-efficacy) on PCAE across classrooms?

7. Do ninth grade Muscat science classrooms in Oman vary in achievement goal orientations (GOAL)?

8. What are some of the student characteristics (e.g., academic self-efficacy) that might have effects on GOAL?

9. Do the effects of student characteristics (e.g., academic self-efficacy) on GOAL vary across classrooms?

10. What are some of the classroom characteristics (e.g., class gender) that might help explain the variability in the GOAL and in the effects of student characteristics (e.g., academic self-efficacy) on GOAL across classrooms?

Limitations of the Study

There would be several limitations to this study:

1. Measurement of the variables would be based on self-report questionnaires. Thus, there would be no control for factors that might affect responses on the questionnaires such as social desirability.

2. Measurement of perceived classroom assessment environment as well as teacher’s frequent uses of traditional assessments, alternative assessments, and recommended classroom assessment practices would be limited to a questionnaire developed by the author for this study based on the literature.

3. Measurement of student’s achievement goal orientations would be limited to a modified version of items drawn and adapted from the Patterns of Adaptive Learning Scales (PALS) developed by Midgley et al. (2000).

4. Measurement of student’s academic self-efficacy would be limited to a modified version of items drawn and adapted from a scale developed by Greene, Miller, Crowson, Duke, and Akey (2004) as well as items from Midgley et al.’s (2000) PALS.
5. There would be no control for variables that might affect student’s perceptions of classroom assessment environment, academic self-efficacy, and adoption of achievement goals such as prior experiences and academic history.

Definition of Terms

For the purpose of clarification, following are definitions of terms related to classroom assessment and motivation used in this study:

Assessment

Assessment is a broad term defined as the process of collecting, synthesizing, interpreting, and using qualitative and/or quantitative information for making educational decisions about students (e.g., credentialing and certifying their competence), curricular (e.g., their effectiveness), and educational policy (e.g., at the local educational region and national level) (AFT, NCME, & NEA, 1990; Airasian, 1997; Cizek, 1997; Gallagher, 1998; McMillan, 1997; Nitko, 2001).

Classroom Assessment

Classroom assessment refers to the process used in the classroom by the teacher to obtain information about students’ performances on assessment tasks, either as a group or individually, using a wide range of assessment methods, to determine the extent to which students are achieving the target instructional outcomes (Gallagher, 1998; Gronlund, 1998).
Assessment Tasks

An assessment task refers to any activity, test, exercise, or question requiring students to demonstrate knowledge and/or skills on a particular topic (National Center for Research on Evaluation, Standards, and Student Teaching [CRESST], 2006). Consistent with Brookhart’s (1997a) model of classroom assessment and student motivation, assessment tasks in this study were considered as one aspect of the classroom assessment environment defined later in this section of the dissertation.

Grading

Grading is defined as “indicating both a student’s level of performance and a teacher’s valuing of that performance” (AFT, NCME, & NEA, 1990, p. 32). It can be (a) absolute, which involves a criterion-referenced interpretation, or (b) relative, which involves a norm-referenced interpretation (Gronlund, 2006). Specifically, absolute grading involves comparing a student’s performance to a specified standard of performance (Gronlund, 2006). In this study, from the perspective of achievement goal theory (Ames, 1992a, 1992b), a strong emphasis on absolute grading was expected to more likely prompt mastery-related outcome beliefs than performance-related outcome beliefs. Relative grading involves comparing a student’s performance to the performance of the group members (Gronlund, 2006). In this study, from the perspective of achievement goal theory (Ames, 1992a, 1992b), a strong emphasis on relative grading was expected to more likely prompt performance-related outcome beliefs than mastery-related outcome beliefs.
Grades

Grades are “symbols (letters, numbers, words) that represent a value judgment concerning the relative quality of a student’s achievement during a specified period of instruction” (Johnson & Johnson, 2002, p. 243). Experts of educational measurement and assessment recommend that nonachievement factors such as student’s neatness of work, interest in learning the subject, class participation, class attendance, and class behavior as well as performance of other students not be used in grading (Stiggins et al., 1989). Given that grading is one of the teachers’ classroom assessment practices (Zhang & Burry-Stock, 2003) that may affect students’ perceptions of their classroom environment and adoption of achievement goals (Ames, 1992a, 1992b; Brookhart, 1997a; McMillan & Workman, 1998), teacher’s adherence to the recommended grading practices was considered in this study and will further be elaborated in the Recommended Classroom Assessment Practices later in this section and in Chapter II of the dissertation.

Traditional Assessments

Traditional assessments refer to assessments that rely heavily on paper-and-pencil tests that require students to choose the correct or best answer as in multiple-choice, true-false, and matching assessment tasks; or to respond with a word, short phrase, or complete essay answer (Gronlund, 1998). They present a single, highly structured task that does not simulate performance in the real world (Gronlund, 1998). More details about traditional assessments will be provided in Chapter II of the dissertation. For now, the point is that in this study, the focus was on teacher’s frequent use of traditional
assessments. It would be reflected by a total rating score across all the items representing traditional assessments on the Teacher’s Questionnaire developed by the author for this study and described in Chapter III of the dissertation. Higher scores would represent a more frequent use of traditional assessments by the teacher.

*Performance Assessments*

Performance assessments refer to assessments that require students to demonstrate their achievement understandings and skills by *actually performing* an assessment task or a set of tasks (Gronlund, 1998). They are distinguished from traditional assessments by greater realism of the tasks (i.e., like real-life problems), greater complexity of the tasks (i.e., involve multiple learning outcomes and integration of ideas and skills from various sources), greater time needed for the assessment, and greater use of judgment in scoring due to originality of the responses and a variety of possible solutions (Gronlund, 1998). Performance tasks can appear in a restricted or extended format (Gronlund, 1998). Restricted performance tasks tend to be highly structured and limited in scope such as constructing a graph from a given set of data or demonstrating how to set up laboratory equipment (Gronlund, 1998). Extended performance tasks tend to be less structured and comprehensive in scope such as writing a science story or conducting a laboratory experiment and defending the findings (Gronlund, 1998). Consistent with Gronlund’s (1998) definitions, performance assessments in this study were considered as one kind of alternative assessments defined as follows.
Alternative Assessments

Alternative assessments are another title used for describing performance assessments to indicate that they are alternative to traditional assessments described above (Gronlund, 1998). More details about alternative assessments will be provided in Chapter II of the dissertation. For now, the point is that in this study, the focus was on teacher’s frequent use of alternative assessments. It would be reflected by a total rating score across all the items representing alternative assessments on the Teacher’s Questionnaire developed by the author for this study and described in Chapter III of the dissertation. Higher scores would represent a more frequent use of alternative assessments by the teacher.

Authentic Assessments

Authentic assessments are another title used for describing performance assessments to emphasize the importance of focusing on the application of understandings and skills to real life-like problems (Gronlund, 1998).

Portfolio Assessments

Gronlund (1998) defined a portfolio assessment as “a preplanned collection of samples of student work, assessment results, and other data that represent the student’s accomplishments” (p. 225). It is viewed as a type of performance assessment (Bol et al., 1998; Gronlund, 1998). Therefore, consistent with Gronlund’s (1998) definitions, portfolio assessments in this study were considered as one kind of alternative assessments.
Performance Criteria and Standards

Criteria refer to a set of dimensions that are used to judge a student’s performance (Gronlund, 1998). Examples of criteria that might be used for an extended science class project are “problem selection”, “report writing”, and “oral presentation” (Gronlund, 1998). Standards refer to “a prespecified level of performance that is considered satisfactory for the use to be made of the assessment results” (Gronlund, 1998, p. 225). Performance standards often indicate the minimum level of acceptable performance (Gronlund, 1998). These might be concerned with, for example, percentage correct such as “solving correctly 80% of the science problems” (Gronlund, 1998). The standards that describe the success on each criterion should be established at the same time as the criteria are developed (Gronlund, 1998). Consistent with Brookhart’s (1997a) model of classroom assessment and student motivation, assessment standards and criteria in this study were considered as one aspect of the classroom assessment environment defined later in this section of the dissertation.

Recommended Classroom Assessment Practices

This term refers to classroom assessment practices endorsed by experts of educational measurement and assessment for their technical adequacy. More details about the recommended practices will be provided in Chapter II of the dissertation. For now, the point is that in this study, these practices are concerned with revision of assessment (e.g., conducting statistical analyses of the quantitative assessment results), communication of assessment (e.g., informing students about the purpose of an
assessment task prior to its administration), performance criteria and standards (e.g., constructing a model answer for scoring responses to essay questions), student-involved assessment (e.g., providing students opportunities to write test questions based on their understanding of instructional objectives), and grading (e.g., avoid comparing student’s performance with other students when determining a student’s grade). In this study, the focus was on teacher’s frequent use of recommended classroom assessment practices. It would be reflected by a total rating score across all the items representing the recommended classroom assessment practices on the Teacher’s Questionnaire developed by the author for this study and described in Chapter III of the dissertation. Higher scores would represent a more frequent use of the recommended classroom assessment practices, and hence a higher level of alignment or a higher degree of agreement with the practices recommended by experts of educational measurement and assessment.

Classroom Assessment Environment

Classroom assessment environment has been defined as the assessment context experienced by students as the classroom teacher establishes assessment purposes, assigns assessment tasks, sets performance criteria and standards, provides feedback, and monitors outcomes (Brookhart, 1997a). In this study, the focus was on an individual student’s and aggregate (i.e., class) perception of the classroom assessment environment as the classroom teacher assigns assessment tasks, sets performance criteria and standards, and provides feedback. For example, students may perceive the classroom assessment environment as a learning-oriented environment if the teacher uses a variety
of interesting and meaningful assessment tasks, provides students with clear performance criteria and standards in advance of an assessment task, and gives continuous and informative feedback about strengths and weaknesses of student’s performance (Wang, 2004).

In this study, the individual student’s perceived classroom assessment environment refers to the self perception of an individual student about the various aspects (i.e., assessment tasks, assessment feedback, and assessment standards and criteria) of the classroom assessment environment. In other words, in light of achievement goal theory, what does an individual student think about the assessment tasks, assessment feedback, and assessment standards and criteria as aspects of the classroom assessment environment established by the teacher in the class? Does the student think that these aspects of the classroom assessment environment as motivating him/her to learn and master the content materials of the subject or discouraging learning and mastery pursuits? In this study, the individual student’s perceived classroom assessment environment would be reflected for each student by a total rating score across all the items that mainly loaded on each dimension that would emerge from principal components/exploratory factor analyses of the responses to the items pertaining to perceived classroom assessment environment on the Student’s Questionnaire developed by the author for this study and described in Chapter III of the dissertation.

The aggregate perceived classroom assessment environment refers to the shared perception of students (i.e., have in common, or taken as a whole, or collective) in a classroom about the various aspects (i.e., assessment tasks, assessment feedback, and
assessment standards and criteria) of the classroom assessment environment. In other words, in light of achievement goal theory, what do students, taken as a whole, in a particular classroom think about the assessment tasks, assessment feedback, and assessment standards and criteria as aspects of the classroom assessment environment established by the teacher in the class? Do the students collectively think that these aspects of their classroom assessment environment as motivating them to learn and master the content materials of the subject or discouraging learning and mastery pursuits? In this study, the aggregate perceived classroom assessment environment would be reflected for each classroom by the average levels of individual students’ perceptions within the classroom (i.e., class’s average) on each dimension that would emerge from the principal components/exploratory factor analyses of the responses to the items pertaining to perceived classroom assessment environment on the Student’s Questionnaire developed by the author for this study and described in Chapter III of the dissertation. Higher scores on each dimension would represent positive perceptions of the classroom assessment environment defined in that dimension.

**Achievement Goal Orientations**

Midgley et al. (2000) defined achievement goal orientations as “students’ reasons or purposes for engaging in academic behavior” (p. 7). For example, one student might approach a science class assignment with a goal of learning as much as possible about the content of the assignment, whereas another student might approach the same assignment with a goal of doing better than others in the class. More details about achievement goals
will be provided in Chapter II of the dissertation. For now, the point is that in this study, a student’s achievement goal would be reflected by an average rating score across all the items that mainly loaded on each dimension that would emerge from the principal components/exploratory factor analyses of the responses to the items pertaining to achievement goals on the Student’s Questionnaire developed by the author for this study and described in Chapter III of the dissertation. Higher scores on each dimension would represent a higher adoption or endorsement of the achievement goal defined in that dimension.

**Academic Self-Efficacy**

Academic self-efficacy has been defined as students’ judgments of their capabilities to complete their class work successfully (Bandura, 1986; Midgley et al., 2000; Ryan et al., 1998; Schunk, 1991). It pertains to answering the question “Can I do this task?” (Stipek, 2002, p. 41). An example of an item describing student’s academic self-efficacy in science is “I’m certain I can figure out how to do the most difficult class work in science” (Midgley et al., 2000, p. 20). In this study, academic self-efficacy would be measured at both the student-level and the class-level. At the student-level, it would be reflected by a total rating score for each student across all the items that mainly loaded on the component that would emerge from the principal components/exploratory factor analyses of the responses to the items pertaining to academic self-efficacy on the Student’s Questionnaire developed by the author for this study and described in Chapter III of the dissertation. At the class-level, it would be reflected by the average academic
self-efficacy levels of students within a particular class (i.e., class’s average for academic self-efficacy). Higher scores would represent a higher level of academic self-efficacy in science.

Summary

Of increasing interest to classroom assessment researchers is the role of classroom assessment on student motivation. Of interest in this study are the effects of teachers’ assessment practices on ninth grade students’ perceptions of classroom assessment environment and achievement goals in Muscat science classrooms in Oman. Chapter I introduced the research problem being investigated in this study. The next chapter of the dissertation includes a detailed review of the literature on classroom assessment and achievement goals to supplement and elaborate the ideas presented in Chapter I.
CHAPTER II

REVIEW OF THE LITERATURE

Research has been conducted to study classroom assessment practices of teachers. The impact of these practices on student motivation has recently received increased attention in the classroom assessment literature. Following is a review of literature discussing research conducted in classroom assessment and student motivation defined in this study as achievement goal orientations. The review is divided into five major sections including achievement goal orientations, academic self-efficacy, traditional and alternative assessments, recommended classroom assessment practices, and classroom assessment environment. The review will be closed with a synthesis linking previous research to the research questions addressed in this study.

Achievement Goal Orientations

A Dichotomous Framework of Achievement Goals

In educational motivation research, achievement goal theory has evolved within a social-cognitive framework (Dweck & Leggett, 1988; Maehr, 1989; Nicholls, 1989; Weiner, 1990). Achievement goal theorists have traditionally identified two types of achievement goals (also called a dichotomous framework of achievement goals): the goal to develop ability and the goal to demonstrate ability or avoid demonstrating the lack of
ability (Elliot, 1999). These two goals have alternatively been labeled learning and performance goals (Dweck, 1986), task-involvement and ego-involvement goals (Nicholls, 1984), and mastery and performance goals (Ames & Archer, 1987, 1988). Dweck and Leggett (1988) proposed that “the goals individuals are pursuing create the framework within which they interpret and react to events” (p. 256). Mastery goals create a framework in which inputs and outputs provide information about one’s learning and mastery, whereas performance goals create a framework in which inputs and outputs are interpreted in terms of one’s ability and its adequacy (Dweck & Leggett, 1988).

According to achievement goal theorists, mastery goals center on the development of competence, whereas performance goals center on the outward showing of competence (Ames & Archer, 1987, 1988; Dweck, 1986; Nicholls, 1984).

Achievement goals are thought to vary across individuals (Maehr, 1983, 1984), and positive and negative patterns of cognition and affect may be elicited by the adoption of a particular achievement goal (Ames, 1992a). From the perspective of achievement goal theory, students who adopt mastery goals are expected to persist in the face of difficult events, seek challenging activities, and have high intrinsic motivation (Ames, 1992b; Dweck, 1986; Nicholls, 1984). In comparison, students who adopt performance goals are expected to minimally persist in the face of difficult events, avoid challenging activities, and have low intrinsic motivation (Ames, 1992b; Dweck, 1986; Nicholls, 1984). Mastery goals have consistently been linked to a positive set of processes and outcomes such as deep processing of studying materials, long term retention of
information, adaptive attributional patterns of success and failure, and appropriate help-seeking behaviors (Ames, 1992b; Elliot, 1999; Weiner, 1990, 1994, 2000). However, the effects of pursuing performance goals are less clear. Some studies have found that adoption of performance goals has negative effects when accompanied by low perceived competence (e.g., Elliot & Church, 1997; Elliot & Dweck, 1988), whereas other studies have not supported these effects (e.g., Elliot & Harackiewicz, 1996; Harackiewicz & Elliot, 1993). As a result, achievement goal theory has undergone a number of theoretical advances.

A Trichotomous Framework of Achievement Goals

Elliot and his colleagues have proposed a trichotomous framework of achievement goals that further differentiates performance goals into approach and avoidance goals (Elliot & Church, 1997; Elliot & Harackiewicz, 1996). In this framework, three types of achievement goals are posited: mastery goals that focus on the development of competence, performance-approach goals that focus on having favorable judgments of competence, and performance-avoidance goals that focus on avoiding unfavorable judgments of competence (Elliot & Church, 1997; Elliot & Harackiewicz, 1996). The validity and utility of this trichotomous framework of achievement goals have been demonstrated for middle school and college level students.

Although results of factor analytic procedures from studies conducted on college level students by Elliot and Church (1997) indicated that the ability-approach (also called performance-approach) and ability-avoidance (also called performance-avoidance) goals are independent orientations with distinct determinants and a divergent set of
consequences, studies conducted on middle school students tended to reveal some overlap between these two types of goals. For example, in their study of 703 sixth grade students, Middleton and Midgley (1997) found that the scales measuring performance-approach and performance-avoidance goals were moderately correlated ($r = .56$) suggesting that these two types of goals are not qualitatively different. As such, the present study sought to clarify this relationship for ninth grade students in Muscat science classrooms in Oman. Specifically, the study adopted the trichotomous framework of achievement goal theory because it was assumed to be the most prevalent goal framework in achievement settings (Elliot, 1999; Elliot & Church, 1997; Elliot & Thrash, 2001). Thus, it was expected that factor analytic procedures in this study would yield the three dimensions of achievement goals (mastery, performance-approach, and performance-avoidance) with (a) a positive correlation between the two types of performance goals and (b) no correlations between mastery goals and each of the two types of performance goals.

In addition to the type of participants, the difference between the aforementioned Elliot and Church’s (1997) study and Middleton and Midgley’s (1997) study can be attributed to their approach to the measurement of achievement goals. Elliot and Church’s approach to the measurement was based on the conceptualization of achievement goals as “cognitive-dynamic manifestations of two underlying competence-relevant motives, the need for achievement and the need to avoid failure” (p. 219). As a result, some of their items assessed affective components such as worries, fears, and concerns rather than reasons or purposes for engaging in academic behaviors, a definition
on which Middleton and Midgley’s scale was based. The present study focused on middle
school students and as such its approach to the measurement of achievement goals was
based on the approach suggested by Middleton and Midgley’s (1997) study.

A Multiple Goal Framework

In studies of achievement goals, there have been positive, negative, or no
correlation between mastery and performance goals (Midgley et al., 1998). The negative
correlation between mastery and performance goals suggests that students pursue one
goal to the exclusion of the other (Barron & Harackiewicz, 2001). The positive
correlation between mastery and performance goals suggests the possibility that students
can and do pursue multiple goals simultaneously and to varying degrees in achievement
settings (Pintrich, 2000b). This possibility has led some achievement goal theorists to
propose a multiple goal perspective, which suggests that students often pursue more than
one achievement goal at schoolwork and that the simultaneous pursuit of both mastery
and performance-approach goals has a more positive effect on student achievement-
related outcomes than the pursuit of a single achievement goal (Barron & Harackiewicz,
2001; Pintrich, 2000b).

This perspective of multiple goals has been supported by experimental research
studies conducted for elementary, middle, high school, and college level students in
which achievement goals were manipulated and induced in the participants (for more
details see Barron & Harackiewicz, 2001; Linnenbrink, 2005 Pintrich, 2000b). Although
these experimental studies point to the reality of achievement settings in that student can
endorse both mastery and performance goals with respect to the class work, the results
still have little ecological validity due to artificial experimental conditions with very brief treatments that were not carried out by the students’ usual teachers (Black & Wiliam, 1998a; Crooks, 1988). In the present correlational research study, where students were surveyed in their normal classroom settings, statistical support for the multiple goal perspective was expected to come from either (a) a moderate positive correlation between mastery and performance-approach goals with low or no correlations between each of these two goals and performance-avoidance goals or (b) a factor model having a good fit to the data in which the mastery and performance-approach items load together on one latent factor whereas performance-avoidance items load on their respective latent factor.

A 2 × 2 Achievement Goal Framework

Elliot (1999) and Pintrich (2000a) have argued that mastery goals need to be separated into approach and avoidance orientations to account for the broad spectrum of competence-based strivings. As such, Elliot (1999) and Pintrich (2000a) have proposed a theoretical 2×2 conceptualization of achievement goals, in which mastery goals are bifurcated to form mastery-approach and mastery-avoidance goals. As a result, the 2×2 achievement goal framework comprises four achievement goals: mastery-approach, mastery-avoidance, performance-approach, and performance-avoidance goals (Elliot, 1999; Pintrich, 2000a). Individuals who adopt mastery-approach goals seek to achieve task mastery or improvement, whereas those who hold mastery-avoidance goals seek to avoid failing achievement of task mastery (Elliot, 1999; Pintrich, 2000a). Individuals who adopt performance-approach goals seek to do better than others, whereas those who hold performance-avoidance goals seek to avoid doing worse than others (Elliot, 1999;
Pintrich, 2000a). To my knowledge, this framework has been validated only for college students (see Elliot & McGregor, 2001). Therefore, the current study extended this line of research by examining the underlying dimensions of achievement goals for ninth grade students in Muscat science classrooms in Oman.

The Conceptualization of Achievement Goals

The definition of achievement goal construct has been approached by two ways. The first approach construes achievement goals in terms of purposes of achievement behavior (Dweck, 1996; Maehr, 1989). Proponents of this approach have conceptualized achievement goals as a combination of both reasons for engaging in achievement situations and aims sought to be achieved (Pintrich, 2000a; Urdan & Maehr, 1995). Accordingly, mastery goals have been described in terms of both development of competence and adoption of task-based evaluation of competence (Elliot & Thrash, 2001). In contrast, performance goals have been described in terms of both displaying competence and adoption of normative evaluation of competence (Elliot & Thrash, 2001). The second approach describes achievement goals as an integrated pattern of beliefs about success, ability, effort, errors, and evaluation standards that collectively lead to an orientation toward achievement goals (Ames, 1992b; Ames & Archer, 1987).

However, Elliot and Thrash (2001) contend that these definitional approaches are limited in four ways. First, these approaches define achievement goals as a network of interrelated variables, which may make it difficult to specifically identify the particular aspect of achievement goals that may be responsible for any observed effects (Elliot & Thrash, 2001). Second, there have been no guidelines for determining the characteristics
of each achievement goal that can be considered for its adoption (Elliot & Thrash, 2001). Third, conceptualizing achievement goals as an interrelated set of motivational constructs implies that research linking these variables should be considered as construct validation rather than testing predictive roles of the achievement goals (Elliot & Thrash, 2001). Fourth, results from research conducted in this area seem to be somewhat ambiguous, in that some studies have focused on a single aspect of the achievement goal definition, whereas other studies have simultaneously focused on several aspects (Elliot & Thrash, 2001).

Consequently, Elliot and Thrash (2001) have theoretically proposed a hierarchical model of achievement goals. In this model, achievement goals are defined as “a cognitive representation of a competence-based possibility that an individual seeks to attain” (Elliot & Thrash, 2001, p. 144). They are conceptualized based on two dimensions: competence definition and competence valence (Elliot & Thrash, 2001). The definition dimension of competence represents the mastery-performance distinction of achievement goals (Elliot & Thrash, 2001). Mastery goals are defined in terms of absolute or intrapersonal standards of competence, whereas performance goals are defined in terms of normative standards of competence (Elliot & Thrash, 2001).

The valence dimension of competence comprises the distinction between approach and avoidance forms of achievement goals (Elliot & Thrash, 2001). Approach goals are construed in terms of achieving positive desirable outcomes, whereas avoidance goals are construed in terms of avoiding negative undesirable outcomes (Elliot & Thrash, 2001). When the definition and valence dimensions of competence are combined, six
types of achievement goals are produced (Elliot & Thrash, 2001). These are absolute-approach, absolute-avoidance, intrapersonal-approach, intrapersonal-avoidance, normative-approach, and normative-avoidance (Elliot & Thrash, 2001). Each of these goals reflects a unique type of competence-based forms of achievement goals, may draw a distinct set of processes, and leads to different outcomes (Elliot & Thrash, 2001). However, research to date has focused on only a subset of these goals (Elliot & Thrash, 2001). Clearly, more empirical work is needed to investigate the prevalence of achievement goals, and hence the need for the present study.

Along similar lines, Brophy (2005) argued against defining performance-goals as focusing on the demonstration of ability relative to others. In particular, Brophy (2005) suggested that goal theorists should phase the social component out of the performance-goals, and characterize these goals as outcome goals or some other term that does not carry the social comparison connotation. To support his argument, Brophy (2005) offered three situations in which performance goals are a low-incidence phenomenon. First, under natural classroom settings, elementary and middle school students rarely generate performance goals that include elements of peer comparison and competition (Brophy, 2005). Second, performance goal orientations generated, by university undergraduates participating in laboratory experiments, in response to induction procedures, involve little if any of the social comparison or competition emphasis usually associated with performance goals (Brophy, 2005). Third, although goal theory research has reported positive relationships between adoption of performance-approach goals and subsequent
task performance, these relationships are correlational and likely to be epiphenomenal rather than causal (Brophy, 2005).

At a practical level, Brophy’s (2005) argument has implications for teachers, in that, outcome goals characterize the target attainment in criterion-referenced rather than norm-referenced terms and orient students toward achievement rather than peer competition. At a theoretical level, however, this argument raises a number of future research questions that will add to the existing literature on achievement goal theory. More specifically, if we remove the need to socially display competence or remove the need of interpersonal comparison from performance goals, do we still have a performance component? In other words, could the idea of intrapersonal approach be intrapersonal competition? How does that relate to mastery? What about self-referent mastery versus task-referent mastery? Finally, what about the avoidance dimension of goal orientations? Clearly, these questions warrant the need for more empirical refinements in the conceptualization of achievement goals, and hence the need for the present study.

Weiner (1990) pointed to achievement goal theory as “a major new direction, one pulling together different aspects of achievement research” (p. 629). Since its origin in the late 1970s and early 1980s (Kaplan et al., 2002b), many researchers have linked achievement goals to a variety of motivation variables. One of these variables that have been prominent in the area of achievement motivation is student academic self-efficacy. For example, Schunk (1991) suggested that academic self-efficacy is an important determinant of achievement goals. Likewise, Skaalvik (1997) noted that perceived
academic ability might influence whether students adopt performance-approach or performance-avoidance goals. Clearly, these researchers point to the importance of considering student academic self-efficacy when studying achievement goals. The present study took this point into account.

Academic Self-Efficacy

Academic self-efficacy refers to student’s judgments of his or her capabilities to successfully perform school-related works (Pintrich & Schunk, 1996; Schunk, 1991). Bandura (1993) has postulated that student’s beliefs about his or her efficacy to manage academic tasks influence level of motivation and affective reactions. According to Bandura’s (1977, 1982) theoretical perspective, students tend to avoid situations they believe exceed their capabilities, whereas they approach other situations they believe that they are capable of handling their demands. From the perspective of Bandura’s (1986) social-cognitive theory and Elliot’s (1999) review of achievement goal theory, self-efficacy is assumed to be a precursor to adoption of an approach achievement goal. Yet, “reciprocal causation among several of these constructs may occur over time (e.g., goals and self-efficacy)” (Roeser, Midgley, & Urdan, 1996, p. 410). Therefore, consistent with Bandura ‘s (1986) and Elliot’s (1999) theoretical perspectives, academic self-efficacy in this study was assumed to have an influence on achievement goals for ninth grade students in Muscat science classrooms in Oman.
Academic Self-efficacy and Achievement Goals

Since academic self-efficacy and achievement goals are thought to differ based on situational demands (Bandura, 1977, 1982; Maehr, 1983, 1984), several studies have investigated relationships between academic self-efficacy and achievement goal orientations. For example, in a study of 291 elementary and 678 middle school students, Midgley, Anderman, and Hicks (1995) found that task goals were positively related to efficacy at both school levels, whereas ability goals (also called performance goals) were unrelated to efficacy at the elementary school level and positively related to efficacy at the middle school level. Schunk (1996) found that a task goal orientation was positively related and an ego orientation was negatively related to 40 fourth grade students’ certainty that they could solve a set of fraction problems. Similarly, in a study of 703 sixth grade students, Middleton and Midgley (1997) found that a task-goal orientation was positively related, an ability-avoidance (also called performance-avoidance) goal orientation was negatively related, and an ability-approach (also called performance-approach) goal orientation was unrelated to students’ self-efficacy. Likewise, Deevers (2005) found that self-efficacy in math was positively related to mastery goal orientation of 1,571 students in grades 5 through 12, and that this relationship varied significantly across classes. Given that academic self-efficacy and achievement goals are postulated to co-vary based on environmental demands (Bandura, 1977, 1982; Maehr, 1983, 1984) and consistent with Deevers’s (2005) study findings, the relationships between academic self-efficacy and achievement goals in this study were expected to vary across ninth grade Muscat science classrooms in Oman.
When the relationship of academic self-efficacy to achievement goal orientations was studied in Arabic speaking countries, similar results were found to those reported in the United States. For example, Benmansour (1999) examined motivational goal orientations and self-efficacy in learning mathematics for 289 high school students in Morocco. Results showed that students with a stronger intrinsic goal orientation reported stronger self-efficacy than are students with a stronger goal orientation toward grades. Unlike Benmansour’s (1999) study that focused on high school mathematics classrooms in Morocco without taking into account the nested nature of the data, the present study focused on ninth grade Muscat science classrooms in Oman taking into account the hierarchical structure of the data.

It seems clear from previous research conducted on middle school students that students’ self-efficacy may affect their adoption of achievement goals and that this effect may vary across classrooms. Thus, academic self-efficacy was considered in this study to examine its potential impact on achievement goal orientations for ninth grade students in Muscat science classrooms in Oman. Armed with the aforementioned research findings and in accord with Elliot and Thrash’s (2001) conceptualization of achievement goals as well as with Bandura’s (1986) and Elliot’s (1999) theoretical perspectives about the relationship between academic self-efficacy and achievement goals, it was expected in this study that student’s academic self-efficacy would on average be positively related to both mastery goals and performance-approach goals, but negatively related to performance-avoidance goals within classes. Given that academic self-efficacy is thought to vary based on classroom environmental demands (Bandura, 1977, 1982), its
relationship with perceived classroom assessment environment was also examined in this study as follows.

Academic Self-efficacy and Classroom Environment

Of interest to educational assessment community is how student’s academic self-efficacy relates to his or her perception of the classroom assessment environment. In this regard, Anderman and Midgely (1997) found a positive relationship between student’s perceived ability and student’s perceived classroom environment as being mastery-oriented for 341 middle school students. Specifically, students reporting a high perception of ability tended to perceive their classroom environment as having a strong emphasis on mastery-oriented outcomes. Similarly, working with 220 high school students, Greene et al. (2004) found that students with a high sense of self-efficacy tended to perceive their classroom environment as supporting autonomy in learning and having mastery-oriented evaluations.

In their theoretical framework of the effect of classroom assessment practices on student motivation, McMillan and Workman (1998) indicated that when teacher’s assessment standards and criteria are clear, fair, and provided to students in advance of the assessment task with examples of previously graded student work, students are more likely to feel capable in performing the task successfully. This suggests that teacher’s classroom assessment practices may influence student motivational beliefs (Nolen & Haladyna, 1990). In this study, an emphasis was placed upon ninth grade student’s academic self-efficacy in science and its relationship to perceived classroom assessment environment. Based on the aforementioned previous research on achievement goal
theory, it was expected in this study that student’s academic self-efficacy would on average be positively related to perceived learning-oriented and negatively related to perceived normative-oriented classroom assessment environment within ninth grade Muscat science classrooms in Oman.

Classroom assessment refers to the process used in the classroom by the teacher to obtain information about students’ performances on assessment tasks, either as a group or individually, using a wide range of assessment methods, to determine the extent to which students are achieving the target instructional outcomes (Gallagher, 1998; Gronlund, 1998). It involves various activities including, but are not limited to, developing assessment methods such as paper-pencil tests and performance measures; administering, scoring, and interpreting assessment results; developing grading procedures; communicating assessment results; and using them in making educational decisions (AFT, NCME, & NEA, 1990; Zhang & Burry-Stock, 2003). Over the past years, considerable amount of qualitative and quantitative research have examined classroom assessment practices of teachers across different subject areas and school levels. The present study focused on ninth grade science teachers and their students. Some studies considered this grade level as an upper elementary or middle school level while others considered it as a high or secondary school level. As such, studies conducted at these school levels will be reviewed.

Traditional and Alternative Assessments

For the most part, assessment methods can be classified as traditional or alternative based on the extent to which they simulate performance in the real world (i.e.,
realism of assessment tasks), the extent to which they measure higher learning outcomes and requires demonstration of multiple skills (i.e., complexity of tasks), the amount of time needed for the assessment, and the amount of judgment involved in scoring (Gronlund, 2006). Traditional assessments such as multiple-choice, true-false, matching, completion, and short-answer items are often lower in realism and complexity of the tasks assessed, but require little time to administer and can be scored quickly and objectively (Gronlund, 2006). Alternative assessments such as portfolios, student self-assessment, observations, and other performance-based assessments are higher in both realism and complexity of the tasks assessed, but require large amounts of time to use, and that scoring is judgmental and less objective than traditional assessments (Gronlund, 2006).

There have been some reactions to the heavy emphasis on traditional assessments (Gronlund, 2006). These reactions highlight the importance of focusing more attention on the assessment of students’ application of understandings and skills to real problems in real-world contextual settings (Gronlund, 2006). More specifically, many educators have promoted a move away from traditional to alternative forms of assessment (e.g., Darling-Hammond, 1994; Shepard, 2000). Alternative 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 in the real world (Darling-Hammond, 1994). In comparison, traditional assessments have been criticized for being focusing on the product of learning rather than on the process of learning (Henning-Stout, 1994; Rudner & Boston, 1994). However, traditional methods of assessment are considered to be more objective and
reliable than alternative assessments (Worthen, 1993). Several studies utilizing survey techniques have examined teachers’ uses of traditional and alternative methods of assessment.

For example, in a survey of 893 teachers in 34 schools, Bol et al. (1998) investigated teachers’ frequent uses of traditional and alternative assessment methods in relation to teaching experience, grade level, and subject area. Results showed that teachers reported using performance and observation-based assessment methods most frequently. The most frequently used traditional methods of assessment were close-ended examinations, quizzes, and other written assignments. The most experienced teachers indicated the use of alternative assessment more often than the least experienced teachers. Also, elementary school teachers reported using alternative assessment more often than did middle and high school teachers. Mathematics teachers reported using traditional assessment methods much less frequently than did teachers in other subject areas.

In a related study of 625 K-12 teachers in the state of Ohio, Mertler (1998) examined the extent to which teachers used traditional versus alternative forms of assessment techniques. Mertler (1998) found no statistically significant group differences for the use of traditional techniques based on teacher’s gender, years of teaching experience, and school setting. However, female teachers tended to use alternative assessment techniques more frequently than male teachers. Also, urban school teachers indicated using alternative assessments more frequently than did rural school teachers. Like Bol et al.’s (1998) study, elementary school teachers reported using alternative assessments more frequently than did high school teachers. Unlike Bol et al.’s (1998)
study, however, the least experienced teachers reported using alternative assessments more frequently than the most experienced teachers. Both Bol et al.’s (1998) and Mertler’s (1998) studies overlooked the effects of teachers’ frequent uses of traditional and alternative assessments on classroom assessment environment as perceived by students. When this topic was considered, the hierarchical nested structure of the data was ignored.

For example, Snow-Renner (1998) studied teachers’ assessment practices in Colorado classrooms relative to students’ opportunities to learn. Survey responses were received from 737 mathematics and science teachers in grades 4, 8, and 10 as well as from 116 elementary school and 223 secondary school students. Elementary school teachers reported a greater emphasis on alternative assessments than did secondary school teachers. Students in different classrooms reported experiencing differential opportunities to learn relative to reform-oriented assessments, suggesting that teachers’ assessment practices may affect students’ perceptions of the classroom environment and that this effect may vary across classrooms. Snow-Renner (1998) attributed such results to fluctuations in teacher’s capacity and knowledge about assessment and to ambiguous policy definitions of assessment reforms in Colorado.

It should be noted that the objects of interest and measurement in Snow-Renner’s (1998) study were teachers and students. However, the analysis failed to take into account the nested structure of the data, in that students were nested within classrooms, thereby ignoring the consequences that clustering of the data might have for the relationships between variables of interest (teacher assessment practices and student
perceptions). Thus, the current study attempted to overcome this issue by not only considering both levels of the data, but also estimating the relations crossing the student-level and class-level. As implied by Snow-Renner’s (1998) study findings, it was expected in this study that significant variation would be found among class means on perceived classroom assessment environment.

In an investigation of classroom assessment practices of 246 third preparatory science teachers from 112 schools in Oman, Alsarimi (2000) found that teachers indicated using short answer, completion, oral exams, extended answer, and multiple-choice item formats with no significant differences based on teacher’s gender and years of teaching experience, thereby contradicting the findings obtained by Bol et al.’s (1998) and Mertler’s (1998) studies conducted in the United States with regard to the use of alternative assessments. These findings also suggest that Omani third preparatory science teachers in Alsarimi’s (2000) study tended to use traditional assessments more frequently than alternative assessments despite the argument that traditional assessments are often less engaging and less motivating for students to learn than alternative assessments (Frederiksen & Collins, 1989). However, Alsarimi (2000) did not examine the effects of teachers’ assessment practices on students’ perceptions of classroom assessment environment and motivational goals. The present study sought to extend Alsarimi’s (2000) study by addressing these effects.

In a survey of 158 teachers in Southern Maine’s schools about their assessment practices, McKenna (2003) found that primary and upper elementary teachers used different assessment techniques. Primary teachers used fewer traditional assessments than
upper elementary teachers. Student involvement in the classroom assessment process was limited at the elementary school level. The current study extended McKenna’s (2003) study by examining the effect of student-involved assessment, as part of the recommended classroom assessment practices, on students’ perceptions of classroom assessment environment and motivational goals in ninth grade Muscat science classrooms in Oman. It has been argued that classrooms in which students are involved and invited as partners in the assessment process tend to be more motivating for students to learn than are classrooms in which assessment is entirely a teacher-centered activity (Black & Wiliam, 1998a, 1998b; Crooks, 1988; Stiggins & Chappuis, 2005).

In a study of 297 teachers, Zhang and Burry-Stock (2003) investigated classroom assessment practices across teaching levels and content areas. Zhang and Burry-Stock (2003) found that middle school and high school teachers used paper-pencil tests more often than did elementary school teachers. In contrast, elementary school teachers reported using performance assessment more often than did middle school and high school teachers. Teachers in language arts, science, and social studies reported using paper-pencil tests more often than did mathematics teachers. Recently, Greenstein (2004) surveyed 115 teachers in two Connecticut high schools and interviewed a group of six teachers at each school about their assessment strategies. Results indicated that although teachers are more inclined to the use of traditional assessments, there was some increase in the use of alternative assessments.

The aforementioned studies reveal that teachers use various traditional and alternative types of assessment depending on their gender, teaching grade level, teaching
area, teaching experience, and school setting. Although the findings of these studies tended to be consistent with regard to teachers’ frequent use of traditional assessments, they were not conclusive regarding teachers’ frequent use of alternative assessments based on teacher’s gender and years of teaching experience. In addition to the context where the study was conducted such as Ohio versus Oman, this contradiction in the findings can be attributed to the lack of consistency in operationalizing the variable teaching experience and to the lack of focus on a specific instructional setting by combining participants from various geographic regions, subject areas, and grade levels. As such, unlike most of the previous studies, the current study focused on one grade level and one subject area within one educational region, which were ninth grade Muscat science classrooms in Oman, in order to have a closer examination of the classroom assessment practices. Nevertheless, like most of the previous research, the present research utilized survey methods in investigating teachers’ frequent uses of traditional and alternative assessments. Also, the current study added to the previous research the relations of teacher’s frequent uses of traditional and alternative assessments to students’ perceptions of classroom assessment environment and achievement goal orientations, considering both levels of the hierarchy, the student-level and the class-level, as crucially important.

Recommended Classroom Assessment Practices

The main purpose of classroom assessment is to improve student learning and motivation to learn (Gronlund, 2006; Harlen & Crick, 2003; Stipek, 2002). In this regard, Gronlund (2006) suggests that a sound classroom assessment requires a clear conception
of all intended learning outcomes of the instruction and a variety of assessment
procedures that are relevant to the instruction, adequately sample student performance,
and fair to everyone. In addition, a sound assessment requires the specifications of
criteria for judging successful performance and timely and detailed feedback to students
emphasizing strengths of their performance and weaknesses to be corrected (Gronlund,
2006).

Teachers are required to develop classroom assessment and grading that align
with practices recommended by experts of educational measurement and assessment. For
example, educational measurement and assessment experts have recommended that
students should clearly be informed about the grading procedure in advance and involved
in the assessment process (Stiggins, 1994; Stiggins & Chapuis, 2005; Stiggins et al.,
1989); student personal characteristics such as ability, effort, motivation, interest, and
neatness of work should not be incorporated into grading due to the lack of objective
measurement (Stiggins et al., 1989); a final grade for borderline cases should be
determined using additional academic achievement data rather than nonachievement data
(Stiggins et al., 1989); and students should be given continuous and informative
assessment feedback rather than judgmental feedback about their academic performance
(Brookhart, 1994).

Moreover, teachers are encouraged to use more than one assessment method in
order to have enough, accurate evidence of student learning (Nitko, 2001; Stiggins et al.,
1989). It is emphasized that the assessment should match the learning targets and provide
meaningful feedback to students (Nitko, 2001). In addition, the American Federation of
Teachers (AFT), the National Council on Measurement in Education (NCME), and the National Education Association (NEA) (1990) have jointly defined seven *Standards for Teacher Competence in Educational Assessment of Students*. The standards emphasized that teachers should competently be able to choose and develop assessment methods appropriate for instructional decisions; administer, score, and interpret results of externally produced and teacher-made assessment; use assessment results when making educational decisions; develop valid assessment-based grading procedures; communicate assessment results; and recognize unethical, illegal, and inappropriate methods and uses of assessment (AFT, NCME, & NEA, 1990).

Unfortunately, findings from past and recent studies of classroom assessment practices have consistently expressed a concern about the adequacy of teachers’ assessment and grading practices. For example, in an earlier survey of statistical analyses of test results for 336 elementary and secondary school teachers, Gullickson (1982) found that a substantial proportion of teachers reported using relatively little statistical information such as means, medians, and standard deviations to describe assessment results. Also, these same teachers did not have an adequate understanding of basic testing concepts such as item difficulty and reliability. Parallel to Gullickson’s (1982) study, Mertler (1998, 1999) found in two studies of 625 K-12 Ohio state teachers that teachers did not spend much time conducting statistical analyses of the assessment data with no significant differences based on teacher’s gender and years of teaching experience. Further, Hills (1991) identified four misuses of classroom assessment in schools including using grades for controlling students’ behavior, assigning grades that are
contingent on improvement, using tests that are technically inadequate, and deviation from established standardized-test administration procedures.

In a review of literature on teachers’ grading practices, Brookhart (1994) located 19 studies that were done since 1985. Seven studies focused on grading practices of secondary school teachers, 11 studies investigated both elementary and secondary school teachers, and one study included only elementary school teachers. Research methods employed in those studies included surveys in which teachers were asked about the components incorporated in term grades, grade distributions, and their beliefs about grading issues and grading scenarios; and observations, interviews, and document analyses. Based on this review, Brookhart (1994) concluded that:

1. Teachers try hard to be fair when assigning grades.
2. Teachers inform students about the components of the grades.
3. Achievement measures, particularly tests, are the main contributors to grades.
4. Teachers take into account student’s effort and ability in grading.
5. Elementary teachers depend on informal evidence and observations, whereas secondary teachers use paper-pencil achievement measures and other written activities in grading.
6. Teachers differ in their perceptions of the meaning and purpose of grades, and consideration of achievement and nonachievement factors in grading.
7. Teachers’ grading practices deviate from the recommendations of educational measurement and assessment experts.

Using survey and interview research techniques, Cizek et al. (1995) investigated sources of grading major assignments and tests and meaning of grades for 143 Midwestern elementary and secondary teachers attending an introductory master’s level
course in measurement and evaluation. Most teachers (83.8%) considered the percentage or number of correct on the assignment or test when assigning grades. Other considerations included student’s ability, performance of the entire class, student’s effort, and difficulty of the assignment or tests, which were endorsed by 51.5%, 43.4%, 41.9%, and 35.3% of the teachers, respectively. When asked about sources of information used to assign final grades, most teachers (89%) indicated using formal achievement measures such as tests, assignments, reports, and quizzes. Other sources included students’ conduct and impressions of their effort and teamwork as well as class attendance and participation that were endorsed by 61% and 52.2% of the teachers, respectively. Interviews with some of the teachers indicated that informal assessments of students based upon many observations would usually be more accurate than a lesser number of formal assessments. Cizek et al. (1995) concluded that teacher’s frequent use of recommended assessment practices did not vary with regard to teacher’s gender and years of teaching experience.

In a related survey research study of grading practices for 91 middle/high school science teachers, Feldman et al. (1998) found that teachers primarily used traditional forms of assessment to determine report card grades. Student’s work habits, class attendance, and behavior were reported as being rarely used. No statistically significant differences were found by teacher’s gender, years of teaching experience, and school’s geographic location on forms of assessment used in grading. Feldman et al. (1998) concluded that reform efforts on classroom assessment had little or no effect on the practices of the participating middle/high school science teachers.
Like both Cizek et al. (1995) and Feldman et al. (1998), Alsarimi (2000) found that science teachers indicated using four main sources of information when assigning grades to students in his survey study of 246 third preparatory science teachers from 112 schools in Oman. These sources were final exams, midterm exams, class participation, and oral questioning. Also, these same teachers tended to incorporate some nonachievement factors such as student’s effort in grading. The teachers commented that the grades reflect student improvement, effort, and knowledge of the subject matter.

Recently, Zhang and Burry-Stock (2003) surveyed 297 teachers across teaching levels and content areas about their classroom assessment practices. Zhang and Burry-Stock (2003) found that mathematics and science teachers reported grading on nonachievement factors more frequently than did teachers in social studies and nonacademic subjects. These findings together with Alsarimi’s (2000) study findings suggest that the expected participating ninth grade Muscat science teachers in this study may use nonachievement factors in grading.

Along similar lines, McMillan and Lawson (2001) investigated classroom assessment and grading practices of 213 secondary science teachers. Results indicated that teachers placed the greatest emphasis on academic performance and academic-enabling behaviors such as effort and much less emphasis on external benchmarks, extra credit, homework, and participation. Teachers of high ability students tended to use types of assessments, cognitive levels of assessments, and grading criteria that reflect the characteristics of performance assessment. In contrast, teachers of low ability students emphasized recall of knowledge and graded homework, and focused less on academic
achievement and higher order thinking when assessing and grading students. McMillan and Lawson (2001) recommended future research relating assessment and grading practices to student motivation be conducted to better understand teachers’ assessment practices. Thus, this study was intended to address McMillan and Lawson’s (2001) recommendation.

In conclusion, the aforementioned studies tend to confirm that classroom assessment practices may be unique from one teaching level and subject area to another as well as from one teacher to another. It is also evident from the classroom assessment literature that there seems to be some contradictions between teachers’ practices and recommendations of educational measurement and assessment experts regarding issues of classroom assessment and grading. Although the recommendations of educational measurement and assessment experts are based on well-founded theories of educational measurement and assessment (e.g., Airasian, 1996; Cureton, 1971; Ebel & Frisbie, 1986; Hopkins, Stanley, & Hopkins, 1990; Nitko, 2001; Stiggins, 1994; Stiggins et al., 1989; Stipek, 2002), investigating the motivational aspects of classroom assessment may help understand the lack of agreement between teachers’ practices and experts’ suggestions. Since teacher’s frequent use of recommended assessment practices has consistently not been found to vary as a function of teacher’s gender, years of teaching experience, grade level, subject area, and geographic region, only main and cross-level interaction effects for teacher’s frequent use of recommended assessment practices on perceived classroom assessment environment and achievement goals would be specified and tested in this
study. The hypothesized directional nature of these effects along with their theoretical and empirical support will further be discussed at the end of this chapter.

Most classroom assessment research has been toward what teachers do. Given that students’ views about their classroom work are important determinants of their motivation to learn (Brophy, 1999; Harlen & Crick, 2003), little empirical quantitative research attention has been devoted to how do students perceive their teachers’ classroom assessment practices. In an earlier review of research on testing practices, Haertel, Ferrara, Korpi, and Prescott (1984) found that little is known about students’ perceptions of classroom assessment. Additionally, in her review of research on teachers’ grading practices, Brookhart (1994) argued that “grading theory and practice will be better connected once the role of classroom assessment and grading practices in student achievement motivation… is understood” (p. 279). Similarly, McMillan and Workman (1998) noted that the new forms of assessment are based on theory and not empirical evidence. McMillan and Workman (1998) called for more research on the effect of classroom assessment on student motivation. As such, the present study was undertaken in an attempt to respond to these calls by investigating the possible relationships of teacher’s classroom assessment practices on student’s perceptions of classroom assessment environment and their achievement goal orientations.

Classroom Assessment Environment

Research on achievement motivation has shown that situational demands can affect the salience and adoption of specific achievement goals, which lead to differential patterns of cognition, affect, and behavior (Ames & Archer, 1988). Students are exposed
to a variety of instructional activities and assessment tasks in the classroom. As students process these events, they develop beliefs about the importance, utility, value, and difficulty of these tasks as well as their personal chance of success (McMillan & Workman, 1998). Educators have long recognized that the tasks used in the classroom communicate important messages to students about what is emphasized there, which in turn may influence their adoption of achievement goals (Ames, 1992b; Ames & Archer, 1988; Linnenbrink & Pintrich, 2001, 2002). For instance, classroom contexts that are organized toward challenge are likely to activate the need for achievement, which in turn may lead to adoption of mastery and/or performance-approach goals; whereas classroom contexts that are organized toward threat are likely to activate fear of failure, which in turn may lead to adoption of performance-avoidance and/or performance-approach goals (Elliot, 1999). For the purpose of clarification, the terms need for achievement and fear of failure represent the primary constructs of Atkinson and McClelland’s achievement motivation theory (Atkinson, 1957; McClelland, Atkinson, Clark, & Lowell, 1953) that distinguishes the need for achievement as an approach motive from the fear of failure as an avoidance motive (for more details see Elliot & Church, 1997).

Ames (1992a) described how aspects of classroom structure related to tasks, authority, recognition, grouping, evaluation, and time (TARGET, as it was originally identified by Epstein, 1988, 1989, as cited in Ames, 1992a) influence the salience of a mastery goal orientation, and as a consequence, elicit positive patterns of behaviors, beliefs, and affects in students. The Task (T) dimension is concerned with the design of learning activities, tasks, and assignments (Ames, 1992a). The Authority (A) dimension
refers to the degree to which teachers involve students in classroom decision making (Ames, 1992a). The dimension of Recognition (R) concerns the use of rewards in the classroom (Ames, 1992a). The Grouping (G) dimension regards the way in which students are divided into learning groups (Ames, 1992a). The Evaluation (E) dimension involves the methods, standards, and criteria used to assess student learning (Ames, 1992a). The Time (T) dimension concerns the appropriateness of the workload, the pace of instruction, and the time designated to complete classroom work (Ames, 1992a).

Ames (1992a) noted that the following classroom practices are likely to encourage adoption of mastery goals: (a) designing classroom tasks that include challenge, variety, novelty, and active involvement; (b) giving students opportunities to make choices and decisions regarding their learning; (c) providing private recognition and rewards that focus on individual student effort and improvement; (d) creating small groups of heterogeneous abilities that encourage working effectively with others on learning tasks and developing a feeling of belongingness; (e) conducting evaluation practices that are private, assess progress, improvement, and mastery, and avoid social comparisons; and (f) allowing for time on the assessment task to vary with the nature of the task and student needs. Conversely, performance-oriented classrooms are created when students are not given varied tasks, the teacher maintains authority, students are recognized for their ability relative to others, homogeneous ability groups are used, evaluation is based on normative practices, and time is inflexible to do the assessment task.
These practices are typically initiated by the classroom teacher. The overall sense or meaning that students make out of the various classroom assessment events constitutes the classroom assessment environment (Brookhart & DeVoge, 1999). Brookhart and her colleagues pointed out that each classroom has its own “assessment ‘character’ or environment” perceived by the students and springs from the teacher’s classroom assessment practices (Brookhart, 2004, p. 444; Brookhart & Bronowicz, 2003). The concept of classroom assessment environment was first introduced by Stiggins and Conklin (1992) as a result of their observations of the assessment practices of four teachers in three sixth grade classrooms. According to Stiggins and Conklin (1992), the classroom assessment environment includes eight key elements. These are assessment purposes, assessment methods, criteria for selecting these methods, quality of assessment, feedback on assessment results, teacher’s assessment background and preparation, teacher’s perception of students, and assessment policy (Stiggins & Conklin, 1992). It should be noted that Stiggins and Conklin’s (1992) conceptualization of classroom assessment environment centers more on teacher’s practices than on student’s perception of these practices (Brookhart & Durkin, 2003).

Of increasing interest to educational assessment community is the effect of classroom assessment on student motivation and learning. In this regard, Brookhart (1997a) developed a theoretical framework for the role of classroom assessment in motivating students. This framework was based on a synthesis of classroom assessment literature and motivation (Brookhart, 1997a). In this framework, classroom assessment environment is construed 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 (Brookhart, 1997a). There are four dimensions that underlie the classroom assessment environment: teacher’s attitudes toward students and subject matter, teacher’s use of assessments, teacher’s preparation in assessment, integration of assessment with instruction, and communication of assessment results (Brookhart, 1997a). Brookhart (1997a) postulated that students’ perceptions of the classroom assessment activities may influence their motivational beliefs and achievement, and that these perceptions are formed based in part on students’ experiences as a group in the classroom (Brookhart, 2004). This suggests that student perceived classroom assessment environment should deserve recognition and investigation as a valuable instructional outcome to be improved in the learning process.

Building on Brookhart’s (1997a) theoretical model and other motivational literature, McMillan and Workman (1998) have illustrated how particular assessment and grading practices increase or decrease student motivation. Specifically, McMillan and Workman (1998) explained that the following teacher’s assessment and grading practices may enhance student motivation to learn (pp. 22 – 23): (a) being clear about how learning will be evaluated, (b) providing specific feedback following an assessment activity, (c) using mistakes to show students how learning can be improved, (d) using moderately difficult assessments, (e) using many assessments rather than a few major tests, (f) using authentic assessment tasks, (g) using preestablished criteria for evaluating student work, (h) providing incremental assessment feedback, and (i) providing scoring criteria prior to administering the assessment task.
Along similar lines, Stiggins (1999) as well as Stiggins and Chappuis (2005) contend that day-to-day classroom assessment can be used in more productive ways to motivate students to learn and increase their learning confidence. Stiggins and Chappuis (2005) described four conditions that together may foster positive motivational patterns for students. These conditions state that classroom assessments should focus on clear purposes, provide accurate reflections of achievement, provide frequent descriptive feedback on work improvement rather than judgmental feedback, and involve students in the assessment process (Stiggins & Chappuis, 2005).

In a study investigating part of Brookhart’s (1997a) theoretical model, Brookart and DeVoge (1999) used observation, survey, and interview techniques to collect data from four classroom assessment events in two third-grade language arts classes. Results revealed positive relationships among assessment task’s characteristics as perceived by students, their perceptions of ability to do the task, effort, and achievement. During the interviews, students expressed the importance of the assessment task in accordance with their goal orientations. Some students indicated the importance of the task in terms of its value for learning, which is consistent with the mastery goal orientation. Other students expressed task importance in terms of getting good grades, which is consistent with the performance goal orientation. At the conclusion, Brookhart and DeVoge (1999) suggested that students’ goal orientations should be considered when studying the impact of classroom assessment on student motivation. Unlike Brookhart and DeVoge’s (1999) study that focused on the third grade language arts, the present study focused on the ninth grade science. The study explored the possible effects of classroom assessment
environment as perceived by ninth grade students in Muscat science classrooms in Oman on achievement goals utilizing survey research methods.

Although the focus was on college level students, Church et al. (2001) conducted two studies to examine the relationships among perceptions of classroom environment and adoption of achievement goals for undergraduate students enrolled in chemistry classes. Perceptions of classroom environment were represented by three dimensions. These were lecture engagement, defined as “the extent to which students perceive that the professor makes the lecture material interesting” (p. 44); evaluation focus, defined as “the degree to which students perceive that the professor emphasizes the importance of grades and performance evaluation” (p. 44); and harsh evaluation, defined as “the extent to which students view the grading structure as so difficult that it minimizes the likelihood of successful performance” (p. 44). Achievement goals were conceptualized in terms of mastery, performance-approach, and performance-avoidance goals.

Results from the two studies conducted by Church et al. (2001) indicated that mastery goals were positively related to lecture engagement and negatively related to both evaluation focus and harsh evaluation, performance-approach goals were positively related to evaluation focus, and performance-avoidance goals were positively related to both evaluation focus and harsh evaluation. When perceived classroom environment and achievement goal variables were tested together as predictors of graded performance and intrinsic motivation, results showed that the perceived classroom environment influenced adoption of achievement goals, which in turn directly influenced graded performance and intrinsic motivation after controlling for student’s gender, competence valuation, and
SAT scores. Church et al. (2001) concluded that these findings suggest that educators need to understand that classroom learning environments featuring stringent evaluative standards may represent a risk factor in the achievement domain.

It should be pointed out that despite a common classroom learning environment, Church et al. (2001) noted that their analyses were based on the notion that each student perceives and responds to the classroom events differently. When objective aspects of the classroom environment were measured by using the mean levels of students’ perceptions within classrooms as a proxy, no average classroom perception variable was found to be a statistically significant predictor of achievement goals (Church et al., 2001). However, this finding may be due to a small sample size of classrooms ($N = 9$). The present study attempted to overcome this issue by having a large sample size of classrooms.

Recently, Greene et al. (2004) used a path analysis technique to test predictions of a model explaining the impact of students’ perceptions of classroom environment on their achievement goal orientations in English classes for 220 high school students. Greene et al. (2004) found that students who perceived the tasks assigned in the class as meaningful and motivating tended to endorse mastery goals. However, mastery goals were not positively predicted by either perceptions of the classroom environment as supporting autonomy and having a mastery-oriented evaluation. Also, students’ perceptions of the classroom environment as having a mastery-oriented evaluation did not have a significant negative relationship with performance-approach goals. Although Brophy (2005) argued that performance-approach goals are counterproductive in the long run, results of Greene et al.’s (2004) study imply that adoption of performance-approach goals may not be
problematic. Apparently, more research is needed to explore the nature of performance goals. Hence, the present study attempted to address this need.

Although the focus was on college level students, factor analyses conducted by Wang (2004) yielded three distinct dimensions of classroom assessment environment as perceived by 503 first year Chinese EFL students: learning-oriented, test-oriented, and praise-oriented classroom assessment environment. In her factor analysis of achievement goals, mastery goals were differentiated to form mastery-external and mastery in-class goals. Wang (2004) found that the learning-oriented classroom assessment environment positively predicted performance-approach goals and both forms of mastery goals. Classroom assessment environment that is test-oriented positively predicted performance-avoidance goals and negatively predicted mastery-external goals. Praise-oriented classroom assessment environment positively predicted performance-approach goals. Student’s gender mediated the relationships of perceived classroom assessment environment to performance-approach and mastery in-class goals. Specifically, 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. Also, mastery in-class 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.

The implication of both Church et al.’ (2001) and Wang’s (2004) studies at the college level for the current study was that student’s gender may also need to be considered in the study of classroom assessment environment and achievement goals at
the middle school level. It should be noted, however, that gender in this study varied across classes because in Oman, students within the same class and their science teacher are of the same gender, either all of them are males or all of them are females. Therefore, unlike previous studies, gender in this study was more appropriately to be treated as an independent variable at the class-level.

As depicted in Brookhart’s (1997a) theoretical framework of classroom assessment and motivation, teacher’s use of different forms of assessment is one component of the classroom assessment environment. As noted earlier, there has been a movement toward the use of more alternative assessments than traditional assessments (Darling-Hammond, 1994; Helgeson, 1992; Shepard, 2000). The arguments in favor of alternative assessments over traditional assessments are based on the notion that alternative assessments are more motivating than traditional assessments. In particular, proponents of alternative assessments claim that alternative assessments are more authentic, engaging, challenging, involve group and individual work, encourage higher cognitive levels, and provide opportunities for student involvement (Baker, O’Neill, & Linn, 1993; Darling-Hammond, 1994; Frederiksen & Collins, 1989; McMillan & Workman, 1998). However, these arguments are based on theory and little empirical research exists about the effects of assessment type on achievement goal orientations and classroom assessment environment as perceived by students.

Maslovaty and Kuzi (2002) examined differences in perception and adoption of ability and task motivational goals among 261 sixth grade students in eight math classrooms in four schools in Israel. Two of the schools mainly applied traditional
assessment and the other two mainly applied alternative assessment. Results showed that students who were assessed through alternative methods reported more enjoyment, challenge, improvement, and interest than students who were assessed through traditional methods. Also, students who experienced alternative assessment class climates tended to adopt task goals, whereas those who experienced traditional assessment class climates tended to adopt ability goals. The implication of Maslovaty and Kuzi’s (2002) study for the current study was that their results support the theoretical perspectives that motivational aspects of alternative assessments may foster adoption of mastery-oriented outcome beliefs (Ames, 1992a; McMillan & Workman, 1998), which was examined in this study for ninth grade students in Muscat science classrooms in Oman.

In a study of classroom assessment and motivation of 96 high school students enrolled in social studies classes, Brookhart and Durkin (2003) examined relations among student perceptions of the assessment tasks, self-efficacy, goal orientations, and type of assessment used for a variety of classroom assessment events. Results indicated that correlations between mastery goal orientations, self-efficacy, and perceived assessment task characteristics were positive and moderate in strength for all classroom assessment events. The more students perceived the assessment task to be important and interesting, the more likely they were to report wanting to learn and able to do it. Perceptions of mastery goal orientations were positively associated with perceptions of performance goal orientations for all performance assessments. However, correlations between performance goal orientation and mastery goal orientation were either weak or negative for traditional paper-pencil tests. Therefore, these results illustrate that personal (i.e.,
individual) adoption of performance goals may not be detrimental as far as the overall classroom assessment climate is perceived as mastery-oriented. It should be noted that goal orientations in Brookhart and Durkin’s (2003) study were subjectively assessed using content analyses of student interviews and implications of students’ responses on other instruments thereby introducing biases in the measurement of achievement goals. The present study attempted to overcome this issue by using objective measures that have been validated for middle school students.

Similarly, Stefanou and Parkes (2003) studied the effects of classroom assessment on students’ motivational goals. Seventy nine students in three fifth grade science classes, taught by the same teacher, were exposed to three different classroom assessment conditions: traditional paper-pencil tests, a laboratory task assessment, and a performance assessment. Results revealed a statistically significant effect for assessment type on goal orientation. Traditional paper-pencil tests and performance assessment fostered more task-focused orientation than the laboratory assessment. Interviews with students showed that the goal orientation fostered by traditional paper-pencil tests seemed to be a performance-approach orientation.

Unlike Stefanou and Parkes’ (2003) study that focused on the fifth grade science classrooms, the current study focused on the ninth grade students in Muscat science classrooms in Oman. Also, although Stefanou and Parkes’ (2003) experimental research design had the strength of having the same teacher across the groups, the findings had a limited ecological validity because the study was not part of the normal classroom work (Black & Wiliam, 1998a; Crooks, 1988). Therefore, the present study examined the
effects of teachers’ frequent uses of traditional and alternative assessments on students’ perceptions of classroom assessment environment and achievement goals using a survey research method within a correlational research design. In light of classroom assessment literature and achievement goal theory (Ames, 1992a, 1992b; McMillan & Workman, 1998), it was expected in this study that classes with a high emphasis on traditional assessments would be associated with performance-related outcomes more than mastery-related outcomes. The opposite was expected to be true for classes with a high emphasis on alternative assessments. In addition, since laboratory assessment tasks are one kind of alternative assessments, their effects on student motivational goals were not clear in Stefanou and Parkes’ (2003) study, thereby casting doubt about the effect of alternative assessments on achievement goals. Therefore, the current study sought to clarify this relationship by including performance assessments, both structured and unstructured, within alternative assessments used by the teacher following Gronlund’s (1998) definitions of performance and alternative assessments provided earlier in Chapter I of the dissertation.

In a recent relevant study of interrelationships among teacher assessment practices, feedback, student self-efficacy, effort, and performance on a large scale assessment, Rodriguez (2004) analyzed data collected from 6963 seventh and eighth grade students nested within 326 mathematics teachers who participated in the Third International Mathematics and Science Study (TIMSS) in the United States. Results indicated that at the classroom-level, teacher’s use of teacher-made objective tests and his or her use of assessment information for grading and evaluation rather than for feedback
and discussion had significant negative relationships to classroom performance. Frequent use of teacher-made objective tests, at the classroom-level, neutralized the positive relationship between self-efficacy and performance at the student-level. Rodriguez (2004) concluded that classroom assessment practices might uniquely interact with student characteristics in their role of motivating student effort and performance. The implication of Rodriguez’s (2004) study for the present study is that teacher’s assessment practices (as class-level characteristics) may affect the relationships of student’s academic self-efficacy to both student’s perceived classroom assessment environment and student’s achievement goal orientations at the student-level. These cross-level interaction effects would be specified in the models explaining student’s perceived classroom assessment environment and student’s achievement goal orientations and would be tested in this study.

Overall, a limited number of studies have demonstrated that classroom assessment practices may influence students’ perceptions of classroom assessment environment and achievement goal orientations. However, most of these studies tended to overlook the hierarchical structure of the data, in that students were nested within classrooms. As might be expected, ignoring the nested nature of the data may lead to statistical and conceptual problems such as unit of analysis, violation of independence assumption, and loss of information (Kasim, 1994; Raudenbush & Bryk, 2002). It may be argued that students within a classroom may share a common experience of the classroom assessment events and as such even though students respond differently to the same classroom assessment process, their responses may have commonality (Kasim, 1994). It should be
noted that the central issue in this area of research is students’ collective perceptions (i.e., class average perception) of the classroom assessment practices. Therefore, their responses should be aggregated at the classroom level, and hence identified as a classroom characteristic.

A number of empirical evidences exist to support this argument. First, in a study of 516 sixth grade students situated in 63 math classrooms, Ryan et al. (1998) found a small within-class variance in students’ perceptions of classroom goal structure, a situation that suggested some degree of consensus among students’ perceptions within classrooms. Hence, it seemed reasonable to average students’ responses about their classroom (Ryan et al., 1998). Similarly, Kaplan, Gheen, and Midgley (2002a) investigated whether classroom goal structure as perceived by students is related to student disruptive behavior. A sample of 388 ninth grade students nested within 60 math classrooms responded to surveys about their classroom goal structure and their involvement in classroom disruptive behavior. Prior to the analyses, Kaplan et al. (2002a) found a large degree of dependency among students’ perceptions of classroom goal structure within classrooms, and hence Kaplan et al. (2002a) aggregated students’ responses to the classroom-level and identified them as classroom-level predictors.

Likewise, in a recent hierarchical study of 1,571 students nested within 84 math teachers in grades 5 through 12, Deevers (2005) found that students’ perceptions of teacher practices that press students academically and endorse mastery goals within the classroom had significant positive effects on student’s adoption of mastery goals. Therefore, classroom assessment research needs to take advantage of the hierarchical
linear modeling (HLM) to examine the impact of classroom assessment practices on student motivation. Hence, the present study was undertaken to address this need.

Summary

Achievement goal theorists believe that the reasons students engage in academic tasks (also called achievement goals) have implications for how and what they learn (Stipek, 2002). The research reviewed in this chapter clearly points to the importance of fostering desired achievement goals for students in the classroom learning environment. Yet, from a socio-cognitive perspective, students “are not social isolates of the influence of those around them” (Bandura, 1997, p. 469). It has been suggested that teachers’ classroom assessment practices and students’ self and aggregate perceptions of these practices are critical to student motivation for learning.

Although there is an abundance of theory and opinion of the role of classroom assessment in student motivation, there has been a paucity of empirical research about the impact of classroom assessment practices on students’ perceptions of classroom assessment environment and achievement goal orientations (Harlen & Crick, 2003). Thus, the current study was expected to address the potential contextual effects of classroom assessment on student’s perceived classroom assessment environment and achievement goal orientations that have been largely not addressed in the previous studies. A multilevel analysis would be employed to avoid conceptual and statistical problems that sometimes may compromise results of single-level analyses of data. Also, the present study attempted to take into account some of the previous studies’ limitations
in terms of small sample size, lack of focus on a specific instructional setting, and inadequate instrumentation.

Synthesis of the Literature and Research Questions

The research problem addressed with this study was the development of models that might explain the variation in student perceived classroom assessment environment and achievement goal orientations for ninth grade students in Muscat science classrooms in Oman. Classroom assessment literature and achievement goal theory would guide the prediction of the final models through the identification of variables that may affect student’s perceived classroom assessment environment and adoption of achievement goals. Following is a summary of the theoretical and empirical support of the predicted final models. The summary is divided into four sections addressing the ten research questions shaping this study.

Research Question One

RQ (1): What are the underlying dimensions of ninth grade students’ perceptions of classroom assessment environment in Muscat science classrooms in Oman?

Brookhart (1997a) developed a theoretical framework for the role of classroom assessment in student motivation. In this framework, classroom assessment environment is construed 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. The present study focused on only three aspects of the classroom assessment environment: assessment tasks, assessment standards and criteria, and assessment feedback. This was because these three aspects were assumed to
be more directly related to students’ experiences of their classroom assessment environment (Crooks, 1988; Wang, 2004). However, with the exception of Wang’s (2004) study, there was no clear operational definition for the perceived classroom assessment environment.

Therefore, in this study, the classroom assessment literature (e.g., McMillan & Workman, 1998) was consulted to develop questionnaire items pertaining to the aforementioned three aspects of classroom assessment environment. In light of the achievement goal theory (e.g., Ames, 1992a, 1992b), it was expected that principal components/exploratory factor analyses (PCA/EFA) would yield two independent dimensions of the perceived classroom assessment environment (a) a learning-oriented and (b) a normative-oriented classroom assessment environment whether rotated orthogonally or obliquely. The perceived learning-oriented classroom assessment environment refers to the extent to which students believe that in their class: (a) many assessment tasks exist that are interesting, authentic, and within students ability; (b) assessment standards and criteria are clear, fair, absolute, attainable, and communicated to students in advance; and (c) assessment feedback is individualized and informative. The perceived normative-oriented classroom assessment environment refers to the extent to which students believe that in their class: (a) little assessment tasks exist that are boring, less meaningful, and either difficult or easy; (b) assessment standards and criteria are not clear, unattainable, relative, and not communicated to students in advance; and (c) assessment feedback is public, judgmental, and emphasizes social comparisons.
Due to the need for orthogonality in subsequent analyses, scale scores as suggested by Hair, Anderson, Tatham, and Black (1998) would be created by summing scores on the items for each dimension of the best factor model revealed by the PCA/EFA. Each student would have a scale score for each dimension of the perceived classroom assessment environment. Higher scores on each dimension would represent positive perceptions of the classroom assessment environment defined in that dimension.

Research Question Two

RQ (2): What are the underlying dimensions of ninth grade students’ achievement goal orientations in Muscat science classrooms in Oman?

Since the origin of achievement goal theory in the late 1970s and early 1980s (Kaplan et al., 2002b), several frameworks of achievement goals have been developed. Examples of these frameworks include a dichotomous framework of mastery and performance goals (Ames & Archer, 1987, 1988; Dweck, 1986; Nicholls, 1984); a trichotomous framework of mastery, performance-approach, and performance-avoidance goals (Elliot & Church, 1997; Elliot & Harackiewicz, 1996; Midgley et al., 1998); and a multiple goal framework of mastery/performance-approach combined and performance-avoidance goals (Barron & Harackiewicz, 2001; Pintrich, 2000b). This study adopted the trichotomous framework of achievement goals because it was assumed to be the most prevalent goal framework in achievement settings (Elliot, 1999; Elliot & Church, 1997; Elliot & Thrash, 2001). The study would use a previously developed and validated achievement goal questionnaire for middle and high school students in the United States (Midgely et al., 2000) to assess each of the three goals in the trichotomous framework.
Due to changes in the rating scale from a 5-point Likert scale to a 4-point Likert scale as recommended by Assor and Connell (1992), which is discussed in Chapter III of the dissertation, as well as changes in the wording, language, and context where the items would be used, a PCA/EFA described in Chapter III of the dissertation would be conducted on the items. Consistent with the trichotomous framework of achievement goal theory, it was expected that the PCA/EFA in this study would yield the three dimensions of achievement goals (mastery, performance-approach, and performance-avoidance) with (a) a positive correlation between the two types of performance goals and (b) no correlations between mastery goals and each of the two types of performance goals whether rotated orthogonally or obliquely.

Due to the need for replication in other studies of achievement goal theory, summated scales as suggested by Hair et al. (1998) would be created by averaging scores on the items for each dimension of the best factor model revealed by the PCA/EFA. Each student would have a summated scale score for each dimension of the achievement goals. Higher scores on each dimension would represent a higher adoption or endorsement of the achievement goal defined in that dimension.

Research Questions Three, Four, Five, and Six

RQ (3): Do ninth grade Muscat science classrooms in Oman vary in perceived classroom assessment environment (PCAE)?

RQ (4): What are some of the student characteristics (e.g., academic self-efficacy) that might have effects on PCAE?

RQ (5): Do the effects of student characteristics (e.g., academic self-efficacy) on PCAE vary across classrooms?
RQ (6): What are some of the classroom characteristics (e.g., class gender) that might help explain the variability in the PCAE and in the effects of student characteristics (e.g., academic self-efficacy) on PCAE across classrooms?

The goal of these research questions is to construct a parsimonious model explaining student’s perceived classroom assessment environment as a function of student-level and class-level characteristics. The data pertaining to these research questions are hierarchically structured, in that students are nested within classes. Therefore, in order to adequately take the nested nature of the data into account that has been ignored in some of the past studies (i.e., Snow-Renner, 1998), hierarchical linear model (HLM) analyses (Raudenbush & Bryk, 2002) would be conducted, one for each dependent variable defined as a dimension of student’s perceived classroom assessment environment that would emerge from the PCA/EFA described in the first research question above. As noted in the first research question, two dimensions of student’s perceived classroom assessment environment were expected to emerge from the PCA/EFA in this study: (a) a learning-oriented and (b) a normative-oriented classroom assessment environment.

The literature review was used to identify the variables that have been theoretically and/or empirically hypothesized to be related to student’s perceived classroom assessment environment. These variables would be used to specify the expected final explanatory model of student’s perceived classroom assessment environment. As will be discussed more fully in Chapter III of the dissertation, the modeling process would be the same for each expected dimension of student’s perceived classroom assessment environment. Thus, for the purpose of illustration, following is a
summary of the theoretical and empirical support for the variables identified in this study from the literature to be related to student’s perceived learning-oriented classroom assessment environment.

1. Among middle and high school students, Anderman and Midgely (1997) and Greene et al. (2004) found that high self-efficacy beliefs held by students tended to be positively associated with perceptions of the classroom environment as being learning-oriented. Armed with these findings and consistent with Bandura’s (1986) social-cognitive theory and Elliot’s (1999) review of achievement goal theory in that high self-efficacy beliefs are related to approach motivation, I hypothesized/predicted in this study that student’s academic self-efficacy would on average be positively related to perceived learning-oriented classroom assessment environment within classes.

2. Brookhart (2004, p. 444) indicated that each class has “an assessment ‘character’ or environment” that springs from the teacher’s assessment practices. Likewise, Brookhart and Bronowicz (2003, p. 222) stated that “Assessments differ from class to class…because of different instructional experiences.” Empirically, Snow-Renner (1998) found that students in different classrooms reported experiencing differential opportunities to learn relative to their respective teacher’s assessment practices, suggesting that student’s perceptions of the classroom assessment environment may vary across classrooms. Given these theoretical and empirical observations along with the aforementioned reported relationship between student’s academic self-efficacy and student’s perceived
classroom learning environment (e.g., Anderman & Midgley, 1997; Greene et al., 2004), it was expected in this study that (a) there would be a significant variation among class means on perceived learning-oriented classroom assessment environment and that (b) the relationship between student’s academic self-efficacy and student’s perceived learning-oriented classroom assessment environment would vary significantly across classes.

3. Among middle school students, research has shown that girls tended to hold higher science self-efficacy beliefs (e.g., Britner & Pajares, 2006) and more positive perceptions of their classroom environment (e.g., Anderman & Midgley, 1997) than boys. Gender in this study was a class-level variable and as such based on this previous research it was expected that holding other variables constant, there would be a two-way interaction effect between class’s gender and class average academic self-efficacy on both (a) the class average perceived learning-oriented classroom assessment environment and (b) the relationship between student’s academic self-efficacy and student’s perceived learning-oriented classroom assessment environment (i.e., the differentiating effect of academic self-efficacy). Specifically, holding other variables constant, the difference in class average perceived learning-oriented classroom assessment environment between high versus low efficacious classes would be weaker in female classes than in male classes. Also, holding other variables constant, the differentiating effect of student’s academic self-efficacy on student’s perceived learning-oriented classroom assessment environment would be weaker in high efficacious female
classes than in low efficacious female classes, whereas the differentiating effect of
student’s academic self-efficacy on student’s perceived learning-oriented
classroom assessment environment would be stronger in high efficacious male
classes than in low efficacious male classes.

4. From a socio-cognitive perspective, students “are not social isolates of the
influence of those around them” (Bandura, 1997, p. 469). As such, Brookhart
(2004, p. 444) maintained that “classes have an assessment ‘character’ or
environment” that springs 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 [i.e., class] experiences” (p. 445).
Further, as mentioned above, middle school girls tend to have more positive
perceptions of their classroom environment as being learning-oriented than boys
(e.g., Anderman & Midgley, 1997). Gender in this study was a class-level
variable and as such based on these theoretical and empirical observations along
with the aforementioned reported relationship between student’s academic self-
efficacy and student perceived classroom environment (e.g., Greene et al., 2004),
it was expected in this study that holding other variables constant, there would be
a two-way interaction effect between class’s gender and class average perceived
normative-oriented classroom assessment environment on (a) the class average
perceived learning-oriented classroom assessment environment and (b) the
relationship between student’s academic self-efficacy and student’s perceived
learning-oriented classroom assessment environment. Specifically, holding other
variables constant, the difference in class average perceived learning-oriented classroom assessment environment between high versus low normative-oriented classes would be weaker in male classes than in female classes. Also, holding other variables constant, the differentiating effect of student’s academic self-efficacy on student’s perceived learning-oriented classroom assessment environment would be stronger in high normative-oriented female classes than in low normative-oriented female classes, whereas the differentiating effect of student’s academic self-efficacy on student’s perceived learning-oriented classroom assessment environment would be weaker in high normative-oriented male classes than in low normative-oriented male classes.

5. Traditional assessments has been theoretically criticized for being focusing on the product of learning rather than on the process of learning (Henning-Stout, 1994) and for assessing superficial rather than deep learning (Harackiewicz, Barron, Tauer, Carter, & Elliot, 2000). As stated by Harackiewicz et al. (1994, p. 317) “Either of these factors may create a performance-oriented classroom environment”. These theoretical arguments have led me to hypothesize/predict that holding other variables constant, teacher’s frequent use of traditional assessments would be negatively related to class average perceived learning-oriented classroom assessment environment. Also, holding other variables constant, the relationship between student’s academic self-efficacy and student’s perceived learning-oriented classroom assessment environment would be weaker
in classes with a high emphasis on traditional assessments than in classes with a low emphasis on traditional assessments.

6. Teachers are required to follow classroom assessment practices recommended by experts of educational measurement and assessment (AFT, NCME, & NEA, 1990). Further, Ames (1992a), McMillan and Workman (1998), Stiggins (1999), and Stiggins and Chappius (2005) have theoretically suggested that teacher’s adherence to the recommended classroom assessment practices concerning assessment revision, communication of assessment, assessment standards and criteria, student-involved assessment, and nonachievement-based grading may enhance student motivation for learning. Consistent with these theoretical perspectives, it was expected that holding other variables constant, teacher’s frequent use of recommended assessments would be positively related to class average perceived learning-oriented classroom assessment environment. Also, holding other variables constant, the relationship between student’s academic self-efficacy and student’s perceived learning-oriented classroom assessment environment would be stronger in classes with a high adherence to the recommended assessments than in classes with a low adherence to the recommended assessments.

7. Alternative assessments have been thought to be more authentic, engaging, challenging, and focusing on deep rather than superficial understanding than traditional assessments (Baker et al., 1993; Shepard, 2000). In light of achievement goal theory, the use of alternative assessments should create a
learning-oriented classroom assessment environment (Ames, 1992a, 1992b; Maslovaty & Kuzi, 2002). However, classroom assessment research has shown significant differences between teachers with regard to teacher’s frequent use of alternative assessment based on teacher’s gender and teacher’s years of teaching experience (for more details see Bol et al., 1998; Mertler, 1998). In this study, it was argued that the most experienced teachers not only might have learned through experience how to enhance student motivation for learning, but also might have observed through experience the hypothesized positive effects of using alternative assessments on the classroom environment, and hence they use them more frequently than the least experienced teachers.

These findings and arguments have led me to hypothesize/predict that holding other variables constant, there would be a three-way interaction effect between class’s gender, teacher’s years of teaching experience, and teacher’s frequent use of alternative assessments on both (a) the class average perceived learning-oriented classroom assessment environment and (b) the relationship between student’s academic self-efficacy and student’s perceived learning-oriented classroom assessment environment. Specifically, holding other variables constant, it was expected that for classes with a high emphasis on alternative assessments, the difference in class average perceived learning-oriented classroom assessment environment between high versus low levels of teaching experience would be weaker in female classes than in male classes. For classes with a low emphasis on alternative assessments, the difference in class average perceived learning-oriented classroom assessment environment between high versus low levels of
teaching experience would be stronger in female classes than in male classes. Also, holding other variables constant, it was expected that for classes with a high emphasis on alternative assessments, the differentiating effect of student’s academic self-efficacy on student’s perceived learning-oriented classroom assessment environment would be (a) weaker in female classes having a high experienced teacher than in female classes having a low experienced teacher and (b) stronger in male classes having a high experienced teacher than in male classes having a low experienced teacher. For classes with a low emphasis on alternative assessments, the differentiating effect of student’s academic self-efficacy on student’s perceived learning-oriented classroom assessment environment would be (a) stronger in female classes having a high experienced teacher than in female classes having a low experienced teacher and (b) weaker in male classes having a high experienced teacher than in male classes having a low experienced teacher.

Taking together, the expected final combined explanatory model of student’s perceived classroom assessment environment was as follows:

\[
(SPACE)_{ij} = \gamma_0 + \gamma_1 (GNDR \times CEFC) + \gamma_2 (GNDR \times CPCAE) + \gamma_3 (TRAD) + \\
\gamma_4 (RECOM) + \gamma_5 (GNDR \times TEXP \times ALTR) + \gamma_6 (SEFC_{ij} - CEFC_{ij}) + \\
\gamma_7 (GNDR \times CEFC) (SEFC_{ij} - CEFC_{ij}) + \gamma_8 (GNDR \times CPCAE) (SEFC_{ij} - CEFC_{ij}) + \\
\gamma_9 (TRAD) (SEFC_{ij} - CEFC_{ij}) + \gamma_{10} (RECOM) (SEFC_{ij} - CEFC_{ij}) + \\
\gamma_{11} (GNDR \times TEXP \times ALTR) (SEFC_{ij} - CEFC_{ij}) + u_{ij} + u_{ij} (SEFC_{ij} - CEFC_{ij}) + \epsilon_{ij}.
\]

Each symbol and term in this model along with the modeling process will be explained in Chapter III of the dissertation.
Research Questions Seven, Eight, Nine, and Ten

RQ (7): Do ninth grade Muscat science classrooms in Oman vary in achievement goal orientations (GOAL)?

RQ (8): What are some of the student characteristics (e.g., academic self-efficacy) that might have effects on GOAL?

RQ (9): Do the effects of student characteristics (e.g., academic self-efficacy) on GOAL vary across classrooms?

RQ (10): What are some of the classroom characteristics (e.g., class gender) that might help explain the variability in the GOAL and in the effects of student characteristics (e.g., academic self-efficacy) on GOAL across classrooms?

The goal of these research questions is to construct a parsimonious model explaining student’s achievement goal orientations as a function of student-level and class-level characteristics. The data pertaining to these research questions are hierarchically structured, in that students are nested within classes. Therefore, in order to adequately take the nested nature of the data into account that has been ignored in some of the past studies (i.e., Wang, 2004), hierarchical linear model (HLM) analyses (Raudenbush & Bryk, 2002) would be conducted, one for each dependent variable defined as a dimension of student’s achievement goal orientations that would emerge from the PCA/EFA described in the second research question above. As noted in the second research question, three dimensions of student’s achievement goal orientations were expected to emerge from the PCA/EFA: (a) mastery, (b) performance-approach, and (c) performance-avoidance goals in accord with the trichotomous framework of achievement goal theory (Elliot & Church, 1997; Midgley et al., 1998, 2000).

The literature review was used to identify the variables that have been theoretically and/or empirically hypothesized to be related to student’s achievement goal
orientations. These variables would be used to specify the expected final explanatory model of student’s achievement goal orientation. As will be discussed more fully in Chapter III of the dissertation, the modeling process would be the same for each dimension of student’s achievement goals. Thus, for the purpose of illustration, following is a summary of the theoretical and empirical support for the variables identified in this study from the literature to be related to student’s achievement goals only in the mastery dimension.

1. According to Bandura’s (1986) social-cognitive theory and Elliot’s (1999) review of achievement goal theory, self-efficacy is assumed to be a precursor to adoption of an approach achievement goal. Among middle and high school students, self-efficacy was found to be positively related to a mastery goal orientation (e.g., Deevers, 2005; Middleton & Midgely, 1997). Consistent with Bandura and Elliot theoretical views along with previous research findings, it was expected in this study that holding other variables constant, student’s academic self-efficacy would on average be positively related to the mastery goal orientation within classes.

2. Theoretically and empirically, it has been repeatedly documented that classroom environment perceived by students as having a strong emphasis on mastery pursuits prompt the adoption of mastery goals for students (e.g., Ames, 1992a, 1992b; Church et al., 2001). Consistent with these findings, it was expected in this study that holding other variables constant, student’s perceived learning-oriented
classroom assessment environment would on average be positively related to student’s mastery goal orientation within classes.

3. Achievement goals and self-efficacy are thought to co-vary based on situational demands (Bandura, 1977, 1982; Maehr, 1983, 1984). In addition, the relationships between student’s academic self-efficacy and student’s adoption of a mastery goal orientation as well as between student’s perceived learning-oriented classroom environment and student’s adoption of a mastery goal orientation were found to vary across classes (Deevers, 2005). Consistent with these perspectives and findings, it was expected that in this study (a) there would be a significant variation among class means on the adoption of a mastery goal orientation, (b) the relationship between student’s academic self-efficacy and student’s adoption of a mastery goal orientation would vary significantly across classes, and that (c) the relationship between student’s perceived learning-oriented classroom assessment environment and student’s adoption of a mastery goal orientation would vary significantly across classes.

4. It has been found that female adolescents are more oriented to mastery goals than male adolescents (e.g., Anderman & Midgley, 1997). Also, middle school girls tended to hold stronger science self-efficacy beliefs than boys (e.g., Britner & Pajares, 2006). Gender in this study was a class-level variable and as such based on these previous research findings it was expected that holding other variables constant, there would be a two-way interaction effect between class’s gender and class average academic self-efficacy on (a) the class average mastery goal
orientation and (b) the relationship between student’s academic self-efficacy and student’s mastery goal orientation (i.e., the differentiating effect of student’s academic self-efficacy). Specifically, holding other variables constant, the difference in class average mastery goal orientation between high versus low efficacious classes would be weaker in female classes than in male classes. Also, holding other variables constant, the differentiating effect of student’s academic self-efficacy on student’s mastery goal orientation would be weaker in high efficacious female classes than in low efficacious female classes, whereas the differentiating effect of student’s academic self-efficacy on student’s mastery goal orientation would be stronger in high efficacious male classes than in low efficacious male classes.

5. Since female adolescents tend to be more oriented to mastery goals and have more positive perceptions of their classroom environment than male adolescents (e.g., Anderman & Midgley, 1997) and that gender was a class-level variable in this study, it was expected that holding other variables constant, there would be a two-way interaction effect between class’s gender and class average perceived learning-oriented classroom assessment environment on (a) the class average mastery goal orientation and (b) the relationship between student’s perceived learning-oriented classroom assessment environment and student’s mastery goal orientation (i.e., the differentiating effect of student’s perceived learning-oriented classroom assessment environment). Specifically, holding other variables constant, the difference in class average mastery goal orientation between high
versus low learning-oriented classes would be weaker in female classes than in male classes. Also, holding other variables constant, the differentiating effect of student’s perceived learning-oriented classroom assessment environment on student’s mastery goal orientation would be weaker in high learning-oriented female classes than in low learning-oriented female classes, whereas the differentiating effect of student’s perceived learning-oriented classroom assessment environment on student’s mastery goal orientation would be stronger in high learning-oriented male classes than in low learning-oriented male classes.

6. In their studies about the effect of assessment type on student’s motivational goals, Maslovaty and Kuzi (2002) and Stefanou and Parkes (2003) found that traditional assessments fostered adoption of performance goals more than adoption of mastery goals. Consistent with these findings and theoretical arguments against traditional assessments (e.g., see Henning-Stout, 1994), it was expected that in this study holding other variables constant teacher’s frequent use of traditional assessments would negatively be related to class average mastery goal orientation. Also, holding other variables constant, the relationship between student’s academic self-efficacy and student’s adoption of a mastery goal orientation as well as the relationship between student’s perceived learning-oriented classroom assessment environment and student’s adoption of a mastery goal orientation would be weaker in classes with a high emphasis on traditional assessments than in classes with a low emphasis on traditional assessments.
7. Teachers are required to follow classroom assessment practices recommended by experts of educational measurement and assessment (AFT, NCME, & NEA, 1990). Further, Ames (1992a), McMillan and Workman (1998), Stiggins (1999), and Stiggins and Chappius (2005) have theoretically suggested that teacher’s adherence to the recommended classroom assessment practices concerning assessment revision, communication of assessment, assessment standards and criteria, student-involved assessment, and nonachievement-based grading may enhance student motivation for learning. Consistent with these theoretical perspectives, it was expected that holding other variables constant, teacher’s frequent use of recommended assessments would be positively related to class average mastery goal orientation. Also, holding other variables constant, the relationship between student’s academic self-efficacy and student’s adoption of a mastery goal orientation as well as the relationship between student’s perceived learning-oriented classroom assessment environment and student’s adoption of a mastery goal orientation would be stronger in classes with a high adherence to the recommended assessments than in classes with a low adherence to the recommended assessments.

8. In their studies about the effect of assessment type on student’s motivational goals, Maslovaty and Kuzi (2002) and Stefanou and Parkes (2003) found that alternative assessments fostered adoption of mastery goals more than adoption of performance goals. In addition, classroom assessment research has shown significant differences between teachers with regard to teacher’s frequent use of
alternative assessment based on teacher’s gender and teacher’s years of teaching experience (for more details see Bol et al., 1998; Mertler, 1998). In this study, it was argued that the most experienced teachers not only might have learned through experience how to enhance student motivation for learning, but also might have observed through experience the hypothesized positive effects of using alternative assessments on student motivation to learn, and hence they use them more frequently than the least experienced teachers.

These findings and arguments have led me to hypothesize/predict that holding other variables constant, there would be a three-way interaction effect between class’s gender, teacher’s years of teaching experience, and teacher’s frequent use of alternative assessments on (a) the class average mastery goal orientation, (b) the relationship between student’s academic self-efficacy and student’s adoption of a mastery goal orientation, and (c) the relationship between student’s perceived learning-oriented classroom assessment environment and student’s adoption of a mastery goal orientation.

Specifically, holding other variables constant, it was expected that for classes with a high emphasis on alternative assessments, the difference in class average mastery goal orientation between high versus low levels of teaching experience would be weaker in female classes than in male classes. For classes with a low emphasis on alternative assessments, the difference in class average mastery goal orientation between high versus low levels of teaching experience would be stronger in female classes than in male classes. Also, holding other variables constant, it was expected that for classes with a high emphasis on alternative assessments, the differentiating effect of student’s academic
self-efficacy on student’s adoption of a mastery goal orientation would be (a) weaker in female classes having a high experienced teacher than in female classes having a low experienced teacher and (b) stronger in male classes having a high experienced teacher than in male classes having a low experienced teacher. For classes with a low emphasis on alternative assessments, the differentiating effect of student’s academic self-efficacy on student’s adoption of a mastery goal orientation would be (a) stronger in female classes having a high experienced teacher than in female classes having a low experienced teacher and (b) weaker in male classes having a high experienced teacher than in male classes having a low experienced teacher. The same was expected to be true for the differentiating effect of student’s perceived learning-oriented classroom assessment environment on student’s adoption of a mastery goal orientation.

Taking together, the expected final combined explanatory model of each achievement goal orientation was as follows:

\[
(\text{GOAL})_j = \gamma_{00} + \gamma_{01}(\text{GNDR} \times \text{CEFC})_j + \gamma_{02}(\text{GNDR} \times \text{CPCAE})_j + \gamma_{03}(\text{TRAD})_j + \\
\gamma_{04}(\text{RECOM})_j + \gamma_{05}(\text{GNDR} \times \text{TEXP} \times \text{ALTR})_j + \gamma_{10}(\text{SEFC}_j - \text{CEFC}_j) + \\
\gamma_{11}(\text{GNDR} \times \text{CEFC})_j(\text{SEFC}_j - \text{CEFC}_j) + \gamma_{12}(\text{TRAD})_j(\text{SEFC}_j - \text{CEFC}_j) + \\
\gamma_{13}(\text{RECOM})_j(\text{SEFC}_j - \text{CEFC}_j) + \gamma_{14}(\text{GNDR} \times \text{TEXP} \times \text{ALTR})_j(\text{SEFC}_j - \text{CEFC}_j) + \\
\gamma_{20}(\text{SPCAE}_j - \text{CPCAE}_j) + \gamma_{21}(\text{GNDR} \times \text{CPCAE})_j(\text{SPCAE}_j - \text{CPCAE}_j) + \\
\gamma_{22}(\text{TRAD})_j(\text{SPCAE}_j - \text{CPCAE}_j) + \gamma_{23}(\text{RECOM})_j(\text{SPCAE}_j - \text{CPCAE}_j) + \\
\gamma_{24}(\text{GNDR} \times \text{TEXP} \times \text{ALTR})_j(\text{SPCAE}_j - \text{CPCAE}_j) + \gamma_{0j} + \gamma_{1j}(\text{SEFC}_j - \text{CEFC}_j) + \\
\gamma_{2j}(\text{SPCAE}_j - \text{CPCAE}_j) + \gamma_{2j}(\text{SPCAE}_j - \text{CPCAE}_j) + r_j.
\]

Each symbol and term in this model along with the modeling process will be explained in Chapter III of the dissertation.
CHAPTER III

METHODS

This chapter specifies the methods that were employed to gather and analyze data regarding the impact of classroom assessment practices on students’ perceptions of classroom assessment environment and achievement goal orientations. Specifically, the chapter addresses details regarding the participants that were included in the study, the instruments that were utilized, the data collection procedures that were followed, the research design that was employed, and the analysis that was carried out.

Participants

The participants in this study were ninth grade students and their science teachers enrolled in public schools within Muscat educational region in Oman during the spring semester 2007. A list of all ninth grade students and their science teachers in Muscat public schools could not be obtained from the Ministry of Education in Oman. Thus, a list of all public schools was utilized to select the participants. The list contained 36 male public schools and 36 female public schools that had the ninth grade. Due to limited finances and time for the study, a sample of 24 (67%) male schools and 20 (56%) female schools were selected. This resulted in a sample of 1,636 students situated in 83 ninth grade science classrooms taught by 83 independent teachers. The number of participating
students in each classroom ranged from 14 to 21, with an average of 20 students. A note about the appropriateness of power and sample size in this study will later be addressed in this section.

Of all participating students (\(N = 1,636\)) in this study, approximately 45% were males and 55% were females. The majority of the students (96.1%) were Omani and the remainders were from other Arabic speaking countries. The ages of the students ranged from 13 to 19 years with an average of 15 and a standard deviation of .91.

Of all participating teachers (\(J = 83\)) in this study, 44.6% were males and 55.4% were females. The majority of the teachers (95.2%) were Omani and the remainders were from other Arabic speaking countries. Ninety-four percent of the teachers had a bachelor degree in education and the remainders had an educational diploma. 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.

It should be noted that a desirable element of a quantitative research study is to have a high probability for rejecting a false null hypothesis (i.e., a high power) for a given sample size, level of significance (\(\alpha\)), and effect size (Cohen, 1988). Although a number of software programs have been designed for calculations of power (e.g., Raudenbush & Liu, 2000; Snijders & Bosker, 1993), there are no specific guidelines regarding power issues and appropriate sample sizes for hierarchical linear models (HLM). Yet, some general recommendations have been discussed. For example, Mok (1995) indicated that less bias and more efficiency would be expected from research designs involving more classrooms and fewer students per classroom than designs
involving fewer classrooms and more students per classroom. After reviewing some simulation studies investigating the power of HLM, Kreft and Leeuw (1998) indicated that 60 classrooms with 25 students per classroom, bringing the total number to 1500, will produce a sufficiently high power. Bassiri (1988) as well as van der Leeden and Busing (1994) showed that at least 30 classrooms and 30 students within each classroom are needed to obtain a sufficient power (e.g., .90) to detect interactions between variables measured at different levels in hierarchically structured data (i.e., cross-level interactions). As previously mentioned, the present study included 83 classrooms with an average of 20 students per classroom, bringing the total number to 1636 students. Therefore, this discussion of power and sample size issues might draw readers’ attention to the level of power and sample size when interpreting HLM results of this study.

Setting

Ninth grade male and female students in Omani public schools are segregated. Male teachers only teach in male students’ schools and female teachers only teach in female students’ schools. The educational system in Oman is centralized. Science textbooks are prescribed for the ninth grade classes by the Ministry of Education. The instructional objectives and teaching time for each unit are also prescribed for the teacher. Facilities in schools located in coastal educational regions and those in mountain or interior educational regions are to a large extent comparable. Due to its proximity to the researcher, this study focused on Muscat educational region, which is one of the coastal regions. Thus, in this study, grade level and educational region were not considered variables because they did not represent non-uniform characteristics of the observational
units, which were students and teachers. In other words, grade level and educational region were not considered variables in this study because they did not vary across the observational units of the study, which were students and teachers.

Instrumentation

Two questionnaires were developed and used by the author for this study, one for students and one for teachers. The student’s questionnaire included items assessing perceived classroom assessment environment, achievement goal orientations, and academic self-efficacy. Assor and Connell (1992) suggested that Likert-type scales of four points provide more valid information on self-report measures designed for elementary, middle, and high school students. Therefore, as recommended by Assor and Connell (1992), the students in this study were asked to indicate the degree to which each item is true for them on a 4-point Likert scale ranging from 1 (completely not true) to 4 (completely true). The teacher’s questionnaire included items assessing teacher’s frequent uses of traditional assessments, alternative assessments, and various classroom assessment practices recommended by experts of educational measurement and assessment. The teachers were asked to indicate the frequency of using each type of assessment practice on a 5-point Likert scale ranging from 1 (never) to 5 (always). The literature was relied upon heavily to guide the development of the items appearing in the student and teacher’s questionnaires. In addition, a section about demographic data was included at the end of each questionnaire. These data were used to describe the sample.

Given that the language of the participants is Arabic, the author translated the questionnaires into Arabic. To verify the accuracy of the translation, the Arabic and
English versions of the questionnaires were given to two professors in the areas of educational measurement and psychology, from Sultan Qaboos University, who were fluent in both Arabic and English. A discussion was held with the professors to verify discrepancies between the original and the translated versions. Few editing modifications were made as a result of the translation.

To establish content validity, the Arabic versions of the questionnaires were then given to five professors in the areas of educational measurement and psychology from Sultan Qaboos University. They were asked to judge the clarity of wording and appropriateness of each item for the use with the targeted participants and its relevance to the construct being measured. Their feedback was used for further refinement of the questionnaires. For both questionnaires, the majority of the consulted judges agreed that the items were clearly worded, appropriate for the participants, and relevant to the constructs being measured. As suggested by Midgley et al. (2000) and one of the reviewers, the items on each questionnaire were mixed. Following is a description of the questionnaires.

**Student’s Questionnaire**

*Classroom assessment environment.* This section of the student’s questionnaire was developed by the author of this study based on classroom assessment literature that addressed certain aspects of the classroom assessment environment (Ames, 1992a, 1992b; Brookhart, 1994, 1997a; Church et al., 2001; Crooks, 1988; Greene et al., 2004; Maslovaty & Kuzi, 2002; McMillan & Workman, 1998; Midgley et al., 2000; Stiggins &
Chappuis, 2005; Stipek, 2002; Wang, 2004). The section consisted of 25 items divided into three areas that were identified from the literature. These areas were assessment tasks (6 items), assessment feedback (14 items), and assessment standards and criteria (5 items). These areas were assumed to be more directly related to students’ experiences of their classroom assessment environment (Crooks, 1988; Wang, 2004). The items of perceived classroom assessment environment are presented in Appendix A.

Following Allen and Yen (2002) as well as Crocker and Algina (1986), construct validity of the perceived classroom assessment environment was assessed by submitting students’ responses to a principal components/exploratory factor analysis (PCA/EFA) described latter in the analysis section of the dissertation. Although the items assessing perceived classroom assessment environment were based on previous research and similar questionnaires, the PCA/EFA was performed because the items have not been validated in the past for use with students in Oman and because of the changes in wording and language of the items and context where they were used. As such, in this study, the purpose of the PCA/EFA was to identify the factor structure of the perceived classroom assessment environment present in the data for the study’s participants.

Due to the need for orthogonality in subsequent analyses, scale scores as suggested by Hair et al. (1998) were created by summing scores on the items for each dimension of the best factor model revealed by the PCA/EFA. Each student had a scale score for each dimension of the perceived classroom assessment environment. Higher scores on each dimension represented positive perceptions of the classroom assessment
environment defined in that dimension. Internal consistency reliability for each dimension’s scores was established through Cronbach’s alpha.

In this study, perceived classroom assessment environment was assessed at two levels: student and class. At the student-level, the individual perceived classroom assessment environment score on each dimension for each student was constructed as the sum of the student’s responses to all items defining that dimension of the classroom assessment environment obtained from the PCA/EFA. With respect to the class-level, the sample size (83 classes) in this study did not provide an appropriate sample size to conduct PCA/EFA at the class-level because it was far less than the recommended minimum sample size of 200 (Kelloway, 1998; Stevens, 2002). Therefore, at the class-level, the aggregate perceived classroom assessment environment score on each dimension for each class was reflected by the average levels of individual students’ perceptions within the class on that dimension of the classroom assessment environment obtained from the PCA/EFA at the student-level. These values were added to the class-level data set with a corresponding class’s identification number as the selection variable. Results of the validity and reliability analyses of perceived classroom assessment environment are presented in Chapter IV of the dissertation.

Achievement goal orientations. This section contained 14 items from PALS (Midgley et al., 2000). In their original version, the items assessed students’ adoption of mastery (5 items), performance-approach (5 items), and performance-avoidance (4 items) goals on a 5-point Likert scale ranging from 1 (not all true) to 5 (very true) (Midgley et al., 2000). Midgley et al. (2000) reported internal consistency reliabilities of .85, .89, and
.74 for mastery, performance-approach, and performance-avoidance goals as measured by Cronbach’s alpha, respectively. In this study, the items were phrased in relation to the science class work using a 4-point Likert scale ranging from 1 (completely not true) to 4 (completely true) as recommended by Assor and Connell (1992). The items of achievement goal orientations are presented in Appendix A.

These items have been validated for use with middle school students in the United States (Midgley et al., 1998; Midgley et al., 2000), but not for use with students in Oman to the best of my knowledge. As such, due to changes in the scales’ anchors, wording and language of the items, and context where they were used, students’ responses were submitted to a principal components/exploratory factor analysis (PCA/EFA) described latter in the analysis section of the dissertation. The purpose of the PCA/EFA was to identify the factor structure of achievement goal orientations present in the data for the study’s participants.

Due to the need for replication in other studies of achievement goal theory, summated scales as suggested by Hair et al. (1998) were created by averaging scores on the items for each dimension of the best factor model revealed by the PCA/EFA. Each student had a summated scale score for each dimension of the achievement goals. Higher scores on each dimension represented a higher adoption or endorsement of the achievement goal defined in that dimension. Internal consistency reliability for each dimension’s scores was established through Cronbach’s alpha. Results of the validity and reliability analyses of achievement goal orientations are presented in Chapter IV of the dissertation.
Academic self-efficacy. The measure of student’s academic self-efficacy was adapted from Greene et al. (2004) and PALS (Midgley et al., 2000). This measure contained six positively worded items assessing students’ perceptions of their competence to do their class work. Although the items did not assess task-specific efficacy beliefs, they were domain-specific and situated (Pajares, 1996), in that they focused on students’ perceptions about their competence to do the work in their current science class according to Bandura’s (1986) theoretical guidelines. The items related to academic self-efficacy are presented in Appendix A. These items have been validated for use with middle and high school students in the United States (Greene et al., 2004; Midgley et al., 2000), but not for use with students in Oman to the best of my knowledge. As such, due to changes in the wording and language of the items and context where they were used, a principal components/exploratory factor analysis (PCA/EFA) was conducted on the six academic self-efficacy items to determine whether they represent a single construct. This analysis indeed yielded a single factor with an eigenvalue of 2.625, and that this uni-factor solution accounted for 43.74% of the total variance. All items loaded higher than .60 on the factor.

In this study, academic self-efficacy was assessed at two levels: student and class. At the student-level, the student’s academic self-efficacy was reflected by a total rating score on the items for the dimension of the best factor model revealed by the PCA/EFA. With respect to the class-level, the sample size (83 classes) in this study did not provide an appropriate sample size to conduct PCA/EFA at the class-level because it was far less than the recommended minimum sample size of 200 (Kelloway, 1998; Stevens, 2002).
Therefore, at the class-level, the class’s academic self-efficacy score for each class was reflected by the average levels of individual students’ academic self-efficacy within the class on that component of the academic self-efficacy obtained from the PCA/EFA at the student-level. These values were added to the class-level data set with a corresponding class’s identification number as the selection variable. Higher scores represented a higher level of academic self-efficacy. Internal consistency reliability of the academic self-efficacy scores was assessed by obtaining Cronbach’s alpha. The reliability coefficient was .74 with 95%CI = [.72, .76] and a standard error of measurement of 1.39.

Demographic data. At the end of the student’s questionnaire, the students were asked to indicate their gender, age, and nationality as either Omani or non-Omani.

Teacher’s Questionnaire

Type of assessment. This section of the teacher’s questionnaire contained 13 items drawn from a questionnaire developed by Alsarimi (2000) to assess the frequency with which the third preparatory science teachers in Oman use various types of traditional and alternative assessments in their classrooms. Traditional assessments included seven items about using true-false, multiple-choice, matching, completion, short-answer, and extended short-answer test items as well as oral exams (Alsarimi, 2000). Alternative assessments included six items about using essay items, research papers, portfolios, models, and structured and unstructured performance assessments (Alsarimi, 2000). These items are presented in Appendix B.
The sample size (83 teachers) in this study did not provide a good sample size to conduct principal components/exploratory factor analysis for the assessment types used by the teachers because it was far less than the recommended minimum sample size of 200 (Kelloway, 1998; Stevens, 2002). Therefore, in this study, two scale scores, one for traditional assessment and one for alternative assessment, were derived to indicate teacher’s 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 95%CI = [.03, .50] and a standard error of measurement of 2.24. The reliability coefficient for scores representing alternative assessments was .41 with 95%CI = [.19, .59] and a standard error of measurement of 2.60.

As might be noted, the reliability coefficients for scores on the scales measuring teachers’ frequent uses of traditional and alternative assessments were low. Yet, their corresponding standard errors of measurement 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; Thompson, 2003). 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. More details about reliability data will later be discussed in the last chapter of the dissertation.

*Recommended classroom assessment practices.* This section of the teacher’s questionnaire was developed by the author of this study to measure teacher’s frequent use of classroom assessment practices recommended by experts of educational measurement and 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 (AFT, NCME, & NEA, 1990; Alkharusi, 2002; Alsarimi, 2000; Ames, 1992a, 1992b; Church et al., 2001; McMillan et al., 2002; McMillan & Workman, 1998; Nava & Loyd, 1992; Nitko, 2001; Stiggins & Chappuis, 2005; Stiggins et al., 1989; Zhang & Burry-Stock, 2003). The areas were revision of assessment (6 items), communicating assessment (9 items), assessment standards and criteria (5 items), student-involved assessment (4 items), and nonachievement-based grading factors (6 items). These items are presented in Appendix B.

Scores for items that reflect un-recommended assessment practices were reversed so that higher scores represent more alignment or agreement or adherence with the recommended classroom assessment practices. The sample size (83 teachers) in this study did not provide a good sample size to conduct principal components/exploratory factor analysis for the recommended assessment practices used by the teachers because it was far less than the recommended minimum sample size of 200 (Kelloway, 1998; Stevens,
Therefore, in this study, a teacher’s frequent use of the recommended classroom assessment practices was reflected by a total rating score across all the items. Internal consistency reliability of the scores representing teacher’s frequent use of recommended classroom assessment practices was assessed by obtaining Cronbach’s alpha. The reliability coefficient was .56 with 95%CI = [.42, .69] and a standard error of measurement of 5.81.

Demographic data. At the end of the teacher’s questionnaire, the teachers were asked to indicate their gender, nationality as either Omani or non-Omani, number of years of teaching experience, and the educational qualification as either a bachelor degree in education or an educational diploma.

Procedures

On July 2006, the author requested permission from Ministry of Education in Oman to collect data from ninth grade students and their science teachers in the selected public schools of the Muscat educational region during the spring semester 2007. The author was granted the permission to collect the requested data (see Appendix C). According to Stiggins and Conklin (1992), teachers’ assessment practices tend to be consistent over time and that the classroom assessment environment is usually established during the first few days of the semester. As such, during approximately the third week of February 2007, the author contacted the principal from each of the selected schools to schedule a meeting time together with ninth grade science teachers in the school.
During the meeting, the author informed the principal and the teachers that a study is being conducted to investigate relationships of classroom assessment to students’ perceptions of the classroom assessment environment and achievement goal orientations in ninth grade science classrooms. At this time, the author requested the participation of the teachers and their permission to collect data from one classroom of their ninth grade students. The teachers were informed that they were not obligated to participate in the study, and that if they wished, their responses would remain anonymous and confidential. Then, teachers who wished to participate in the study were provided with consent forms (see Appendix D) to obtain parents’ permission to collect data from the students. After that, the author scheduled another meeting with the teachers to collect parents’ consent forms and to administer the questionnaires during a regular class meeting.

During the first week of March 2007, the author provided the teachers who wished to participate in the study a consent form (see Appendix D), a cover letter (see Appendix E), and the teacher’s questionnaire. Each teacher was allowed approximately five minutes to read and sign the consent form. Brief instructions were provided by the author about the information that was requested in the questionnaire, how to respond to the items, and where to find directions that were also included both on the cover letter and the questionnaire. The requested information took approximately ten minutes to complete. This presentation was made during each teacher’s regular scheduled class meeting. At this time, the teacher was requested to leave the class to avoid a potential influence of the teacher’s presence on students’ responses.
Then, the author informed the students that a study is being conducted to investigate the relationship of classroom assessment to students’ perceptions of the classroom assessment environment and achievement goals in ninth grade science classrooms. At this time, the author requested the participation of the students who were given parent’s permission. The author explained to the students that participation in the study would not influence the student’s grade or relation with the teacher in any way. Also, the students were informed that their responses would remain anonymous and confidential. In addition, they were informed that they were not obligated to participate in the study.

Students who were granted parent’s permission and wished to participate in the study were given a consent form (see Appendix D), a cover letter (see Appendix E), and the student’s questionnaire. The students were allowed approximately five minutes to read and sign the consent form. According to administration’s instructions stated in Midgley et al.’s (2000) PALS, the students were told that the questionnaire was not a test and that there was no right or wrong answer. They were also told that the information in the questionnaire was confidential and that no one at home or at school would ever see the responses (Midgley et al., 2000). Instructions were given by the author about the information that was requested in the questionnaire, how to respond to the items, and where to find directions that were also included both on the cover letter and the questionnaire. This process took approximately twenty minutes to complete.

To maintain the data in classroom units, completed teacher and student questionnaires were placed together in sealed envelopes. Each envelop was labeled with
an identification number to link students’ data with respective classrooms’ data in the analysis.

Research Design

A correlational research design that is relational in nature (Arnold, 1992) was employed in this study to examine the effects of certain student-level and class-level characteristics on perceived classroom assessment environment and achievement goal orientations for ninth grade students in Muscat science classrooms in Oman. Although the word effects may imply casual thinking, there was no manipulation of variables in the study, which is a necessary condition for causality (Pedhazur, 1997). As such, definitive causal inferences cannot be made from the correlational nature of this study’s results (Arnold, 1992). The objects of measurement and interest in this study were students and classrooms, making the resulting data set hierarchical (Kreft & Leeuw, 1998). Thus, the level of inference from this study is ninth grade students within Muscat science classrooms in Oman during the spring semester 2007.

Analysis

Prior to the analysis, all variables were examined for accuracy of data entry and missing values. All the items in the questionnaires were checked to ensure that they were conceptually coded in the appropriate direction. Descriptive statistics such as means, minimum and maximum values, standard deviations, kurtosis and skewness values, histograms, normal probability (Q-Q) plots, and boxplots were carried out to examine
response patterns of the participants. The statistical analyses were performed at the .05 level of significance. Then, the research questions were addressed as follows.

**Research Questions One and Two**

RQ (1): What are the underlying dimensions of ninth grade students’ perceptions of classroom assessment environment in Muscat science classrooms in Oman?

RQ (2): What are the underlying dimensions of ninth grade students’ achievement goal orientations in Muscat science classrooms in Oman?

These research questions were concerned with the underlying dimensions of perceived classroom assessment environment and achievement goal orientations, respectively, for ninth grade students in Muscat science classrooms in Oman. Due to changes in scales’ anchors, wording, language, and study context, principal components/exploratory factor analyses (PCA/EFA) were conducted for each topic separately to identify the underlying dimensions and to reduce the number of variables (i.e., items) for subsequent analyses. Data pertaining to each topic were examined for outliers; normality; linearity; intercorrelations among the items; and factorability through statistical significance of the intercorrelations, Bartlett test of sphericity, Kasier’s measure of sampling adequacy, and anti-image correlation matrix (Hair et al., 1998; Stevens, 2002; Tabachnick & Fidell, 2001).

When deciding whether to rotate the components (i.e., dimensions) orthogonally or obliquely, Pedhazur and Schmelkin (1991, p. 615) stated that “The preferred course of action is…to rotate both orthogonally and obliquely. When, on the basis of the latter, it is concluded that the correlations among the factors are negligible, the interpretation of the simpler orthogonal solution becomes tenable.” As such, in this study, the principal
components analyses were performed with both orthogonal and oblique rotations for the perceived classroom assessment environment and achievement goal orientations separately. Varimax rotation was used for the orthogonal solution, and oblimin rotation was used for the oblique solution.

No particular number of components (i.e., dimensions) was hypothesized and the criterion was set to eigenvalues greater than one (Hair et al., 1998; Stevens, 2002; Tabachnick & Fidell, 2001). As suggested by Stevens (2002) and Tabachnick and Fidell (2001), the cutoff for size of loading to be interpreted was set at .32. The number of components (i.e., dimensions) that best fits the sample was determined through scree plots, eigenvalues, and the criteria of having smallest residual correlation matrix and largest amount of total variance explained (Hair et al., 1998; Stevens, 2002; Tabachnick & Fidell, 2001). The best factor solution based on conceptual expectations and statistical criteria was then examined to reveal the dimensions of perceived classroom assessment environment and achievement goal orientations adopted by the participating students. Each dimension was named and interpreted by examining the items that mainly loaded higher than .32 (Stevens, 2002; Tabachnick & Fidell, 2001) onto that dimension.

Due to the need for orthogonality in subsequent analyses, scale scores as suggested by Hair et al. (1998) were created by summing scores on the items for each dimension of the best factor model of the perceived classroom assessment environment revealed by the PCA/EFA. Each student had a scale score for each dimension of the perceived classroom assessment environment. Higher scores on each dimension
represented positive perceptions of the classroom assessment environment defined in that dimension.

Due to the need for replication in other studies of achievement goal theory, summated scales as suggested by Hair et al. (1998) were created by averaging scores on the items for each dimension of the best factor model revealed by the PCA/EFA. Each student had a summated scale score for each dimension of the achievement goals. Higher scores on each dimension represented a higher adoption or endorsement of the achievement goal defined in that dimension.

Research Questions Three, Four, Five, and Six

RQ (3): Do ninth grade Muscat science classrooms in Oman vary in perceived classroom assessment environment (PCAE)?

RQ (4): What are some of the student characteristics (e.g., academic self-efficacy) that might have effects on PCAE?

RQ (5): Do the effects of student characteristics (e.g., academic self-efficacy) on PCAE vary across classrooms?

RQ (6): What are some of the classroom characteristics (e.g., class gender) that might help explain the variability in the PCAE and in the effects of student characteristics (e.g., academic self-efficacy) on PCAE across classrooms?

Based on the literature reviewed in Chapter II of the dissertation, these research questions had the following dependent and independent variables:

Dependent variable:

- An individual student’s perceived classroom assessment environment (SPCAE)
Independent variable at the student-level:

- An individual student’s academic self-efficacy (SEFC)

Independent variables at the class-level:

- Class’s gender (GNDR)
- Teacher’s years of teaching experience (TEXP)
- Teacher’s frequent use of traditional assessments (TRAD)
- Teacher’s frequent use of alternative assessments (ALTR)
- Teacher’s frequent use of recommended assessment practices (REC)
- Class’s average for academic self-efficacy (CEFC); as discussed in Chapters I and II of the dissertation, the variable CEFC was included in order to study its contextual effect (Kreft & Leeuw, 1998; Raudenbush & Bryk, 2002) on perceived classroom assessment environment. Kreft and Leeuw (1998) stated that “contextual models are defined as regression models containing two types of variables: individual-level variables and aggregated context variables, such as group means” (p. 8). When applying Kreft and Leeuw’s (1998) definition of contextual models in the present study, the variable SEFC was used twice: once as an individual student variable and once as an aggregated class characteristic.
- Class’s average for perceived classroom assessment environment (CPCAE); the variable CPCAE was included in order to study its compositional effect on an individual student’s perception of the classroom assessment environment because as discussed in Chapters I and II of the dissertation, it was assumed that “students construct their own meaning [of the classroom assessment environment] based in part on their group experiences” (Brookhart, 2004, p. 445).

To answer research questions from three to six, hierarchical linear model (HLM) analyses were conducted, one for each dependent variable defined as a dimension of the individual student’s perceived classroom assessment environment that emerged from the principal components/exploratory factor analyses of students’ perceptions of classroom assessment environment. Under the HLM framework, a clear conceptual distinction is made between student-level and class-level variables and effects (Raudenbush & Bryk, 2002). This conception is reflected in the two models that make up a two-level HLM. The first model captures the primary effects at the student-level within each class
(Raudenbush & Bryk, 2002). The second model attempts to explain these student-level effects in terms of class-level independent variables (Raudenbush & Bryk, 2002).

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. As a result, the magnitude of the class-level coefficients (i.e., the $\gamma$ s) can be directly compared when assessing the relative contributions of the class-level variables in each of the class effects’ models (Arnold, 1992). Univariate shape and frequency distribution of each variable at each level and bivariate relationships at each level were examined to identify discrepant cases and unusual observations that could lead to implausible results (Raudenbush & Bryk, 2002).

A fully unconditional model. Each HLM analysis began with a fully unconditional model (also called a random ANOVA model) as follows:

Student-level:

$$(SPCAE)_{ij} = \beta_{0j} + r_{ij}$$

where $r_{ij} \sim iidN(0, \sigma^2)$; $i = 1, \ldots, n_j$ students in class $j$, and $j = 1, \ldots, J$ classes; and $\sigma^2$ is the within-class variance.

Class-level:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

where $u_{0j} \sim iidN(0, \tau_{00})$; and $\tau_{00}$ is the variance of the true outcome class means ($\beta_{0j}$).

Combined-model:

$$(SPCAE)_{ij} = \gamma_{00} + u_{0j} + r_{ij}.$$
A random-coefficient regression model. The next step in the analysis was to proceed with a random-coefficient regression model in order to test the effect of student-level independent variable (student’s academic self-efficacy) on the individual student’s perceived classroom assessment environment. In this analysis, the student-level
independent variable was centered around its class mean. The generic form of the models tested in this step was as follows:

**Student-level:**

\[(SPCAE)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + r_{ij}, \text{ where } r_{ij} \sim iidN(0, \sigma^2).\]

**Class-level:**

\[
\begin{align*}
\beta_{0j} &= \gamma_{00} + u_{0j}, \\
\beta_{1j} &= \gamma_{10} + u_{1j},
\end{align*}
\]

Where \(u_{0j}\) and \(u_{1j}\) are bivariate normally distributed with means 0 and variance-covariance matrix \(T = \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix}\).

**Combined-model:**

\[(SPCAE)_{ij} = \gamma_{00} + \gamma_{10}(SEFC_{ij} - CEFC_{j}) + u_{0j} + u_{1j}(SEFC_{ij} - CEFC_{j}) + r_{ij}.\]

Where:

\(\beta_{1j}\) is the effect of student’s academic self-efficacy on the individual perceived classroom assessment environment in class \(j\) (i.e., academic self-efficacy slope). Specifically, \(\beta_{1j}\) represents the differentiating effect of academic self-efficacy in class \(j\), that is, the degree to which differences in students’ academic self-efficacy relate to individual perceived classroom assessment environment in class \(j\).

\((SEFC)_{ij}\) is the academic self-efficacy score for student \(i\) in class \(j\).

\((CEFC_{j})\) is the average academic self-efficacy for class \(j\).

\(r_{ij}\) is the residual associated with the individual perceived classroom assessment environment score for student \(i\) in class \(j\).

\(\gamma_{10}\) is the average effect of academic self-efficacy on the individual perceived classroom assessment environment across all classes (i.e., average academic self-efficacy slope across all classes).

\(u_{0j}\) is the unique part associated with the intercept for class \(j\).
$u_{ij}$ is the unique part associated with the academic self-efficacy slope for class j.

This analysis of the random-coefficient regression model provides the following information:

1. Whether on average student’s academic self-efficacy is significantly related to individual perceived classroom assessment environment within classes ($H_0: \gamma_{10} = 0$).

2. The proportion of the student-level variance in the individual perceived classroom assessment environment accounted for by student’s academic self-efficacy as follows $\left(\hat{\sigma}^2 (\text{base model}) - \hat{\sigma}^2 (\text{random regression})\right)/\hat{\sigma}^2 (\text{base model})$ (Raudenbush & Bryk, 2002). The random ANOVA model provides the appropriate base model for this application.

3. Whether statistically significant differences exist among class means on perceived classroom assessment environment (i.e., variances of the intercepts $H_0: \tau_{00} = 0$).

4. Whether the relationship between student’s academic self-efficacy and individual perceived classroom assessment environment varies significantly across classes (i.e., variances of the academic self-efficacy slopes $H_0: \tau_{11} = 0$).

5. The correlation coefficient $\hat{p}(\hat{\beta}_{0j}, \hat{\beta}_{1j}) = \hat{\tau}_{01}/(\hat{\tau}_{00}\hat{\tau}_{11})$ (Raudenbush & Bryk, 2002) between class means on the individual perceived classroom assessment environment and academic self-efficacy effects on the individual perceived classroom assessment environment. In other words, do classes with high means on perceived classroom assessment environment tend also to have a strong relationship between student’s academic self-efficacy and individual perceived classroom assessment environment?

6. Reliability estimates, on average, of the class’s intercept ($\hat{\beta}_{0j}$) and slope ($\hat{\beta}_{1j}$).

A final explanatory model of perceived classroom assessment environment. The final step in the analysis involved modeling perceived classroom assessment environment as a function of student-level and class-level characteristics. Having estimated the variability in the class means (i.e., intercepts) and student’s academic self-efficacy effects
I then proceeded with intercepts-and-slopes-as-outcomes regression models to explain the variability in these intercepts and slopes.

The student-level model remained the same as in the random-coefficient regression model. To build the class-level model, Raudenbush and Bryk (2002, p. 267) stated that “we suggest dividing the level-2 predictors into conceptually distinct subsets and fitting a submodel for each. The strongest predictors from these submodels might then be combined in an overall model.” As such, the class-level independent 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’s gender. The second set represented the joint effects of class’s gender and class’s average for perceived classroom assessment environment. The third set represented the joint effects of class’s 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. I next proceeded to develop an overall intercepts-and-slopes-as-outcomes regression model based on the statistically significant variables from the three sets of class-level variables. The three submodels of the intercepts-and-slopes-as-outcomes regression model were as follows:

**Submodel (1): Contextual-effect model of academic self-efficacy along with its differential contextual effect by class’s gender.** Within HLM, a contextual effect is represented by including the class aggregate of a student-level variable in the between-class model for that differentiating effect (Lee & Bryk, 1989; Raudenbush & Bryk,
Thus, in this submodel, the contextual effect of academic self-efficacy was represented by including class’s average academic self-efficacy (CEFC) in the model that investigated the distributive effect of student’s academic self-efficacy. As indicated by Lee and Bryk (1989), the differential contextual effect by class’s gender (GNDR) was represented by the inclusion of (GNDR × CEFC) interaction term in the between-class model. This submodel (1) was as follows:

**Student-level:**

\[(SPCAE)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + r_{ij}, \text{ where } r_{ij} \sim iidN(0, \sigma^2).\]

**Class-level:**

\[
\begin{align*}
\beta_{0j} &= \gamma_{00} + \gamma_{01}(CEFC)_{j} + \gamma_{02}(GNDR)_{j} + \gamma_{03}(GNDR \times CEFC)_{j} + u_{0j}, \\
\beta_{1j} &= \gamma_{10} + \gamma_{11}(CEFC)_{j} + \gamma_{12}(GNDR)_{j} + \gamma_{13}(GNDR \times CEFC)_{j} + u_{1j}. 
\end{align*}
\]

Where \(u_{0j}\) and \(u_{1j}\) are bivariate normally distributed with means 0 and variance-covariance matrix \(T = \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix} \).

**Combined-model:**

\[
\begin{align*}
(SPCE)_{ij} &= \gamma_{00} + \gamma_{01}(CEFC)_{j} + \gamma_{02}(GNDR)_{j} + \gamma_{03}(GNDR \times CEFC)_{j} + \\
&\gamma_{10}(SEFC_{ij} - CEFC_{j}) + \gamma_{11}(CEFC)_{j}(SEFC_{ij} - CEFC_{j}) + \\
&\gamma_{12}(GNDR)_{j}(SEFC_{ij} - CEFC_{j}) + \gamma_{13}(GNDR \times CEFC)_{j}(SEFC_{ij} - CEFC_{j}) + \\
&u_{0j} + u_{1j}(SEFC_{ij} - CEFC_{j}) + r_{ij}. 
\end{align*}
\]

Where:

\(\gamma_{01}\) and \(\gamma_{11}\) are the main effects of class’s mean for academic self-efficacy on (a) the class average perceived classroom assessment environment and (b) the relationship between student’s academic self-efficacy and individual perceived classroom assessment environment, holding other variables constant, respectively.

\(\gamma_{02}\) and \(\gamma_{12}\) are the differences between female classes and male classes on (a) the class average perceived classroom assessment environment and (b) the
relationship between student’s academic self-efficacy and individual perceived classroom assessment environment, holding other variables constant, respectively.

\(\gamma_{03}\) and \(\gamma_{13}\) are the two-way interaction effects of class’s gender-by-class’s mean for academic self-efficacy on (a) the class average perceived classroom assessment environment and (b) the relationship between student’s academic self-efficacy and individual perceived classroom assessment environment, holding other variables constant, respectively.

\(r_{ij}, u_{0ij},\) and \(u_{1ij}\) are the residuals associated with the individual perceived classroom assessment environment score for student \(i\) in class \(j\), the average perceived classroom assessment environment for class \(j\), and the relationship between student’s academic self-efficacy and individual perceived classroom assessment environment for class \(j\), respectively.

Submodel (2): Joint effects model of class’s gender and class’s average for perceived classroom assessment environment. This submodel included class’s gender (GNDR) and class’s mean for perceived classroom assessment environment (CPCAE) along with their interaction term (GNDR × CPCAE) within the class-level model as follows:

**Student-level:**

\[
(SPACE)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_j) + r_{ij}, \text{ where } r_{ij} \sim iidN(0, \sigma^2).
\]

**Class-level:**

\[
\beta_{0j} = \gamma_{00} + \gamma_{01}(CPCAE)_j + \gamma_{02}(GNDR)_j + \gamma_{03}(GNDR \times CPCAE)_j + u_{0j}.
\]

\[
\beta_{1j} = \gamma_{10} + \gamma_{11}(CPCAE)_j + \gamma_{12}(GNDR)_j + \gamma_{13}(GNDR \times CPCAE)_j + u_{1j}.
\]

Where \(u_{0j}\) and \(u_{1ij}\) are bivariate normally distributed with means 0 and variance-covariance matrix \(T = \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix} \).
Combined-model:

\[
(SPCA)_{ij} = \gamma_0 + \gamma_1(CPAE_j) + \gamma_2(GNDR_j) + \gamma_3(GNDR \times CPAE_j) + \\
\gamma_{10}(SEFC_{ij} - CEFC_j) + \gamma_{11}(CPAAE_j)(SEFC_{ij} - CEFC_j) + \\
\gamma_{12}(GNDR_j)(SEFC_{ij} - CEFC_j) + \gamma_{13}(GNDR \times CPAE_j)(SEFC_{ij} - CEFC_j) + \\
u_{0_j} + u_{1_j}(SEFC_{ij} - CEFC_j) + r_{ij}.
\]

Where:

\(\gamma_1\) and \(\gamma_{11}\) are the main effects of class’s mean for perceived classroom assessment environment defined by one dimension on (a) the class average perceived classroom assessment environment defined by another dimension and (b) the relationship between student’s academic self-efficacy and individual perceived classroom assessment environment, holding other variables constant, respectively.

\(\gamma_{12}\) and \(\gamma_{13}\) are the differences between female classes and male classes on (a) the class average perceived classroom assessment environment and (b) the relationship between student’s academic self-efficacy and individual perceived classroom assessment environment, holding other variables constant, respectively.

\(\gamma_{03}\) and \(\gamma_{13}\) are the two-way interaction effects of class’s gender-by-class’s mean for perceived classroom assessment environment defined by one dimension on (a) the class average perceived classroom assessment environment defined by another dimension and (b) the relationship between student’s academic self-efficacy and individual perceived classroom assessment environment, holding other variables constant, respectively.

Submodel (3): Joint effects model of class’s gender and teacher’s assessment practices and years of teaching experience. This submodel included teacher’s frequent use of traditional assessments (TRAD), teacher’s frequent use of alternative assessments (ALTR), and teacher’s frequent use of recommended classroom assessment practices (RECOM) along with the interactions of class’s gender (GNDR), teacher’s years of teaching experience (TEXP), and teacher’s frequent use of alternative assessments within the class-level model as follows:
Student-level:

\[(SPCAE)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + r_{ij}, \text{ where } r_{ij} \sim iidN(0, \sigma^2).\]

Class-level:

\[\beta_{0j} = \gamma_{00} + \gamma_{01}(TRAD)_{j} + \gamma_{02}(ALTR)_{j} + \gamma_{03}(RECOM)_{j} + \gamma_{04}(TEXP)_{j} + \gamma_{05}(GNDR) + \gamma_{06}(TEXP \times GNDR)_{j} + \gamma_{07}(TEXP \times ALTR)_{j} + \gamma_{08}(GNDR \times ALTR)_{j} + \gamma_{09}(TEXP \times GNDR \times ALTR)_{j} + u_{0j}.\]

\[\beta_{1j} = \gamma_{10} + \gamma_{11}(TRAD)_{j} + \gamma_{12}(ALTR)_{j} + \gamma_{13}(RECOM)_{j} + \gamma_{14}(TEXP)_{j} + \gamma_{15}(GNDR) + \gamma_{16}(TEXP \times GNDR)_{j} + \gamma_{17}(TEXP \times ALTR)_{j} + \gamma_{18}(GNDR \times ALTR)_{j} + \gamma_{19}(TEXP \times GNDR \times ALTR)_{j} + u_{1j}.\]

Where \(u_{0j}\) and \(u_{1j}\) are bivariate normally distributed with means 0 and variance-covariance matrix \(T = \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix}\).

Combined-model:

\[(SPCAE)_{ij} = \gamma_{00} + \gamma_{01}(TRAD)_{j} + \gamma_{02}(ALTR)_{j} + \gamma_{03}(RECOM)_{j} + \gamma_{04}(TEXP)_{j} + \gamma_{05}(GNDR) + \ldots + \gamma_{09}(SEFC_{ij} - CEFC_{j}) + \gamma_{10}(SEFC_{ij} - CEFC_{j}) + \gamma_{11}(TRAD)_{j}(SEFC_{ij} - CEFC_{j}) + \gamma_{12}(ALTR)_{j}(SEFC_{ij} - CEFC_{j}) + \gamma_{13}(RECOM)_{j}(SEFC_{ij} - CEFC_{j}) + \gamma_{14}(TEXP)_{j}(SEFC_{ij} - CEFC_{j}) + \gamma_{15}(GNDR)_{j}(SEFC_{ij} - CEFC_{j}) + \ldots + \gamma_{19}(TEXP \times GNDR \times TRAD)_{j}(SEFC_{ij} - CEFC_{j}) + u_{0j} + u_{1j}(SEFC_{ij} - CEFC_{j}) + r_{ij}.\]

Where:

\(\gamma_{01}, \gamma_{02}, \gamma_{03}\), and \(\gamma_{04}\) are the main effects of teacher’s frequent use of traditional assessments, teacher’s frequent use of alternative assessments, teacher’s frequent use of recommended assessment practices, and teacher’s years of teaching experience on class average perceived classroom assessment environment holding other variables constant, respectively.
\( \gamma_{11}, \gamma_{12}, \gamma_{13}, \text{ and } \gamma_{14} \) are the main effects of teacher’s frequent use of traditional assessments, teacher’s frequent use of alternative assessments, teacher’s frequent use of recommended assessment practices, and teacher’s years of teaching experience on the relationship between student’s academic self-efficacy and individual perceived classroom assessment environment holding other variables constant, respectively.

\( \gamma_{05} \) and \( \gamma_{15} \) are the differences between female classes and male classes on (a) the class average perceived classroom assessment environment and (b) the relationship between student’s academic self-efficacy and individual perceived classroom assessment environment, holding other variables constant, respectively.

\( \gamma_{06} \) and \( \gamma_{16} \) are the two-way interaction effects of teacher’s years of teaching experience-by-class’s gender on (a) the class average perceived classroom assessment environment and (b) the relationship between student’s academic self-efficacy and individual perceived classroom assessment environment, holding other variables constant, respectively.

\( \gamma_{07} \) and \( \gamma_{17} \) are the two-way interaction effects of teacher’s years of teaching experience-by-teacher’s frequent use of alternative assessments on (a) the class average perceived classroom assessment environment and (b) the relationship between student’s academic self-efficacy and individual perceived classroom assessment environment, holding other variables constant, respectively.

\( \gamma_{08} \) and \( \gamma_{18} \) are the two-way interaction effects of class’s gender-by-teacher’s frequent use of alternative assessments on (a) the class average perceived classroom assessment environment and (b) the relationship between student’s academic self-efficacy and individual perceived classroom assessment environment, holding other variables constant, respectively.

\( \gamma_{09} \) and \( \gamma_{19} \) are the three-way interaction effects of teacher’s years of teaching experience-by-class’s gender-by-teacher’s frequent use of alternative assessments on (a) the class average perceived classroom assessment environment and (b) the relationship between student’s academic self-efficacy and individual perceived classroom assessment environment, holding other variables constant, respectively.

In each of the aforementioned submodels, hypotheses tests for the effect of each class-level variable and interaction on class average perceived classroom assessment environment (i.e., intercept \( \beta_{0,i} \)) and on the relationship between student’s academic self-
efficacy and individual perceived classroom assessment environment (i.e., slope $\beta_{ij}$), (i.e., $H_0: \gamma_{qs} = 0$, where $q = 0$ and 1; and $s = 1, \ldots, S$ class-level variable and interaction), were examined to determine the statistically significant class-level variables and interactions. These statistically significant class-level variables and interactions were then combined together to produce a parsimonious overall intercepts-and-slopes-as-outcomes regression model explaining the variability in the aforementioned intercepts and slopes. The proportion of variance in these intercepts and slopes was computed as follows $\left(\frac{\hat{\tau}_{qq}(\text{base model}) - \hat{\tau}_{qq}(\text{overall model})}{\hat{\tau}_{qq}(\text{base model})}\right)$, where $q = 0$ and 1 (Raudenbush & Bryk, 2002). The random-coefficient regression model provides the appropriate base model for this application.

The validity of inferences based on the final model was assessed by verifying the tenability of the assumptions of a two-level HLM (Raudenbush & Bryk, 2002, p. 255). On the basis of residual analyses, variables that were deleted in the early steps of the analysis were reconsidered for inclusion in the final composite model to obtain an adequate fitted model. Across the tested models, the student-level coefficients were specified as fixed, random, or non-randomly varying at the class-level model by examining the following indices (Raudenbush & Bryk, 2002):

1. The point estimate of $\tau_{qq}$. As $\tau_{qq}$ becomes negligible, its corresponding coefficient may be specified as either fixed or non-randomly varying.

2. The $\chi^2$ homogeneity test of $H_0$: $\tau_{qq} = 0$. As $H_0$ becomes tenable, its corresponding coefficient may be specified as either fixed or non-randomly varying.
3. The likelihood ratio test (i.e., deviance) of the variance-covariance components. If there is no statistically significant difference between the deviances of two models, then the reduced model will typically be preferred.

4. The reliability of $\beta_{qj}$ because it tells us how much of the observed variation in the estimated $\beta_{qj}$ is potentially explainable. As the reliability drops below .05, that coefficient may be specified as either fixed or non-randomly varying. Also, the reliability of each class-level random effect provides some insight about the available power in the data to detect statistically significant effects of class characteristics on the outcome. A small reliability coefficient suggests caution in inferring about the effects of class-level characteristics.

5. Correlations among $\beta_{qj}$, as the correlations are high (.90 and above), one or more of the coefficients may be specified as either fixed or non-randomly varying because the two random effects are carrying the same variation across level-2 units.

Research Questions Seven, Eight, Nine, and Ten

RQ (7): Do ninth grade Muscat science classrooms in Oman vary in achievement goal orientations (GOAL)?

RQ (8): What are some of the student characteristics (e.g., academic self-efficacy) that might have effects on GOAL?

RQ (9): Do the effects of student characteristics (e.g., academic self-efficacy) on GOAL vary across classrooms?

RQ (10): What are some of the classroom characteristics (e.g., class gender) that might help explain the variability in the GOAL and in the effects of student characteristics (e.g., academic self-efficacy) on GOAL across classrooms?

Based on the literature reviewed in Chapter II of the dissertation, these research questions had the following dependent and independent variables:

Dependent variable:

- Student’s achievement goal orientation (GOAL)
Independent variables at the student-level:

- An individual student’s academic self-efficacy (SEFC)
- An individual student’s perceived classroom assessment environment (SPCAE)

Independent variables at the class-level:

- Class’s gender (GNDR)
- Teacher’s years of teaching experience (TEXP)
- Teacher’s frequent use of traditional assessments (TRAD)
- Teacher’s frequent use of alternative assessments (ALTR)
- Teacher’s frequent use of recommended assessment practices (RECOM)

As discussed in Chapters I and II of the dissertation, the following two variables class’s average for academic self-efficacy (CEFC) and class’s average for perceived classroom assessment environment (CPCAE) were included in order to study their contextual effects (Kreft & Leeuw, 1998; Raudenbush & Bryk, 2002) on achievement goal orientations. Kreft and Leeuw (1998) stated that “contextual models are defined as regression models containing two types of variables: individual-level variables and aggregated context variables, such as group means” (p. 8). When applying Kreft and Leeuw’s (1998) definition of contextual models in the present study, the variables SEFC and SPCAE were used twice: once as individual student variables and once as aggregated class characteristics as follows:
- Class’s average for academic self-efficacy (CEFC)
- Class’s average for perceived classroom assessment environment (CPCAE)

To answer research questions from seven to ten, hierarchical linear model (HLM) analyses were conducted, one for each dependent variable defined as a dimension of student’s achievement goal orientations that emerged from the principal components/exploratory factor analyses of students’ achievement goal orientations. As noted earlier, the conceptual strength of the HLM analysis is the clear distinction made between student- and class-level variables and effects (Raudenbush & Bryk, 2002). While the dependent variable is specified at the student-level, the independent variables include both student- and class-level measures (Raudenbush & Bryk, 2002).
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. As a result, the magnitude of the class-level coefficients (i.e., the $\gamma$s) can be directly compared when assessing the relative contributions of the class-level variables in each of the class effects’ models (Arnold, 1992). Univariate shape and frequency distribution of each variable at each level and bivariate relationships at each level were examined to identify discrepant cases and unusual observations that may lead to implausible results (Raudenbush & Bryk, 2002).

A fully unconditional model. Each HLM analysis began with a fully unconditional model (also called a random ANOVA model) as follows:

Student-level:

$ (GOAL)_{ij} = \beta_{0j} + r_{ij}, \text{ where } r_{ij} \sim iidN(0, \sigma^2); \ i = 1, \ldots, n_j \text{ students in class } j, \text{ and } j = 1, \ldots, J \text{ classes; and } \sigma^2 \text{ is the within-class variance.} $

Class-level:

$ \beta_{0j} = \gamma_{00} + u_{0j}, \text{ where } u_{0j} \sim iidN(0, \tau_{00}); \text{ and } \tau_{00} \text{ is the variance of the true outcome class means (} \beta_{0j} \text{).} $

Combined-model:

$ (GOAL)_{ij} = \gamma_{00} + u_{0j} + r_{ij}. $

Where:

$ (GOAL)_{ij} \text{ is the achievement goal orientation score for student } i \text{ in class } j. $
\( \beta_{0j} \) is the average achievement goal orientation for class \( j \) (i.e., intercept).

\( r_i \) is the error associated with the achievement goal orientation score for student \( i \) in class \( j \).

\( \gamma_{00} \) is the grand mean of achievement goal orientation across all classes.

\( u_{0j} \) is the unique effect for class \( j \) on average achievement goal orientation.

This fully unconditional model determines the following:

1. The pooled within-class or student-level variance \( (\hat{\sigma}^2) \) in achievement goal orientation.
2. Whether a statistically significant variation exists among class means on achievement goal orientation (H\(_0\): \( \tau_{00} = 0 \)).
3. The proportion of variance in achievement goal orientation between classes, also called the intraclass correlation, which is equal to \( \hat{\tau}_{00}/(\hat{\tau}_{00} + \hat{\sigma}^2) \) (Raudenbush & Bryk, 2002).
4. An overall measure of the reliability of the class sample means as indicators of the true class means.

A random-coefficient regression model. The next step in the analysis was to proceed with a random-coefficient regression model in order to test the effect of student-level independent variables (student’s academic self-efficacy and student’s perceived classroom assessment environment) on achievement goal orientation. In this analysis, the student-level independent variables were centered around their respective class means.

The generic form of the models tested in this step was as follows:

\[
(GOAL)_i = \beta_{0j} + \beta_{1j} (SEFC_y - CEFC_j) + \beta_{2j} (SPCAE_y - CPACAE_j) + r_i, \quad \text{where} \quad r_i \sim iidN(0, \sigma^2).
\]
Class-level:

\[ \beta_{0j} = \gamma_{00} + u_{0j}, \]
\[ \beta_{1j} = \gamma_{10} + u_{1j}, \]
\[ \beta_{2j} = \gamma_{20} + u_{2j}. \]

Where \( u_{0j}, u_{1j}, \) and \( u_{2j} \) are multivariate normally distributed with means 0 and variance-covariance matrix \( T = \begin{bmatrix} \tau_{00} & \tau_{01} & \tau_{02} \\ \tau_{10} & \tau_{11} & \tau_{12} \\ \tau_{20} & \tau_{21} & \tau_{22} \end{bmatrix} \).

Combined-model:

\[(GOAL)_{ij} = \gamma_{00} + \gamma_{10} (SEFC_{ij} - CEFC_{j}) + \gamma_{20} (SPCAE_{ij} - CPCAE_{j}) + u_{0j} + u_{1j} (SEFC_{ij} - CEFC_{j}) + u_{2j} (SPCAE_{ij} - CPCAE_{j}) + r_{ij},\]

\( \beta_{1j} \) and \( \beta_{2j} \) are the effects of student’s academic self-efficacy and individual student’s perceived classroom assessment environment on student’s achievement goal orientation in class \( j \) (i.e., academic self-efficacy slope and individual perceived classroom assessment environment slope, respectively), each one is net of the other, respectively. Specifically, \( \beta_{1j} \) represents the degree to which differences in students’ academic self-efficacy relate to student’s achievement goal orientation in class \( j \) (i.e., the differentiating effect of academic self-efficacy in class \( j \)). Likewise, \( \beta_{2j} \) represents the degree to which differences in individual perceived classroom assessment environment relate to student’s achievement goal orientation in class \( j \) (i.e., the differentiating effect of individual perceived classroom assessment environment in class \( j \)).

\((SEFC)_{ij} \) and \((SPCAE)_{ij}\) are the academic self-efficacy score and the individual perceived classroom assessment environment score for student \( i \) in class \( j \), respectively.

\((CEFC_{j}) \) and \((CPCAE_{j}) \) are the average academic self-efficacy and the average perceived classroom assessment environment for class \( j \), respectively.

\( r_{ij} \) is the residual associated with the achievement goal orientation score for student \( i \) in class \( j \).
\( \gamma_{10} \) and \( \gamma_{20} \) are the average effect of student’s academic self-efficacy and the average effect of individual perceived classroom assessment environment on student’s achievement goal orientation across all classes, respectively (i.e., average academic self-efficacy slope and average individual perceived classroom assessment environment slope across all classes, respectively).

\( u_{0j} \), \( u_{1j} \), and \( u_{2j} \) are the unique parts associated with the intercept, the academic self-efficacy slope, and the individual perceived classroom assessment environment slope for class \( j \), respectively.

This analysis of the random-coefficient regression model provides the following information:

1. Whether on average student’s academic self-efficacy is significantly related to student’s achievement goal orientation within classes \( (H_0: \gamma_{10} = 0) \).

2. Whether on average individual perceived classroom assessment environment is significantly related to student’s achievement goal orientation within classes \( (H_0: \gamma_{20} = 0) \).

3. The proportion of the student-level variance in achievement goal orientation accounted for by student’s academic self-efficacy and individual perceived classroom assessment environment as follows:\[
\frac{\hat{\sigma}^2 (\text{base model}) - \hat{\sigma}^2 (\text{random regression})}{\hat{\sigma}^2 (\text{base model})}\]
(Raudenbush & Bryk, 2002). The random ANOVA model provides the appropriate base model for this application.

4. Whether statistically significant differences exist among class means on achievement goal orientation (i.e., variances of the intercepts \( H_0: \tau_{00} = 0 \)).

5. Whether the relationship between student’s academic self-efficacy and student’s achievement goal orientation varies significantly across classes (i.e., variances of the academic self-efficacy slopes \( H_0: \tau_{11} = 0 \)).

6. Whether the relationship between individual perceived classroom assessment environment and student’s achievement goal orientation varies significantly across classes (i.e., variances of the individual perceived classroom assessment environment slopes \( H_0: \tau_{22} = 0 \)).
7. The correlation coefficient \( p(\hat{\beta}_{0j}, \hat{\beta}_{1j}) = \hat{\tau}_{01}/(\hat{\tau}_{00}\hat{\tau}_{11}) \) (Raudenbush & Bryk, 2002) between class means on achievement goal orientation and academic self-efficacy effects on student’s achievement goal orientation. In other words, do classes with high means on achievement goal orientation tend also to have a strong relationship between student’s academic self-efficacy and student’s achievement goal orientation?

8. The correlation coefficient \( p(\hat{\beta}_{0j}, \hat{\beta}_{2j}) = \hat{\tau}_{02}/(\hat{\tau}_{00}\hat{\tau}_{22}) \) (Raudenbush & Bryk, 2002) between class means on achievement goal orientation and individual perceived classroom assessment environment effects on student’s achievement goal orientation. In other words, do classes with high means on achievement goal orientation tend also to have a strong relationship between individual perceived classroom assessment environment and student’s achievement goal orientation?

9. The correlation coefficient \( p(\hat{\beta}_{1j}, \hat{\beta}_{2j}) = \hat{\tau}_{12}/(\hat{\tau}_{11}\hat{\tau}_{22}) \) (Raudenbush & Bryk, 2002) between academic self-efficacy effects on student’s achievement goal orientation and individual perceived classroom assessment environment effects on student’s achievement goal orientation. In other words, do classes with a positive relationship between student’s academic self-efficacy and student’s achievement goal orientation tend also to have a strong relationship between individual perceived classroom assessment environment and student’s achievement goal orientation?

10. Reliability estimates, on average, of the class’s intercept (\( \hat{\beta}_{0j} \)) and slope (\( \hat{\beta}_{1j} \) and \( \hat{\beta}_{2j} \)).

**A final explanatory model of achievement goal orientation.** The final step in the analysis involved modeling achievement goal orientation as a function of student-level and class-level characteristics. Having estimated the variability in the class means (i.e., intercepts) and student’s academic self-efficacy effects and individual perceived classroom assessment environment effects (i.e., slopes), I then proceeded with intercepts-and-slopes-as-outcomes regression models to explain the variability in these intercepts and slopes.
The student-level model remained the same as in the random-coefficient regression model. To build the class-level model, Raudenbush and Bryk (2002, p. 267) stated that “we suggest dividing the level-2 predictors into conceptually distinct subsets and fitting a submodel for each. The strongest predictors from these submodels might then be combined in an overall model.” As such, the class-level independent variables were divided into two sets. The first set represented the contextual-effects model of academic self-efficacy and perceived classroom assessment environment along with their differential contextual effects by class’s gender. The second set represented the joint effects model of class’s 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. I next proceeded to develop an overall intercepts-and-slopes-as-outcomes regression model based on the statistically significant variables from the two sets of class-level variables. The two submodels of the intercepts-and-slopes-as-outcomes regression model were as follows:

**Submodel (1): Contextual-effects model of academic self-efficacy and perceived classroom assessment environment along with their differential contextual effects by class’s gender.** Within HLM, a contextual effect is represented by including the class aggregate of a student-level variable in the between-class model for that differentiating effect (Lee & Bryk, 1989; Raudenbush & Bryk, 2002). Differential contextual effects by class’s gender are represented by the inclusion of a class aggregate variable-by-class’s gender interaction term in the between-class model (Lee & Bryk, 1989). Therefore, this
submodel (1) included class’s gender (GNDR), class’s mean for academic self-efficacy (CEFC), and class’s mean for perceived classroom assessment environment (CPCAE) along with the interaction terms (GNDR × CEFC) and (GNDR × CPCAE) within the class-level model representing the differential contextual effects by class’s gender as follows:

**Student-level:**

\[
(GOAL)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + \beta_{2j}(SPCAE_{ij} - CPCAE_{j}) + r_{ij}, \text{ where } r_{ij} \sim iidN(0, \sigma^2).
\]

**Class-level:**

\[
\begin{align*}
\beta_{0j} &= \gamma_{00} + \gamma_{01}(CEFC)_{j} + \gamma_{02}(CPCAE)_{j} + \gamma_{03}(GNDR)_{j} + \\
\beta_{1j} &= \gamma_{10} + \gamma_{11}(CEFC)_{j} + \gamma_{12}(GNDR)_{j} + \gamma_{13}(GNDR \times CEFC)_{j} + u_{0j}, \\
\beta_{2j} &= \gamma_{20} + \gamma_{21}(CPCAE)_{j} + \gamma_{22}(GNDR)_{j} + \gamma_{23}(GNDR \times CPCAE)_{j} + u_{1j}.
\end{align*}
\]

Where \(u_{0j}, u_{1j}, \text{ and } u_{2j}\) are multivariate normally distributed with means 0 and variance-covariance matrix \(T = \begin{bmatrix} \tau_{00} & \tau_{01} & \tau_{02} \\ \tau_{10} & \tau_{11} & \tau_{12} \\ \tau_{20} & \tau_{21} & \tau_{22} \end{bmatrix} \).

**Combined-model:**

\[
(GOAL)_{ij} = \gamma_{00} + \gamma_{01}(CEFC)_{j} + \gamma_{02}(CPCAE)_{j} + \gamma_{03}(GNDR \times CEFC)_{j} + \\
\gamma_{10}(SEFC_{ij} - CEFC_{j}) + \gamma_{11}(CEFC)_{j}(SEFC_{ij} - CEFC_{j}) + \gamma_{12}(GNDR \times CEFC)_{j} + \\
\gamma_{13}(GNDR \times CEFC)_{j}(SEFC_{ij} - CEFC_{j}) + \gamma_{20}(SPCAE_{ij} - CPCAE_{j}) + \\
\gamma_{21}(CPCAE)_{j}(SPCAE_{ij} - CPCAE_{j}) + \gamma_{22}(GNDR \times CPCAE)_{j}(SPCAE_{ij} - \\
CPCAE_{j}) + u_{0j} + u_{1j}(SEFC_{ij} - CEFC_{j}) + u_{2j}(SPCAE_{ij} - CPCAE_{j}) + r_{ij}.
\]

Where:
\( \gamma_{01} \) and \( \gamma_{02} \) are the main effects of class’s mean for academic self-efficacy and class’s mean for perceived classroom assessment environment on class average achievement goal orientation holding other variables constant, respectively.

\( \gamma_{03} \), \( \gamma_{12} \), and \( \gamma_{22} \) are the differences between female classes and male classes on (a) the class average achievement goal orientation, (b) the relationship between student’s academic self-efficacy and student’s achievement goal orientation, and (c) the relationship between individual perceived classroom assessment environment and student’s achievement goal orientation, holding other variables constant, respectively.

\( \gamma_{04} \) and \( \gamma_{13} \) are the two-way interaction effects of class’s gender-by-class’s mean for academic self-efficacy on (a) the class average achievement goal orientation and (b) the relationship between student’s academic self-efficacy and student’s achievement goal orientation holding other variables constant, respectively.

\( \gamma_{05} \) and \( \gamma_{23} \) are the two-way interaction effects of class’s gender-by-class’s mean for perceived classroom assessment environment on (a) the class average achievement goal orientation and (b) the relationship between individual perceived classroom assessment environment and achievement goal orientation holding other variables constant, respectively.

\( r_{i}, u_{0j}, u_{1j}, \) and \( u_{2j} \) are the residuals associated with the achievement goal orientation score for student \( i \) in class \( j \), the average achievement goal orientation for class \( j \), the relationship between student’s academic self-efficacy and student’s achievement goal orientation for class \( j \), and the relationship between individual perceived classroom assessment environment and achievement goal orientation for class \( j \), respectively.

**Submodel (2): Joint effects model of class’s gender and teacher’s assessment practices and years of teaching experience.** This submodel included teacher’s frequent use of traditional assessments (TRAD), teacher’s frequent use of alternative assessments (ALTR), teacher’s frequent use of recommended classroom assessment practices (RECOM), teacher’s years of teaching experience (TEXP), and class’s gender (GNDR) along with the two-and-three-way interaction terms of class’s gender, teacher’s years of
teaching experience, and teacher’s frequent use of alternative assessments within the
class-level model as follows:

**Student-level:**

\[(GOAL)_g = \beta_{0j} + \beta_{1j} (SEFC_g - CEFC_j) + \beta_{2j} (SPCAE_g - CPCAE_j) + r_g,\]

where \( r_g \sim iidN(0, \sigma^2) \).

**Class-level:**

\[
\begin{align*}
\beta_{0j} &= \gamma_{00} + \gamma_{01} (TRAD)_j + \gamma_{02} (ALTR)_j + \gamma_{03} (RECOM)_j + \gamma_{04} (TEXP)_j + \gamma_{05} (GNDR) + \gamma_{06} (TEXP \times GNDR)_j + \gamma_{07} (TEXP \times ALTR)_j + \gamma_{08} (GNDR \times ALTR)_j + \gamma_{09} (TEXP \times GNDR \times ALTR)_j + u_{0j}, \\
\beta_{1j} &= \gamma_{10} + \gamma_{11} (TRAD)_j + \gamma_{12} (ALTR)_j + \gamma_{13} (RECOM)_j + \gamma_{14} (TEXP)_j + \gamma_{15} (GNDR) + \gamma_{16} (TEXP \times GNDR)_j + \gamma_{17} (TEXP \times ALTR)_j + \gamma_{18} (GNDR \times ALTR)_j + \gamma_{19} (TEXP \times GNDR \times ALTR)_j + u_{1j}, \\
\beta_{2j} &= \gamma_{20} + \gamma_{21} (TRAD)_j + \gamma_{22} (ALTR)_j + \gamma_{23} (RECOM)_j + \gamma_{24} (TEXP)_j + \gamma_{25} (GNDR) + \gamma_{26} (TEXP \times GNDR)_j + \gamma_{27} (TEXP \times ALTR)_j + \gamma_{28} (GNDR \times ALTR)_j + \gamma_{29} (TEXP \times GNDR \times ALTR)_j + u_{2j}.
\end{align*}
\]

Where \( u_{0j}, u_{1j}, \) and \( u_{2j} \) are multivariate normally distributed with means 0 and variance-
covariance matrix \( T = \begin{bmatrix} \tau_{00} & \tau_{01} & \tau_{02} \\ \tau_{10} & \tau_{11} & \tau_{12} \\ \tau_{20} & \tau_{21} & \tau_{22} \end{bmatrix} \).

**Combined-model:**

\[
\begin{align*}
(GOAL)_g &= \gamma_{00} + \gamma_{01} (TRAD)_j + \gamma_{02} (ALTR)_j + ... + \gamma_{09} (TEXP \times GNDR \times ALTR)_j + \gamma_{10} (SEFC_g - CEFC_j) + \gamma_{11} (TRAD)_j (SEFC_g - CEFC_j) + ... + \gamma_{19} (TEXP \times GNDR \times ALTR)_j (SEFC_g - CEFC_j) + \gamma_{20} (SPCAE_g - CPCAE_j) + \gamma_{21} (TRAD)_j (SPCAE_g - CPCAE_j) + ... + \gamma_{29} (TEXP \times GNDR \times ALTR)_j (SPCAE_g - CPCAE_j) + u_{0j} + u_{1j} (SEFC_g - CEFC_j) + u_{2j} (SPCAE_g - CPCAE_j) + r_g.
\end{align*}
\]
Where:

$\gamma_{01}, \gamma_{02}, \gamma_{03}, \text{ and } \gamma_{04}$ are the main effects of teacher’s frequent use of traditional assessments, teacher’s frequent use of alternative assessments, teacher’s frequent use of recommended assessment practices, and teacher’s years of teaching experience on class average achievement goal orientation holding other variables constant, respectively.

$\gamma_{11}, \gamma_{12}, \gamma_{13}, \text{ and } \gamma_{14}$ are the main effects of teacher’s frequent use of traditional assessments, teacher’s frequent use of alternative assessments, teacher’s frequent use of recommended assessment practices, and teacher’s years of teaching experience on the relationship between student’s academic self-efficacy and student’s achievement goal orientation holding other variables constant, respectively.

$\gamma_{21}, \gamma_{22}, \gamma_{23}, \text{ and } \gamma_{24}$ are the main effects of teacher’s frequent use of traditional assessments, teacher’s frequent use of alternative assessments, teacher’s frequent use of recommended assessment practices, and teacher’s years of teaching experience on the relationship between individual perceived classroom assessment environment and student’s achievement goal orientation holding other variables constant, respectively.

$\gamma_{05}, \gamma_{15}, \text{ and } \gamma_{25}$ are the differences between female classes and male classes on (a) the class average achievement goal orientation, (b) the relationship between student’s academic self-efficacy and student’s achievement goal orientation, and (c) the relationship between individual perceived classroom assessment environment and student’s achievement goal orientation, holding other variables constant, respectively.

$\gamma_{06}, \gamma_{16}, \text{ and } \gamma_{26}$ are the two-way interaction effects of teacher’s years of teaching experience and class’s gender on (a) the class average achievement goal orientation, (b) the relationship between student’s academic self-efficacy and student’s achievement goal orientation, and (c) the relationship between individual perceived classroom assessment environment and student’s achievement goal orientation, holding other variables constant, respectively.

$\gamma_{07}, \gamma_{17}, \text{ and } \gamma_{27}$ are the two-way interaction effects of teacher’s years of teaching experience and teacher’s frequent use of alternative assessments on (a) the class average achievement goal orientation, (b) the relationship between student’s academic self-efficacy and student’s achievement goal orientation, and (c) the relationship between individual perceived classroom assessment environment and
student’s achievement goal orientation, holding other variables constant, respectively.

\( \gamma_{08}, \gamma_{18}, \) and \( \gamma_{28} \) are the two-way interaction effects of class’s gender and teacher’s frequent use of alternative assessments on (a) the class average achievement goal orientation, (b) the relationship between student’s academic self-efficacy and student’s achievement goal orientation, and (c) the relationship between individual perceived classroom assessment environment and student’s achievement goal orientation, holding other variables constant, respectively.

\( \gamma_{09}, \gamma_{19}, \) and \( \gamma_{29} \) are the three-way interaction effects of teacher’s years of teaching experience-by-class’s gender-by teacher’s frequent use of alternative assessments on (a) the class average achievement goal orientation, (b) the relationship between student’s academic self-efficacy and student’s achievement goal orientation, and (c) the relationship between individual perceived classroom assessment environment and student’s achievement goal orientation, holding other variables constant, respectively.

In each of the aforementioned submodels, hypotheses tests for the effect of each class-level variable and interaction on class average achievement goal orientation (i.e., intercept \( \beta_{0j} \)), on the relationship between student’s academic self-efficacy and student’s achievement goal orientation (i.e., slope \( \beta_{1j} \)), and on the relationship between individual perceived classroom assessment environment and student’s achievement goal orientation (i.e., slope \( \beta_{2j} \)), (i.e., \( H_0: \gamma_{qs} = 0 \), where \( q = 0, 1, \) and \( 2; \) and \( s = 1, \ldots, S \) class-level variable and interaction), were examined to determine the statistically significant class-level variables and interactions. These statistically significant class-level variables and interactions were then combined together to produce a parsimonious overall intercepts-and-slopes-as-outcomes regression model explaining the variability in the aforementioned intercepts and slopes. The proportion of variance in these intercepts and slopes was computed as follows
\[ \left| \hat{\tau}_{qq} (\text{base model}) - \hat{\tau}_{qq} (\text{overall model}) \right| / \hat{\tau}_{qq} (\text{base model}) \], where \( q = 0, 1, \text{ and } 2 \) (Raudenbush & Bryk, 2002). The random-coefficient regression model provides the appropriate base model for this application.

The validity of inferences based on the final model was assessed by verifying the tenability of the assumptions of a two-level HLM (Raudenbush & Bryk, 2002, p. 255). On the basis of residual analyses, variables that were deleted in the early steps of the analysis were reconsidered for inclusion in the final composite model to obtain an adequate fitted model. Across all tested models, the student-level coefficients were specified as fixed, random, or non-randomly varying at the class-level model by examining the following indices (Raudenbush & Bryk, 2002):

1. The point estimate of \( \tau_{qq} \). As \( \tau_{qq} \) becomes negligible, its corresponding coefficient may be specified as either fixed or non-randomly varying.

2. The \( \chi^2 \) homogeneity test of \( H_0: \tau_{qq} = 0 \). As \( H_0 \) becomes tenable, its corresponding coefficient may be specified as either fixed or non-randomly varying.

3. The likelihood ratio test (i.e., deviance) of the variance-covariance components. If there is no statistically significant difference between the deviances of two models, then the reduced model will typically be preferred.

4. The reliability of \( \beta_{qj} \) because it tells us how much of the observed variation in the estimated \( \beta_{qj} \) is potentially explainable. As the reliability drops below .05, that coefficient may be specified as either fixed or non-randomly varying. Also, the reliability of each class-level random effect provides some insight about the available power in the data to detect statistically significant effects of class characteristics on the outcome. A small reliability coefficient suggests caution in inferring about the effects of class-level characteristics.
5. Correlations among $\beta_{\omega}$, as the correlations are high (.90 and above), one or more of the coefficients may be specified as either fixed or non-randomly varying depending on the literature because the two random effects are carrying the same variation across level-2 units.

Summary

The participants in this study were 1,636 ninth grade students and their corresponding 83 science teachers enrolled in public schools within Muscat educational region in Oman during the spring semester 2007. Two questionnaires were developed and used, one for students and one for teachers. The student’s questionnaire focused on students’ perceived classroom assessment environment, achievement goal orientations, and academic self-efficacy. The teacher’s questionnaire focused on teachers’ frequent uses of traditional assessments, alternative assessments, and classroom assessment practices recommended by experts of educational measurement and assessment. Principal components/exploratory factor analyses were conducted to identify the underlying dimensions of students’ perceptions of classroom assessment environment and achievement goal orientations. Hierarchical linear model (HLM) analyses were employed to examine the effects of certain student-level and class-level variables on students’ perceptions of classroom assessment environment and achievement goal orientations. Results of these analyses are presented in the next chapter of the dissertation.
CHAPTER IV

RESULTS

The research problem addressed in this study was the development of models that might explain the variation in student perceived classroom assessment environment and achievement goal orientations in ninth grade Muscat science classrooms in Oman. Within the framework of achievement goal theory and classroom assessment literature, ten general research questions shaped this problem as follows:

1. What are the underlying dimensions of ninth grade students’ perceptions of classroom assessment environment in Muscat science classrooms in Oman?

2. What are the underlying dimensions of ninth grade students’ achievement goal orientations in Muscat science classrooms in Oman?

3. Do ninth grade Muscat science classrooms in Oman vary in perceived classroom assessment environment (PCAE)?

4. What are some of the student characteristics (e.g., academic self-efficacy) that might have effects on PCAE?

5. Do the effects of student characteristics (e.g., academic self-efficacy) on PCAE vary across classrooms?

6. What are some of the classroom characteristics (e.g., class gender) that might help explain the variability in the PCAE and in the effects of student characteristics (e.g., academic self-efficacy) on PCAE across classrooms?

7. Do ninth grade Muscat science classrooms in Oman vary in achievement goal orientations (GOAL)?
Following are results pertaining to these questions.

Research Question One

RQ (1): What are the underlying dimensions of ninth grade students’ perceptions of classroom assessment environment in Muscat science classrooms in Oman?

The first research question was concerned with identifying the underlying dimensions of classroom assessment environment as perceived by ninth grade students in Muscat science classrooms in Oman. The classroom assessment literature was consulted to develop 25 questionnaire items reflecting three aspects of the classroom assessment assumed to be more directly related to students’ experience of their classroom assessment environment (Crooks, 1988; Wang, 2004). These aspects were: assessment tasks, assessment standards and criteria, and assessment feedback.

Due to changes in scales’ anchors, wording, language, and study context, principal components/exploratory factor analyses (PCA/EFA) were conducted on the items to identify their underlying dimensions. In light of achievement goal theory and classroom assessment literature (e.g., Ames, 1992a, 1992b; McMillan & Workman, 1998), it was expected that the PCA/EFA would yield two independent dimensions of perceived classroom assessment environment (a) a learning-oriented and (b) a normative-oriented
classroom assessment environment whether rotated orthogonally or obliquely. Prior to the analyses, students’ responses to the 25 items of the perceived classroom assessment environment questionnaire were screened for accuracy of data entry, missing values, normality, linearity, outliers, multicollinearity and singularity, and factorability. Following is a detailed description of the results.

Data Screening and Assumptions

Accuracy of data entry and missing values. Initially data were available from responses of 1636 students on 25 items assessing perceived classroom assessment environment using a 4-point Likert scale ranging from 1 (completely not true) to 4 (completely true). Inspection of the frequency distribution of each item (i.e., variable) showed that all data were plausible with no missing values. Therefore, using Comrey and Lee’s (1992, as cited in Tabachnick and Fidell, 2001, p. 588) guidelines, 1636 students provided an “excellent” sample size for PCA/EFA.

Outliers. As suggested by Tabachnick and Fidell (2001, p.67), “Cases with standardized scores in excess of 3.29 (p < .001, two-tailed test) are potential outliers.” Using this criterion, 39, 55, 46, and 87 cases were identified as outliers on the following variables (i.e., items): In this class, teacher’s oral questions encourage thinking; In this class, the tests match what we learn in class; In this class, the teacher uses more than one way to determine grades such as tests, projects…etc, and In this class, my teacher encourages viewing mistakes as learning opportunities; respectively. Further inspection of
these cases revealed no outlying response pattern on other variables, indicating that they were part of the target population, and thereby as suggested by Tabachnick and Fidell (2001), they should remain in the analysis. Using Molloy and Newfields’s (2005) recommendation for reducing the influence of outliers, these outlying cases were assigned a raw score on the aforementioned offending variables that was three standard deviations below the mean as shown in Table 1.

Table 1

<table>
<thead>
<tr>
<th>Variable (item)</th>
<th>$M$</th>
<th>$SD$</th>
<th>Assigned raw score = $M - (3 \times SD)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. In this class, teacher’s oral questions encourage thinking.</td>
<td>3.64</td>
<td>.69</td>
<td>$3.64 - (3 \times .69) = 1.57$</td>
</tr>
<tr>
<td>2. In this class, the tests match what we learn in class.</td>
<td>3.63</td>
<td>.72</td>
<td>$3.63 - (3 \times .72) = 1.47$</td>
</tr>
<tr>
<td>3. In this class, the teacher uses more than one way to determine grades such as tests, projects…etc.</td>
<td>3.62</td>
<td>.71</td>
<td>$3.62 - (3 \times .71) = 1.49$</td>
</tr>
<tr>
<td>4. In this class, my teacher encourages viewing mistakes as learning opportunities.</td>
<td>3.58</td>
<td>.80</td>
<td>$3.58 - (3 \times .80) = 1.18$</td>
</tr>
</tbody>
</table>

Normality. After dealing with outliers, distributions of the 25 items of perceived classroom assessment environment were examined for skewness and kurtosis. Using George and Mallery’s (2001) guidelines, results showed that the skewness and kurtosis values for all variables were within acceptable range of ± 2. Therefore, there was no concern about deviation from normality.
**Linearity.** With 25 variables, an examination of all pairwise scatter plots (about 600 plots) was impractical. Therefore, a spot check on a few plots was run. Results showed no threat to linearity.

**Multicollinearity and singularity.** Using Tabachnick and Fidell’s (2001) guidelines, the original nonrotated PCA revealed that the smallest eigenvalue was .538, not dangerously close to zero. Simultaneously, it was observed that the squared multiple correlations (SMCs) between the variables where each, in turn, served as a dependent variable for the others, did not approach 1. The largest SMC among the variables was .276. Therefore, multicollinearity and singularity were not a threat in this data set.

**Factorability of the correlation matrix (R).** Inspection of the correlation matrix of the 25 variables revealed that the correlations when taken overall were statistically significant as indicated by the Bartlett’s test of sphericity, $\chi^2 (300) = 4202.865, p < .001$. According to Tabachnick and Fidell (2001), the Kaiser’s measure of sampling adequacy (MSA) fell within acceptable range (values of .60 and above) with a value of .794. Each variable also exceeded the threshold value (.60) of MSA. Finally, most of the partial correlations were small as indicated by the anti-image correlation matrix. These measures all led to the conclusion that the set of 25 items of perceived classroom assessment environment was appropriate for PCA.
Principal Components Analyses

Principal components analyses (PCA) were conducted on the 25 items of perceived classroom assessment environment to determine their underlying dimensions. The initial unrotated PCA resulted in a factor model of four dimensions as suggested by the scree plot and eigenvalues exceeding unity. However, based on its pattern of factor loadings, this unrotated factor model was theoretically less meaningful and as such was difficult to interpret. Therefore, the analysis proceeded to rotate the factor matrix both orthogonally and obliquely to achieve a simple and theoretically more meaningful solution. A total of nine orthogonal and nine oblique rotations were run. Varimax rotation was used for the orthogonal solution, and oblimin rotation was used for the oblique solution. In each time, the pattern of factor loadings was examined, and items loading highly on multiple factors (i.e., complex items) were identified. These items were candidates for deletion. The purpose of item deletion was to develop a more parsimonious scale of perceived classroom assessment environment and to approximate the simple structure as closely as possible. Twelve items were deleted through this process. It should be noted that, in the classroom assessment literature, the perceived classroom assessment environment construct has not been well defined in operational terms (Wang, 2004). Therefore, according to DeVellis’s (2003) guidelines in scale development, “It would not be unusual to begin with a pool of items that is three of four times as large as the final scale…If items are particularly difficult to generate for a given content area [like the area in classroom assessment environment being considered in the present study] or if empirical data [like the PCA/EFA output in the present study] indicate that numerous
items are not needed…., then the initial pool may be as small as 50% larger than the final scale” (p. 66).

The analyses yielded three factors as suggested by the eigenvalue rule and scree plot. Although the pattern of factor loadings was similar in both the orthogonal and the oblique solutions, the orthogonal solution was decided to be retained for interpretation because, as suggested by Tabachnick and Fidell (2001), the correlations among the factors in the factor correlation matrix were smaller than .24. Table 2 displays the factor loadings for this orthogonal three-factor model of perceived classroom assessment environment. All items loaded above .35 on their primary factor. The minimum difference between the primary and secondary loadings was .18. Together the three factors accounted for 41.02% of the total variance. The first factor accounted for 16.51% of the variance (eigenvalue = 2.15) and consisted of five items. According to the content of its items shown in Table 2 and in light of achievement goal theory and classroom assessment literature (Ames, 1992a, 1992b; McMillan & Workman, 1998; Wang, 2004, p. 59), this factor was labeled a perceived “learning” classroom assessment environment because its items focused on classroom assessment practices that improve student learning and mastery of content materials. The second factor accounted for 12.31% of the variance (eigenvalue = 1.60) and consisted of three items. According to the content of its items presented in Table 2 and in light of achievement goal theory and classroom assessment literature (Church et al., 2001, p. 44), this factor was named a perceived
Table 2

Summary of Factor Loadings by Principal Components Analysis for the Orthogonal Three-Factor Model of Perceived Classroom Assessment Environment

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. In this class, my teacher helps me identify the places where I</td>
<td>.71</td>
</tr>
<tr>
<td>need more effort in future.</td>
<td></td>
</tr>
<tr>
<td>2. In this class, my teacher encourages viewing mistakes as</td>
<td>.69</td>
</tr>
<tr>
<td>learning opportunities.</td>
<td></td>
</tr>
<tr>
<td>3. In this class, students are given a chance to correct their</td>
<td>.60</td>
</tr>
<tr>
<td>mistakes.</td>
<td></td>
</tr>
<tr>
<td>4. In this class, teacher’s oral questions encourage thinking.</td>
<td>.56</td>
</tr>
<tr>
<td>5. In this class, my teacher’s grading is clear.</td>
<td>.55</td>
</tr>
<tr>
<td>6. The science tests in this class are difficult to students.</td>
<td>.77</td>
</tr>
<tr>
<td>7. It is difficult to achieve high grades in this class.</td>
<td>.71</td>
</tr>
<tr>
<td>8. In this class, the homework is boring.</td>
<td>.62</td>
</tr>
<tr>
<td>9. In this class, students who do poorly are criticized in front of</td>
<td>.64</td>
</tr>
<tr>
<td>the whole class.</td>
<td></td>
</tr>
<tr>
<td>10. In this class, my teacher compares my performance with</td>
<td>.62</td>
</tr>
<tr>
<td>performance of other students.</td>
<td></td>
</tr>
<tr>
<td>11. In this class, students who do well are praised in front of</td>
<td>.60</td>
</tr>
<tr>
<td>the whole class.</td>
<td></td>
</tr>
<tr>
<td>12. In this class, the grades are used to penalize students who</td>
<td>.49</td>
</tr>
<tr>
<td>disturb the class.</td>
<td></td>
</tr>
<tr>
<td>13. In this class, the teacher is more concerned with our grades</td>
<td>.39</td>
</tr>
<tr>
<td>than our learning.</td>
<td></td>
</tr>
</tbody>
</table>

Note. Factor 1 = perceived learning classroom assessment environment. Factor 2 = perceived harsh classroom assessment environment. Factor 3 = perceived public classroom assessment environment.
“harsh” classroom assessment environment because its items focused on harshness of assessment and grading. The third factor accounted for 12.21% of the variance (eigenvalue = 1.59) and consisted of five items. According to the content of its items presented in Table 2 and in light of achievement goal theory and classroom assessment literature (Ames, 1992a, 1992b), this factor was labeled a perceived “public” classroom assessment environment because its items focused on public evaluation and recognition practices.

Having found that a three-factor model of perceived classroom assessment environment was the best model to represent the data, three scales as suggested by Hair et al. (1998) were developed by summing item responses in each dimension of the three-factor model of perceived classroom assessment environment. Table 3 presents descriptive statistics and reliabilities for, and zero-order intercorrelations among the three scales of perceived classroom assessment environment. As shown in Table 3, on average, students tended to favorably perceive their classroom assessment environment as being more learning-oriented than harsh- or public-oriented. Also, there was more variability in students’ perceptions of their classroom assessment environment for being public-oriented than for being harsh- or learning-oriented.

As also presented in Table 3, the magnitude and sign of the zero-order correlations among the dimensions of perceived classroom assessment environment suggested that these dimensions represented related, yet unique, aspects of the classroom assessment environment as perceived by the participating students. Perceived learning classroom assessment environment was negatively related to perceived harsh and perceived public classroom assessment environments, whereas perceived public and perceived harsh
classroom assessment environments were positively correlated with each other. These
correlational patterns suggest that students who perceived their classroom assessment
environment as being learning-oriented were less likely to perceive it as being either
harsh- or public-oriented, whereas students who perceived their classroom assessment
environment as being public-oriented tended also to perceive it as being harsh.

Table 3

Descriptive Statistics and Reliabilities for, and Zero-Order Correlations among the Three Scales of Perceived Classroom Assessment Environment

<table>
<thead>
<tr>
<th>Factor</th>
<th>No. of items</th>
<th>Expected score range</th>
<th>M</th>
<th>SD</th>
<th>Reliability</th>
<th>SEM</th>
<th>Zero-order correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>5 - 20</td>
<td>17.18</td>
<td>2.53</td>
<td>.63</td>
<td>1.54</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>3 - 12</td>
<td>6.86</td>
<td>2.10</td>
<td>.52</td>
<td>1.45</td>
<td>-.25***</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>5 - 20</td>
<td>12.07</td>
<td>3.10</td>
<td>.45</td>
<td>2.30</td>
<td>-.12*** .21***</td>
</tr>
</tbody>
</table>

According to Table 3, the scores from the scales measuring perceived classroom assessment environment did not show a high level of internal consistency as indicated by Cronbach’s alpha (DeVellis, 2003). Crocker and Algina (1986, p. 135) stated that “the internal consistency coefficient is an index of both content homogeneity and item quality.” Based on the content validation process described in Chapter III of the dissertation, the initial pool of perceived classroom assessment environment items was judged to be adequately representing the content domain of the perceived classroom assessment environment construct. However, the reliability data suggested little consistency in students’ responses to the items. One possible explanation for this situation could be that some items were not well written or the instructions were not clear, causing students to misinterpret the items and respond on some basis unrelated to the content being measured, thereby lowering the internal consistency (Crocker & Algina, 1986). More details about reliability will be discussed in the next chapter of the dissertation.

Research Question Two

RQ (2): What are the underlying dimensions of ninth grade students’ achievement goal orientations in Muscat science classrooms in Oman?

The second research question was concerned with identifying the underlying dimensions of achievement goal orientations for ninth grade students in Muscat science classrooms in Oman. A previously 14-items achievement goal questionnaire developed and validated by Midgley et al. (2000) for middle and high school students in the United States was used in this study. However, due to changes in scales’ anchors, wording,
language, and study context, principal components/exploratory factor analyses (PCA/EFA) were conducted on the items to identify their underlying dimensions. Consistent with the trichotomous framework of achievement goal theory described more fully in Chapter II of the dissertation, it was expected that the PCA/EFA would yield the three dimensions of achievement goals (mastery, performance-approach, and performance-avoidance) with (a) a positive correlation between the two types of performance goals and (b) no correlations between mastery goals and each of the two types of performance goals whether rotated orthogonally or obliquely. Prior to the analyses, students’ responses to the 14 items of the achievement goal questionnaire were screened for accuracy of data entry, missing values, normality, linearity, outliers, multicollinearity and singularity, and factorability. Following is a detailed description of the results.

Data Screening and Assumptions

Accuracy of data entry and missing values. Initially data were available from responses of 1636 students on 14 items assessing achievement goal orientations using a 4-point Likert scale ranging from 1 (completely not true) to 4 (completely true). Inspection of the frequency distribution of each item (i.e., variable) showed that all data were plausible with no missing values. Therefore, using Comrey and Lee’s (1992, as cited in Tabachnick and Fidell, 2001, p. 588) guidelines, 1636 students provided an “excellent” sample size for PCA/EFA.
Outliers. As suggested by Tabachnick and Fidell (2001, p.67), “Cases with standardized scores in excess of 3.29 ($p < .001$, two-tailed test) are potential outliers.” Using this criterion, between 22 and 57 cases were identified as outliers on the five mastery goal items. Further inspection of these cases revealed no outlying response pattern on other variables, indicating that they were part of the target population, and thereby as suggested by Tabachnick and Fidell (2001), they should remain in the analysis. Using Molloy and Newfields’s (2005) recommendation for reducing the influence of outliers, these outlying cases were assigned a raw score on the aforementioned offending variables that was three standard deviations below the mean as shown in Table 4.

Table 4

Assigned Raw Scores to the Outlying Cases on the Offending Variables (Items) of Achievement Goal Orientations

<table>
<thead>
<tr>
<th>Variable (item)</th>
<th>$M$</th>
<th>$SD$</th>
<th>Assigned raw score = $M - (3 \times SD)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. It is important to me that I learn a lot of new science concepts this semester.</td>
<td>3.70</td>
<td>.63</td>
<td>3.70 – (3 $\times$ .63) = 1.81</td>
</tr>
<tr>
<td>2. One of my goals is to master a lot of new science skills this semester.</td>
<td>3.56</td>
<td>.75</td>
<td>3.56 – (3 $\times$ .75) = 1.31</td>
</tr>
<tr>
<td>3. One of my goals in science class is to learn as much as I can.</td>
<td>3.51</td>
<td>.72</td>
<td>3.51 – (3 $\times$ .72) = 1.35</td>
</tr>
<tr>
<td>4. It is important to me that I thoroughly understand my science class work.</td>
<td>3.69</td>
<td>.65</td>
<td>3.69 – (3 $\times$ .65) = 1.74</td>
</tr>
<tr>
<td>5. It is important to me that I improve my science skills this semester.</td>
<td>3.75</td>
<td>.60</td>
<td>3.75 – (3 $\times$ .60) = 1.95</td>
</tr>
</tbody>
</table>
Normality. After dealing with outliers, distributions of the 14 items of achievement goal orientations were examined for skewness and kurtosis. Using George and Mallery’s (2001) guidelines, results showed that the skewness and kurtosis values for all variables were within acceptable range of ±2. Therefore, there was no concern about deviation from normality.

Linearity. With 14 variables, an examination of all pairwise scatter plots (about 182 plots) was impractical. Therefore, a spot check on a few plots was run. Results showed no threat to linearity.

Multicollinearity and singularity. Using Tabachnick and Fidell’ (2001) guidelines, the original nonrotated PCA revealed that the smallest eigenvalue was .497, not dangerously close to zero. Simultaneously, it was observed that the squared multiple correlations (SMCs) between the variables where each, in turn, served as a dependent variable for the others, did not approach 1. The largest SMC among the variables was .361. Therefore, multicollinearity and singularity were not a threat in this data set.

Factorability of the correlation matrix (R). Inspection of the correlation matrix of the 14 variables revealed that the correlations when taken overall were statistically significant as indicated by the Bartlett’s test of sphericity, $\chi^2(91) = 3863.283, p < .001$. According to Tabachnick and Fidell (2001), the Kaiser’s measure of sampling adequacy (MSA) fell within acceptable range (values of .60 and above) with a value of .867. Each of the variables also exceeded the threshold value (.60) of MSA. Finally, most of the partial
correlations were small as indicated by the anti-image correlation matrix. These measures all led to the conclusion that the set of 14 items of achievement goal orientations was appropriate for PCA.

Principal Components Analyses

Principal components analyses (PCA) were conducted on the 14 items of achievement goal orientations to determine their underlying dimensions. The initial unrotated PCA resulted in a factor model of three dimensions as indicated by the scree plot and eigenvalues exceeding unity. However, based on its pattern of factor loadings, this unrotated factor model was theoretically less meaningful and as such was difficult to interpret. Therefore, the analysis proceeded to rotate the factor matrix both orthogonally and obliquely to achieve a simple and theoretically more meaningful solution. One orthogonal rotation and one oblique rotation were run. Varimax rotation was used for the orthogonal solution, and oblimin rotation was used for the oblique solution. Both rotations resulted in a factor model of three dimensions as suggested by the scree plot and eigenvalues exceeding unity. There were no remarkable differences between the orthogonal solution and the oblique solution in terms of factor structure and pattern of factor loadings. However, as suggested by Tabachnick and Fidell (2001), the oblique solution was retained because the correlation between the first factor and the third factor was .38 as indicated by the factor correlation matrix.
Table 5 displays the factor loadings for the oblique three-factor model of achievement goal orientations. All items loaded .59 and above on their primary factor. The minimum difference between the primary and secondary loadings was .22. Together the three factors accounted for 45.82% of the total variance. The first factor accounted for 18.13% of the variance (eigenvalue = 2.538) and consisted of five performance-approach goal orientation items. The second factor accounted for 15.50% of the variance (eigenvalue = 2.169) and consisted of five mastery goal orientation items. The third factor accounted for 12.19% of the variance (eigenvalue = 1.707) and consisted of four performance-avoidance goal orientation items.

Having found that the trichotomous model of achievement goals was the best model to represent the data, three summated scales as suggested by Hair et al. (1998) were developed by averaging item responses in each dimension of the oblique three-factor model of achievement goals. Table 6 presents descriptive statistics and reliabilities for, and zero-order intercorrelations among the three scales of achievement goals. As shown in Table 6, the means for the achievement goals tended to be higher than their respective scales’ midpoints, with small variability among students’ responses suggesting that together the three goals could be operative in the achievement setting at the same time. The three measures of achievement goals were significantly moderately and positively correlated with each other suggesting that they could be measuring the same construct. The positive correlations among achievement goals suggest that the participating students
Table 5

*Summary of Factor Loadings by Principal Components Analysis for the Oblique Three-Factor Model of Achievement Goal Orientations*

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. It is important to me that I look smart compared to others in my class.</td>
<td>.76</td>
</tr>
<tr>
<td>2. One of my goals is to show others that I’m good at science class work.</td>
<td>.74</td>
</tr>
<tr>
<td>3. It is important to me that other students in my class think I am good at my science class work.</td>
<td>.70</td>
</tr>
<tr>
<td>4. One of my goals is to show others that science class work is easy for me.</td>
<td>.67</td>
</tr>
<tr>
<td>5. One of my goals is to look smart in comparison to the other students in my class.</td>
<td>.67</td>
</tr>
<tr>
<td>6. It is important to me that I thoroughly understand my science class work.</td>
<td>.68</td>
</tr>
<tr>
<td>7. It is important to me that I improve my science skills this semester.</td>
<td>.66</td>
</tr>
<tr>
<td>8. One of my goals is to master a lot of new science skills this semester.</td>
<td>.64</td>
</tr>
<tr>
<td>9. It is important to me that I learn a lot of new science concepts this semester.</td>
<td>.63</td>
</tr>
<tr>
<td>10. One of my goals in science class is to learn as much as I can.</td>
<td>.63</td>
</tr>
<tr>
<td>11. One of my goals in science class is to avoid looking like I have trouble doing the work.</td>
<td>.70</td>
</tr>
<tr>
<td>12. One of my goals is to keep others from thinking I’m not smart in science class.</td>
<td>.66</td>
</tr>
<tr>
<td>13. It is important to me that my teacher doesn’t think that I know less than others in science.</td>
<td>.61</td>
</tr>
<tr>
<td>14. It is important to me that I don’t look stupid in science class.</td>
<td>.59</td>
</tr>
</tbody>
</table>

**Factor correlations**

<table>
<thead>
<tr>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Factor 1 = performance-approach goal orientation. Factor 2 = mastery goal orientation. Factor 3 = performance-avoidance goal orientation.
Table 6

Descriptive Statistics and Reliabilities for, and Zero-Order Correlations among the Three Scales of Achievement Goal Orientations

<table>
<thead>
<tr>
<th>Factor</th>
<th>No. of items</th>
<th>Expected score range</th>
<th>M</th>
<th>SD</th>
<th>Reliability</th>
<th>SEM</th>
<th>Zero-order correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>1 - 4</td>
<td>3.21</td>
<td>.69</td>
<td>.75</td>
<td>.35</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>1 - 4</td>
<td>3.66</td>
<td>.38</td>
<td>.66</td>
<td>.22</td>
<td>.32***</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>1 - 4</td>
<td>3.12</td>
<td>.69</td>
<td>.54</td>
<td>.47</td>
<td>.45*** .33*** -</td>
</tr>
</tbody>
</table>

*Note.* Factor 1 = performance-approach goal orientation. Factor 2 = mastery goal orientation. Factor 3 = performance-avoidance goal orientation. ***p < .001.

In this study may pursue more than one goal at the same time in the achievement setting (Pintrich, 2000b).

In his 1967’s book of *Psychometric Theory*, Nunnally stated that in the early-stage research, “reliabilities of .60 or .50 will suffice” (p. 226). In this study, the Midgley et al.’s (2000) achievement goal orientation scales were modified in terms of scales’ anchors and language and were used for the first time with Omani students to the best of my knowledge. Taking this relatively new research context into account, the Cronbach’s alpha reliability coefficients of .75, .66, and .54 found in this study fell within the aforementioned acceptable range specified by Nunnally (1967). However, more details about reliability will be discussed in the next chapter of the dissertation.
Research Questions Three, Four, Five, and Six

RQ (3):  Do ninth grade Muscat science classrooms in Oman vary in perceived classroom assessment environment (PCAE)?

RQ (4):  What are some of the student characteristics (e.g., academic self-efficacy) that might have effects on PCAE?

RQ (5):  Do the effects of student characteristics (e.g., academic self-efficacy) on PCAE vary across classrooms?

RQ (6):  What are some of the classroom characteristics (e.g., class gender) that might help explain the variability in the PCAE and in the effects of student characteristics (e.g., academic self-efficacy) on PCAE across classrooms?

The goal of these research questions was to construct a parsimonious model explaining student’s perceived classroom assessment environment as a function of student-level and class-level characteristics. The data pertaining to these questions were hierarchically structured, in that students were nested within classes. Therefore, hierarchical linear model (HLM) analyses (Raudenbush & Bryk, 2002) were conducted, one for each dependent variable defined as a dimension of student’s perceived classroom assessment environment that emerged from the PCA/EFA described in the first research question. As discussed in the first research question, results of the PCA/EFA showed three dimensions of perceived classroom assessment environment: (a) learning-, (b) harsh-, and (c) public-oriented classroom assessment environments. Therefore, based on the PCA/EFA results and the literature reviewed in Chapter II of the dissertation, the research questions from three to six had the following dependent and independent variables.
Dependent variables:

- Student’s perceived learning classroom assessment environment (SPLCAE)
- Student’s perceived harsh classroom assessment environment (SPHCAE)
- Student’s perceived public classroom assessment environment (SPPCAE)

Independent variable at the student-level:

- Student’s academic self-efficacy (SEFC)

Independent variables at the class-level:

- Class gender (GNDR)
- Teacher’s years of teaching experience (TEXP)
- Teacher’s frequent use of traditional assessments (TRAD)
- Teacher’s frequent use of alternative assessments (ALTR)
- Teacher’s frequent use of recommended assessment practices (RECOM)
- Class’s average for academic self-efficacy (CEFC); as discussed in Chapters I and II of the dissertation, the variable CEFC was included in order to study its contextual effect (Kreft & Leeuw, 1998; Raudenbush & Bryk, 2002) on perceived classroom assessment environment. Kreft and Leeuw (1998) stated that “contextual models are defined as regression models containing two types of variables: individual-level variables and aggregated context variables, such as group means” (p. 8). When applying Kreft and Leeuw’s (1998) definition of contextual models in the present study, the variable SEFC was used twice: once as an individual student variable and once as an aggregated class characteristic.

Moreover, the following three independent variables class’s average for perceived learning classroom assessment environment (CPLCAE), class’s average for perceived harsh classroom assessment environment (CPHCAE), and class’s average for perceived public classroom assessment environment (CPPCAE) were included in order to study their compositional effects on an individual student’s perception of the classroom assessment environment because as discussed in Chapters I and II of the dissertation, it was assumed that “students construct their own meaning [of the classroom assessment environment] based in part on their group experiences” (Brookhart, 2004, p. 445):

- Class’s average for perceived learning assessment environment (CPLCAE)
- Class’s average for perceived harsh assessment environment (CPHCAE)
- Class’s average for perceived public assessment environment (CPPCAE)
In order to facilitate interpretation of the HLM results, all dependent and independent 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 modeling process for each dependent variable began with a fully unconditional model followed by a random-coefficient regression model to test the effect of student-level independent variable. To build the class-level model, the class-level independent variables were divided into three sets following Raudenbush and Bryk’s (2002, p. 267) suggestion. 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’s 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 the statistically significant class-level variables from the aforementioned submodels to produce a parsimonious overall intercepts-and-slopes-as-outcomes regression model of each dimension of perceived classroom assessment environment. Finally, the validity of inferences based on the final model was assessed by verifying the tenability of two-level HLM assumptions (Raudenbush & Bryk, 2002, p. 255). Prior to the HLM analyses, univariate shape and frequency distribution of each variable at each level and bivariate relationships at each
level were examined to identify discrepant cases and unusual observations that could have led to implausible results. Following is a detailed description of the results.

**Data Screening at the Student-Level**

*Accuracy of data entry and missing values.* Inspection of the frequency distribution of each student-level variable showed that all data seemed plausible with no missing values.

*Outliers.* As suggested by Tabachnick and Fidell (2001, p. 67), “Cases with standardized scores in excess of 3.29 ($p < .001$, two-tailed test) are potential outliers.” Using this criterion, 10, 19, and 22 cases were identified as outliers on the following variables: *student academic self-efficacy, student perceived learning classroom assessment environment,* and *student mastery goal orientation,* respectively. Further inspection of the data revealed no outlying response pattern of these cases on other variables suggesting that they were part of the target population, and thereby as suggested by Tabachnick and Fidell (2001), they should remain in the analysis. Using Molloy and Newfields’s (2005) recommendation for reducing the influence of outliers, these outlying cases were assigned a raw score on the aforementioned offending variables that was three standard deviations below the mean as shown Table 7.

*Descriptive statistics.* After dealing with outliers, Table 8 presents descriptive statistics of student-level variables for the overall sample of students and for males and
Table 7

Assigned Raw Scores to the Outlying Cases on the Offending Student-Level Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>Assigned raw score = M – (3 × SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Student academic self-efficacy</td>
<td>20.22</td>
<td>2.78</td>
<td>20.22 – (3 × 2.78) = 11.88</td>
</tr>
<tr>
<td>2. Student perceived learning classroom</td>
<td>17.16</td>
<td>2.62</td>
<td>17.16 – (3 × 2.62) = 9.30</td>
</tr>
<tr>
<td>assessment environment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Student mastery goal orientation</td>
<td>3.65</td>
<td>.40</td>
<td>3.65 – (3 × .40) = 2.45</td>
</tr>
</tbody>
</table>

females. Overall, with the exception of perceived harsh and perceived public classroom assessment environments, the means for student-level variables tended to be higher than their respective scales’ midpoints, with small variability among students’ responses.

Further, using George and Mallery’s (2001) guidelines, the skewness and kurtosis values of each student-level variable for the overall data set were within acceptable range of ±2, suggesting no concern about deviation from normality.

When classified by gender, female students’ scores tended on average to be higher than male students’ scores across all student-level variables, except for perceived harsh and perceived public classroom assessment environments where male students’ scores were on average higher than female students’ scores. Although the distribution of female students’ scores pertaining to mastery goal orientations tended to be flatter than the normal distribution (kurtosis = 2.67), the skewness and kurtosis values of each student-level variable for males and females were within acceptable range of ±2 as suggested by George and Mallery’s (2001) guidelines. Therefore, there was no concern about deviation from normality for student-level variables within each group.
Table 8

Descriptive Statistics of Student-Level Variables for Overall Sample of and for Male and Female Students

<table>
<thead>
<tr>
<th>Variable</th>
<th>Expected score range</th>
<th>$M$</th>
<th>$SD$</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall sample ($N = 1636$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEFC</td>
<td>6 – 24</td>
<td>20.23</td>
<td>2.73</td>
<td>-.67</td>
<td>.15</td>
</tr>
<tr>
<td>SPLCAE</td>
<td>5 – 20</td>
<td>17.18</td>
<td>2.53</td>
<td>-1.15</td>
<td>1.04</td>
</tr>
<tr>
<td>SPHCAE</td>
<td>3 – 12</td>
<td>6.86</td>
<td>2.10</td>
<td>.18</td>
<td>-.52</td>
</tr>
<tr>
<td>SPPCAE</td>
<td>5 - 20</td>
<td>12.07</td>
<td>3.10</td>
<td>.03</td>
<td>-.52</td>
</tr>
<tr>
<td>SMASG</td>
<td>1 – 4</td>
<td>3.66</td>
<td>.38</td>
<td>-1.34</td>
<td>1.39</td>
</tr>
<tr>
<td>SPAPG</td>
<td>1 – 4</td>
<td>3.21</td>
<td>.69</td>
<td>-1.04</td>
<td>.65</td>
</tr>
<tr>
<td>SPAVG</td>
<td>1 – 4</td>
<td>3.12</td>
<td>.69</td>
<td>-.64</td>
<td>-.11</td>
</tr>
<tr>
<td>Male students ($n = 735$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEFC</td>
<td>6 – 24</td>
<td>19.80</td>
<td>2.79</td>
<td>-.57</td>
<td>-.03</td>
</tr>
<tr>
<td>SPLCAE</td>
<td>5 – 20</td>
<td>16.94</td>
<td>2.60</td>
<td>-1.08</td>
<td>.78</td>
</tr>
<tr>
<td>SPHCAE</td>
<td>3 – 12</td>
<td>6.96</td>
<td>2.14</td>
<td>.21</td>
<td>-.42</td>
</tr>
<tr>
<td>SPPCAE</td>
<td>5 - 20</td>
<td>12.48</td>
<td>3.01</td>
<td>-.11</td>
<td>-.33</td>
</tr>
<tr>
<td>SMASG</td>
<td>1 – 4</td>
<td>3.58</td>
<td>.40</td>
<td>-1.06</td>
<td>.53</td>
</tr>
<tr>
<td>SPAPG</td>
<td>1 – 4</td>
<td>3.15</td>
<td>.68</td>
<td>-.95</td>
<td>.59</td>
</tr>
<tr>
<td>SPAVG</td>
<td>1 – 4</td>
<td>3.03</td>
<td>.69</td>
<td>-.46</td>
<td>-.38</td>
</tr>
<tr>
<td>Female students ($n = 901$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEFC</td>
<td>6 – 24</td>
<td>20.58</td>
<td>2.64</td>
<td>-.76</td>
<td>.37</td>
</tr>
<tr>
<td>SPLCAE</td>
<td>5 – 20</td>
<td>17.38</td>
<td>2.46</td>
<td>-1.21</td>
<td>1.31</td>
</tr>
<tr>
<td>SPHCAE</td>
<td>3 – 12</td>
<td>6.78</td>
<td>2.07</td>
<td>.15</td>
<td>-.26</td>
</tr>
<tr>
<td>SPPCAE</td>
<td>5 - 20</td>
<td>11.73</td>
<td>3.13</td>
<td>.15</td>
<td>-.57</td>
</tr>
<tr>
<td>SMASG</td>
<td>1 – 4</td>
<td>3.73</td>
<td>.34</td>
<td>-1.64</td>
<td>2.67</td>
</tr>
<tr>
<td>SPAPG</td>
<td>1 – 4</td>
<td>3.26</td>
<td>.70</td>
<td>-1.14</td>
<td>.79</td>
</tr>
<tr>
<td>SPAVG</td>
<td>1 – 4</td>
<td>3.19</td>
<td>.69</td>
<td>-.82</td>
<td>.26</td>
</tr>
</tbody>
</table>

Note. SEFC = academic self-efficacy. SPLCAE = perceived learning classroom assessment environment. SPHCAE = perceived harsh classroom assessment environment. SPPCAE = perceived public classroom assessment environment. SMASG = mastery goal. SPAPG = performance-approach goal. SPAVG = performance-avoidance goal.
Bivariate relationships. Table 9 presents zero-order correlations among student-level variables. As shown in Table 9, although perceived harsh classroom assessment environment was negatively related to mastery goal orientation, it was unrelated to any type of performance goal orientation. The rest of the correlations were statistically significant at levels less than .05 and .01, and ranged from a low of .05 between perceived public classroom assessment environment and mastery and performance-avoidance goal orientations to a high of .46 between academic self-efficacy and mastery goal orientation. For the most part, the correlations among the variables tended to be in the expected directions. When considering the set of student-level independent variables for each dependent variable, the magnitude of the correlations posed no extreme threat to multicollinearity.

Table 9

Zero-Order Correlations among Student-Level Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. SEFC</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. SPLCAE</td>
<td>.31**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. SPHCAE</td>
<td>-.25**</td>
<td>-.25**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. SPPCAE</td>
<td>-.04</td>
<td>-.12**</td>
<td>.21**</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. SMASG</td>
<td>.46**</td>
<td>.33**</td>
<td>-.18**</td>
<td>-.05*</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>6. SPAPG</td>
<td>.29**</td>
<td>.10**</td>
<td>-.04</td>
<td>.14**</td>
<td>.32**</td>
<td>-</td>
</tr>
<tr>
<td>7. SPAVG</td>
<td>.23**</td>
<td>.07**</td>
<td>-.02</td>
<td>.05*</td>
<td>.33**</td>
<td>.45**</td>
</tr>
</tbody>
</table>

Note. SEFC = academic self-efficacy. SPLCAE = perceived learning classroom assessment environment. SPHCAE = perceived harsh classroom assessment environment. SPPCAE = perceived public classroom assessment environment. SMASG = mastery goal. SPAPG = performance-approach goal. SPAVG = performance-avoidance goal. *p < .05. **p < .01.
Data Screening at the Class-Level

Accuracy of data entry and missing values. Inspection of the frequency distribution of each class-level variable showed that all data seemed plausible with no missing values.

Outliers. As suggested by Tabachnick and Fidell (2001, p. 67), “Cases with standardized scores in excess of 3.29 (p < .001, two-tailed test) are potential outliers.” Using this criterion, two cases were identified as outliers, one on each of the following variables: teacher’s teaching experience and class average mastery goal orientation. Further inspection of the data revealed no outlying response pattern of these cases on other variables suggesting that they were part of the target population, and thereby as suggested by Tabachnick and Fidell (2001), they should remain in the analysis. Using Molloy and Newfields’s (2005) recommendation for reducing the influence of outliers, these outlying cases were assigned a raw score on the aforementioned offending variables that was three standard deviations above the mean for teaching experience and three standard deviations below mean for class average mastery goal orientation as shown in Table 10.

Table 10

Assigned Raw Scores to the Outlying Cases on the Offending Class-Level Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>Assigned raw score = $M \pm (3 \times SD)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Teacher’s teaching experience</td>
<td>5.23</td>
<td>2.75</td>
<td>$5.23 + (3 \times 2.75) = 13.48$</td>
</tr>
<tr>
<td>3. Class average mastery goal orientation</td>
<td>3.66</td>
<td>.14</td>
<td>$3.66 - (3 \times .14) = 3.24$</td>
</tr>
</tbody>
</table>
Descriptive statistics. After dealing with outliers, Table 11 presents descriptive statistics of class-level variables for the overall sample of classes and for males and females. Overall, the participating teachers had an average of 5.20 years of teaching experience. With the exception of class average perceived harsh and perceived public classroom assessment environments, the means for class-level variables were higher than their respective scales’ midpoints. Further, using George and Mallery’s (2001) guidelines, the skewness and kurtosis values of each class-level variable for the overall data set were within acceptable range of ±2, suggesting no concern about deviation from normality.

When classified by gender, female classes tended on average to have higher scores than male classes on frequent uses of alternative and recommended assessments, academic self-efficacy, perceived learning classroom assessment environment, and all types of achievement goal orientations. Although the distribution of male teachers’ years of teaching experience tended to be flatter than the normal distribution (kurtosis = 3.69), the skewness and kurtosis values of each class-level variable for males and females were within acceptable range of ±2 as suggested by George and Mallery’s (2001) guidelines. Therefore, there was no concern about deviation from normality for class-level variables within each group.
Table 11

Descriptive Statistics of Class-Level Variables for Overall Sample of and for Male and Female Classes

<table>
<thead>
<tr>
<th>Variable</th>
<th>Expected score range</th>
<th>Overall sample (N = 83)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
<td>Skewness</td>
<td>Kurtosis</td>
</tr>
<tr>
<td>TEXP</td>
<td>-</td>
<td>5.20</td>
<td>2.64</td>
<td>.74</td>
<td>.81</td>
</tr>
<tr>
<td>TRAD</td>
<td>7 – 35</td>
<td>25.77</td>
<td>2.66</td>
<td>-.27</td>
<td>.30</td>
</tr>
<tr>
<td>ALTR</td>
<td>6 – 30</td>
<td>20.20</td>
<td>3.39</td>
<td>.13</td>
<td>-.48</td>
</tr>
<tr>
<td>RECOM</td>
<td>30 – 150</td>
<td>97.70</td>
<td>8.76</td>
<td>-.31</td>
<td>-.09</td>
</tr>
<tr>
<td>CEFC</td>
<td>6 – 24</td>
<td>20.26</td>
<td>1.03</td>
<td>-.50</td>
<td>.26</td>
</tr>
<tr>
<td>CPLCAE</td>
<td>5 - 20</td>
<td>17.19</td>
<td>1.00</td>
<td>-.66</td>
<td>-.13</td>
</tr>
<tr>
<td>CPHCAE</td>
<td>3 - 12</td>
<td>6.86</td>
<td>.94</td>
<td>.53</td>
<td>.65</td>
</tr>
<tr>
<td>CPPCAE</td>
<td>5 - 20</td>
<td>12.07</td>
<td>1.44</td>
<td>-.29</td>
<td>-.50</td>
</tr>
<tr>
<td>CMASG</td>
<td>1 – 4</td>
<td>3.66</td>
<td>.13</td>
<td>-.92</td>
<td>1.52</td>
</tr>
<tr>
<td>CPAPG</td>
<td>1 – 4</td>
<td>3.21</td>
<td>.22</td>
<td>-.24</td>
<td>-.47</td>
</tr>
<tr>
<td>CPAVG</td>
<td>1 – 4</td>
<td>3.11</td>
<td>.22</td>
<td>-.53</td>
<td>.26</td>
</tr>
</tbody>
</table>

Male classes (n = 37)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Expected score range</th>
<th>Overall sample (N = 83)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
<td>Skewness</td>
<td>Kurtosis</td>
</tr>
<tr>
<td>TEXP</td>
<td>-</td>
<td>4.76</td>
<td>2.40</td>
<td>1.44</td>
<td>3.69</td>
</tr>
<tr>
<td>TRAD</td>
<td>7 – 35</td>
<td>25.68</td>
<td>2.57</td>
<td>.17</td>
<td>.25</td>
</tr>
<tr>
<td>ALTR</td>
<td>6 – 30</td>
<td>19.73</td>
<td>3.56</td>
<td>.53</td>
<td>-.43</td>
</tr>
<tr>
<td>RECOM</td>
<td>30 – 150</td>
<td>96.78</td>
<td>8.45</td>
<td>-.23</td>
<td>.33</td>
</tr>
<tr>
<td>CEFC</td>
<td>6 – 24</td>
<td>19.85</td>
<td>1.11</td>
<td>-.34</td>
<td>-.07</td>
</tr>
<tr>
<td>CPLCAE</td>
<td>5 - 20</td>
<td>16.94</td>
<td>.95</td>
<td>-.71</td>
<td>-.01</td>
</tr>
<tr>
<td>CPHCAE</td>
<td>3 - 12</td>
<td>6.95</td>
<td>.91</td>
<td>.86</td>
<td>1.09</td>
</tr>
<tr>
<td>CPPCAE</td>
<td>5 - 20</td>
<td>12.47</td>
<td>1.21</td>
<td>-.55</td>
<td>.17</td>
</tr>
<tr>
<td>CMASG</td>
<td>1 – 4</td>
<td>3.58</td>
<td>.13</td>
<td>-.88</td>
<td>1.47</td>
</tr>
<tr>
<td>CPAPG</td>
<td>1 – 4</td>
<td>3.15</td>
<td>.21</td>
<td>.01</td>
<td>-.82</td>
</tr>
<tr>
<td>CPAVG</td>
<td>1 – 4</td>
<td>3.02</td>
<td>.21</td>
<td>-.57</td>
<td>.11</td>
</tr>
</tbody>
</table>

Female classes (n = 46)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Expected score range</th>
<th>Overall sample (N = 83)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
<td>Skewness</td>
<td>Kurtosis</td>
</tr>
<tr>
<td>TEXP</td>
<td>-</td>
<td>5.55</td>
<td>2.80</td>
<td>.33</td>
<td>.02</td>
</tr>
<tr>
<td>TRAD</td>
<td>7 – 35</td>
<td>25.85</td>
<td>2.75</td>
<td>-.57</td>
<td>.54</td>
</tr>
<tr>
<td>ALTR</td>
<td>6 – 30</td>
<td>20.59</td>
<td>3.24</td>
<td>-.21</td>
<td>-.05</td>
</tr>
<tr>
<td>RECOM</td>
<td>30 – 150</td>
<td>98.43</td>
<td>9.02</td>
<td>-.40</td>
<td>-.17</td>
</tr>
<tr>
<td>CEFC</td>
<td>6 – 24</td>
<td>20.59</td>
<td>.85</td>
<td>-.20</td>
<td>-.23</td>
</tr>
<tr>
<td>CPLCAE</td>
<td>5 - 20</td>
<td>17.39</td>
<td>1.00</td>
<td>-.79</td>
<td>.04</td>
</tr>
</tbody>
</table>
Table 11 (Continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Expected score range</th>
<th>$M$</th>
<th>$SD$</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPHCAE</td>
<td>3 - 12</td>
<td>6.78</td>
<td>.97</td>
<td>.36</td>
<td>.44</td>
</tr>
<tr>
<td>CPPCAE</td>
<td>5 - 20</td>
<td>11.74</td>
<td>1.54</td>
<td>.02</td>
<td>-.62</td>
</tr>
<tr>
<td>CMASG</td>
<td>1 – 4</td>
<td>3.73</td>
<td>.09</td>
<td>-.38</td>
<td>-.06</td>
</tr>
<tr>
<td>CPAPG</td>
<td>1 – 4</td>
<td>3.27</td>
<td>.23</td>
<td>-.55</td>
<td>.19</td>
</tr>
<tr>
<td>CPAVG</td>
<td>1 – 4</td>
<td>3.18</td>
<td>.21</td>
<td>-.65</td>
<td>.84</td>
</tr>
</tbody>
</table>

*Note.* TEXP = teacher’s years of teaching experience. TRAD = teacher’s frequent use of traditional assessments. ALTR = teacher’s frequent use of alternative assessments. RECOM = teacher’s frequent use of recommended assessments. CEFC = class average academic self-efficacy. CPLCAE = class average perceived learning classroom assessment environment. CPHCAE = class average perceived harsh classroom assessment environment. CPPCAE = class average perceived public classroom assessment environment. CMASG = class average mastery goals. CPAPG = class average performance-approach goals. CPAVG = class average performance-avoidance goals.

*Bivariate relationships.* Table 12 presents zero-order correlations among class-level variables. As shown in Table 12, neither teaching experience nor teacher’s assessment practices were related to academic self-efficacy, perceived classroom assessment environment, and achievement goal orientations at the class-level. Although the direction of the correlations among academic self-efficacy, perceived classroom assessment environment, and achievement goal orientations was similar to that at the student-level presented in Table 9, the magnitude of these correlations seemed to be higher at the class-level than at the student-level. When considering the set of class-level independent variables for each dependent variable, the magnitude of the correlations posed no extreme threat to multicollinearity.
Table 12

Zero-Order Correlations among Class-Level Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEXP</td>
<td>-</td>
<td>-.19</td>
<td>-</td>
<td>-.19</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TRAD</td>
<td>-</td>
<td>-.19</td>
<td>-</td>
<td>-.19</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ALTR</td>
<td>.11</td>
<td>.19</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>RECOM</td>
<td>.19</td>
<td>.06</td>
<td>.50***</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CEFC</td>
<td>-.16</td>
<td>.07</td>
<td>.01</td>
<td>.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CPLCAE</td>
<td>.08</td>
<td>.04</td>
<td>-.03</td>
<td>-.04</td>
<td>.47***</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CPHCAE</td>
<td>-.04</td>
<td>.11</td>
<td>.08</td>
<td>-.03</td>
<td>-.40***</td>
<td>-.46***</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CPPCAE</td>
<td>-.18</td>
<td>.12</td>
<td>.06</td>
<td>-.02</td>
<td>-.21</td>
<td>-.32**</td>
<td>.38***</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CMASG</td>
<td>.00</td>
<td>.02</td>
<td>.06</td>
<td>-.04</td>
<td>.63**</td>
<td>.45***</td>
<td>-.16</td>
<td>-.16</td>
<td>-</td>
</tr>
<tr>
<td>CPAPG</td>
<td>.00</td>
<td>-.08</td>
<td>-.16</td>
<td>-.06</td>
<td>.40**</td>
<td>.13</td>
<td>-.06</td>
<td>.10</td>
<td>.41***</td>
</tr>
<tr>
<td>CPAVG</td>
<td>.04</td>
<td>-.09</td>
<td>-.02</td>
<td>-.04</td>
<td>.38**</td>
<td>.09</td>
<td>-.01</td>
<td>-.03</td>
<td>.64***</td>
</tr>
</tbody>
</table>

Note. TEXP = teacher’s years of teaching experience. TRAD = teacher’s frequent use of traditional assessments. ALTR = teacher’s frequent use of alternative assessments. RECOM = teacher’s frequent use of recommended assessments. CEFC = class average academic self-efficacy. CPLCAE = class average perceived learning classroom assessment environment. CPHCAE = class average perceived harsh classroom assessment environment. CPPCAE = class average perceived public classroom assessment environment. CMASG = class average mastery goals. CPAPG = class average performance-approach goals. CPAVG = class average performance-avoidance goals. **p < .01. ***p < .001.
Modeling Perceived Learning Classroom Assessment Environment

A fully unconditional model. The HLM analyses to model perceived learning classroom assessment environment (SPLCAE) began with a fully unconditional model, which is equivalent to a one-way random effects ANOVA, to examine how much variation in perceived learning classroom assessment environment lay within and between classrooms. The outcome variable (SPLCAE) was standardized to a mean of zero and a standard deviation of one. The initial model that was tested in this step of the analyses was as follows.

**Student-level:**

\[(SPLCAE)_{ij} = \beta_0 + r_{ij} .\]

**Class-level:**

\[\beta_{0j} = \gamma_{00} + u_{0j} ,\]

where all terms were defined in the Analysis section of Chapter III of the dissertation.

Based on this model, a statistically significant variation was found among class means on perceived learning classroom assessment environment; \( \hat{\tau}_{00} = .1107 ,\)

\( \chi^2(82) = 283.745 , p < .001 .\) The estimated within-class variance (\( \hat{\sigma}^2 \)) was .8901. Hence, the intraclass correlation (\( \hat{\rho} \)) was estimated as .1106, indicating that approximately 11% of the variance in perceived learning classroom assessment environment was between classrooms. The average reliability (\( \hat{\lambda} \)) of the class means was .710, suggesting that the sample means tended to be quite reliable as indicators of their true class means.
A random-coefficient regression model. The next step in the analyses involved posing a random-coefficient regression model to examine the relationship between student academic self-efficacy (SEFC) and perceived learning classroom assessment environment (SPLCAE), and whether this relationship varied significantly across classes. All variables were standardized to a mean of zero and a standard deviation of one. The student-level independent variable (student academic self-efficacy) was group-mean centered. The initial model that was tested in this step of the analyses was as follows.

**Student-level:**

\[
(SPLCAE)_{ij} = \beta_{0j} + \beta_{ij}(SEFC_{ij} - CEFC_{j}) + r_{ij}.
\]

**Class-level:**

\[
\beta_{0j} = \gamma_{00} + u_{0j},
\]

\[
\beta_{ij} = \gamma_{10} + u_{ij},
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation.

Table 13 presents results of the random-coefficient regression model of perceived learning classroom assessment environment. As shown in Table 13, on average, student academic self-efficacy was positively related to perceived learning classroom assessment environment within classrooms; \( \hat{\gamma}_{10} = .263 \), \( t(82) = 9.512, p < .001 \); suggesting that a one standard deviation increase in student academic self-efficacy was on average associated with a .263 standard deviation increase in perceived learning classroom assessment
Table 13

*Random-Coefficient Regression Model of Perceived Learning Classroom Assessment Environment*

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PLCAE mean, $\gamma_{00}$</td>
<td>.001</td>
<td>.043</td>
<td>.982</td>
</tr>
<tr>
<td>SEFC slope mean, $\gamma_{10}$</td>
<td>.263</td>
<td>.028</td>
<td>9.512***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLCAE mean, $u_{0j}$</td>
<td>.1149</td>
<td>82</td>
<td>312.707***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0155</td>
<td>82</td>
<td>109.703*</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.8077</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Correlation among class effects

<table>
<thead>
<tr>
<th></th>
<th>$\beta_{0j}$</th>
<th>$\beta_{1j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLCAE mean, $\beta_{0j}$</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td>-.477</td>
<td>-</td>
</tr>
</tbody>
</table>

Reliability of OLS regression-coefficient estimates

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PLCAE mean, $\beta_{0j}$</td>
<td>.737</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td>.239</td>
</tr>
</tbody>
</table>

*Note.* PLCAE = perceived learning classroom assessment environment. SEFC = student academic self-efficacy.

* $p < .05$. *** $p < .001$. 
environment. This relationship between academic self-efficacy and perceived learning classroom assessment environment varied significantly across classrooms; \( \hat{\tau}_{11} = .0155 \), \( \chi^2(82) = 109.703, p < .05 \). After taking student academic 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 academic self-efficacy accounted for about 9% of the within-class variance in perceived learning classroom assessment environment.

Also, as shown in Table 13, classes with a high average perceived learning classroom assessment environment tended to have a weaker positive effect of student academic self-efficacy on perceived learning classroom assessment environment than did classes with a low average perceived learning classroom assessment environment, \( \hat{\beta} = -.477 \). The estimated effects of academic self-efficacy on perceived learning classroom assessment environment for each class (i.e., slopes \( \hat{\beta}_{ij} \)) were on average less reliable than the estimates of class means for perceived learning classroom assessment environment (i.e., intercepts \( \hat{\beta}_{0j} \)), \( \hat{\lambda}(\hat{\beta}_{0j}) = .737 \) and \( \hat{\lambda}(\hat{\beta}_{ij}) = .239 \). This result suggested that there was more observed variation to be explained in the intercepts than in the slopes using class characteristics.

Having estimated the variability in the class means for perceived learning classroom assessment environment (i.e., intercepts) and student academic self-efficacy effects (i.e., academic self-efficacy slopes), the analyses proceeded with intercepts-and-slopes-as-outcomes regression models to explain the variability in these 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. Following are results pertaining to this process.

Submodel (1): Contextual-effect model of academic self-efficacy along with its
differential contextual effect by class’s gender. Following suggestions of Lee and Bryk
(1989) and Raudenbush and Bryk (2002), the contextual effect of academic self-efficacy
was represented by including class average for academic self-efficacy (CEFC) in the
model that investigated the distributive effect of student academic self-efficacy. The
differential contextual effect by class gender was represented by the inclusion of (GNDR
× CEFC) interaction term in the between-class model. Specifically, the initial model that
was tested in this step of the analyses was as follows.

Student-level:

\[ (SPLCAE)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_j) + r_{ij}. \]

Class-level:

\[ \beta_{0j} = \gamma_{00} + \gamma_{01}(GNDR)_j + \gamma_{02}(CEFC)_j + \gamma_{03}(GNDR \times CEFC)_j + u_{0j}, \]

\[ \beta_{1j} = \gamma_{10} + \gamma_{11}(GNDR)_j + \gamma_{12}(CEFC)_j + \gamma_{13}(GNDR \times CEFC)_j + u_{1j}, \]
where all terms were defined in the Analysis section of Chapter III of the dissertation. 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 was group-mean centered. Following Raudenbush and Bryk’s (2002) suggestion, some class-level variables were deleted because they had t-ratios near or less than one, and the model was re-estimated. Table 14 presents results of the reduced contextual effect model of academic self-efficacy on perceived learning classroom assessment environment.

With regard to class mean perceived learning classroom assessment environment, as shown in Table 14, the average academic self-efficacy of students was positively related to class mean perceived learning classroom assessment environment; $\hat{\gamma}_{01} = .184$, $t(81) = 4.949$, $p < .001$; suggesting that a one standard deviation increase in average academic self-efficacy of students was associated with a .184 standard deviation increase in class mean for perceived learning classroom assessment environment. Using the random-coefficient regression model presented in Table 13 as the base model, approximately 29% of the variance among classrooms in average perceived learning classroom assessment environment was explained once class average academic self-efficacy was taken into account. With regard to academic self-efficacy differentiation, however, there was no evidence of gender, context, or gender-by-context effects.
Table 14

*Reduced Contextual Effect Model of Academic Self-Efficacy on Perceived Learning Classroom Assessment Environment*

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PLCAE mean, $\beta_{0j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{00}$</td>
<td>.001</td>
<td>.038</td>
<td>.034</td>
</tr>
<tr>
<td>CEFC, $\gamma_{01}$</td>
<td>.184</td>
<td>.037</td>
<td>4.949***</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{10}$</td>
<td>.266</td>
<td>.027</td>
<td>9.568***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLCAE mean, $u_{0j}$</td>
<td>.0817</td>
<td>81</td>
<td>242.993***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0156</td>
<td>82</td>
<td>109.592*</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.8077</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* PLCAE = perceived learning classroom assessment environment. CEFC = class average academic self-efficacy. SEFC = student academic self-efficacy. *$p < .05$. ***$p < .001$.

*Submodel (2): Joint effects of class gender and class average for perceived classroom assessment environment.* This submodel examined the main effects of class gender (GNDR), class average perceived harsh classroom assessment environment (CPHCAE), and class average perceived public classroom assessment environment (CPPCAE) as well as interaction effects of (GNDR × CPHCAE) and (GNDR × CPPCAE) on (a) class mean for perceived learning classroom assessment environment and (b) relationship of student academic self-efficacy to perceived learning classroom assessment environment. Specifically, the initial model that was tested in this step of the analyses was as follows.
\[ (SPLCAE)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{ij}) + r_{ij}. \]

**Class-level:**

\[
\begin{align*}
\beta_{0j} &= \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(CPHCAE)_{j} + \gamma_{03}(CPPCAE)_{j} + \\
&\quad \gamma_{04}(GNDR \times CPHCAE)_{j} + \gamma_{05}(GNDR \times CPPCAE)_{j} + u_{0j}, \\
\beta_{1j} &= \gamma_{10} + \gamma_{11}(GNDR)_{j} + \gamma_{12}(CPHCAE)_{j} + \gamma_{13}(CPPCAE)_{j} + \\
&\quad \gamma_{14}(GNDR \times CPHCAE)_{j} + \gamma_{15}(GNDR \times CPPCAE)_{j} + u_{1j},
\end{align*}
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation. 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 was group-mean centered. Following Raudenbush and Bryk’ (2002) suggestion, some class-level variables were deleted because they had t-ratios near or less than one, and the model was re-estimated. Table 15 presents results of the reduced joint effects model of class gender and class averages perceived harsh and perceived public classroom assessment environments on perceived learning classroom assessment environment.

With regard to class mean perceived learning classroom assessment environment, as shown in Table 15, holding other factors constant, classes with a high average perceived harsh classroom assessment environment tended to have a smaller average
Table 15

Reduced Joint Effects Model of Class Gender and Class Averages for Perceived Harsh and Public Classroom Assessment Environments on Perceived Learning Classroom Assessment Environment

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PLCAE mean, $\beta_{0j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{00}$</td>
<td>.062</td>
<td>.044</td>
<td>1.402</td>
</tr>
<tr>
<td>GNDR, $\gamma_{01}$</td>
<td>-.558</td>
<td>.282</td>
<td>-1.982</td>
</tr>
<tr>
<td>CPHCAE, $\gamma_{02}$</td>
<td>-.148</td>
<td>.038</td>
<td>-3.853***</td>
</tr>
<tr>
<td>CPPCAE, $\gamma_{03}$</td>
<td>-.073</td>
<td>.039</td>
<td>-1.876</td>
</tr>
<tr>
<td>GNDR×CPPCAE, $\gamma_{04}$</td>
<td>.620</td>
<td>.279</td>
<td>2.220*</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{10}$</td>
<td>.266</td>
<td>.028</td>
<td>9.576***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLCAE mean, $u_{0j}$</td>
<td>.0741</td>
<td>78</td>
<td>218.983***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0155</td>
<td>82</td>
<td>109.595*</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.8078</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. PLCAE = perceived learning classroom assessment environment. GNDR = class gender (1 = female and -1 = male). CPHCAE = class average perceived harsh classroom assessment environment. CPPCAE = class average perceived public classroom assessment environment. SEFC = student academic self-efficacy.

* $p < .05$. *** $p < .001$. 
perceived learning classroom assessment environment than were classes with a low average perceived harsh classroom assessment environment; $\gamma_{02} = -.148$, $t(78) = -3.853$, $p < .001$. Specifically, holding other factors constant, a one standard deviation increase in class mean for perceived harsh classroom assessment environment was associated with a .148 standard deviation decrease in class mean for perceived learning classroom assessment environment. According to Table 15, holding other factors constant, the effect of average perceived public classroom assessment environment on average perceived learning classroom assessment environment varied significantly across male and female classrooms. In male classrooms, classes with a high average perceived public classroom assessment environment tended to have a lower average perceived learning classroom assessment environment than were classes with a low average perceived public classroom assessment environment; $\gamma_{03} + (GNDR)_{j}\gamma_{04} = -.073 + (-1)(.620) = -.693$; holding other factors constant. In female classrooms, the opposite was true, in that classes with a high average perceived public classroom assessment environment tended to have a higher average perceived learning classroom assessment environment than were classes with a low average perceived public classroom assessment environment; $\gamma_{03} + (GNDR)_{j}\gamma_{04} = -.073 + (1)(.620) = .547$; holding other factors constant.

Using the random-coefficient regression model presented in Table 13 as the base model, approximately 36% of the variance among classrooms in average perceived learning classroom assessment environment was explained by class gender, class average perceived harsh classroom assessment environment, class average perceived public
classroom assessment environment, and interaction of class gender-by-class average perceived public classroom assessment environment. With regard to academic self-efficacy differentiation, there was no evidence of class gender, class average perceived harsh classroom assessment environment, class average perceived public classroom assessment environment, gender-by-class average perceived harsh classroom assessment environment, or gender-by-class average perceived public classroom assessment environment interaction effects on the relationship between student academic self-efficacy and perceived learning classroom assessment environment.

*Submodel (3): Joint effect of class’s gender and teacher’s assessment practices and teaching experience.* The focus in this submodel was on the main effects of teacher’s teaching experience (TEXP); and teacher’s frequent uses of traditional assessments (TRAD), alternative assessments (ALTR), and recommended assessment practices (RECOM) along with three-way interaction effect of GNDR-by-TEXP-by-ALTR on (a) class mean perceived learning classroom assessment environment and (b) relationship between student academic self-efficacy and perceived learning classroom assessment environment. Specifically, the initial model that was tested in this step of the analyses was as follows.

**Student-level:**

\[
(SPLCAE)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_j) + r_{ij}.
\]
Class-level:

\[
\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{GNDR})_j + \gamma_{02}(\text{TEXP})_j + \gamma_{03}(\text{TRAD})_j + \gamma_{04}(\text{ALTR})_j + \\
\gamma_{05}(\text{RECOM})_j + \gamma_{06}(\text{GNDR} \times \text{TEXP})_j + \gamma_{07}(\text{GNDR} \times \text{ALTR})_j + \\
\gamma_{08}(\text{TEXP} \times \text{ALTR})_j + \gamma_{09}(\text{GNDR} \times \text{TEXP} \times \text{ALTR})_j + u_{0j}
\]

\[
\beta_{1j} = \gamma_{10} + \gamma_{11}(\text{GNDR})_j + \gamma_{12}(\text{TEXP})_j + \gamma_{13}(\text{TRAD})_j + \gamma_{14}(\text{ALTR})_j + \\
\gamma_{15}(\text{RECOM})_j + \gamma_{16}(\text{GNDR} \times \text{TEXP})_j + \gamma_{17}(\text{GNDR} \times \text{ALTR})_j + \\
\gamma_{18}(\text{TEXP} \times \text{ALTR})_j + \gamma_{19}(\text{GNDR} \times \text{TEXP} \times \text{ALTR})_j + u_{1j}
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation. 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 was group-mean centered. Following Raudenbush
and Bryk’s (2002) suggestion, some class-level variables were deleted because they had t-
ratios near or less than one, and the model was re-estimated. Table 16 presents results of
the reduced joint effects model of class gender and teacher’s assessment practices on
perceived learning classroom assessment environment.

With regard to class mean perceived learning classroom assessment environment, as
shown in Table 16, a statistically significant interaction effect was detected between class
gender and teacher’s frequent use of alternative assessments; \( \hat{\gamma}_{03} = .682, t(79) = 2.955, p < .01 \). The presence of such an interaction effect means that the magnitude of gender effect
on average perceived learning classroom assessment environment depended on teacher’s
frequent use of alternative assessments in the classes compared. In general,
Table 16

*Reduced Joint Effects Model of Class Gender and Teacher’s Assessment Practices on Perceived Learning Classroom Assessment Environment*

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PLCAE mean, $\beta_{0j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{00}$</td>
<td>.064</td>
<td>.048</td>
<td>1.352</td>
</tr>
<tr>
<td>GNDR, $\gamma_{01}$</td>
<td>-.582</td>
<td>.229</td>
<td>-2.540*</td>
</tr>
<tr>
<td>ALTR, $\gamma_{02}$</td>
<td>-.005</td>
<td>.038</td>
<td>-.129</td>
</tr>
<tr>
<td>GNDR $\times$ ALTR, $\gamma_{03}$</td>
<td>.682</td>
<td>.231</td>
<td>2.955**</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{10}$</td>
<td>.263</td>
<td>.028</td>
<td>9.488***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLCAE mean, $u_{0j}$</td>
<td>.1000</td>
<td>79</td>
<td>273.274***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0156</td>
<td>82</td>
<td>109.746*</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.8075</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. PLCAE = perceived learning classroom assessment environment. GNDR = class gender (1 = female and -1 = male). ALTR = teacher’s frequent use of alternative assessments. SEFC = student academic self-efficacy. *p < .05. **p < .01. ***p < .001.*
the gender effect on average perceived learning classroom assessment environment can be computed as follows:

Female prediction – male prediction

\[
= [(1)\hat{b}_{01} + (1)(ALTR)\hat{b}_{03}] - [(-1)\hat{b}_{01} + (-1)(ALTR)\hat{b}_{03}]
\]

\[
= 2 \left[ \hat{b}_{01} + (ALTR)\hat{b}_{03} \right].
\]

Accordingly, in classes using alternative assessments more frequently (ALTR = 1), the average perceived learning classroom assessment environment was higher in female classrooms. In classes using alternative assessments less frequently (ALTR = -1), the average perceived learning classroom assessment environment was higher in male classrooms. Using the random-coefficient regression model presented in Table 13 as the base model, approximately 13% of the variance among classrooms in average perceived learning classroom assessment environment was explained once class gender, frequent use of alternative assessments, and interaction of class gender-by-frequent use of alternative assessments were taken into account. With regard to academic self-efficacy differentiation, however, there was no evidence of class gender, teaching experience, or teacher’s assessment practices effects on the relationship between student academic self-efficacy and perceived learning classroom assessment environment.

A final explanatory model of perceived learning classroom assessment environment. The final step in the analyses involved combining together statistically significant class-level variables detected in the early steps of the analyses to produce a parsimonious overall intercepts-and-slopes-as-outcomes regression model explaining the
variability in (a) class mean perceived learning classroom assessment environment and (b) relationship between student academic self-efficacy and perceived learning classroom assessment environment. Based on the early steps of the analyses, the initial model that was tested in this step of the analyses was as follows.

**Student-level:**

\[ (SPLCAE)_g = \beta_{0j} + \beta_{1j} (SEFC_g - CEFC_j) + r_g. \]

**Class-level:**

\[ \beta_{0j} = \gamma_{00} + \gamma_{01}(GNDR)_j + \gamma_{02}(ALTR)_j + \gamma_{03}(CEFC)_j + \gamma_{04}(CPHCAE)_j + \gamma_{05}(CPPCAE)_j + \gamma_{06}(GNDR \times ALTR)_j + \gamma_{07}(GNDR \times CPPCAE)_j + u_{0j}, \]

\[ \beta_{1j} = \gamma_{10} + u_{1j}, \]

where all terms were defined in the Analysis section of Chapter III of the dissertation. 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 was group-mean centered. Following Raudenbush and Bryk’s (2002) suggestion, some class-level variables were deleted because they had t-ratios near or less than one, and the composite model was re-estimated. Further, the empirical Bayes residuals from the reduced composite model were regressed on the excluded variables. On the basis of this residual analysis, no variable was needed to be added to the model.

However, the data pertaining to this model displayed a heterogeneous residual variance at the student-level; \( \chi^2(82) = 158.943, p < .001 \). Further inspection of the data
showed neither extreme observations than normally expected nor units having bad data. Further inspection of the model revealed that the effect of student-level independent variable was appropriately specified. As such, the reason for heterogeneity of student-level residual variances in this data set might be the omission of other important student-level independent variables from the model, and thus this might need to be considered in future research. Although estimation of the fixed effects and their standard errors is robust to violations of the homogeneity assumption about student-level residual variance (Kasim & Raudenbush, 1998), the final model was estimated both with and without the heterogeneous variance specification at the student-level. The student-level variance was modeled as a function of student academic self-efficacy as follows:

$$\ln(\sigma_{ij}^2) = \alpha_0 + \alpha_1(SEFC)_{ij}.$$  

Results indicated that the model with the heterogeneous variance specification at the student-level appeared to fit the data better than the model without the heterogeneous variance specification at the student-level; $$\chi^2(1) = 11.448, p < .01.$$ After considering the heterogeneity of student-level variance, the effect of student academic self-efficacy on perceived learning classroom assessment environment did not vary significantly across classrooms; $$\hat{\tau}_{11} = .0109, \chi^2(82) = 100.045, p = .086.$$ Therefore, the deviance statistic for the unrestricted model (4360.952 with 12 df) where the academic self-efficacy slope ($$\beta_{1j}$$) was specified as varying randomly across classrooms was compared with the deviance statistic for a restricted model (4363.955 with 10 df) where $$\beta_{1j}$$ was specified as fixed. The reduction in deviance, which was 3.003, was not statistically significant when
compared against the \( \chi^2 \) distribution with two degrees of freedom. Therefore, the restricted model with \( \beta_j \) specified as fixed appeared sufficient. Table 17 presents results of this final reduced composite model of perceived learning classroom assessment environment.

With regard to class mean perceived learning classroom assessment environment, as shown in Table 17, the average academic self-efficacy of students was positively related to class mean perceived learning classroom assessment environment; \( \hat{\gamma}_{03} = .129, t(77) = 3.325, p < .01 \); holding other factors constant. This suggested that holding other factors constant, a one standard deviation increase in average academic self-efficacy of students was associated with a .129 standard deviation increase in class mean for perceived learning classroom assessment environment. Also, holding other factors constant, class average perceived harsh classroom assessment environment was negatively related to class mean perceived learning classroom assessment environment; \( \hat{\gamma}_{04} = -.163, t(77) = -4.181, p < .001 \). This suggested that holding other factors constant, a one standard deviation increase in class mean perceived harsh classroom assessment environment was associated with a .143 standard deviation decrease in class mean for perceived learning classroom assessment environment.

Further, as shown in Table 17, holding other factors constant, the effect of class gender on class average perceived learning classroom assessment environment depended significantly on teacher’s frequent use of alternative assessments; \( \hat{\gamma}_{05} = .844, t(77) = \)
Table 17

*Final Fitted Composite Model of Perceived Learning Classroom Assessment Environment with Heterogeneous Level-1 Variance*

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PLCAE mean, $\beta_{0,j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{00}$</td>
<td>.089</td>
<td>.037</td>
<td>2.368*</td>
</tr>
<tr>
<td>GNDR, $\gamma_{01}$</td>
<td>-.802</td>
<td>.172</td>
<td>-4.662***</td>
</tr>
<tr>
<td>ALTR, $\gamma_{02}$</td>
<td>.007</td>
<td>.029</td>
<td>.243</td>
</tr>
<tr>
<td>CEFC, $\gamma_{03}$</td>
<td>.129</td>
<td>.039</td>
<td>3.325**</td>
</tr>
<tr>
<td>CPHCAE, $\gamma_{04}$</td>
<td>-.143</td>
<td>.034</td>
<td>-4.181***</td>
</tr>
<tr>
<td>GNDR $\times$ ALTR, $\gamma_{05}$</td>
<td>.844</td>
<td>.170</td>
<td>4.958***</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{10}$</td>
<td>.265</td>
<td>.028</td>
<td>9.423***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLCAE mean, $u_{0j}$</td>
<td>.0417</td>
<td>77</td>
<td>167.541***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model for level-1 variance Parameter</th>
<th>Coefficient</th>
<th>SE</th>
<th>z-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept, $\alpha_0$</td>
<td>.206</td>
<td>.036</td>
<td>5.742***</td>
</tr>
<tr>
<td>SEFC, $\alpha_1$</td>
<td>-.151</td>
<td>.039</td>
<td>-3.916***</td>
</tr>
</tbody>
</table>

*Note. PLCAE = perceived learning classroom assessment environment. GNDR = class gender (1 = female and -1 = male). ALTR = teacher’s frequent use of alternative assessments. CEFC = class average academic self-efficacy. CPHCAE = class average perceived harsh classroom assessment environment. SEFC = student academic self-efficacy.  
*p < .05. **p < .01. ***p < .001.*
4.958, \( p < .001 \). This gender effect on average perceived learning classroom assessment environment can be found as follows:

Female prediction – male prediction

\[
= [(1)\hat{\gamma}_{01} + (1)(ALTR) + \hat{\gamma}_{05}] - [(-1)\hat{\gamma}_{01} + (-1)(ALTR) + \hat{\gamma}_{05}]
\]

\[
= 2 \ [\hat{\gamma}_{01} + (ALTR) + \hat{\gamma}_{05}].
\]

Accordingly, in classes using alternative assessments more frequently (\( ALTR = 1 \)), the average perceived learning classroom assessment environment was higher in female classrooms. The opposite was true in classes using alternative assessments less frequently (\( ALTR = -1 \)), in that, the average perceived learning classroom assessment environment was higher in male classrooms.

Using the random-coefficient regression model presented in Table 13 as the base model, approximately 64\% of the variance among classrooms in average perceived learning classroom assessment environment was explained once class gender, class average academic self-efficacy, class average perceived harsh classroom 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 academic self-efficacy slope, as shown in Table 17, perceptions of high efficacious students about their classroom assessment environment as being learning-oriented were on average not only higher; \( \hat{\gamma}_{10} = .265, t(1629) = 9.423, p < .001 \); but also less variable; \( \hat{\alpha}_1 = -.151, z = -3.916, p < .001 \); than those for less efficacious students.
Modeling Perceived Harsh Classroom Assessment Environment

A fully unconditional model. The HLM analyses to model perceived harsh classroom assessment environment (SPHCAE) began with a fully unconditional model, which is equivalent to a one-way random effects ANOVA, to examine how much variation in perceived harsh classroom assessment environment lay within and between classrooms. The outcome variable (SPHCAE) was standardized to a mean of zero and a standard deviation of one. The initial model that was tested in this step of the analyses was as follows.

\[ (SPHCAE)_{ij} = \beta_{0j} + r_{ij}. \]

\[ \beta_{0j} = \gamma_{00} + u_{0j}, \]

where all terms were defined in the Analysis section of Chapter III of the dissertation.

Based on this model, a statistically significant variation was found among class means on perceived harsh classroom assessment environment; \( \hat{\tau}_{00} = .1600, \)

\( \chi^2(2) = 391.099, p < .001. \) The estimated within-class variance (\( \hat{\sigma}^2 \)) was .8409. Hence, the intraclass correlation (\( \hat{\rho} \)) was estimated as .1599, indicating that approximately 16% of the variance in perceived harsh classroom assessment environment was between classrooms. The average reliability (\( \hat{\lambda} \)) of the class means was .789, suggesting that the sample means tended to be quite reliable as indicators of their true class means.
A random-coefficient regression model. The next step in the analyses involved posing a random-coefficient regression model to examine the relationship between student academic self-efficacy (SEFC) and perceived harsh classroom assessment environment (SPHCAE), and whether this relationship varied significantly across classes. All variables were standardized to a mean of zero and a standard deviation of one. The student-level independent variable (student academic self-efficacy) was group-mean centered. The initial model that was tested in this step of the analyses was as follows.

Student-level:

\[
(SPHCAE)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + r_{ij},
\]

Class-level:

\[
\beta_{0j} = \gamma_{00} + u_{0j}, \\
\beta_{1j} = \gamma_{10} + u_{1j},
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation.

Table 18 presents results of the random-coefficient regression model of perceived harsh classroom assessment environment. As shown in Table 18, on average, student academic self-efficacy was negatively related to perceived harsh classroom assessment environment within classrooms; \( \hat{\gamma}_{10} = -0.216 \), \( t(82) = -8.130 \), \( p < .001 \); suggesting that a one standard deviation increase in student academic self-efficacy was on average associated with a .216 standard deviation decrease in perceived harsh classroom assessment environment. This relationship between academic self-efficacy and perceived harsh classroom assessment environment varied significantly across classrooms;
\[ \hat{\tau}_{11} = .0115, \quad \chi^2(82) = 107.944, \quad p < .05. \]

After taking student academic self-efficacy into account, the estimated within-class variance (\( \hat{\sigma}^2 \)) was reduced from .8409 in the random-effects ANOVA model to .7923. Hence, student academic self-efficacy accounted for about 6% of the within-class variance in perceived harsh classroom assessment environment.

Also, as shown in Table 18, classes with a high average perceived harsh classroom assessment environment tended to have a stronger negative effect of student academic self-efficacy on perceived harsh classroom assessment environment than did classes with a low average perceived harsh classroom assessment environment, \( \hat{p} = .343 \). The estimated effects of academic self-efficacy on perceived harsh classroom assessment environment for each class (i.e., slopes \( \hat{\beta}_{ij} \)) were on average less reliable than the estimates of class means for perceived harsh classroom assessment environment (i.e., intercepts \( \hat{\beta}_{0j} \)), \( \hat{\lambda}(\hat{\beta}_{0j}) = .801 \) and \( \hat{\lambda}(\hat{\beta}_{ij}) = .193 \). This result suggested that there was more observed variation to be explained in the intercepts than in the slopes using class characteristics.

Having estimated the variability in the class means for perceived harsh classroom assessment environment (i.e., intercepts) and student academic self-efficacy effects (i.e., academic self-efficacy slopes), the analyses proceeded with intercepts-and-slopes-as-
Table 18

*Random-Coefficient Regression Model of Perceived Harsh Classroom Assessment Environment*

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PHCAE mean, $\gamma_{00}$</td>
<td>-.001</td>
<td>.049</td>
<td>-.030</td>
</tr>
<tr>
<td>SEFC slope mean, $\gamma_{10}$</td>
<td>-.216</td>
<td>.027</td>
<td>-8.130***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHCAE mean, $u_{0j}$</td>
<td>.1625</td>
<td>82</td>
<td>415.089***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0115</td>
<td>82</td>
<td>107.944*</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.7923</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Correlation among class effects</th>
<th>$\beta_{0j}$</th>
<th>$\beta_{1j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHCAE mean, $\beta_{0j}$</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td>.343</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reliability of OLS regression-coefficient estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHCAE mean, $\beta_{0j}$</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
</tr>
</tbody>
</table>

*Note.* PHCAE = perceived harsh classroom assessment environment. SEFC = student academic self-efficacy.

*p < .05. ***p < .001.*
outcomes regression models to explain the variability in these 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. Following are results pertaining to this process.

Submodel (1): Contextual-effect model of academic self-efficacy along with its differential contextual effect by class’s gender. Following suggestions of Lee and Bryk (1989) and Raudenbush and Bryk (2002), the contextual effect of academic self-efficacy was represented by including class average for academic self-efficacy (CEFC) in the model that investigated the distributive effect of student academic self-efficacy. The differential contextual effect by class gender was represented by the inclusion of (GNDR × CEFC) interaction term in the between-class model. Specifically, the initial model that was tested in this step of the analyses was as follows.

Student-level:

\[(SPHCAE)_{ij} = \beta_{0ij} + \beta_{1ij}(SEFC_{ij} - CEFC_{j}) + r_{ij}.\]
Class-level:

\[ \beta_{0j} = \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(CEFC)_{j} + \gamma_{03}(GNDR \times CEFC)_{j} + u_{0j}, \]
\[ \beta_{1j} = \gamma_{10} + \gamma_{11}(GNDR)_{j} + \gamma_{12}(CEFC)_{j} + \gamma_{13}(GNDR \times CEFC)_{j} + u_{1j}, \]

where all terms were defined in the Analysis section of Chapter III of the dissertation. 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 was group-mean centered. Following Raudenbush and Bryk’s (2002) suggestion, some class-level variables were deleted because they had t-ratios near or less than one, and the model was re-estimated. Table 19 presents results of the reduced contextual effect model of academic self-efficacy on perceived harsh classroom assessment environment.

With regard to class mean perceived harsh classroom assessment environment, as shown in Table 19, the average academic self-efficacy of students was negatively related to class mean perceived harsh classroom assessment environment; \( \hat{\gamma}_{01} = -.174, t(81) = -3.283, p < .01; \) suggesting that a one standard deviation increase in average academic self-efficacy of students was associated with a .174 standard deviation decrease in class mean for perceived harsh classroom assessment environment. Using the random-coefficient regression model presented in Table 18 as the base model, approximately 19% of the variance among classrooms in average perceived harsh classroom assessment environment was explained once class average academic self-efficacy was taken into account. With
Table 19

Reduced Contextual Effect Model of Academic Self-Efficacy on Perceived Harsh Classroom Assessment Environment

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PHCAE mean, $\beta_{0j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{00}$</td>
<td>-.002</td>
<td>.045</td>
<td>-.036</td>
</tr>
<tr>
<td>CEFC, $\gamma_{01}$</td>
<td>-.174</td>
<td>.053</td>
<td>-3.283**</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{10}$</td>
<td>-.186</td>
<td>.027</td>
<td>6.867***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHCAE mean, $u_{0j}$</td>
<td>.1321</td>
<td>81</td>
<td>347.507***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0111</td>
<td>82</td>
<td>107.839*</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.7927</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. PHCAE = perceived harsh classroom assessment environment. CEFC = class average academic self-efficacy. SEFC = student academic self-efficacy. *$p < .05$. **$p < .01$. ***$p < .001$. 
regard to academic self-efficacy differentiation, however, there was no evidence of
gender, context, or gender-by-context effects.

Submodel (2): Joint effects of class gender and class average for perceived
classroom assessment environment. This submodel examined the main effects of class
gender (GNDR), class average perceived learning classroom assessment environment
(CPLCAE), and class average perceived public classroom assessment environment
(CPPCAE) as well as interaction effects of (GNDR × CPLCAE) and (GNDR × CPPCAE)
on (a) class mean for perceived harsh classroom assessment environment and (b)
relationship of student academic self-efficacy to perceived harsh classroom assessment
environment. Specifically, the initial model that was tested in this step of the analyses was
as follows.

\[
\begin{align*}
\text{Student-level:} \quad (SPHCAE)_{ij} & = \beta_{0j} + \beta_{1j} (SEFC_{ij} - CEFC_{j}) + r_{ij}.
\end{align*}
\]

\[
\begin{align*}
\text{Class-level:} \quad \beta_{0j} & = \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(CPLCAE)_{j} + \gamma_{03}(CPPCAE)_{j} + \\
& \quad \gamma_{04}(GNDR \times CPLCAE)_{j} + \gamma_{05}(GNDR \times CPPCAE)_{j} + u_{0j}, \\
\beta_{1j} & = \gamma_{10} + \gamma_{11}(GNDR)_{j} + \gamma_{12}(CPLCAE)_{j} + \gamma_{13}(CPPCAE)_{j} + \\
& \quad \gamma_{14}(GNDR \times CPLCAE)_{j} + \gamma_{15}(GNDR \times CPPCAE)_{j} + u_{1j},
\end{align*}
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation. 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 was group-mean centered. Following Raudenbush and Bryk’s (2002) suggestion, some class-level variables were deleted because they had t-ratios near or less than one, and the model was re-estimated. Table 20 presents results of the reduced joint effects model of class averages perceived learning and perceived public classroom assessment environments on perceived harsh classroom assessment environment.

With regard to class mean perceived harsh classroom assessment environment, as shown in Table 20, holding other factors constant, classes with a high average perceived learning classroom assessment environment tended to have a smaller average perceived harsh classroom assessment environment than were classes with a low average perceived learning classroom assessment environment; $\hat{\gamma}_{01} = -.167, t(80) = -3.336, p < .01$. This suggested that holding other factors constant, a one standard deviation increase in class mean for perceived learning classroom assessment environment was associated with a .167 standard deviation decrease in class mean for perceived harsh classroom assessment environment. Also, holding other factors constant, the average perceived public classroom assessment environment was positively related to class mean perceived harsh classroom assessment environment; $\hat{\gamma}_{02} = .115, t(80) = 2.440, p < .05$. This suggested that holding other factors constant, a one standard deviation increase in class mean for perceived public classroom assessment environment was associated with a .115 standard deviation increase in class mean for perceived harsh classroom assessment environment.
Table 20

Reduced Joint Effects Model of Class Averages for Perceived Learning and Public Classroom Assessment Environments on Perceived Harsh Classroom Assessment Environment

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PHCAE mean, $\beta_{0j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{00}$</td>
<td>-.002</td>
<td>.042</td>
<td>-.039</td>
</tr>
<tr>
<td>CPLCAE, $\gamma_{01}$</td>
<td>-.167</td>
<td>.050</td>
<td>-3.336**</td>
</tr>
<tr>
<td>CPPCAE, $\gamma_{02}$</td>
<td>.115</td>
<td>.047</td>
<td>2.440*</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{10}$</td>
<td>-.214</td>
<td>.026</td>
<td>-8.081***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHCAE mean, $u_{0j}$</td>
<td>.1104</td>
<td>80</td>
<td>299.640***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0117</td>
<td>82</td>
<td>107.900*</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.7923</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. PHCAE = perceived harsh classroom assessment environment. CPLCAE = class average perceived learning classroom assessment environment. CPPCAE = class average perceived public classroom assessment environment. SEFC = student academic self-efficacy.

*p < .05. **p < .01. ***p < .001.
Using the random-coefficient regression model presented in Table 18 as the base model, approximately 32% of the variance among classrooms in average perceived harsh classroom assessment environment was explained by class averages perceived learning and public classroom assessment environments. With regard to academic self-efficacy differentiation, there was no evidence of class gender, class average perceived learning classroom assessment environment, class average perceived public classroom assessment environment, gender-by-class average perceived learning classroom assessment environment, or gender-by-class average perceived public classroom assessment environment interaction effects on the relationship between student academic self-efficacy and perceived harsh classroom assessment environment.

Submodel (3): Joint effect of class’s gender and teacher’s assessment practices and teaching experience. The focus in this submodel was on the main effects of teacher’s teaching experience (TEXP); and teacher’s frequent uses of traditional assessments (TRAD), alternative assessments (ALTR), and recommended assessment practices (RECOM) along with three-way interaction effect of GNDR-by-TEXP-by-ALTR on (a) class mean perceived harsh classroom assessment environment and (b) relationship between student academic self-efficacy and perceived harsh classroom assessment environment. Specifically, the initial model that was tested in this step of the analyses was as follows.

Student-level:

\[
(SPHCAE)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + r_{ij}.
\]
Class-level:

\[
\beta_{0j} = \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(TEXP)_{j} + \gamma_{03}(TRAD)_{j} + \gamma_{04}(ALTR)_{j} + \\
\gamma_{05}(RECOM)_{j} + \gamma_{06}(GNDR \times TEXP)_{j} + \gamma_{07}(GNDR \times ALTR)_{j} + \\
\gamma_{08}(TEXP \times ALTR)_{j} + \gamma_{09}(GNDR \times TEXP \times ALTR)_{j} + u_{0j}
\]

\[
\beta_{1j} = \gamma_{10} + \gamma_{11}(GNDR)_{j} + \gamma_{12}(TEXP)_{j} + \gamma_{13}(TRAD)_{j} + \gamma_{14}(ALTR)_{j} + \\
\gamma_{15}(RECOM)_{j} + \gamma_{16}(GNDR \times TEXP)_{j} + \gamma_{17}(GNDR \times ALTR)_{j} + \\
\gamma_{18}(TEXP \times ALTR)_{j} + \gamma_{19}(GNDR \times TEXP \times ALTR)_{j} + u_{1j}
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation. 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 was group-mean centered. Following Raudenbush and Bryk’s (2002) suggestion, some class-level variables were deleted because they had t-ratios near or less than one, and the model was re-estimated. Table 21 presents results of the reduced joint effects model of class gender, teacher’s teaching experience, and assessment practices on perceived harsh classroom assessment environment.

With regard to class mean perceived harsh classroom assessment environment, as shown in Table 21, there was no evidence of class gender, teaching experience, or teacher’s assessment practices effects. However, as shown in Table 21, holding other factors constant, the differentiating effect of student academic self-efficacy on perceived harsh classroom assessment environment within a classroom depended jointly on class gender, teacher’s teaching experience, and teacher’s frequent use of alternative assessments; \( \hat{\gamma}_{15} = -.300, t(77) = -2.180, p < .05 \). This can be seen by computing the
Table 21

Reduced Joint Effects Model of Class Gender, Teaching Experience, and Teacher’s Assessment Practices on Perceived Harsh Classroom Assessment Environment

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PHCAE mean, $\beta_{0j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{00}$</td>
<td>-.002</td>
<td>.049</td>
<td>-.031</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{10}$</td>
<td>-.221</td>
<td>.026</td>
<td>-8.399***</td>
</tr>
<tr>
<td>GNDR, $\gamma_{11}$</td>
<td>.040</td>
<td>.053</td>
<td>.755</td>
</tr>
<tr>
<td>ALTR, $\gamma_{12}$</td>
<td>.021</td>
<td>.025</td>
<td>.847</td>
</tr>
<tr>
<td>TEXP, $\gamma_{13}$</td>
<td>.024</td>
<td>.023</td>
<td>1.018</td>
</tr>
<tr>
<td>GNDR $\times$ TEXP, $\gamma_{14}$</td>
<td>.280</td>
<td>.133</td>
<td>2.107*</td>
</tr>
<tr>
<td>GNDR $\times$ TEXP $\times$ ALTR, $\gamma_{15}$</td>
<td>-.300</td>
<td>.138</td>
<td>-2.180*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHCAE mean, $u_{0j}$</td>
<td>.1624</td>
<td>82</td>
<td>414.688***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0100</td>
<td>77</td>
<td>100.716*</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.7931</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. PHCAE = perceived harsh classroom assessment environment. SEFC = student academic self-efficacy. GNDR = class gender (1 = female and -1 = male). ALTR = teacher’s frequent use of alternative assessments. TEXP = teacher’s teaching experience. *$p < .05$. ***$p < .001$. 
differentiating effect of academic 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:

\[ \hat{y}_{10} + \hat{y}_{11}(GNDR)_j + \hat{y}_{12}(ALTR)_j + \hat{y}_{13}(TEXP)_j + \hat{y}_{14}(GNDR \times TEXP)_j + \hat{y}_{15}(GNDR \times TEXP \times ALTR)_j. \]

Accordingly, for male classrooms using alternative assessments less frequently, classes having a high experienced teacher were less differentiating with regard to student academic self-efficacy than were classes having a low experienced teacher. The opposite was true in male classrooms using alternative assessments more frequently, in that, classes having a low experienced teacher were less differentiating with regard to student academic self-efficacy than were classes having a high experienced teacher. For female classrooms using alternative assessments less frequently, classes having a high experienced teacher were more differentiating with regard to student academic self-efficacy than were classes having a low experienced teacher. The opposite was true in female classrooms using alternative assessments more frequently, in that, classes having a low experienced teacher were more differentiating with regard to student academic self-efficacy than were classes having a high experienced teacher. Using the random-coefficient regression model presented in Table 18 as the base model, approximately 13% of the variance among classrooms in student academic self-efficacy effect on perceived harsh classroom assessment environment was explained once class gender, teaching experience, frequent use of alternative assessments, interaction of class gender-by-teaching experience, and
interaction of class gender-by-teaching experience-by-frequent use of alternative assessments were taken into account.

A final explanatory model of perceived harsh classroom assessment environment. The final step in the analyses involved combining together statistically significant class-level variables detected in the early steps of the analyses to produce a parsimonious overall intercepts-and-slopes-as-outcomes regression model explaining the variability in (a) class mean perceived harsh classroom assessment environment and (b) relationship between student academic self-efficacy and perceived harsh classroom assessment environment. Based on the early steps of the analyses, the initial model that was tested in this step of the analyses was as follows.

Student-level:

\[(SPHCAE)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + r_{ij}.\]

Class-level:

\[
\begin{align*}
\beta_{0j} &= \gamma_{00} + \gamma_{01}(CEFC)_{j} + \gamma_{02}(CPLCAE)_{j} + \gamma_{03}(CPPCAE)_{j} + u_{0j}, \\
\beta_{1j} &= \gamma_{10} + \gamma_{11}(GNDR)_{j} + \gamma_{12}(ALTR)_{j} + \gamma_{13}(TEXP)_{j} + \\
&\gamma_{14}(GNDR \times TEXP)_{j} + \gamma_{15}(GNDR \times TEXP \times ALTR)_{j} + u_{1j},
\end{align*}
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation. 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 was group-mean centered. Following Raudenbush and Bryk’s (2002) suggestion, class-level variables having t-ratios near or less than one
were deleted and the model was re-estimated. Variables that were deleted in the early steps of the analyses were re-considered. The empirical Bayes residuals from the reduced composite model were regressed on the excluded variables. On the basis of this residual analysis, no variable was needed to be added to the model.

Based on the reduced composite model, the data showed no significant violation of the homogeneity assumption about student-level residual variance; \( \chi^2 (82) = 80.598, p > .500 \). In addition, the HLM model-based and robust standard errors were in quite close agreement with each other thereby giving no signal that the final fitted model of perceived harsh classroom assessment environment was misspecified. Therefore, Table 22 presents results of the final fitted explanatory model of perceived harsh classroom assessment environment.

With regard to class mean perceived harsh classroom assessment environment, as shown in Table 22, holding other factors constant, the average perceived learning classroom assessment environment was negatively related to class mean perceived harsh classroom assessment environment; \( \hat{\gamma}_{01} = -.167, t(80) = -3.306, p < .01 \). This suggested that holding other factors constant, a one standard deviation increase in class mean perceived learning classroom assessment environment was associated with a .167 standard deviation decrease in class mean for perceived harsh classroom assessment environment. Also, holding other factors constant, the average perceived public classroom assessment environment was positively related to class mean perceived harsh classroom assessment environment; \( \hat{\gamma}_{02} = .115, t(80) = 2.457, p < .05 \). This suggested that holding other factors
Table 22

**Final Fitted Composite Model of Perceived Harsh Classroom Assessment Environment**

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PHCAE mean, $\beta_{0j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{00}$</td>
<td>-.002</td>
<td>.042</td>
<td>-.040</td>
</tr>
<tr>
<td>CPLCAE, $\gamma_{01}$</td>
<td>-.167</td>
<td>.050</td>
<td>-3.306**</td>
</tr>
<tr>
<td>CPPCAE, $\gamma_{02}$</td>
<td>.115</td>
<td>.047</td>
<td>2.457*</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{10}$</td>
<td>-.219</td>
<td>.026</td>
<td>-8.344***</td>
</tr>
<tr>
<td>GNDR, $\gamma_{11}$</td>
<td>.041</td>
<td>.054</td>
<td>.755</td>
</tr>
<tr>
<td>ALTR, $\gamma_{12}$</td>
<td>.023</td>
<td>.025</td>
<td>.890</td>
</tr>
<tr>
<td>TEXP, $\gamma_{13}$</td>
<td>.023</td>
<td>.024</td>
<td>.963</td>
</tr>
<tr>
<td>GNDR $\times$ TEXP, $\gamma_{14}$</td>
<td>.310</td>
<td>.134</td>
<td>2.319*</td>
</tr>
<tr>
<td>GNDR $\times$ TEXP $\times$ ALTR, $\gamma_{15}$</td>
<td>-.337</td>
<td>.139</td>
<td>-2.423**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHCAE mean, $u_{0j}$</td>
<td>.1103</td>
<td>80</td>
<td>299.417***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0105</td>
<td>77</td>
<td>100.865*</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.7929</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. PHCAE = perceived harsh classroom assessment environment. CPLCAE = class average perceived learning classroom assessment environment. CPPCAE = class average perceived public classroom assessment environment. SEFC = student academic self-efficacy. GNDR = class gender (1 = female and -1 = male). ALTR = teacher’s frequent use of alternative assessments. TEXP = teacher’s teaching experience. 
*p < .05. **p < .01. ***p < .001.*
constant, a one standard deviation increase in class mean perceived public classroom assessment environment was associated with a .115 standard deviation increase in class mean for perceived harsh classroom assessment environment. Using the random-coefficient regression model presented in Table 18 as the base model, approximately 32% of the variance among classrooms in average perceived harsh classroom assessment environment was explained by class average perceived learning classroom assessment environment and class average perceived public classroom assessment environment.

With regard to academic self-efficacy slope, as shown in Table 22, on average, student academic self-efficacy was negatively related to perceived harsh classroom assessment environment within classrooms; \( \hat{\gamma}_{10} = -.219, t(77) = -8.344, p < .001 \); holding other factors constant. This suggested that holding other factors constant, a one standard deviation increase in student academic self-efficacy was on average associated with a .219 standard deviation decrease in perceived harsh classroom assessment environment within a classroom. Also, holding other factors constant, the differentiating effect of student academic self-efficacy on perceived harsh classroom assessment environment within a classroom depended jointly on class gender, teacher’s teaching experience, and teacher’s frequent use of alternative assessments; \( \hat{\gamma}_{15} = -.337, t(77) = -2.423, p < .05 \). This can be seen by computing the differentiating effect of academic 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:
Accordingly, for male classrooms using alternative assessments less frequently, classes having a high experienced teacher were less differentiating with regard to student academic self-efficacy than were classes having a low experienced teacher. The opposite was true in male classrooms using alternative assessments more frequently, in that, classes having a low experienced teacher were less differentiating with regard to student academic self-efficacy than were classes having a high experienced teacher. For female classrooms using alternative assessments less frequently, classes having a high experienced teacher were more differentiating with regard to student academic self-efficacy than were classes having a low experienced teacher. The opposite was true in female classrooms using alternative assessments more frequently, in that, classes having a low experienced teacher were more differentiating with regard to student academic self-efficacy than were classes having a high experienced teacher. Using the random-coefficient regression model presented in Table 18 as the base model, approximately 9% of the variance among classrooms in student self-efficacy effect on perceived harsh classroom assessment environment was explained once class gender, frequent use of alternative assessments, teaching experience, interaction of class gender-by-teaching experience, and interaction of class gender-by-teaching experience-by-frequent use of alternative assessments were taken into account.
Modeling Perceived Public Classroom Assessment Environment

A fully unconditional model. The HLM analyses to model perceived public classroom assessment environment (SPPCAE) began with a fully unconditional model, which is equivalent to a one-way random effects ANOVA, to examine how much variation in perceived public classroom assessment environment lay within and between classrooms. The outcome variable (SPPCAE) was standardized to a mean of zero and a standard deviation of one. The initial model that was tested in this step of the analyses was as follows.

Student-level:

\[(SPPCAE)_{ij} = \beta_{0j} + r_{ij}.\]

Class-level:

\[\beta_{0j} = \gamma_{00} + u_{0j},\]

where all terms were defined in the Analysis section of Chapter III of the dissertation.

Based on this model, a statistically significant variation was found among class means on perceived public classroom assessment environment; \(\hat{\tau}_{00} = .1748,\)

\[\chi^2(82) = 426.165, p < .001.\] The estimated within-class variance (\(\hat{\sigma}^2\)) was .8260. Hence, the intraclass correlation (\(\hat{p}\)) was estimated as .1747, indicating that approximately 17% of the variance in perceived public classroom assessment environment was between classrooms. The average reliability (\(\hat{\lambda}\)) of the class means was .806, suggesting that the sample means tended to be quite reliable as indicators of their true class means.
A random-coefficient regression model. The next step in the analyses involved posing a random-coefficient regression model to examine the relationship between student academic self-efficacy (SEFC) and perceived public classroom assessment environment (SPPCAE), and whether this relationship varied significantly across classes. All variables were standardized to a mean of zero and a standard deviation of one. The student-level independent variable (student academic self-efficacy) was group-mean centered. The initial model that was tested in this step of the analyses was as follows.

**Student-level:**

\[
(SPPCAE)_{ij} = \beta_{0j} + \beta_{1j} (SEFC_{ij} - CEFC_{j}) + r_{ij}.
\]

**Class-level:**

\[
\beta_{0j} = \gamma_{00} + u_{0j},
\beta_{1j} = \gamma_{10} + u_{1j},
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation.

Results indicated that, on average, student academic self-efficacy was not related to perceived public classroom assessment environment within classrooms; \( \hat{\gamma}_{10} = .003, t(82) = .106, p = .916 \). In addition, this relationship did not vary significantly across classrooms; \( \hat{\gamma}_{11} = .0063, \chi^2(82) = 97.005, p = .124 \). According to Raudenbush and Bryk (2002), these two conditions suggested that the variable student academic self-efficacy (SEFC) could be dropped from the model.

Having estimated the variability in the class means for perceived public classroom assessment environment (i.e., intercepts), the analyses proceeded with means-as-outcomes
regression models to explain this variability 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 joint effects of class gender, class average for academic self-efficacy, and class average for perceived classroom assessment environment. The second set represented the joint effects of class gender, teacher’s teaching experience, and teacher’s assessment practices. Then, two submodels of the means-as-outcomes regression model were fitted, one for each of the two sets of the class-level variables. Following are results pertaining to this process.

**Submodel (1): Joint effects of class gender, average academic self-efficacy, and average perceived classroom assessment environment.** This submodel examined the main effects of class gender (GNDR), class average academic self-efficacy (CEFC), class average perceived learning classroom assessment environment (CPLCAE), and class average perceived harsh classroom assessment environment (CPHCAE) as well as interaction effects of (GNDR × CEFC), (GNDR × CPLCAE), and (GNDR × CPHCAE) on class mean for perceived public classroom assessment environment. The initial model that was tested in this step of the analyses was as follows.

*Student-level:*

\[(SPPCAE)_{ij} = \beta_{0j} + r_{ij}.\]

*Class-level:*

\[\beta_{0j} = \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(CEFC)_{j} + \gamma_{03}(CPLCAE)_{j} + \gamma_{04}(CPHCAE)_{j} + \gamma_{05}(GNDR \times CEFC)_{j} + \gamma_{06}(GNDR \times CPLCAE)_{j} + \gamma_{07}(GNDR \times CPHCAE)_{j} + u_{0j} .\]
where all terms were defined in the Analysis section of Chapter III of the dissertation. 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. Following Raudenbush and Bryk’s (2002) suggestion, some class-level variables were deleted because they had t-ratios near or less than one, and the model was re-estimated. Results of this re-estimated model showed that only the main effect of average perceived harsh classroom assessment environment was statistically significant; \( \hat{\gamma}_{01} = .175, t(81) = 3.739, p < .01; \) suggesting that a one standard deviation increase in average perceived harsh classroom assessment environment was associated with a .175 standard deviation increase in class mean for perceived public classroom assessment environment. After taking average perceived harsh classroom assessment environment into account, the estimated between-class variance in perceived public classroom assessment environment \( (\hat{\tau}_{oo}) \) was reduced from .1748 in the fully unconditional model to .1464. Hence, using the fully unconditional model as the base model, about 16% of the variance among classrooms in average perceived public classroom assessment environment was accounted for by average perceived harsh classroom assessment environment.

*Submodel (2): Joint effects of class’s gender and teacher’s assessment practices and teaching experience.* The focus in this submodel was on the main effects of teacher’s teaching experience (TEXP); and teacher’s frequent uses of traditional assessments (TRAD), alternative assessments (ALTR), and recommended assessment practices (RECOM) along with three-way interaction effect of GNDR-by-TEXP-by-ALTR on class
mean perceived public classroom assessment environment. Specifically, the initial model that was tested in this step of the analyses was as follows.

**Student-level:**

\[
(SPPCAE)_{ij} = \beta_{0j} + r_{ij}. 
\]

**Class-level:**

\[
\beta_{0j} = \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(TEXP)_{j} + \gamma_{03}(TRAD)_{j} + \gamma_{04}(ALTR)_{j} + \\
\gamma_{05}(RECOM)_{j} + \gamma_{06}(GNDR \times TEXP)_{j} + \gamma_{07}(GNDR \times ALTR)_{j} + \\
\gamma_{08}(TEXP \times ALTR)_{j} + \gamma_{09}(GNDR \times TEXP \times ALTR)_{j} + u_{0j}
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation. 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. Results showed no statistically significant effects for class gender, teacher’s assessment practices, and teacher’s teaching experience on class mean perceived public classroom assessment environment.

*A final explanatory model of perceived public classroom assessment environment.*

The final step in the analyses involved combining together statistically significant class-level variables detected in the early steps of the analyses to produce a parsimonious overall means-as-outcomes regression model explaining the variability in class mean perceived public classroom assessment environment. Based on the early steps of the analyses, the initial model that was tested in this step of the analyses was as follows.
Student-level:

\[(SPPCAE)_{ij} = \beta_{0j} + r_{ij},\]

Class-level:

\[\beta_{0j} = \gamma_{00} + \gamma_{01}(CPHCAE)_{j} + u_{0j},\]

where all terms were defined in the Analysis section of Chapter III of the dissertation. All variables were standardized to a mean of zero and a standard deviation of one. Variables that were deleted in the early steps of the analyses were re-considered. The empirical Bayes residuals from the initial composite model were regressed on the excluded variables. On the basis of this residual analysis, a number of variables were added to the model as follows.

Student-level:

\[(SPPCAE)_{ij} = \beta_{0j} + r_{ij}.\]

Class-level:

\[
\begin{align*}
\beta_{0j} &= \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(ALTR)_{j} + \gamma_{03}(CPHCAE)_{j} + \\
&\quad \gamma_{04}(GNDR \times ALTR)_{j} + u_{0j},
\end{align*}
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation. 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.

Based on this composite model, the data showed no significant violation of the homogeneity assumption about student-level residual variance; \(\chi^2(82) = 99.966, p = .086\). In addition, the HLM model-based and robust standard errors were in quite close
agreement with each other thereby giving no signal that the final fitted model of perceived public classroom assessment environment was misspecified. Therefore, Table 23 presents results of the final fitted explanatory model of perceived public classroom assessment environment.

As shown in Table 23, holding other factors constant, the average perceived harsh classroom assessment environment was positively related to class mean perceived public classroom assessment environment; \( \hat{\gamma}_{03} = .176, t(78) = 4.008, p < .001 \). This suggested that holding other factors constant, a one standard deviation increase in class mean perceived harsh classroom assessment environment was associated with a .176 standard deviation increase in class mean for perceived public classroom assessment environment. Further, holding other factors constant, the effect of class gender on class average perceived public classroom assessment environment depended significantly on teacher’s frequent use of alternative assessments; \( \hat{\gamma}_{04} = -.615, t(78) = -2.522, p < .01 \). This gender effect on average perceived public classroom assessment environment can be found as follows:

Female prediction – male prediction

\[
= [(1)\hat{\gamma}_{01} + (1)(ALTR)_j\hat{\gamma}_{04}] - [(-1)\hat{\gamma}_{01} + (-1)(ALTR)_j\hat{\gamma}_{04}]
= 2 [\hat{\gamma}_{01} + (ALTR)_j\hat{\gamma}_{04}].
\]

Accordingly, in classes using alternative assessments more frequently (\( ALTR = 1 \)), the average perceived public classroom assessment environment was higher in male
Table 23

*Final Fitted Composite Model of Perceived Public Classroom Assessment Environment*

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PPCAE mean, $\beta_{0j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{00}$</td>
<td>-.055</td>
<td>.052</td>
<td>-1.061</td>
</tr>
<tr>
<td>GNDR, $\gamma_{01}$</td>
<td>.500</td>
<td>.251</td>
<td>1.993*</td>
</tr>
<tr>
<td>ALTR, $\gamma_{02}$</td>
<td>.030</td>
<td>.040</td>
<td>.762</td>
</tr>
<tr>
<td>CPHCAE, $\gamma_{03}$</td>
<td>.176</td>
<td>.044</td>
<td>4.008*</td>
</tr>
<tr>
<td>GNDR $\times$ ALTR, $\gamma_{04}$</td>
<td>-.615</td>
<td>.244</td>
<td>-2.522*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPCAE mean, $u_{0j}$</td>
<td>.1308</td>
<td>78</td>
<td>322.170***</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.8261</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* PPCAE = perceived public classroom assessment environment. GNDR = class gender (1 = female and -1 = male). ALTR = teacher’s frequent use of alternative assessments. CPHCAE = class average perceived harsh classroom assessment environment.

*p < .05. ***p < .001.
classrooms. The opposite was true in classes using alternative assessments less frequently (ALTR = -1), in that, the average perceived public classroom assessment environment was higher in female classrooms. Using the fully unconditional model as the base model, approximately 25% of the variance among classrooms in average perceived public classroom assessment environment was explained once class gender, class average perceived harsh classroom assessment environment, frequent use of alternative assessments, and interaction of class gender-by-frequent use of alternative assessments were taken into account.

Research Questions Seven, Eight, Nine, and Ten

RQ (7): Do ninth grade Muscat science classrooms in Oman vary in achievement goal orientations (GOAL)?

RQ (8): What are some of the student characteristics (e.g., academic self-efficacy) that might have effects on GOAL?

RQ (9): Do the effects of student characteristics (e.g., academic self-efficacy) on GOAL vary across classrooms?

RQ (10): What are some of the classroom characteristics (e.g., class gender) that might help explain the variability in the GOAL and in the effects of student characteristics (e.g., academic self-efficacy) on GOAL across classrooms?

The goal of these research questions was to construct a parsimonious model explaining student’s achievement goal orientations as a function of student-level and class-level characteristics. The data pertaining to these questions were hierarchically structured, in that students were nested within classes. Therefore, hierarchical linear model (HLM) analyses (Raudenbush & Bryk, 2002) were conducted, one for each dependent variable defined as a dimension of student’s achievement goal orientations that emerged
from the PCA/EFA described in the second research question. As discussed in the second research question, results of the PCA/EFA showed three dimensions of achievement goal orientations: (a) mastery, (b) performance-approach, and (c) performance-avoidance goals. Therefore, based on the PCA/EFA results and the literature reviewed in Chapter II of the dissertation, the research questions from seven to ten had the following dependent and independent variables

**Dependent variables:**
- Mastery goal orientation (SMASG)
- Performance-approach goal orientation (SPAPG)
- Performance-avoidance goal orientation (SPAVG)

**Independent variables at the student-level:**
- Student’s academic self-efficacy (SEFC)
- Student’s perceived learning classroom assessment environment (SPLCAE)
- Student’s perceived harsh classroom assessment environment (SPHCAE)
- Student’s perceived public classroom assessment environment (SPPCAE)

**Independent variables at the class-level:**
- Class’s gender (GNDR)
- Teacher’s years of teaching experience (TEXP)
- Teacher’s frequent use of traditional assessments (TRAD)
- Teacher’s frequent use of alternative assessments (ALTR)
- Teacher’s frequent use of recommended assessment practices (RECOM)

As discussed in Chapters I and II of the dissertation, the following four variables class’s average academic self-efficacy (CEFC), class’s average perceived learning classroom assessment environment (CPLCAE), class’s average perceived harsh classroom assessment environment (CPHCAE), and class’s average perceived public classroom assessment environment (CPPCAE) were included in order to study their contextual effects (Kreft & Leeuw, 1998; Raudenbush & Bryk, 2002) on achievement goal orientations. Kreft and Leeuw (1998) stated that “contextual
models are defined as regression models containing two types of variables: individual-level variables and aggregated context variables, such as group means” (p. 8). When applying Kreft and Leeuw’s (1998) definition of contextual models in the present study, the variables SEFC, SPLCAE, SPHCAE, and SPPCAE were used twice: once as individual student variables and once as aggregated class characteristics as follows:

- Class’s average for academic self-efficacy (CEFC)
- Class’s average for perceived learning classroom assessment environment (CPLCAE)
- Class’s average for perceived harsh classroom assessment environment (CPHCAE)
- Class’s average for perceived public classroom assessment environment (CPPCAE)

In order to facilitate interpretation of the HLM results, all dependent and independent 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 modeling process for each dependent variable began with a fully unconditional model followed by a random-coefficient regression model to test the effect of student-level independent variables. To build the class-level model, the class-level independent variables were divided into two sets following Raudenbush and Bryk’s (2002, p. 267) suggestion. The first set represented the contextual-effects of academic self-efficacy and perceived classroom 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 the statistically significant class-level variables from the aforementioned
submodels to produce a parsimonious overall intercepts-and-slopes-as-outcomes regression model of each dimension of achievement goal orientations. Finally, the validity of inferences based on the final model was assessed by verifying the tenability of two-level HLM assumptions (Raudenbush & Bryk, 2002, p. 255). Prior to the HLM analyses, univariate shape and frequency distribution of each variable at each level and bivariate relationships at each level were examined to identify discrepant cases and unusual observations that could have led to implausible results. Results of data screening were discussed early in this chapter. Following is a detailed description of the results pertaining to the HLM analyses.

Modeling Mastery Goal Orientation

A fully unconditional model. The HLM analyses to model mastery goal orientation (SMASG) began with a fully unconditional model, which is equivalent to a one-way random effects ANOVA, to examine how much variation in mastery goal orientation lay within and between classrooms. The outcome variable (SMASG) was standardized to a mean of zero and a standard deviation of one. The initial model that was tested in this step of the analyses was as follows.

\[ (SMASG)_j = \beta_{0j} + r_{ij}. \]

Student-level:

\[ \beta_{0j} = \gamma_{00} + u_{0j}, \]

Class-level:
where all terms were defined in the Analysis section of Chapter III of the dissertation.

Based on this model, a statistically significant variation was found among class means on mastery goal orientation; \( \hat{\tau}_{00} = .0851, \chi^2(82) = 231.861, p < .001 \). 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 orientation was between classrooms. The average reliability (\( \hat{\lambda} \)) of the class means was .646, suggesting that the sample means tended to be quite reliable as indicators of their true class means.

A random-coefficient regression model. The next step in the analyses involved posing a random-coefficient regression model to examine the relationships of student-level independent variables to mastery goal orientation (SMASG), and whether these relationships varied significantly across classes. The student-level independent variables were student academic self-efficacy (SEFC), perceived learning classroom assessment environment (SPLCAE), perceived harsh classroom assessment environment (SPHCAE), and perceived public classroom assessment environment (SPPCAE). All variables were standardized to a mean of zero and a standard deviation of one. The student-level independent variables were group-mean centered. The initial model that was tested in this step of the analyses was as follows.

Student-level:

\[
(SMASG)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_j) + \beta_{2j}(SPLCAE_{ij} - CPLCAE_j) + \\
\beta_{3j}(SPHCAE_{ij} - CPHCAE_j) + \beta_{4j}(SPPCAE_{ij} - CPPCAE_j) + r_{ij}
\]
Class-level:

\[
\begin{align*}
\beta_{0j} &= \gamma_{00} + u_{0j}, \\
\beta_{1j} &= \gamma_{10} + u_{1j}, \\
\beta_{2j} &= \gamma_{20} + u_{2j}, \\
\beta_{3j} &= \gamma_{30} + u_{3j}, \\
\beta_{4j} &= \gamma_{40} + u_{4j},
\end{align*}
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation.

Results of the initial random-coefficient regression model showed that, on average, the effect of student perceived public classroom assessment environment on mastery goal orientation tended to be null within classrooms; \( \hat{\gamma}_{40} = -.012, t(82) = -0.488, p = .627; \) and that this effect did not vary significantly across classrooms; \( \hat{\tau}_{44} = .0052, t(82) = 0.380, p = .296. \) According to Raudenbush and Bryk (2002), these results suggested that the variable student perceived public classroom assessment environment (SPPCAE) was a candidate for deletion from the model. In addition, although on average student perceived harsh classroom assessment environment was negatively related to mastery goal orientation within classrooms; \( \hat{\gamma}_{30} = -.075, t(82) = -2.914, p < .01; \) this relationship did not vary significantly across classrooms; \( \hat{\tau}_{33} = .0081, \chi^2(82) = 86.199, p = .354. \) Therefore, an alternative restricted random-coefficient regression model was estimated in which the variable student perceived public classroom assessment environment (SPPCAE) was dropped from the student-level model and that the student perceived harsh classroom assessment environment slope was specified as fixed as follows.
**Student-level:**

\[(\text{SMASG})_{ij} = \beta_{0j} + \beta_{1j}(\text{SEFC}_{ij} - \text{CEFC}_{ij}) + \beta_{2j}(\text{SPLCAE}_{ij} - \text{CPLCAE}_{ij}) + \beta_{3j}(\text{SPHCAE}_{ij} - \text{CPHCAE}_{ij}) + r_{ij}\]

**Class-level:**

\[
\begin{align*}
\beta_{0j} &= \gamma_{00} + u_{0j}, \\
\beta_{1j} &= \gamma_{10} + u_{1j}, \\
\beta_{2j} &= \gamma_{20} + u_{2j}, \\
\beta_{3j} &= \gamma_{30},
\end{align*}
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation.

The deviance statistic for the alternative reduced model (4109.648 with 11 df) where the variable student perceived public classroom assessment environment was dropped from the student-level model and that the student perceived harsh classroom assessment environment slope was specified as fixed was compared with the deviance statistic for the initial unrestricted model (4104.344 with 21 df) where the variable student perceived public classroom assessment environment was included in the student-level model and that the student perceived harsh classroom assessment environment slope was specified as having both a random effect and a fixed effect. The difference between these two deviance statistics (5.304) was not statistically significant when compared against the \(\chi^2\) distribution with 10 degrees of freedom. Therefore, the reduced model appeared sufficient.

Table 24 presents results of this reduced random-coefficient regression model of mastery goal orientation.
Table 24

Reduced Random-Coefficient Regression Model of Mastery Goal Orientation

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class MASG mean, $\gamma_{00}$</td>
<td>.000</td>
<td>.040</td>
<td>-.002</td>
</tr>
<tr>
<td>SEFC slope mean, $\gamma_{10}$</td>
<td>.333</td>
<td>.031</td>
<td>10.845***</td>
</tr>
<tr>
<td>SPLCAE slope, $\gamma_{20}$</td>
<td>.200</td>
<td>.032</td>
<td>6.302***</td>
</tr>
<tr>
<td>SPHCAE slope, $\gamma_{30}$</td>
<td>-.081</td>
<td>.025</td>
<td>-3.254**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MASG mean, $u_{0j}$</td>
<td>.0989</td>
<td>82</td>
<td>327.834***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0317</td>
<td>82</td>
<td>137.071***</td>
</tr>
<tr>
<td>SPLCAE slope, $u_{2j}$</td>
<td>.0364</td>
<td>82</td>
<td>144.992***</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.6479</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Correlations among class effects</th>
<th>$\beta_{0j}$</th>
<th>$\beta_{1j}$</th>
<th>$\beta_{2j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MASG mean, $\beta_{0j}$</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td>-.905</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SPLCAE slope, $\beta_{2j}$</td>
<td>.023</td>
<td>-.359</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reliability of OLS regression-coefficient estimates</th>
<th>$\beta_{0j}$</th>
<th>$\beta_{1j}$</th>
<th>$\beta_{2j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MASG mean, $\beta_{0j}$</td>
<td>.750</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td>.401</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPLCAE slope, $\beta_{2j}$</td>
<td>.420</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. MASG = mastery goal orientation. SEFC = student academic self-efficacy. SPLCAE = student perceived learning classroom assessment environment. SPHCAE = student perceived harsh classroom assessment environment. **$p < .01$. ***$p < .001$. 
As shown in Table 24, on average, student academic self-efficacy was positively related to mastery goal orientation within classrooms; $\hat{\gamma}_{10} = .333, t(82) = 10.845, p < .001$; suggesting that a one standard deviation increase in student academic self-efficacy was on average associated with a .333 standard deviation increase in mastery goal orientation within classrooms. This relationship between academic self-efficacy and mastery goal orientation varied significantly across classrooms; $\hat{r}_{11} = .0317, \chi^2(82) = 137.071, p < .001$. Also, on average, student perceived learning classroom assessment environment was positively related to mastery goal orientation within classrooms; $\hat{\gamma}_{20} = .200, t(82) = 6.302, p < .001$; signifying that a one standard deviation increase in student perceived learning classroom assessment environment was on average associated with a .200 standard deviation increase in mastery goal orientation within classrooms. This relationship between perceived learning classroom assessment environment and mastery goal orientation varied significantly across classrooms; $\hat{r}_{22} = .0364, \chi^2(82) = 144.992, p < .001$. Although, on average, student perceived harsh classroom assessment environment was negatively related to mastery goal orientation; $\hat{\gamma}_{30} = -.081, t(1632) = -3.254, p < .01$; this relationship was invariant across classrooms.

After taking student academic self-efficacy and perceived classroom assessment environment into account, the estimated within-class variance ($\hat{\sigma}^2$) was reduced from .9161 in the random-effects ANOVA model to .6479. Hence, student academic self-efficacy and perceived classroom assessment environment accounted for about 29% of the within-class variance in mastery goal orientation. As shown in Table 24, the
correlation between class mean mastery goal orientation and academic self-efficacy slope
\( \hat{p}_{01} = -.905 \) suggested that classes with high levels of mastery goal orientation tended to
be less differentiating with regard to student academic self-efficacy than were classes with
low levels of mastery goal orientation. The correlation between the random effects \( \hat{p}_{12} =
-.359 \) suggested that there was sufficient independent variation to treat each of them as
separate class effects. The estimates of the average class mastery goal orientation as well
as of the differentiating effects of academic self-efficacy and perceived learning classroom
assessment environment were moderately reliable, \( \lambda(\hat{\beta}_{0j}) = .750, .401, \) and \( .420 \) where \( q
= 0, 1 \) and \( 2; \) respectively, suggesting sufficient observed variation to be explained in the
intercepts \( (\hat{\beta}_{0j}) \) and slopes \( (\hat{\beta}_{qj}, \text{ where } q = 1 \text{ and } 2) \) using class characteristics.

Having estimated the variability in the class mastery goal orientation means (i.e.,
intercepts), student academic self-efficacy effects (i.e., academic self-efficacy slopes), and
student perceived classroom assessment environment effects (i.e., perceived classroom
assessment environment slopes), the analyses proceeded with intercepts-and-slopes-as-
outcomes regression models to explain the variability in these 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 academic self-efficacy and perceived classroom 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. Following are results pertaining to this process.

Submodel (1): Contextual-effects model of academic self-efficacy and perceived classroom assessment environment along with their differential contextual effects by class gender. Within HLM, a contextual effect is represented by including the class aggregate of a student-level variable in the between-class model for that differentiating effect (Lee & Bryk, 1989; Raudenbush & Bryk, 2002). Differential contextual effects by class gender are represented by the inclusion of a class aggregate variable-by-class gender interaction term in the between-class model (Lee & Bryk, 1989). Therefore, this submodel included class gender (GNDR), class average for academic self-efficacy (CEFC), class average for perceived learning classroom assessment environment (CPLCAE), and class average for perceived harsh classroom assessment environment (CPHCAE). In addition, the interaction terms (GNDR × CEFC), (GNDR × CPLCAE), and (GNDR × CPHCAE) were included in the class-level model to represent the differential contextual effects by class gender. The initial model that was tested in this step of the analyses was as follows.

**Student-level:**

\[
(SMASG)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + \beta_{2j}(SPLCAE_{ij} - CPLCAE_{j}) + \beta_{3j}(SPHCAE_{ij} - CPHCAE_{j}) + r_{ij}
\]

**Class-level:**

\[
\begin{align*}
\beta_{0j} &= \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(CEFC)_{j} + \gamma_{03}(CPLCAE)_{j} + \gamma_{04}(CPHCAE)_{j} + \\
\gamma_{05}(GNDR \times CEFC)_{j} + \gamma_{06}(GNDR \times CPLCAE)_{j} + \gamma_{07}(GNDR \times CPHCAE)_{j} + u_{0j},
\end{align*}
\]
\[ \beta_{1j} = \gamma_{10} + \gamma_{11}(GNDR)_{j} + \gamma_{12}(CEFC) + \gamma_{13}(GNDR \times CEFC)_{j} + u_{1j}, \]
\[ \beta_{2j} = \gamma_{20} + \gamma_{21}(GNDR)_{j} + \gamma_{22}(CPLCAE)_{j} + \gamma_{23}(GNDR \times CPLCAE)_{j} + u_{2j}, \]
\[ \beta_{3j} = \gamma_{30} + \gamma_{31}(GNDR)_{j} + \gamma_{32}(CPHCAE)_{j} + \gamma_{33}(GNDR \times CPHCAE)_{j}, \]

where all terms were defined in the Analysis section of Chapter III of the dissertation. 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. Following Raudenbush and Bryk’s (2002) suggestion, some class-level variables were deleted because they had t-ratios near or less than one, and the model was re-estimated as follows.

**Student-level:**
\[ (SMASG)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + \beta_{2j}(SPLCAE_{ij} - CPLCAE_{j}) + \beta_{3j}(SPHCAE_{ij} - CPHCAE_{j}) + r_{ij}. \]

**Class-level:**
\[ \beta_{0j} = \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(CEFC)_{j} + u_{0j}, \]
\[ \beta_{1j} = \gamma_{10} + u_{1j}, \]
\[ \beta_{2j} = \gamma_{20} + \gamma_{21}(GNDR)_{j} + \gamma_{22}(CPLCAE)_{j} + \gamma_{23}(GNDR \times CPLCAE)_{j} + u_{2j}, \]
\[ \beta_{3j} = \gamma_{30} + \gamma_{32}(CPHCAE)_{j}, \]

where all terms were defined in the Analysis section of Chapter III of the dissertation.
After deleting class-level variables with t-ratios near or less than one, the effect of student perceived learning classroom assessment environment ($\hat{\beta}_{2j}$) within classrooms became non-statistically significant; $\hat{\gamma}_{20} = .091, t(79) = 1.545, p = .126$. Yet, this effect varied significantly across classrooms; $\hat{\gamma}_{22} = .0271, \chi^2 (79) = 124.973, p < .01$. Therefore, an alternative restricted model was estimated in which the $\beta_{2j}$ was specified without an intercept (i.e., without a fixed effect) as follows.

**Student-level:**

$$(SMASG)_j = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + \beta_{2j}(SPLCAE_{ij} - CPLCAE_{j}) + \beta_{3j}(SPHCAE_{ij} - CPHCAE_{j}) + r_{ij},$$

**Class-level:**

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(CEFC)_{j} + u_{0j},$$

$$\beta_{1j} = \gamma_{10} + u_{1j},$$

$$\beta_{2j} = \gamma_{21}(GNDR)_{j} + \gamma_{22}(CPLCAE)_{j} + \gamma_{23}(GNDR \times CPLCAE)_{j} + u_{2j},$$

$$\beta_{3j} = \gamma_{30j},$$

where all terms were defined in the Analysis section of Chapter III of the dissertation.

The deviance statistic for this alternative restricted model (4042.217 with 15 df) where the perceived learning classroom assessment environment slope ($\beta_{2j}$) was specified without an intercept was compared with the deviance statistic for the unrestricted model (4039.829 with 16 df) where $\beta_{2j}$ was specified with an intercept. The difference between these two deviance statistics (2.388) was not statistically significant when
compared against the $\chi^2$ distribution with one degree of freedom. Therefore, the restricted model with $\beta_{2,j}$ specified without an intercept appeared sufficient. Table 25 presents results of the reduced contextual effects model of academic self-efficacy and perceived classroom assessment environment along with the differential contextual effects by class gender on mastery goal orientation.

With regard to class mean mastery goal orientation, as shown in Table 25, holding other factors constant, female classrooms had significantly higher average mastery goal orientation than did male classrooms; $\hat{\gamma}_{01} = .111, t(80) = 4.626, p < .001$. Also, holding other factors constant, the average academic self-efficacy of students was positively related to class mean mastery goal orientation; $\hat{\gamma}_{02} = .164, t(80) = 5.258, p < .001$. This suggested that holding other factors constant, a one standard deviation increase in average academic self-efficacy of students was associated with a .164 standard deviation increase in class mean mastery goal orientation. Using the random-coefficient regression model presented in Table 24 as the base model, approximately 62% of the variance among classrooms in average mastery goal orientation was explained by class gender and class average academic self-efficacy.

There was no evidence of gender, context, or gender-by-context effects for academic self-efficacy differentiation and perceived harsh classroom assessment environment differentiation. As shown in Table 25, the only statistically significant contextual effect in the data was for the relationship between student perceived learning classroom assessment environment and mastery goal orientation. This contextual effect,
Table 25

*Reduced Contextual Effects Model of Academic Self-Efficacy and Perceived Classroom Assessment Environment along with the Differential Contextual Effects by Class Gender on Mastery Goal Orientation*

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class MASG mean, $\beta_{0j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{00}$</td>
<td>-.008</td>
<td>.028</td>
<td>-.280</td>
</tr>
<tr>
<td>GNDR, $\gamma_{01}$</td>
<td>.111</td>
<td>.024</td>
<td>4.626***</td>
</tr>
<tr>
<td>CEFC, $\gamma_{02}$</td>
<td>.164</td>
<td>.031</td>
<td>5.258***</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{10}$</td>
<td>.348</td>
<td>.031</td>
<td>11.201***</td>
</tr>
<tr>
<td>SPLCAE slope, $\beta_{2j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GNDR, $\gamma_{21}$</td>
<td>1.673</td>
<td>.229</td>
<td>7.313***</td>
</tr>
<tr>
<td>CPLCAE, $\gamma_{22}$</td>
<td>.065</td>
<td>.028</td>
<td>2.304*</td>
</tr>
<tr>
<td>GNDR $\times$ CPLCAE, $\gamma_{23}$</td>
<td>-1.693</td>
<td>.228</td>
<td>-7.412***</td>
</tr>
<tr>
<td>SPHCAE slope, $\beta_{3j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{30}$</td>
<td>-.082</td>
<td>.025</td>
<td>-3.237**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MASG mean, $u_{0j}$</td>
<td>.0374</td>
<td>80</td>
<td>163.232***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0365</td>
<td>82</td>
<td>138.124***</td>
</tr>
<tr>
<td>SPLCAE slope, $u_{2j}$</td>
<td>.0276</td>
<td>80</td>
<td>128.722**</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.6417</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* MASM = mastery goal orientation. GNDR = class gender (1 = female and -1 = male). CEFC = class average academic self-efficacy. SEFC = student academic self-efficacy. SPLCAE = student perceived learning classroom assessment environment. CPLCAE = class average perceived learning classroom assessment environment. SPHCAE = student perceived harsh classroom assessment environment.

*p < .05. **p < .01. ***p < .001.*
however, worked differently in male and female classrooms; \( \hat{\gamma}_{23} = -1.693, t(80) = -7.412, p < .001 \). This can be seen by computing the effect of average perceived learning classroom assessment environment on perceived learning classroom assessment environment differentiation separately for male and female classrooms based on \( \hat{\gamma}_{22} + (GNDR) \hat{\gamma}_{23} \). For male classrooms, classes with a high average perceived learning classroom assessment environment were more differentiating with regard to student perceived learning classroom assessment environment than were classes with a low average perceived learning classroom assessment environment; \( \hat{\gamma}_{22} + (GNDR) \hat{\gamma}_{23} = .065 + (-1)(-1.693) = 1.758 \). For female classrooms, the opposite was true, in that, classes with a high average perceived learning classroom assessment environment were less differentiating with regard to student perceived learning classroom assessment environment than were classes with a low average perceived learning classroom assessment environment; \( \hat{\gamma}_{22} + (GNDR) \hat{\gamma}_{23} = .065 + (1)(-1.693) = -1.628 \). Using the random-coefficient regression model presented in Table 24 as the base model, approximately 24% of the variance among classrooms in the relationship between student perceived learning classroom assessment environment and mastery goal orientation was explained by class gender, class average perceived learning classroom assessment environment, and interaction of class gender-by-class average perceived learning classroom assessment environment.

Submodel (2): Joint effects model of class gender and teacher’s assessment practices and years of teaching experience. The focus in this submodel was on the main
effects of teacher’s years of teaching experience (TEXP); and teacher’s frequent uses of traditional assessments (TRAD), alternative assessments (ALTR), and recommended assessment practices (RECOM) along with three-way interaction effect of GNDR-by-TEXP-by-ALTR on (a) class mean mastery goal orientation, (b) relationship between student academic self-efficacy and mastery goal orientation, and (c) relationship between student perceived classroom assessment environment and mastery goal orientation.

Specifically, the initial model that was tested in this step of the analyses was as follows.

**Student-level:**

\[
(SMASG)_{ij} = \beta_{0j} + \beta_{ij}(SEFC_{ij} - CEFC_{ij}) + \beta_{2j}(SPLCAE_{ij} - CPLCAE_{ij}) + \\
\beta_{3j}(SPHCAE_{ij} - CPHCAE_{ij}) + r_{ij}.
\]

**Class-level:**

\[
\beta_{0j} = \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(TEXP)_{j} + \gamma_{03}(TRAD)_{j} + \gamma_{04}(ALTR)_{j} + \\
\gamma_{05}(RECOM)_{j} + \gamma_{06}(GNDR \times TEXP)_{j} + \gamma_{07}(GNDR \times ALTR)_{j} + \\
\gamma_{08}(TEXP \times ALTR)_{j} + \gamma_{09}(GNDR \times TEXP \times ALTR)_{j} + u_{0j}.
\]

\[
\beta_{1j} = \gamma_{10} + \gamma_{11}(GNDR)_{j} + \gamma_{12}(TEXP)_{j} + \gamma_{13}(TRAD)_{j} + \gamma_{14}(ALTR)_{j} + \\
\gamma_{15}(RECOM)_{j} + \gamma_{16}(GNDR \times TEXP)_{j} + \gamma_{17}(GNDR \times ALTR)_{j} + \\
\gamma_{18}(TEXP \times ALTR)_{j} + \gamma_{19}(GNDR \times TEXP \times ALTR)_{j} + u_{1j}.
\]

\[
\beta_{2j} = \gamma_{20} + \gamma_{21}(GNDR)_{j} + \gamma_{22}(TEXP)_{j} + \gamma_{23}(TRAD)_{j} + \gamma_{24}(ALTR)_{j} + \\
\gamma_{25}(RECOM)_{j} + \gamma_{26}(GNDR \times TEXP)_{j} + \gamma_{27}(GNDR \times ALTR)_{j} + \\
\gamma_{28}(TEXP \times ALTR)_{j} + \gamma_{29}(GNDR \times TEXP \times ALTR)_{j} + u_{2j}.
\]

\[
\beta_{3j} = \gamma_{30} + \gamma_{31}(GNDR)_{j} + \gamma_{32}(TEXP)_{j} + \gamma_{33}(TRAD)_{j} + \gamma_{34}(ALTR)_{j} + \\
\gamma_{35}(RECOM)_{j} + \gamma_{36}(GNDR \times TEXP)_{j} + \gamma_{37}(GNDR \times ALTR)_{j} + \\
\gamma_{38}(TEXP \times ALTR)_{j} + \gamma_{39}(GNDR \times TEXP \times ALTR)_{j}.
\]
where all terms were defined in the Analysis section of Chapter III of the dissertation. 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. Results showed no statistically significant effects for class gender, teacher’s assessment practices, and teacher’s teaching experience on (a) class mean mastery goal orientation, (b) relationship between student academic self-efficacy and mastery goal orientation, and (c) relationship between student perceived classroom assessment environment and mastery goal orientation.

A final explanatory model of mastery goal orientation. The final step in the analyses involved combining together statistically significant class-level variables detected in the early steps of the analyses to produce a parsimonious overall intercepts-and-slopes-as-outcomes regression model explaining the variability in (a) class mean mastery goal orientation, (b) relationship between student academic self-efficacy and mastery goal orientation, and (c) relationship between student perceived classroom assessment environment and mastery goal orientation. Based on the early steps of the analyses, the initial model that was tested in this step of the analyses was as follows.

Student-level:

\[(SMASG)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + \beta_{2j}(SPLCAE_{ij} - CPLCAE_{j}) + \beta_{3j}(SPHCAE_{ij} - CPHCAE_{j}) + r_{ij}\]
Class-level:

\[ \beta_{0j} = \gamma_{00} + \gamma_{01}(GNDR)_j + \gamma_{02}(CEFC)_j + u_{0j}, \]
\[ \beta_{1j} = \gamma_{10} + u_{1j}, \]
\[ \beta_{2j} = \gamma_{21}(GNDR)_j + \gamma_{22}(CPLCAE)_j + \gamma_{23}(GNDR \times CPLCAE)_j + u_{2j}, \]
\[ \beta_{3j} = \gamma_{30}j, \]

where all terms were defined in the Analysis section of Chapter III of the dissertation. 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.

Variables that were deleted in the early steps of the analyses were re-considered. The empirical Bayes residuals from the initial composite model were regressed on the excluded variables. On the basis of this residual analysis, class gender was added to the model as follows.

Student-level:

\[ (SMASG)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_fj) + \beta_{2j}(SPLCAE_{ij} - CPLCAE_fj) + \beta_{3j}(SPHCAE_{ij} - CPHCAE_fj) + r_{ij}. \]

Class-level:

\[ \beta_{0j} = \gamma_{00} + \gamma_{01}(GNDR)_j + \gamma_{02}(CEFC)_j + u_{0j}, \]
\[ \beta_{1j} = \gamma_{10} + \gamma_{11}(GNDR)_j + u_{1j}, \]
\[ \beta_{2j} = \gamma_{21}(GNDR)_j + \gamma_{22}(CPLCAE)_j + \gamma_{23}(GNDR \times CPLCAE)_j + u_{2j}, \]
\[ \beta_{3j} = \gamma_{30j}, \]

where all terms were defined in the Analysis section of Chapter III of the dissertation.

However, the data pertaining to this model displayed a heterogeneous residual variance at the student-level; \( \chi^2(82) = 129.968, p < .01 \). Further inspection of the data showed neither extreme observations than normally expected nor units having bad data. Further inspection of the model revealed that the effects of student-level independent variables were appropriately specified. As such, the reason for heterogeneity of student-level residual variances in this data set might be the omission of other important student-level independent variables from the model, and thus this might need to be considered in future research. Although estimation of the fixed effects and their standard errors is robust to violations of the homogeneity assumption about student-level residual variance (Kasim & Raudenbush, 1998), the final model was estimated both with and without the heterogeneous variance specification at the student-level. The student-level variance was modeled as a function of student academic self-efficacy, student perceived learning classroom assessment environment, and student perceived harsh classroom assessment environment as follows:

\[
\ln(\sigma_{ij}^2) = \alpha_0 + \alpha_1(SEFC)_i + \alpha_2(SPLCAE)_i + \alpha_3(SPHCAE)_i.
\]

Results indicated that the model with the heterogeneous variance specification at the student-level appeared to fit the data better than the model without the heterogeneous variance specification at the student-level; \( \chi^2(3) = 116.301, p < .001 \). Therefore, Table 26...
presents results of the final fitted explanatory model of mastery goal orientation with the heterogeneous variance specification at the student-level.

With regard to class mean mastery goal orientation, as shown in Table 26, female classrooms had significantly higher average mastery goal orientations than did male classrooms; $\hat{\gamma}_{01} = .139, t(80) = 5.262, p < .001$; controlling for class average academic self-efficacy. Also, holding other factors constant, average academic self-efficacy of students was positively related to class mean mastery goal orientation; $\hat{\gamma}_{02} = .168, t(80) = 5.664, p < .001$; suggesting that holding other factors constant, a one standard deviation increase in average academic self-efficacy of students was associated with a .168 standard deviation increase in class mean mastery goal orientation. Using the random coefficient regression model presented in Table 24 as the base model, approximately 69% of the variance among classrooms in average mastery goal orientation was explained by class gender and class average academic self-efficacy.

With regard to academic self-efficacy slope, as shown in Table 26, mastery goal orientation 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; $\hat{\alpha}_i = -.259, z = -5.927, p < .001$. Also, the relationship between student academic self-efficacy and mastery goal orientation 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 presented in Table 24 as the base model, class gender explained
approximately 2% of the variance in the relationship between student academic self-efficacy and mastery goal orientation.

With regard to perceived learning classroom assessment environment slope, as shown in Table 26, the differentiating effect of perceived learning classroom assessment environment within a classroom depended jointly on class average perceived learning classroom 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 classroom assessment environment were more differentiating with regard to student perceived learning classroom assessment environment than were classes with a low average perceived learning classroom assessment environment;

$$\hat{\gamma}_{22} + (GNDR) \hat{\gamma}_{23} = 0.059 + (-1)(-1.571) = 1.631.$$  In female classrooms, however, classes with a high average perceived learning classroom assessment environment were less differentiating with regard to student perceived learning classroom assessment environment than were classes with a low average perceived learning classroom assessment environment; $\hat{\gamma}_{22} + (GNDR) \hat{\gamma}_{23} = 0.059 + (1) (-1.572) = -1.513$. In addition, mastery goal orientation levels of students with a high perceived learning classroom assessment environment were less variable than those for students with low perceived learning classroom assessment environment levels; $\hat{\alpha}_2 = -0.181, z = -4.057, p < .001$. Using the random-coefficient regression model presented in Table 24 as the base model, approximately 46% of the variance among classrooms in the perceived learning classroom assessment environment differentiation effect was explained by class gender, class
average perceived learning classroom assessment environment, and interaction of class
gender-by-class average perceived learning classroom assessment environment.

With regard to perceived harsh classroom assessment environment slope, as shown
in Table 26, mastery goal orientation levels of students with a high perceived harsh
classroom assessment environment were on average lower; $\hat{\gamma}_{30} = -.090$, $t(1627) = -3.720$,
$p < .001$; and more variable than those for students with low perceived harsh classroom
assessment environment levels; $\hat{\alpha}_3 = .177$, $z = 4.172$, $p < .001$.

**Modeling Performance-Approach Goal Orientation**

*A fully unconditional model.* The HLM analyses to model performance-approach
goal orientation (SPAPG) began with a fully unconditional model, which is equivalent to a
one-way random effects ANOVA, to examine how much variation in performance-
approach goal orientation lay within and between classrooms. The outcome variable
(SPAPG) was standardized to a mean of zero and a standard deviation of one. The initial
model that was tested in this step of the analyses was as follows.

**Student-level:**

$$(SPAPG)_{ij} = \beta_{0j} + r_{ij}.$$  

**Class-level:**

$$\beta_{0j} = \gamma_{00} + u_{0j}.$$
Table 26

Final Fitted Composite Model of Mastery Goal Orientation with Heterogeneous Level-1 Variance

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class MASG mean, $\beta_{0j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{00}$</td>
<td>-.006</td>
<td>.029</td>
<td>-.216</td>
</tr>
<tr>
<td>GNDR, $\gamma_{01}$</td>
<td>.139</td>
<td>.026</td>
<td>5.262***</td>
</tr>
<tr>
<td>CEFC, $\gamma_{02}$</td>
<td>.168</td>
<td>.030</td>
<td>5.664***</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{10}$</td>
<td>.334</td>
<td>.031</td>
<td>10.743***</td>
</tr>
<tr>
<td>GNDR, $\gamma_{11}$</td>
<td>-.071</td>
<td>.031</td>
<td>-2.292*</td>
</tr>
<tr>
<td>SPLCAE slope, $\beta_{2j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GNDR, $\gamma_{21}$</td>
<td>1.557</td>
<td>.216</td>
<td>7.214***</td>
</tr>
<tr>
<td>CPLCAE, $\gamma_{22}$</td>
<td>.059</td>
<td>.022</td>
<td>2.747**</td>
</tr>
<tr>
<td>GNDR $\times$ CPLCAE, $\gamma_{23}$</td>
<td>-1.572</td>
<td>.222</td>
<td>-7.082***</td>
</tr>
<tr>
<td>SPHCAE slope, $\beta_{3j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{30}$</td>
<td>-.090</td>
<td>.024</td>
<td>-3.720***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MASG mean, $u_{0j}$</td>
<td>.0308</td>
<td>80</td>
<td>157.817***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0311</td>
<td>81</td>
<td>138.965***</td>
</tr>
<tr>
<td>SPLCAE slope, $u_{2j}$</td>
<td>.0198</td>
<td>80</td>
<td>120.549**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model for level-1 variance Parameter</th>
<th>Coefficient</th>
<th>SE</th>
<th>z-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept, $\alpha_0$</td>
<td>.504</td>
<td>.038</td>
<td>13.389***</td>
</tr>
<tr>
<td>SEFC, $\alpha_1$</td>
<td>-.259</td>
<td>.044</td>
<td>-5.927***</td>
</tr>
<tr>
<td>SPLCAE, $\alpha_2$</td>
<td>-.181</td>
<td>.045</td>
<td>-4.057***</td>
</tr>
<tr>
<td>SPHCAE, $\alpha_3$</td>
<td>.177</td>
<td>.042</td>
<td>4.172***</td>
</tr>
</tbody>
</table>

Note. MASG = mastery goal orientation. GNDR = class gender (1 = female and -1 = male). CEFC = class average academic self-efficacy. SEFC = student academic self-efficacy. SPLCAE = student perceived learning classroom assessment environment. CPLCAE = class average perceived learning classroom assessment environment. SPHCAE = student perceived harsh classroom assessment environment. 

*p < .05. **p < .01. ***p < .001.
where all terms were defined in the Analysis section of Chapter III of the dissertation.

Based on this model, a statistically significant variation was found among class means on performance-approach goal orientation; \( \hat{\tau}_{00} = .0571, \chi^2(82) = 179.7803, p < .001 \). The estimated within-class variance (\( \hat{\sigma}^2 \)) was .9436. Hence, the intraclass correlation (\( \hat{p} \)) was estimated as .0571, indicating that about 6% of the variance in performance-approach goal orientation was between classrooms. The average reliability (\( \hat{\lambda} \)) of the class means was .544, suggesting that the sample means were moderately reliable as indicators of their true class means.

A random-coefficient regression model. The next step in the analyses involved posing a random-coefficient regression model to examine the relationships of student-level independent variables to performance-approach goal orientation (SPAPG), and whether these relationships varied significantly across classes. The student-level independent variables were student academic self-efficacy (SEFC), perceived learning classroom assessment environment (SPLCAE), perceived harsh classroom assessment environment (SPHCAE), and perceived public classroom assessment environment (SPPCAE). All variables were standardized to a mean of zero and a standard deviation of one. The student-level independent variables were group-mean centered. The initial model that was tested in this step of the analyses was as follows.

**Student-level:**

\[
(\text{SPAPG})_{ij} = \beta_{0j} + \beta_{1j}(\text{SEFC}_{ij} - \text{CEFC}_{ij}) + \beta_{2j}(\text{SPLCAE}_{ij} - \text{CPLCAE}_{ij}) + \\
\beta_{3j}(\text{SPHCAE}_{ij} - \text{CPHCAE}_{ij}) + \beta_{4j}(\text{SPPCAE}_{ij} - \text{CPPCAE})_{ij} + r_{ij}
\]
Class-level:

\[
\begin{align*}
\beta_{0j} &= \gamma_{00} + u_{0j}, \\
\beta_{1j} &= \gamma_{10} + u_{1j}, \\
\beta_{2j} &= \gamma_{20} + u_{2j}, \\
\beta_{3j} &= \gamma_{30} + u_{3j}, \\
\beta_{4j} &= \gamma_{40} + u_{4j},
\end{align*}
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation.

Results of the initial random-coefficient regression model showed that, on average, both the effects of student perceived learning and perceived harsh classroom assessment environments on performance-approach goal orientation tended to be null within classrooms; \( \hat{\gamma}_{20} = .031, t(82) = 1.118, p = .267 \) and \( \hat{\gamma}_{30} = -.003, t(82) = -.102, p = .920; \) respectively. Also, each of these effects did not vary significantly across classrooms; \( \hat{\tau}_{22} = .0136, \chi^2(82) = 88.493, p = .292 \) and \( \hat{\tau}_{33} = .0059, \chi^2(82) = 83.956, p = .419; \) respectively. According to Raudenbush and Bryk (2002), these results suggested that the variables student perceived learning (SPLCAE) and perceived harsh (SPHCAE) classroom assessment environments were candidates for deletion from the model. Therefore, an alternative reduced random-coefficient regression model was estimated in which the variables SPLCAE and SPHCAE were dropped from the student-level model as follows.

Student-level:

\[
(SPAPG)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + \beta_{2j}(SPPCAE_{ij} - CPPCAE_{j}) + r_{ij}.
\]
Class-level:

\[
\begin{align*}
\beta_{0j} &= \gamma_{00} + u_{0j}, \\
\beta_{1j} &= \gamma_{10} + u_{1j}, \\
\beta_{2j} &= \gamma_{20} + u_{2j},
\end{align*}
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation.

The deviance statistic for the restricted alternative model (4421.132 with 10 df) where the variables SPLCAE and SPHCAE were dropped from the student-level model was compared with the deviance statistic for the unrestricted initial model (4409.784 with 21 df) where the variables SPLCAE and SPHCAE were included in the student-level model. The difference between these two deviance statistics (11.348) was not statistically significant when compared against the \( \chi^2 \) distribution with 11 degrees of freedom. Therefore, the restricted alternative model appeared sufficient.

Table 27 presents results of this reduced random-coefficient regression model of performance-approach goal orientation. As shown in Table 27, on average, student academic self-efficacy was positively related to performance-approach goal orientation within classrooms; \( \hat{\beta}_{10} = .274, t(82) = 9.291, p < .001 \); suggesting that a one standard deviation increase in student academic self-efficacy was on average associated with a .274 standard deviation increase in performance-approach goal orientation within classrooms. This relationship between academic self-efficacy and performance-approach goal orientation varied significantly across classrooms; \( \hat{\gamma}_{11} = .0229, \chi^2(82) = 124.412, p < .01 \). Also, on average, student perceived public classroom assessment environment was positively related to performance-approach goal orientation within classrooms;
\[ \hat{\beta}_{20} = .166, \ t(82) = 5.284, \ p < .001; \] indicating that a one standard deviation increase in student perceived public classroom assessment environment was on average associated with a .166 standard deviation increase in performance-approach goal orientation within classrooms. This relationship between perceived public classroom assessment environment and performance-approach goal orientation varied significantly across classrooms; \[ \hat{\tau}_{22} = .0277, \ \chi^2(82) = 110.714, \ p < .05. \]

After taking student academic self-efficacy and perceived public classroom assessment environment into account, the estimated within-class variance \( (\hat{\sigma}^2) \) was reduced from .9436 in the random-effects ANOVA model to .8052. Hence, student academic self-efficacy and perceived public classroom assessment environment accounted for about 15% of the within-class variance in performance-approach goal orientation. As also shown in Table 27, the correlation between class mean performance-approach goal orientation and academic self-efficacy slope \( (\hat{\rho}_{01} = -.589) \) suggested that classes with high levels of performance-approach goal orientation tended to be less differentiating with regard to student academic self-efficacy than were classes with low levels of performance-approach goal orientation. The correlation between the random effects \( (\hat{\rho}_{12} = -.527) \) suggested that there was sufficient independent variation to treat each of them as separate class effects. The estimates of the average class performance-approach goal orientation as well as of the differentiating effects of academic self-efficacy and perceived public classroom assessment environment were moderately
Table 27

*Reduced Random-Coefficient Regression Model of Performance-Approach Goal Orientation*

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PAPG mean, $\gamma_{00}$</td>
<td>.000</td>
<td>.035</td>
<td>-.005</td>
</tr>
<tr>
<td>SEFC slope mean, $\gamma_{10}$</td>
<td>.274</td>
<td>.029</td>
<td>9.291***</td>
</tr>
<tr>
<td>SPPCAE slope, $\gamma_{20}$</td>
<td>.166</td>
<td>.031</td>
<td>5.284***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAPG mean, $u_{0j}$</td>
<td>.0644</td>
<td>82</td>
<td>210.665***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0229</td>
<td>82</td>
<td>124.412**</td>
</tr>
<tr>
<td>SPPCAE slope, $u_{2j}$</td>
<td>.0277</td>
<td>82</td>
<td>110.714*</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.8052</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Correlations among class effects

<table>
<thead>
<tr>
<th>$\beta_{0j}$</th>
<th>$\beta_{1j}$</th>
<th>$\beta_{2j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAPG mean</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SEFC slope</td>
<td>-.589</td>
<td>-</td>
</tr>
<tr>
<td>SPPCAE slope</td>
<td>-.091</td>
<td>-.527</td>
</tr>
</tbody>
</table>

Reliability of OLS regression-coefficient estimates

<table>
<thead>
<tr>
<th>$\beta_{0j}$</th>
<th>$\beta_{1j}$</th>
<th>$\beta_{2j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAPG mean</td>
<td>.611</td>
<td></td>
</tr>
<tr>
<td>SEFC slope</td>
<td>.302</td>
<td></td>
</tr>
<tr>
<td>SPPCAE slope</td>
<td>.323</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* PAPG = performance-approach goal orientation. SEFC = student academic self-efficacy. SPPCAE = student perceived public classroom assessment environment.

*p < .05. **p < .01. ***p < .001.*
reliable; \( \lambda(\hat{\beta}_{qj}) = 0.611, 0.302, \) and 0.323 where \( q = 0, 1, \) and 2, respectively; suggesting sufficient observed variation to be explained in the intercepts \( (\hat{\beta}_{0j}) \) and slopes \( (\hat{\beta}_{qj}) \) where \( q = 1 \) and 2) using class characteristics.

Having estimated the variability in the class performance-approach goal orientation means (i.e., intercepts), student academic self-efficacy effects (i.e., academic self-efficacy slopes), and student perceived classroom assessment environment effects (i.e., perceived classroom assessment environment slopes), the analyses proceeded with intercepts-and-slopes-as-outcomes regression models to explain the variability in these 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 academic self-efficacy and perceived classroom 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. Following are results pertaining to this process.

Submodel (1): Contextual-effects model of academic self-efficacy and perceived classroom assessment environment along with their differential contextual effects by class gender. Within HLM, a contextual effect is represented by including the class aggregate of a student-level variable in the between-class model for that differentiating effect (Lee & Bryk, 1989; Raudenbush & Bryk, 2002). Differential contextual effects by class gender
are represented by the inclusion of a class aggregate variable-by-class gender interaction term in the between-class model (Lee & Bryk, 1989). Therefore, this submodel included class gender (GNDR), class average for academic self-efficacy (CEFC), and class average for perceived public classroom assessment environment (CPPCAE) as well as the interaction terms of (GNDR × CEFC) and (GNDR × CPPCAE). The initial model that was tested in this step of the analyses was as follows.

**Student-level:**

\[(SPAPG)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + \beta_{2j}(CPPCAE_{ij} - CPPCAE_{j}) + r_{ij} \].

**Class-level:**

\[
\beta_{0j} = \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(CEFC)_{j} + \gamma_{03}(CPPCAE)_{j} + \\
\gamma_{04}(GNDR \times CEFC)_{j} + \gamma_{05}(GNDR \times CPPCAE)_{j} + u_{0j}, \\
\beta_{1j} = \gamma_{10} + \gamma_{11}(GNDR)_{j} + \gamma_{12}(CEFC)_{j} + \gamma_{13}(GNDR \times CEFC)_{j} + u_{1j}, \\
\beta_{2j} = \gamma_{20} + \gamma_{21}(GNDR)_{j} + \gamma_{22}(CPPCAE)_{j} + \gamma_{23}(GNDR \times CPPCAE)_{j} + u_{2j},
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation. 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. Following Raudenbush and Bryk’s (2002) suggestion, some class-level variables were deleted because they had t-ratios near or less than one, and the model was re-estimated. Table 28 presents results of the reduced contextual effects model of academic self-efficacy and perceived classroom assessment environment on performance-approach goal orientation.
With regard to class mean performance-approach goal orientation, as shown in Table 28, holding other factors constant, the average academic self-efficacy of students was positively related to class mean performance-approach goal orientation; \( \hat{\gamma}_{02} = .115, t(79) = 3.623, p < .01 \). This suggested that holding other factors constant, a one standard deviation increase in average academic self-efficacy of students was associated with a .115 standard deviation increase in class mean performance-approach goal orientation.

Also, holding other factors constant, the average perceived public classroom assessment environment was positively related to class mean performance-approach goal orientation; \( \hat{\gamma}_{03} = .067, t(79) = 2.273, p < .05 \). This suggested that holding other factors constant, a one standard deviation increase in average perceived public classroom assessment environment was associated with a .067 standard deviation increase in class mean performance-approach goal orientation.

Using the random-coefficient regression model presented in Table 27 as the base model, approximately 30% of the variance among classrooms in average performance-approach goal orientation was explained by class gender, class average academic self-efficacy, and class average perceived public classroom assessment environment. However, there was no evidence of gender, context, or gender-by-context effects for the relationships of student academic self-efficacy and perceived public classroom assessment environment to performance-approach goal orientation.
Table 28

**Reduced Contextual Effects Model of Academic Self-Efficacy and Perceived Classroom Assessment Environment on Performance-Approach Goal Orientation**

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PAPG mean, $\beta_{0j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{00}$</td>
<td>-.006</td>
<td>.032</td>
<td>-.180</td>
</tr>
<tr>
<td>GNDR, $\gamma_{01}$</td>
<td>.055</td>
<td>.034</td>
<td>1.633</td>
</tr>
<tr>
<td>CEFC, $\gamma_{02}$</td>
<td>.115</td>
<td>.032</td>
<td>3.623**</td>
</tr>
<tr>
<td>CPPCAE, $\gamma_{03}$</td>
<td>.067</td>
<td>.030</td>
<td>2.273*</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{10}$</td>
<td>.278</td>
<td>.030</td>
<td>9.310***</td>
</tr>
<tr>
<td>SPPCAE slope, $\beta_{2j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{20}$</td>
<td>.167</td>
<td>.031</td>
<td>5.357***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAPG mean, $u_{0j}$</td>
<td>.0450</td>
<td>79</td>
<td>166.075***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0244</td>
<td>82</td>
<td>124.396**</td>
</tr>
<tr>
<td>SPPCAE slope, $u_{2j}$</td>
<td>.0270</td>
<td>82</td>
<td>110.832*</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.8049</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. PAPG = performance-approach goal orientation. GNDR = class gender (1 = female and -1 = male). CEFC = class average academic self-efficacy. CPPCAE = class average perceived public classroom assessment environment. SEFC = student academic self-efficacy. SPPCAE = student perceived public classroom assessment environment. *$p < .05$. **$p < .01$. ***$p < .001$. 

Submodel (2): Joint effects model of class gender and teacher’s assessment practices and years of teaching experience. The focus in this submodel was on the main effects of teacher’s years of teaching experience (TEXP); and teacher’s frequent uses of traditional assessments (TRAD), alternative assessments (ALTR), and recommended assessment practices (RECOM) along with three-way interaction effect of GNDR-by-TEXP-by-ALTR on (a) class mean performance-approach goal orientation, (b) relationship between student academic self-efficacy and performance-approach goal orientation, and (c) relationship between student perceived classroom assessment environment and performance-approach goal orientation. Specifically, the initial model that was tested in this step of the analyses was as follows.

**Student-level:**

\[(SPAPG)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{ij}) + \beta_{2j}(SPPCAE_{ij} - CPPCAE_{ij}) + r_{ij} \cdot \]

**Class-level:**

\[\beta_{0j} = \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(TEXP)_{j} + \gamma_{03}(TRAD)_{j} + \gamma_{04}(ALTR)_{j} + \gamma_{05}(RECOM)_{j} + \gamma_{06}(GNDR \times TEXP)_{j} + \gamma_{07}(GNDR \times ALTR)_{j} + \gamma_{08}(TEXP \times ALTR)_{j} + \gamma_{09}(GNDR \times TEXP \times ALTR)_{j} + u_{0j} \]

\[\beta_{1j} = \gamma_{10} + \gamma_{11}(GNDR)_{j} + \gamma_{12}(TEXP)_{j} + \gamma_{13}(TRAD)_{j} + \gamma_{14}(ALTR)_{j} + \gamma_{15}(RECOM)_{j} + \gamma_{16}(GNDR \times TEXP)_{j} + \gamma_{17}(GNDR \times ALTR)_{j} + \gamma_{18}(TEXP \times ALTR)_{j} + \gamma_{19}(GNDR \times TEXP \times ALTR)_{j} + u_{1j} \]

\[\beta_{2j} = \gamma_{20} + \gamma_{21}(GNDR)_{j} + \gamma_{22}(TEXP)_{j} + \gamma_{23}(TRAD)_{j} + \gamma_{24}(ALTR)_{j} + \gamma_{25}(RECOM)_{j} + \gamma_{26}(GNDR \times TEXP)_{j} + \gamma_{27}(GNDR \times ALTR)_{j} + \gamma_{28}(TEXP \times ALTR)_{j} + \gamma_{29}(GNDR \times TEXP \times ALTR)_{j} + u_{2j} \]
where all terms were defined in the Analysis section of Chapter III of the dissertation. 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. Following Raudenbush and Bryk’s (2002) suggestion, some class-level variables were deleted because they had t-ratios near or less than one, and the model was re-estimated. Table 29 presents results of the reduced joint effects model of class gender, teacher’s teaching experience, and assessment practices on performance-approach goal orientation.

Results showed no statistically significant effects for class gender, teacher’s teaching experience, and teacher’s assessment practices on class mean performance-approach goal orientation. However, as shown in Table 29, holding other factors constant, the differentiating effect of academic self-efficacy within a classroom depended jointly on class gender, teacher’s teaching experience, and teacher’s frequent use of alternative assessments; 
\[ \hat{\gamma}_{16} = -0.613, t(76) = -2.261, p < .05. \]
This can be seen by computing the differentiating effect of academic 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:

\[
\hat{\gamma}_{10} + \hat{\gamma}_{11}(GNDR) + \hat{\gamma}_{12}(ALTR) + \hat{\gamma}_{13}(TEXP) + \hat{\gamma}_{14}(GNDR \times TEXP) + \hat{\gamma}_{15}(GNDR \times ALTR) + \hat{\gamma}_{16}(GNDR \times TEXP \times ALTR). 
\]

Accordingly, for male classrooms using alternative assessments less frequently, classes having a high experienced teacher were less differentiating with regard to student
academic self-efficacy than were classes having a low experienced teacher. The opposite
was true in male classrooms using alternative assessments more frequently, in that, classes
having a low experienced teacher were less differentiating with regard to student academic
self-efficacy than were classes having a high experienced teacher. For female classrooms
using alternative assessments less frequently, classes having a high experienced teacher
were more differentiating with regard to student academic self-efficacy than were classes
having a low experienced teacher. The opposite was true in female classrooms using
alternative assessments more frequently, in that, classes having a low experienced teacher
were more differentiating with regard to student academic self-efficacy than were classes
having a high experienced teacher. Using the random-coefficient regression model
presented in Table 27 as the base model, approximately 21% of the variance among
classrooms in the academic self-efficacy differentiating effect was accounted once 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 were taken into account.

Likewise, as shown in Table 29, holding other factors constant, the differentiating
effect of perceived public classroom assessment environment within a classroom
depended jointly on class gender, teacher’s teaching experience, and teacher’s frequent
Table 29

Reduced Joint Effects Model of Class Gender, Teaching Experience, and Teacher’s Assessment Practices on Performance-Approach Goal Orientation

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PAPG mean, $\beta_{0j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{00}$</td>
<td>.000</td>
<td>.035</td>
<td>.001</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{10}$</td>
<td>.368</td>
<td>.043</td>
<td>8.642***</td>
</tr>
<tr>
<td>GNDR, $\gamma_{11}$</td>
<td>-.843</td>
<td>.305</td>
<td>-2.767**</td>
</tr>
<tr>
<td>ALTR, $\gamma_{12}$</td>
<td>.007</td>
<td>.026</td>
<td>.284</td>
</tr>
<tr>
<td>TEXP, $\gamma_{13}$</td>
<td>-.003</td>
<td>.020</td>
<td>-.163</td>
</tr>
<tr>
<td>GNDR × TEXP, $\gamma_{14}$</td>
<td>.539</td>
<td>.252</td>
<td>2.136*</td>
</tr>
<tr>
<td>GNDR × ALTR, $\gamma_{15}$</td>
<td>.883</td>
<td>.323</td>
<td>2.732**</td>
</tr>
<tr>
<td>GNDR × TEXP × ALTR, $\gamma_{16}$</td>
<td>-.613</td>
<td>.271</td>
<td>-2.261*</td>
</tr>
<tr>
<td>SPPCAE slope, $\beta_{2j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{20}$</td>
<td>.225</td>
<td>.043</td>
<td>5.169***</td>
</tr>
<tr>
<td>GNDR, $\gamma_{21}$</td>
<td>-.477</td>
<td>.301</td>
<td>-1.582</td>
</tr>
<tr>
<td>ALTR, $\gamma_{22}$</td>
<td>-.121</td>
<td>.052</td>
<td>-2.310*</td>
</tr>
<tr>
<td>TEXP, $\gamma_{23}$</td>
<td>-.299</td>
<td>.142</td>
<td>-2.109*</td>
</tr>
<tr>
<td>GNDR × TEXP, $\gamma_{24}$</td>
<td>.565</td>
<td>.289</td>
<td>1.956</td>
</tr>
<tr>
<td>GNDR × ALTR, $\gamma_{25}$</td>
<td>.611</td>
<td>.302</td>
<td>2.022*</td>
</tr>
<tr>
<td>TEXP × ALTR, $\gamma_{26}$</td>
<td>.368</td>
<td>.159</td>
<td>2.310*</td>
</tr>
<tr>
<td>GNDR × TEXP × ALTR, $\gamma_{27}$</td>
<td>-.655</td>
<td>.307</td>
<td>-2.131*</td>
</tr>
<tr>
<td>PAPG mean, $u_{0j}$</td>
<td>.0642</td>
<td>82</td>
<td>210.664***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0180</td>
<td>76</td>
<td>107.463*</td>
</tr>
<tr>
<td>SPPCAE slope, $u_{2j}$</td>
<td>.0248</td>
<td>75</td>
<td>96.336*</td>
</tr>
<tr>
<td>Level-1 effect, $r_g$</td>
<td>.8053</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. PAPG = performance-approach goal orientation. SEFC = student academic self-efficacy. GNDR = class gender (1 = female and -1 = male). ALTR = teacher’s frequent use of alternative assessments. TEXP = teacher’s teaching experience. SPPCAE = perceived public classroom assessment environment.
*p < .05. **p < .01. ***p < .001.
use of alternative assessments; $\hat{\gamma}_{27} = -.655$, $t(75) = -2.131, p < .05$. This can be seen by computing the differentiating effect of perceived public classroom assessment environment separately for male and female classrooms having high ($\text{TEXP} = 1$) and low ($\text{TEXP} = -1$) experienced teachers using alternative assessments more ($\text{ALTR} = 1$) or less ($\text{ALTR} = -1$) frequently based on the following equation:

$$
\hat{\gamma}_{20} + \hat{\gamma}_{21}(\text{GNDR})_j + \hat{\gamma}_{22}(\text{ALTR})_j + \hat{\gamma}_{23}(\text{TEXP})_j + \hat{\gamma}_{24}(\text{GNDR} \times \text{TEXP})_j + \\
\hat{\gamma}_{25}(\text{GNDR} \times \text{ALTR})_j + \hat{\gamma}_{26}(\text{TEXP} \times \text{ALTR})_j + \hat{\gamma}_{27}(\text{GNDR} \times \text{TEXP} \times \text{ALTR})_j.
$$

Accordingly, for male classrooms using alternative assessments less frequently, classes having a high experienced teacher were less differentiating with regard to student perceived public classroom assessment environment than were classes having a low experienced teacher. The opposite was true in male classrooms using alternative assessments more frequently, in that, classes having a low experienced teacher were less differentiating with regard to student perceived public classroom assessment environment than were classes having a high experienced teacher. For female classrooms using alternative assessments less frequently, classes having a high experienced teacher were more differentiating with regard to student perceived public classroom assessment environment than were classes having a low experienced teacher. The opposite was true in female classrooms using alternative assessments more frequently, in that, classes having a low experienced teacher were more differentiating with regard to student perceived public classroom assessment environment than were classes having a high experienced teacher. Using the random-coefficient regression model presented in Table 27 as the base model, approximately 10% of the variance among classrooms in the perceived public classroom
assessment environment differentiating effect was accounted once class gender, frequent use of alternative assessments, teaching experience, and two- and three-way interactions among class gender, frequent use of alternative assessments, and teaching experience were taken into account.

_A final explanatory model of performance-approach goal orientation._ The final step in the analyses involved combining together statistically significant class-level variables detected in the early steps of the analyses to produce a parsimonious overall intercepts-and-slopes-as-outcomes regression model explaining the variability in (a) class mean performance-approach goal orientation, (b) relationship between student academic self-efficacy and performance-approach goal orientation, and (c) relationship between student perceived classroom assessment environment and performance-approach goal orientation. Based on the early steps of the analyses, the initial model that was tested in this step of the analyses was as follows.

**Student-level:**

\[
(SPAPG)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + \beta_{2j}(SPPCAE_{ij} - CPPCAE_{j}) + r_{ij}.
\]

**Class-level:**

\[
\begin{align*}
\beta_{0j} &= \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(CEFC)_{j} + \gamma_{03}(CPPCAE)_{j} + u_{0j}, \\
\beta_{1j} &= \gamma_{10} + \gamma_{11}(GNDR)_{j} + \gamma_{12}(ALTR)_{j} + \gamma_{13}(TEXP)_{j} + \gamma_{14}(GNDR \times TEXP)_{j} + \gamma_{15}(GNDR \times ALTR)_{j} + \gamma_{16}(GNDR \times TEXP \times ALTR)_{j} + u_{1j}, \\
\beta_{2j} &= \gamma_{20} + \gamma_{21}(GNDR)_{j} + \gamma_{22}(ALTR)_{j} + \gamma_{23}(TEXP)_{j} + \gamma_{24}(GNDR \times TEXP)_{j} + \gamma_{25}(GNDR \times ALTR)_{j} + \gamma_{26}(TEXP \times ALTR)_{j} + \gamma_{27}(GNDR \times TEXP \times ALTR)_{j} + u_{2j},
\end{align*}
\]
where all terms were defined in the Analysis section of Chapter III of the dissertation. 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. Variables that were deleted in the early steps of the analyses were re-considered. The empirical Bayes residuals from the initial composite model were regressed on the excluded variables. On the basis of this residual analysis, one variable was added to the model as follows.

**Student-level:**

\[
(SPAPG)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + \beta_{2j}(SPPCAE_{ij} - CPPCAE_{j}) + r_{ij}.
\]

**Class-level:**

\[
\begin{align*}
\beta_{0j} &= \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(ALTR)_{j} + \gamma_{03}(CEFC)_{j} + \gamma_{04}(CPPCAE)_{j} + u_{0j}, \\
\beta_{1j} &= \gamma_{10} + \gamma_{11}(GNDR)_{j} + \gamma_{12}(ALTR)_{j} + \gamma_{13}(TEXP)_{j} + \gamma_{14}(GNDR \times TEXP)_{j} + \gamma_{15}(GNDR \times ALTR)_{j} + \gamma_{16}(GNDR \times TEXP \times ALTR)_{j} + u_{1j}, \\
\beta_{2j} &= \gamma_{20} + \gamma_{21}(GNDR)_{j} + \gamma_{22}(ALTR)_{j} + \gamma_{23}(TEXP)_{j} + \gamma_{24}(GNDR \times TEXP)_{j} + \gamma_{25}(GNDR \times ALTR)_{j} + \gamma_{26}(TEXP \times ALTR)_{j} + \gamma_{27}(GNDR \times TEXP \times ALTR)_{j} + u_{2j},
\end{align*}
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation.

However, the data pertaining to this composite model displayed a heterogeneous residual variance at the student-level; \(\chi^2(82) = 130.766, p < .001\). Further inspection of the data showed neither extreme observations than normally expected nor units having bad data. Further inspection of the model revealed that the effects of student-level independent variables were appropriately specified. As such, the reason for heterogeneity of student-
level residual variances in this data set might be the omission of other important student-level independent variables from the model, and thus this might need to be considered in future research. Although estimation of the fixed effects and their standard errors is robust to violations of the homogeneity assumption about student-level residual variance (Kasim & Raudenbush, 1998), the final model was estimated both with and without the heterogeneous variance specification at the student-level. The student-level variance was modeled as a function of student perceived public classroom assessment environment as follows:

$$\ln(\sigma_{ij}^2) = \alpha_0 + \alpha_1(SPPCAE)_{ij}.$$ 

Results indicated that the model with the heterogeneous variance specification at the student-level appeared to fit the data better than the model without the heterogeneous variance specification at the student-level; $$\chi^2(1) = 24.674, p < .001$$. After considering the heterogeneity of student-level variance, a number of issues were found in the model. First, the correlation between academic self-efficacy slope ($$\beta_{ij}$$) and perceived public classroom assessment environment slope ($$\beta_{2j}$$) was -.949, suggesting that these two random effects were carrying the same variation across the classrooms. According to Raudenbush and Bryk (2002), this issue warrants a reduction of the model by specifying one of these effects as fixed or non-randomly varying. Second, many of the class-level variables became non-statistically significant thereby making them candidates for deletion from the model. Third, the effect of perceived public classroom assessment environment on performance-approach goal orientation did not vary significantly across classrooms;
\( \hat{r}_{22} = .0117, \chi^2(75) = 86.517, p = .171; \) suggesting that this effect may need to be specified as fixed or non-randomly varying.

Therefore, the deviance statistic for the unrestricted model (4352.973 with 28 df) where the perceived public classroom assessment environment slope \((\beta_{2j})\) was specified as having both random and non-random components was compared with the deviance statistic for a reduced model (4367.170 with 19 df) where \(\beta_{2j}\) was specified as varying strictly as a function of class gender with no additional random component. The reduction in deviance (14.197) was not statistically significant when compared against the \(\chi^2\) distribution with nine degrees of freedom. Therefore, the reduced model appeared sufficient. Table 30 presents results of this final reduced composite model of performance-approach goal orientation.

With regard to class mean performance-approach goal orientation, as shown in Table 30, holding other factors constant, female classrooms had significantly higher average performance-approach goal orientation than did male classrooms; \(\hat{\gamma}_{01} = .069, t(78) = 2.049, p < .05\). Also, 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 orientation 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 academic self-efficacy of students was positively related to class mean performance-approach goal orientation; \(\hat{\gamma}_{03} = .114, t(78) = 3.914, p < .001\); suggesting that holding other factors constant, a one
standard deviation increase in average academic self-efficacy of students was associated with a .114 standard deviation increase in class mean performance-approach goal orientation. Moreover, holding other factors constant, the average perceived public classroom assessment environment was positively related to class mean performance-approach goal orientation; $\hat{\gamma}_{04} = .075, t(78) = 2.697, p < .01$; suggesting that holding other factors constant, a one standard deviation increase in average perceived public classroom assessment environment was associated with a .075 standard deviation increase in class mean performance-approach goal orientation. Using the random coefficient regression model presented in Table 27 as the base model, approximately 42% of the variance among classrooms in average performance-approach goal orientation was explained once class gender, teacher’s frequent use of alternative assessments, class average academic self-efficacy, and class average perceived public classroom assessment environment were taken into account.

With regard to academic self-efficacy slope, as shown in Table 30, holding other factors constant, on average, student academic self-efficacy was positively related to performance-approach goal orientation within classrooms; $\hat{\gamma}_{10} = .382, t(76) = 8.476, p < .001$; suggesting that holding other factors constant, a one standard deviation increase in student academic self-efficacy was on average associated with a .382 standard deviation increase in performance-approach goal orientation within classrooms. Also, holding other factors constant, the differentiating effect of academic 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 academic 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:

\[
\hat{\gamma}_{10} + \hat{\gamma}_{11}(GNDR)_{j} + \hat{\gamma}_{12}(ALTR)_{j} + \hat{\gamma}_{13}(TEXP)_{j} + \hat{\gamma}_{14}(GNDR \times TEXP)_{j} + \\
\hat{\gamma}_{15}(GNDR \times ALTR)_{j} + \hat{\gamma}_{16}(GNDR \times TEXP \times ALTR)_{j}
\]

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

Using the random-coefficient regression model presented in Table 27, approximately 32% of the variance among classrooms in academic 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 also shown in Table 30, performance-approach goal orientation levels of students with a high perceived public classroom 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 classroom assessment environment. Also, the positive relationship between perceived public classroom assessment environment and performance-approach goal orientation 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 Orientation**

A fully unconditional model. The HLM analyses to model performance-avoidance goal orientation (SPAVG) began with a fully unconditional model, which is equivalent to a one-way random effects ANOVA, to examine how much variation in performance-avoidance goal orientation lay within and between classrooms. The outcome variable
Table 30

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

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PAPG mean, $\beta_{0j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{00}$</td>
<td>-.006</td>
<td>.031</td>
<td>-.190</td>
</tr>
<tr>
<td>GNDR, $\gamma_{01}$</td>
<td>.069</td>
<td>.034</td>
<td>2.049*</td>
</tr>
<tr>
<td>ALTR, $\gamma_{02}$</td>
<td>-.068</td>
<td>.036</td>
<td>-1.895</td>
</tr>
<tr>
<td>CEFC, $\gamma_{03}$</td>
<td>.114</td>
<td>.029</td>
<td>3.914***</td>
</tr>
<tr>
<td>CPPCAE, $\gamma_{04}$</td>
<td>.075</td>
<td>.028</td>
<td>2.697**</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{10}$</td>
<td>.382</td>
<td>.045</td>
<td>8.476***</td>
</tr>
<tr>
<td>GNDR, $\gamma_{11}$</td>
<td>-.902</td>
<td>.321</td>
<td>-2.806**</td>
</tr>
<tr>
<td>ALTR, $\gamma_{12}$</td>
<td>.005</td>
<td>.026</td>
<td>.203</td>
</tr>
<tr>
<td>TEXP, $\gamma_{13}$</td>
<td>.008</td>
<td>.021</td>
<td>.405</td>
</tr>
<tr>
<td>GNDR × TEXP, $\gamma_{14}$</td>
<td>.617</td>
<td>.263</td>
<td>2.342*</td>
</tr>
<tr>
<td>GNDR × ALTR, $\gamma_{15}$</td>
<td>.955</td>
<td>.341</td>
<td>2.800**</td>
</tr>
<tr>
<td>GNDR × TEXP × ALTR, $\gamma_{16}$</td>
<td>-.715</td>
<td>.286</td>
<td>-2.502*</td>
</tr>
<tr>
<td>SPPCAE slope, $\beta_{2j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{20}$</td>
<td>.144</td>
<td>.028</td>
<td>5.079***</td>
</tr>
<tr>
<td>GNDR, $\gamma_{21}$</td>
<td>.064</td>
<td>.028</td>
<td>2.269*</td>
</tr>
<tr>
<td>PAPG mean, $u_{0j}$</td>
<td>.0372</td>
<td>78</td>
<td>160.706***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0155</td>
<td>76</td>
<td>117.481**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model for level-1 variance</th>
<th>Coefficient</th>
<th>SE</th>
<th>z-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept, $\alpha_0$</td>
<td>.213</td>
<td>.037</td>
<td>5.802***</td>
</tr>
<tr>
<td>SPPCAE, $\alpha_3$</td>
<td>-.196</td>
<td>.041</td>
<td>-4.789***</td>
</tr>
</tbody>
</table>

*Note.* PAPG = performance-approach goal orientation. GNDR = class gender (1 = female and -1 = male). ALTR = teacher’s frequent use of alternative assessments. CEFC = class average academic self-efficacy. CPPCAE = class average perceived public classroom assessment environment. SEFC = student academic-efficacy. TEXP = teacher’s teaching experience. SPPCAE = student perceived public classroom assessment environment. *$p < .05$. **$p < .01$. ***$p < .001$. 
(SPAVG) was standardized to a mean of zero and a standard deviation of one. The initial model that was tested in this step of the analyses was as follows.

**Student-level:**

\[(SPAVG)_{ij} = \beta_{0j} + r_{ij}.\]

**Class-level:**

\[\beta_{0j} = \gamma_{00} + u_{0j},\]

where all terms were defined in the Analysis section of Chapter III of the dissertation.

Based on this model, a statistically significant variation was found among class means on performance-avoidance goal orientation; \(\hat{\tau}_{00} = .0536, \chi^2(82) = 173.270, p < .001\). The estimated within-class variance (\(\hat{\sigma}^2\)) was .9474. Hence, the intraclass correlation (\(\hat{p}\)) was estimated as .0535, indicating that 5% of the variance in performance-avoidance goal orientation was between classrooms. The average reliability (\(\hat{\lambda}\)) of the class means was .527, suggesting that the sample means were moderately reliable as indicators of their true class means.

*A random-coefficient regression model.* The next step in the analyses involved posing a random-coefficient regression model to examine the relationships of student-level independent variables to performance-avoidance goal orientation (SPAVG), and whether these relationships varied significantly across classes. The student-level independent variables were student academic self-efficacy (SEFC), perceived learning classroom assessment environment (SPLCAE), perceived harsh classroom assessment
environment (SPHCAE), and perceived public classroom assessment environment (SPPCAE). All variables were standardized to a mean of zero and a standard deviation of one. The student-level independent variables were group-mean centered. The initial model that was tested in this step of the analyses was as follows.

**Student-level:**

\[
(SPAVG)_{ij} = \beta_0 + \beta_1(SEFC_{ij} - CEFC_{ij}) + \beta_2(SPLCAE_{ij} - CPLCAE_{ij}) + \\
\beta_3(SPHCAE_{ij} - CPHCAE_{ij}) + \beta_4(SPPCAE_{ij} - CPPCAE_{ij}) + r_{ij}.
\]

**Class-level:**

\[
\beta_{0j} = \gamma_0 + u_{0j}, \\
\beta_{1j} = \gamma_1 + u_{1j}, \\
\beta_{2j} = \gamma_2 + u_{2j}, \\
\beta_{3j} = \gamma_3 + u_{3j}, \\
\beta_{4j} = \gamma_4 + u_{4j},
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation.

Results of the initial random-coefficient regression model showed that, on average, the effects of student perceived learning and perceived harsh classroom assessment environments on performance-avoidance goal orientation each tended to be null within classrooms; \(\hat{\gamma}_{20} = .007, t(82) = .248, p = .804\) and \(\hat{\gamma}_{30} = .013, t(82) = .431, p = .667\); respectively. Also, each of these effects did not vary significantly across classrooms; \(\hat{\tau}_{22} = .0023, \chi^2(82) = 76.428, p > .500\) and \(\hat{\tau}_{33} = .0130, \chi^2(82) = 86.057, p = .358\); respectively. According to Raudenbush and Bryk (2002), these results suggested that the variables student perceived learning (SPLCAE) and perceived harsh (SPHCAE) classroom
assessment environments were candidates for deletion from the model. Therefore, an alternative reduced random-coefficient regression model was estimated in which the variables student perceived learning (SPLCAE) and perceived harsh (SPHCAE) classroom assessment environments were dropped from the student-level model as follows.

**Student-level:**

\[
(SPAVG)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + \beta_{2j}(SPPCAE_{ij} - CPPCAE_{j}) + r_{ij}.
\]

**Class-level:**

\[
\begin{align*}
\beta_{0j} &= \gamma_{00} + u_{0j}, \\
\beta_{1j} &= \gamma_{10} + u_{1j}, \\
\beta_{2j} &= \gamma_{20} + u_{2j},
\end{align*}
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation.

The deviance statistic for the reduced alternative model (4528.094 with 10 df) where both the variables student perceived learning (SPLCAE) and perceived harsh (SPHCAE) classroom assessment environments were dropped from the student-level model was compared with the deviance statistic for the unrestricted initial model (4521.979 with 21 df) where both the variables student perceived learning (SPLCAE) and perceived harsh (SPHCAE) classroom assessment environments were included in the student-level model. The difference between these two deviance statistics (6.115) was not statistically significant when compared against the \( \chi^2 \) distribution with 11 degrees of freedom. Therefore, the reduced model appeared sufficient.

After dropping the variables student perceived learning (SPLCAE) and perceived harsh (SPHCAE) classroom assessment environments from the model, there was a trend
for the effect of perceived public classroom assessment environment on performance-avoidance goal orientation not to vary significantly across classrooms; $\hat{\tau}_{22} = .0166$, $\chi^2(82) = 104.029$, $p = .051$. Therefore, the deviance statistic for the unrestricted model (4541.533 with 7 df) where the perceived public classroom assessment environment slope ($\beta_{2j}$) was specified as varying randomly across classrooms was compared with the deviance statistic for a restricted model (4548.530 with 4 df) where $\beta_{2j}$ was specified as fixed. The reduction in deviance (6.997) was not statistically significant when compared against the $\chi^2$ distribution with three degrees of freedom. Therefore, the restricted model where $\beta_{2j}$ specified as fixed appeared sufficient. Table 31 presents results of this reduced random-coefficient regression model of performance-avoidance goal orientation.

As shown in Table 31, on average, student academic self-efficacy was positively related to performance-avoidance goal orientation within classrooms; $\hat{\gamma}_{10} = .205$, $t(82) = 6.823$, $p < .001$. This suggests that a one standard deviation increase in student academic self-efficacy was on average associated with a .205 standard deviation increase in performance-avoidance goal orientation within classrooms. This relationship between academic self-efficacy and performance-avoidance goal orientation varied significantly across classrooms; $\hat{\tau}_{11} = .0215$, $\chi^2(82) = 122.739$, $p < .01$. Although on average student perceived public classroom assessment environment was positively related to performance-avoidance goal orientation; $\hat{\gamma}_{20} = .068$, $t(1633) = 2.276$, $p < .05$; this relationship was invariant across classrooms. After taking student academic self-efficacy
Table 31

*Reduced Random-Coefficient Regression Model of Performance-Avoidance Goal Orientation*

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PAVG mean, $\gamma_{00}$</td>
<td>-.001</td>
<td>.035</td>
<td>-.041</td>
</tr>
<tr>
<td>SEFC slope mean, $\gamma_{10}$</td>
<td>.205</td>
<td>.030</td>
<td>6.823***</td>
</tr>
<tr>
<td>SPPCAE slope mean, $\gamma_{20}$</td>
<td>.068</td>
<td>.030</td>
<td>2.276*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAVG mean, $u_{0j}$</td>
<td>.0569</td>
<td>82</td>
<td>185.572***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0215</td>
<td>82</td>
<td>122.739**</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.8846</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Correlations among class effects

$\beta_{0j}$ $\beta_{1j}$

| PAVG mean, $\beta_{0j}$           | -                    |
| SEFC slope, $\beta_{1j}$          | -.167                |

Reliability of OLS regression-coefficient estimates

| PAVG mean, $\beta_{0j}$ | .559 |
| SEFC slope, $\beta_{1j}$ | .280 |

*Note.* PAVG = performance-avoidance goal orientation. SEFC = student academic self-efficacy. SPPCAE = student perceived public classroom assessment environment. *p < .05. **p < .01. ***p < .001.
and perceived public classroom assessment environment into account, the estimated
within-class variance (\( \hat{\sigma}^2 \)) was reduced from .9474 in the random-effects ANOVA model
to .8846. Hence, student academic self-efficacy and perceived public classroom
assessment environment accounted for about 7% of the within-class variance in
performance-avoidance goal orientation.

As also shown in Table 31, the correlation between class mean performance-
avoidance goal orientation and academic self-efficacy slope (\( \hat{p}_{0i} = -.167 \)) suggested that
classes with high levels of performance-avoidance goal orientation tended to be less
differentiating with regard to student academic self-efficacy than were classes with low
levels of performance-avoidance goal orientation. The reliability estimates of the average
class performance-avoidance goal orientation, \( \hat{\lambda}(\hat{\beta}_{0j}) = .559 \), and of the academic self-
efficacy differentiating effect, \( \hat{\lambda}(\hat{\beta}_{ij}) = .280 \), suggested more observed variation to be
explained in the intercepts (\( \hat{\beta}_{0j} \)) than in the slopes (\( \hat{\beta}_{ij} \)) using class characteristics.

Having estimated the variability in the class performance-avoidance goal
orientation means (i.e., intercepts) and student academic self-efficacy effects (i.e.,
academic self-efficacy slopes), the analyses proceeded with intercepts-and-slopes-as-
outcomes regression models to explain the variability in these 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 academic self-efficacy and perceived classroom 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. Following are results pertaining to this process.

Submodel (1): Contextual-effects model of academic self-efficacy and perceived classroom assessment environment along with their differential contextual effects by class gender. Within HLM, a contextual effect is represented by including the class aggregate of a student-level variable in the between-class model for that differentiating effect (Lee & Bryk, 1989; Raudenbush & Bryk, 2002). Differential contextual effects by class gender are represented by the inclusion of a class aggregate variable-by-class gender interaction term in the between-class model (Lee & Bryk, 1989). Therefore, this submodel included class gender (GNDR), class average for academic self-efficacy (CEFC), and class average for perceived public classroom assessment environment (CPPCAE) as well as the interaction terms of (GNDR × CEFC) and (GNDR × CPPCAE). The initial model that was tested in this step of the analyses was as follows.

**Student-level:**

\[
(SPAVG)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + \beta_{2j}(SPPCAE_{ij} - CPPCAE_{j}) + r_{ij}.
\]

**Class-level:**

\[
\beta_{0j} = \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(CEFC)_{j} + \gamma_{03}(CPPCAE)_{j} + \gamma_{04}(GNDR \times CEFC)_{j} + \gamma_{05}(GNDR \times CPPCAE)_{j} + u_{0j},
\]

\[
\beta_{1j} = \gamma_{10} + \gamma_{11}(GNDR)_{j} + \gamma_{12}(CEFC)_{j} + \gamma_{13}(GNDR \times CEFC)_{j} + u_{1j}.
\]
\[ \beta_{2j} = \gamma_{20} + \gamma_{21} (GNDR)_{j} + \gamma_{22} (CPPCAE) + \gamma_{23} (GNDR \times CPPCAE)_{j}, \]

where all terms were defined in the Analysis section of Chapter III of the dissertation. 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. Following Raudenbush and Bryk’s (2002) suggestion, some class-level variables were deleted because they had t-ratios near or less than one, and the model was re-estimated. Table 32 presents results of the reduced contextual effect model of academic self-efficacy on performance-avoidance goal orientation.

With regard to class mean performance-avoidance goal orientation, as shown in Table 32, the average academic self-efficacy of students was positively related to class mean performance-avoidance goal orientation; \( \hat{\gamma}_{01} = .123, t(81) = 4.012, p < .001. \) This suggested that a one standard deviation increase in average academic self-efficacy of students was associated with a .123 standard deviation increase in class mean performance-avoidance goal orientation. Using the random-coefficient regression model presented in Table 31 as the base model, approximately 25% of the variance among classrooms in average performance-avoidance goal orientation was explained by class average academic self-efficacy. However, there was no evidence of gender, context, or gender-by-context effects for the relationships of student academic self-efficacy and perceived classroom assessment environment to performance-avoidance goal orientation.
Table 32

*Reduced Contextual Effect Model of Academic Self-Efficacy on Performance-Avoidance Goal Orientation*

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PAVG mean, $\beta_{0j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{00}$</td>
<td>-.001</td>
<td>.032</td>
<td>-.028</td>
</tr>
<tr>
<td>CEFC, $\gamma_{01}$</td>
<td>.123</td>
<td>.031</td>
<td>4.012***</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{10}$</td>
<td>.206</td>
<td>.030</td>
<td>6.845***</td>
</tr>
<tr>
<td>SPPCAE slope, $\beta_{2j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{20}$</td>
<td>.069</td>
<td>.030</td>
<td>2.284*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAVG mean, $u_{0j}$</td>
<td>.0428</td>
<td>81</td>
<td>158.269***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0215</td>
<td>82</td>
<td>122.714**</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.8846</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* PAVG = performance-avoidance goal orientation. CEFC = class average academic self-efficacy. SEFC = student academic self-efficacy. SPPCAE = student perceived public classroom assessment environment.

*$p < .05$. **$p < .01$. ***$p < .001$. 
Submodel (2): Joint effects model of class gender and teacher’s assessment practices and years of teaching experience. The focus in this submodel was on the main effects of teacher’s years of teaching experience (TEXP); and teacher’s frequent uses of traditional assessments (TRAD), alternative assessments (ALTR), and recommended assessment practices (RECOM) along with three-way interaction effect of GNDR-by-TEXP-by-ALTR on (a) class mean performance-avoidance goal orientation, (b) relationship between student academic self-efficacy and performance-avoidance goal orientation, and (c) relationship between student perceived public classroom assessment environment and performance-avoidance goal orientation. Specifically, the initial model that was tested in this step of the analyses was as follows.

**Student-level:**

\[
(SPAVG)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{ij}) + \beta_{2j}(SPPCAE_{ij} - CPPCAE_{ij}) + r_{ij}.
\]

**Class-level:**

\[
\begin{align*}
\beta_{0j} &= \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(TEXP)_{j} + \gamma_{03}(TRAD)_{j} + \gamma_{04}(ALTR)_{j} + \\
&\quad \gamma_{05}(RECOM)_{j} + \gamma_{06}(GNDR \times TEXP)_{j} + \gamma_{07}(GNDR \times ALTR)_{j} + \\
&\quad \gamma_{08}(TEXP \times ALTR)_{j} + \gamma_{09}(GNDR \times TEXP \times ALTR)_{j} + u_{0j},
\end{align*}
\]

\[
\begin{align*}
\beta_{1j} &= \gamma_{10} + \gamma_{11}(GNDR)_{j} + \gamma_{12}(TEXP)_{j} + \gamma_{13}(TRAD)_{j} + \gamma_{14}(ALTR)_{j} + \\
&\quad \gamma_{15}(RECOM)_{j} + \gamma_{16}(GNDR \times TEXP)_{j} + \gamma_{17}(GNDR \times ALTR)_{j} + \\
&\quad \gamma_{18}(TEXP \times ALTR)_{j} + \gamma_{19}(GNDR \times TEXP \times ALTR)_{j} + u_{1j},
\end{align*}
\]

\[
\begin{align*}
\beta_{2j} &= \gamma_{20} + \gamma_{21}(GNDR)_{j} + \gamma_{22}(TEXP)_{j} + \gamma_{23}(TRAD)_{j} + \gamma_{24}(ALTR)_{j} + \\
&\quad \gamma_{25}(RECOM)_{j} + \gamma_{26}(GNDR \times TEXP)_{j} + \gamma_{27}(GNDR \times ALTR)_{j} + \\
&\quad \gamma_{28}(TEXP \times ALTR)_{j} + \gamma_{29}(GNDR \times TEXP \times ALTR)_{j},
\end{align*}
\]
where all terms were defined in the Analysis section of Chapter III of the dissertation. 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. Following
Raudenbush and Bryk’s (2002) suggestion, some class-level variables were deleted
because they had t-ratios near or less than one, and the model was re-estimated. Table 33
presents results of the reduced joint effects model of class gender and teacher’s assessment
practices on performance-avoidance goal orientation.

As shown in Table 33, female classrooms had significantly higher average
performance-avoidance goal orientation than did male classrooms; \( \hat{\gamma}_{01} = .119, t(81) =
3.647, p < .01 \). Using the random-coefficient regression model presented in Table 31 as
the base model, approximately 23% of the variance among classrooms in class mean
performance-avoidance goal orientation was accounted for by class gender. As also shown
in Table 33, although not statistically significant, there was a trend for classes with a high
emphasis on traditional assessments to have a stronger positive effect of student academic
self-efficacy on performance-avoidance goal orientation than did classes with a low
emphasis on traditional assessments; \( \hat{\gamma}_{11} = .051, t(81) = 1.947, p = .055 \). Using the
random-coefficient regression model presented in Table 31 as the base model, teacher’s
frequent use of traditional assessments accounted for about 4% of the variance among
classrooms in the relationship between student academic self-efficacy and performance-
avoidance goal orientation. However, there was no evidence of class gender, teacher’s
teaching experience, and teacher’s assessment practices effects on the relationship between student perceived public classroom assessment environment and performance-avoidance goal orientation.

Table 33

Reduced Joint Effects Model of Class Gender and Teacher’s Assessment Practices on Performance-Avoidance Goal Orientation

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PAVG mean, $\beta_{0j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{00}$</td>
<td>-.014</td>
<td>.033</td>
<td>-.419</td>
</tr>
<tr>
<td>GNDR, $\gamma_{01}$</td>
<td>.119</td>
<td>.033</td>
<td>3.647**</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{10}$</td>
<td>.206</td>
<td>.030</td>
<td>6.948***</td>
</tr>
<tr>
<td>TRAD, $\gamma_{11}$</td>
<td>.051</td>
<td>.026</td>
<td>1.947</td>
</tr>
<tr>
<td>SPPCAE slope, $\beta_{2j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{20}$</td>
<td>.070</td>
<td>.030</td>
<td>2.330*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAVG mean, $u_{0j}$</td>
<td>.0436</td>
<td>81</td>
<td>159.926***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0206</td>
<td>81</td>
<td>118.961**</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.8872</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* PAVG = performance-avoidance goal orientation. GNDR = class gender (1 = female and -1 = male). SEFC = student academic self-efficacy. TRAD = teacher’s frequent use of traditional assessments. SPPCAE = student perceived public classroom assessment environment.

*p < .05. **p < .01. ***p < .001.
**A final explanatory model of performance-avoidance goal orientation.** The final step in the analyses involved combining together statistically significant class-level variables detected in the early steps of the analyses to produce a parsimonious overall intercepts-and-slopes-as-outcomes regression model explaining the variability in (a) class mean performance-avoidance goal orientation, (b) relationship between student academic self-efficacy and performance-avoidance goal orientation, and (c) relationship between student perceived public classroom assessment environment and performance-avoidance goal orientation. Based on the early steps of the analyses, the initial model that was tested in this step of the analyses was as follows.

**Student-level:**

\[(SPAVG)_{ij} = \beta_0 + \beta_{1i} (SEFC_{ij} - CEFC_j) + \beta_{2j} (SPPCAE_{ij} - CPPCAE_j) + r_{ij}.\]

**Class-level:**

\[\beta_{0j} = \gamma_0 + \gamma_{01} (GNDR)_j + \gamma_{02} (CEFC)_j + u_{0j},\]

\[\beta_{1j} = \gamma_{10} + \gamma_{11} (TRAD)_{ij} + u_{1j},\]

\[\beta_{2j} = \gamma_{20},\]

where all terms were defined in the Analysis section of Chapter III of the dissertation. 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. Variables that were deleted in the early steps of the analyses were re-considered. The empirical Bayes residuals from the initial composite model were regressed on the excluded variables. On
the basis of this residual analysis, a number of variables were added to the model as follows.

**Student-level:**

\[
(SPAVG)_{ij} = \beta_{0j} + \beta_{1j}(SEFC_{ij} - CEFC_{j}) + \beta_{2j}(SPPCAE_{ij} - CPPCAE_{j}) + r_{ij}.
\]

**Class-level:**

\[
\beta_{0j} = \gamma_{00} + \gamma_{01}(GNDR)_{j} + \gamma_{02}(CEFC)_{j} + u_{0j},
\]

\[
\beta_{1j} = \gamma_{10} + \gamma_{11}(GNDR)_{j} + \gamma_{12}(TRAD)_{j} + \gamma_{13}(TEXP)_{j} + \\
\gamma_{14}(GNDR \times TEXP)_{j} + u_{1j},
\]

\[
\beta_{2j} = \gamma_{20},
\]

where all terms were defined in the Analysis section of Chapter III of the dissertation.

After considering the residual analysis, the data showed no significant violation of the homogeneity assumption about student-level residual variance; \( \chi^2(82) = 90.503, p = .244 \). In addition, the HLM model-based and robust standard errors were in quite close agreement with each other thereby giving no signal that the final fitted model of performance-avoidance goal orientation was misspecified. Therefore, Table 34 presents results of the final fitted explanatory model of performance-avoidance goal orientation.

As shown in Table 34, holding other factors constant, female classrooms had significantly higher average performance-avoidance goal orientation than did male classrooms; \( \hat{\gamma}_{01} = .086, t(80) = 2.512, p < .05 \). Also, holding other factors constant, the average academic self-efficacy of students was positively related to class mean performance-avoidance goal orientation; \( \hat{\gamma}_{02} = .093, t(80) = 2.789, p < .01 \); suggesting
that holding other factors constant, a one standard deviation increase in average academic self-efficacy of students was associated with a .093 standard deviation increase in class mean performance-avoidance goal orientation. Using the random coefficient regression model presented in Table 31 as the base model, approximately 35% of the variance among classrooms in average performance-avoidance goal orientation was explained by class gender and class average academic self-efficacy.

According to Table 34, holding other factors constant, on average, student academic self-efficacy was positively related to performance-avoidance goal orientation within classrooms; \( \hat{\gamma}_{10} = .196, t(78) = 6.877, p < .001 \); suggesting that holding other factors constant, a one standard deviation increase in student academic self-efficacy was on average associated with a .196 standard deviation increase in performance-avoidance goal orientation within classrooms. Also, holding other factors constant, classes with a high frequent use of traditional assessments had a stronger positive effect of student academic self-efficacy on performance-avoidance goal orientation 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 academic 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 high experienced teacher were less differentiating with regard to student academic self-efficacy than were classes having a low experienced teacher; \( \hat{\gamma}_{13} + (GNDR) \hat{\gamma}_{14} = -.001 + (1)(-.135) = -.136 \); holding other factors constant. In male
Table 34

*Final Fitted Composite Model of Performance-Avoidance Goal Orientation*

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class PAVG mean, $\beta_{0j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{00}$</td>
<td>-.010</td>
<td>.031</td>
<td>-.317</td>
</tr>
<tr>
<td>GNDR, $\gamma_{01}$</td>
<td>.086</td>
<td>.034</td>
<td>2.512*</td>
</tr>
<tr>
<td>CEFC, $\gamma_{02}$</td>
<td>.093</td>
<td>.033</td>
<td>2.789**</td>
</tr>
<tr>
<td>SEFC slope, $\beta_{1j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{10}$</td>
<td>.196</td>
<td>.029</td>
<td>6.877***</td>
</tr>
<tr>
<td>GNDR, $\gamma_{11}$</td>
<td>.083</td>
<td>.065</td>
<td>1.293</td>
</tr>
<tr>
<td>TRAD, $\gamma_{12}$</td>
<td>.054</td>
<td>.027</td>
<td>2.031*</td>
</tr>
<tr>
<td>TEXP, $\gamma_{13}$</td>
<td>-.001</td>
<td>.023</td>
<td>-.049</td>
</tr>
<tr>
<td>GNDR $\times$ TEXP, $\gamma_{14}$</td>
<td>-.135</td>
<td>.054</td>
<td>-2.529*</td>
</tr>
<tr>
<td>SPPCAE slope, $\beta_{2j}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base, $\gamma_{20}$</td>
<td>.071</td>
<td>.029</td>
<td>2.394*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effect</th>
<th>Variance component</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAVG mean, $u_{0j}$</td>
<td>.0369</td>
<td>80</td>
<td>146.244***</td>
</tr>
<tr>
<td>SEFC slope, $u_{1j}$</td>
<td>.0165</td>
<td>78</td>
<td>109.237*</td>
</tr>
<tr>
<td>Level-1 effect, $r_{ij}$</td>
<td>.8850</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

*p < .05. **p < .01. ***p < .001.
classrooms, the opposite was true, in that classes having a high experienced teacher were more differentiating with regard to student academic self-efficacy than were classes having a low experienced teacher; 

\[ \hat{y}_{13} + (GNDR) \hat{y}_{14} = -.001 + (-1)(-1.35) = .134; \]

holding other factors constant. Using the random-coefficient regression model presented in Table 31 as the base model, approximately 23% of the variance in the academic self-efficacy differentiating effect was accounted once class gender, frequent use of traditional assessments, teaching experience, and interaction of class gender-by-teaching experience were taken into account. Finally, although student perceived public classroom assessment environment was on average positively related to performance-avoidance goal orientation; 

\[ \hat{y}_{20} = .071, t(1627) = 2.394, p < .05; \]

this relationship was invariant across classrooms.
CHAPTER V

DISCUSSION

This study utilized classroom assessment literature and achievement goal theory to examine the effects of teachers’ classroom assessment practices on ninth grade students’ perceptions of the classroom assessment environment and achievement goal orientations in Muscat science classrooms in Oman. The study first sought to identify the underlying dimensions of students’ perceptions of the classroom assessment environment and achievement goal orientations utilizing factor analytic techniques. The study then tested the effects of certain student-level and class-level characteristics on student perceived classroom assessment environment and achievement goal orientations utilizing hierarchical linear modeling (HLM) techniques.

This chapter is composed of three sections. In the first two sections, findings from this study pertaining to classroom assessment environment and achievement goal orientations are discussed and compared with findings from previous studies along with their implications for theory, research, and practice. The third section is a conclusion of the study in which limitations of the present study findings and recommendations for future research are discussed.
Theoretically, some educational perspectives in the classroom assessment literature (e.g., Ames, 1992a, 1992b; McMillan & Workman, 1998) have tended to structure classroom assessment environment around two dimensions: learning-oriented and normative-oriented described more fully in Chapters II and III of the dissertation. Empirically, the findings from this study showed that the participating ninth grade students’ perceptions of the classroom assessment environment in Muscat science classrooms in Oman centered around three facets: learning-oriented, harsh-oriented, and public-oriented. The learning-oriented assessment environment focused on classroom assessment practices that enhance student learning and mastery of content materials such as asking students moderately challenging assessment tasks, communicating with them clear assessment standards and criteria, giving them opportunities to improve their performance, and providing them informative assessment feedback. The harsh-oriented assessment environment focused on harshness of assessment and grading by providing students difficult and less meaningful assessment tasks with unattainable assessment standards and criteria. The public-oriented assessment environment emphasized the importance of grades rather than learning and focused on public rather than private evaluation and recognition practices by normatively comparing and socially recognizing student performance. Thus, it appears that in this study the theoretically expected factor of normative-oriented assessment environment was bifurcated into two dimensions: harsh- and public-oriented assessment environments. These findings parallel those of
previous studies exploring students’ perceptions of their classroom environment (e.g., Wang, 2004).

Furthermore, the perceived learning classroom assessment environment correlated negatively with both the perceived harsh and the perceived public classroom assessment environments, $r_s = -.25$ and -.12, respectively. The perceived harsh and the perceived public classroom assessment environments were positively related to each other, $r = .21$, suggesting that these two types of environments may share some aspects of the classroom assessment practices. The magnitude ($r_s$ ranged from .12 to .25) and direction of the zero-order correlations among the scores from the three scales of perceived classroom assessment environment found in this study tended to be similar to those reported by Church et al. (2001, $r_s$ ranged from .23 to .32) and Wang (2004, $r_s$ ranged from .05 to .23) utilizing different measures of conceptually related constructs.

As discussed in Chapter III of the dissertation, the classroom assessment literature was consulted to develop items for the perceived classroom assessment environment scales and as such these scales could be considered as relatively new to the research context of the present study. Nunnally (1967) suggested that for early-stage research with new measures, “reliabilities of .60 or .50 will suffice” (p. 226). However, when comparing the score reliability coefficients for the perceived classroom assessment environment scales ($\alpha$ ranged from .45 to .63) found in this study to those reported in studies conducted by Church et al. (2001, $\alpha$ ranged from .57 to .88) and Wang (2004, $\alpha$ ranged from .57 to .82), the reliability coefficients from the present study seem to be low. As might be expected, one direct consequence of this situation is the underestimation of
the true relationships between the variables (Thompson, 2003). There are a number of possible explanations for the low reliability coefficients found in this study when considering those found in the previous studies using scales for conceptually similar constructs.

One possible explanation involves the issue that reliability is a function of the scale as well as of the sample composition and variability (Dawis, 1987). With regard to the sample composition and variability, previous studies in this research area (e.g., Church et al., 2001; Wang, 2004) sampled a wider population of college undergraduates with diverse academic backgrounds; whereas the current study sampled a narrower population of ninth grade students with similar academic backgrounds. As a consequence of this sample homogeneity, the variance among the present study participants’ true scores would be low, and so would the reliability coefficients (Crocker & Algina, 1986, p. 144).

It should be noted however that according to the classical measurement theory, reliability estimates are useful for interpreting the score of an individual participant (Crocker & Algina, 1986). The accuracy of the individuals’ scores is represented by standard error of measurement (SEM) defined as “the standard deviation of discrepancies between a typical examinee’s true score and observed scores over an infinite number of repeated testings” (Crocker & Algina, 1986, p. 150). A large SEM indicates that there are large discrepancies between an individual participant’s true score and the observed scores over repeated administrations of the scales (Crocker & Algina, 1986). Therefore, although the score reliability coefficients for the perceived classroom assessment
environment scales used in this study tended to be lower than those used in previous studies (e.g., Church et al., 2001; Wang, 2004), the corresponding standard errors of measurement (SEMs ranged from 1.45 to 2.30) found in the present study were approximately equal to those found in the previous studies conducted by Church et al. (2001, SEMs ranged from 1.64 to 2.31) and Wang (2004, SEMs ranged from 1.16 to 2.58), thereby confirming the classical measurement theory that a lower reliability estimate does not necessarily imply little accuracy around participants’ scores (Crocker & Algina, 1986).

With regard to the scale itself, previous studies in classroom assessment environment were based on either 5-point scales (e.g., Wang, 2004) or 7-point scales (e.g., Church et al., 2001) whereas the current study was based on 4-point scales. The reason behind using 4-point scales in this study was that it was assumed that ninth grade students might not be able to meaningfully discriminate among more response options (5 or more) and simultaneously handle them when responding to the scale items. Hence, the study followed Assor and Connell’s (1992) recommendation that scales of 4-points provide more valid information on self-report measures designed for elementary, middle, and high schools students. However, having fewer response options might have reduced the variability and as a consequence might have contributed to the low score reliability. Future research might consider increasing the number of items or the number of response options within items in order to increase the variability (DeVellis, 2003).

Moreover, Crocker and Algina (1986) stated that “the internal consistency coefficient is an index of both item content homogeneity and item quality” (p. 135).
Based on the content validation process described in Chapter III of the dissertation, the initial pool of items were judged to be adequately representing the content domain of the perceived classroom assessment environment construct. However, the reliability data in the present study seem to suggest that the participating students did not respond consistently across the items. One possible reason for this situation could be that some items were not well written or that the instructions were not clear, causing students to misinterpret the items and respond on some basis unrelated to the content being measured, thereby lowering the internal consistency (Crocker & Algina, 1986). As suggested by Crocker and Algina (1986), one way therefore to make scale scores more reliable is by writing the items clearly and making the instructions easily understood.

Furthermore, another possible explanation for the low internal consistency estimates found in this study for the scores from the perceived classroom assessment environment scales involved the problem of inapplicable items. Waller (1989) has noted that self-report scales sometimes contain items that are inapplicable for some participants in specific contexts thereby affecting psychometric properties of the scales. In this study, the items for the perceived classroom assessment environment scales were based on theory and research developed in the United States and published in English-language journals. As might be expected, some items may not operate in contexts with native Arabic speaking participants, thereby making the content of the items less relevant, and as such a small number of inapplicable item responses might have affected the internal consistency.
Scale score validity refers to the usefulness of inferences drawn from the scale scores (Crocker & Algina, 1986, p. 238). As discussed in Chapter III of the dissertation, content validity concerns the extent to which the scale items adequately represent the content domain of the construct being measured, and it could be established by asking a group of experts to judge whether the items adequately sample the content domain of the construct being measured (Crocker & Algina, 1986). Construct validity refers to the extent to which the scale measures the construct it was designed to measure, and it could be established through correlational and factor analyses (Allen & Yen, 2002; Crocker & Algina, 1986). As such, combined with results of the zero-order correlations and principal components/exploratory factor analyses conducted in this study, the review of findings from the present study and past studies indicates that the scores from the perceived classroom assessment environment scales demonstrate reasonable levels of construct validity. Hence, these scales may prove to be useful tools in helping teachers to identify classroom assessment practices targeted at enhancing student learning.

The results of this study also 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). First, 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 across classrooms. On one hand, these findings tend 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 effects on students (Alsarimi, 2000). On the other hand, the findings of the present study highlight the shared common experience and thus perception of students within the same class about their teacher’s classroom assessment practices. Therefore, the implication of the findings from the current study 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).

Next, the results of this study provided empirical evidence that the variation between classrooms in student perceived classroom assessment environment could be explained by student characteristics, class contextual features, and teacher characteristics and assessment practices. Specifically, students’ perceptions of their classroom assessment environment as being learning-oriented were significantly shaped by student characteristics such as academic self-efficacy and class characteristics such as class average academic self-efficacy, class average perceived harsh classroom assessment
environment, and interaction of class gender by teacher’s frequent use of alternative
easessments. Likewise, students’ perceptions of their classroom assessment environment
as being harsh-oriented were significantly influenced by student characteristics such as
academic self-efficacy and class characteristics such as class averages perceived learning
and perceived public classroom assessment environment as well as interaction between
student characteristics like academic self-efficacy and class characteristics like class
gender and teacher’s teaching experience and frequent use of alternative assessments.
Also, students’ perceptions of their classroom assessment environment as being public-
oriented were primarily influenced by class characteristics such as class average
perceived harsh classroom assessment environment and interaction of class gender by
teacher’s frequent use of alternative assessments. Taking together, these findings
theoretically support educational perspectives of Ames (1992a, 1992b) and McMillan and

The present study not only showed that student academic self-efficacy was
significantly related to perceived classroom assessment environment, but also revealed
that this relationship varied significantly across classrooms. The findings that academic
self-efficacy was positively related to perceived learning classroom assessment
environment and negatively related to perceived harsh classroom assessment
environment seem plausible and agree with previous research findings (e.g., Brookhart &
Bronowicz, 2003; Brookhart & DeVoge, 1999). According to the social cognitive theory,
high efficacious students tend to persist in the face of difficulty, seek moderately
challenging learning situations, and view failures as learning opportunities (Bandura,
1986; Deemer, 2004). All these aspects are theoretically consistent with the perceived learning classroom assessment environment found 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 (Bandura, 1986; Deemer, 2004). All these aspects are theoretically consistent with the perceived harsh classroom assessment environment found 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 perhaps 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 relation between student self-efficacy and perceived classroom environment (e.g., Anderman & Midgley, 1997); it may be insensible to make causal inferences from the correlational nature of these results. Reciprocal causal relations between these two constructs may happen in the long run. As such, future research should be conducted to further examine the relationship between perceptions of the classroom assessment environment and self-efficacy.

Nevertheless, the present study findings extend 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 classroom assessment environment. For example, the findings of this study showed that
in male classrooms where alternative assessments were more prevalent, the negative
effect of academic self-efficacy on perceived classroom assessment environment tended
to be weakened in classes having a low experienced teacher than in classes having a high
experienced teacher. The opposite was found true in female classrooms where alternative
assessments were more prevalent. In light of classroom assessment literature, alternative
assessments are designed to foster meaningful, productive, and challenging learning
environments (Hargreaves, Early, & Schmidt, 2002; Shepard, 2000). In light of self-
efficacy theory and research, teachers with higher self-efficacy were more likely to
develop challenging instructional activities for their students and as a result they have
positive classroom environments (e.g., Ashton & Webb, 1986; Bandura, 1997;
Daugherty, 2005; Deemer, 2004). Therefore, the sampled high experienced female
teachers and low experienced male teachers in this study might have high levels of
teaching efficacy, defined as strong beliefs about their capabilities to help students learn
(Guskey & Passaro, 1994), which in turn might have been responsible for using
alternative assessments more frequently, and thereby having positive classroom
environments, which perhaps could diminish the potential negative effects of student
academic self-efficacy on perceived classroom assessment environment. The implications
of these findings for theory and practice are that given that the classroom assessment
environment is often structured by the teacher (Brookhart, 1997a) 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 teachers’ teaching experience and self-beliefs and goals for students’ learning (Ames, 1992b) as well as their views about alternative assessment methods (Snow-Renner, 1998).

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, whereas the higher levels of perceptions about classroom assessment environment as being public-oriented were salient for male students in classes using alternative assessments more 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.

**Achievement Goal Orientations**

As expected, the findings from factor analytic techniques in this study showed the prevalence of three types of achievement goal orientations for the participating ninth grade students in Muscat science classrooms in Oman: mastery, performance-approach, and performance-avoidance goal orientations. The mastery goal orientation focused on the development of competence. The performance-approach goal orientation focused on the demonstration of competence to others. The performance-avoidance goal orientation focused on avoiding the demonstration of incompetence to others. Therefore, from a theoretical point of view, these findings add support to the trichotomous conceptualization of achievement goal theory (Elliot & Church, 1997; Midgley et al., 2000).

However, the present study findings differ from those of previous research in terms of inter-correlations among the goals. In this study, the mastery, performance-approach, and performance-avoidance goals all correlated positively with each other ($r$s ranged from .32 to .45); whereas in both studies conducted by Kaplan et al. (2002a) and Middleton and Midgley (1997) using the same scales for middle school students in the United States, the mastery and performance-avoidance goals were not correlated with each other. Thus, as indicated by Senko and Harackiewicz (2005), the participating students in the current study who strongly adopt any one goal tended to adopt the other goals to a modest degree as well. From a practical point of view, Black and Wiliam
(1998a) stated that “the ecological validity of studies is clearly important in determining the applicability of the results to normal classroom work” (p. 8). Accordingly, unlike previous studies (e.g., Barron & Harackiewicz, 2001), the current study has some ecological validity in the sense that it presents findings related to the impact of a real classroom assessment context, which has become the normal practice carried out by the students’ usual teacher without imposing any artificial experimental conditions. Specifically, the positive correlations found in this study among achievement goal orientations seem to point to the reality of achievement settings in that classrooms often provide students opportunities to pursue more than one goal (Wentzel, 1992). Therefore, the participating teachers in this study seem to expect their classes to not only master the learning materials, but also to achieve higher grades than other classes.

Although the reliability estimates for the scores on the achievement goal orientation scales (α ranged from .54 to .75) found in this study were lower than those reported by Midgely et al. (2000) for the original versions of the scales (α ranged from .74 to .89), the corresponding standard errors of measurement in this study (SEMs ranged from .22 to .47) were approximately similar to their original versions (SEMs ranged from .34 to .35) reported in Midgley et al.’s (2000) study. Therefore, in accord with classical measurement theory, this suggests that a lower reliability estimate does not necessarily mean that there is less accuracy around individuals’ scores (Crocker & Algina, 1986). One possible explanation for the difference in score reliability for achievement goal orientation scales used in this study and those used in Midgley et al.’s (2000) study is the reduction in the scales’ anchors. In this study, achievement goal orientations were
measured using a 4-point scale following Assor and Connell’s (1992) recommendation when using self-report measures designed for elementary, middle, and high school students. In contrast, Midgley et al. (2000) measured achievement goal orientations using a 5-point scale. The reduction in the number of response options might have reduced the variability and as a consequence might have contributed to the lower reliability estimates (DeVellis, 2003).

Another possible explanation for the low internal consistency estimates found in this study for the scores from the achievement goal orientation scales when compared to their original versions (Midgley et al., 2000) was that the items of these scales 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 into Arabic was verified by bilingual professors and that the translated items were subjected to a content validation process described in Chapter III of the dissertation, some of the items and/or instructions might not have been clear, causing students to misinterpret the items and respond on some basis unrelated to the content being measured, thereby lowering the internal consistency of the responses (Crocker & Algina, 1986).

To sum, the present study represents an attempt to testify the applicability of the achievement goal theory in a culture that is different from where it was originally developed. Although the datametric properties in terms of zero-order correlations and score reliability of the achievement goal orientation scales for the ninth grade students in Oman contrasted to those reported for middle school students in the United States, the general components of the trichotomous framework of achievement goal theory (Midgley
et al., 2000) seem to apply equally in both cultures. Therefore, reliability generalization studies (Vacha-Haase, 1998) might need to be conducted in the future to empirically examine the factors that might influence score reliability for achievement goal orientation scales across diverse samples.

Of primary importance in this research study was the final hierarchical linear models that were constructed for each achievement goal orientation. First, the study showed significant variations among classrooms in achievement goal orientations as well as in the effect of certain student characteristics such as academic self-efficacy on achievement goal orientations. These findings support the theoretical perspectives that achievement goals and academic self-efficacy might vary across environmental situations (Bandura, 1977, 1982; Maehr, 1983, 1984). Second, the findings from this study suggested that student characteristics like academic self-efficacy and perceived harsh classroom assessment environment as well as class characteristics like class gender, class average academic self-efficacy, and class average perceived learning classroom assessment environment provided the best model for explaining differences in mastery goal orientation. With respect to performance-approach goal orientation, the findings indicated that student characteristics like academic self-efficacy and perceived public classroom assessment environment as well as class characteristics like class gender, class average academic self-efficacy, class average perceived public classroom assessment environment, and teacher’s teaching experience and frequent use of alternative assessments provided the best model for explaining differences in performance-approach goal orientation. Furthermore, the findings revealed that student characteristics like
academic self-efficacy and perceived public classroom assessment environment as well as class characteristics like class gender, class average academic self-efficacy, and teacher’s teaching experience and frequent use of traditional assessments provided the best model for explaining differences in performance-avoidance goal orientation. Taking together, these findings support many of the educational perspectives, described in Chapter II of the dissertation, regarding classroom assessment and student achievement motivation (e.g., Ames, 1992a, 1992b; Brookhar, 1997a; McMillan & Workman, 1998).

Given that both mastery and performance-approach goal orientations represent forms of approach achievement motivation (Elliot, 1999), it was not surprising that academic self-efficacy in this study was positively related to mastery and performance-approach goals both at the student-level and class-level. These findings not only are consistent with previous studies in academic self-efficacy (e.g., Greene at al., 2004; Kaplan et al., 2002a), but also confirm Bandura’s (1986) social-cognitive theory and Elliot’s (1999) review of achievement goal theory regarding the relation between self-efficacy and approach achievement goals. As stated by Midgley et al. (1998), “construct validity is based on the degree to which the goal orientation scales…are associated with other constructs in ways that are predicted by theory and in ways that are consistent with other research” (p. 119). Therefore, the present study findings provide evidence of the construct validity for the scores on the mastery and performance-approach goal orientation scales used in the study.

In addition, the current study findings showed that the effect of academic self-efficacy on mastery goal orientation was stronger in male classrooms than in female
classrooms. This means that male classrooms tended to be less egalitarian with regard to student’s mastery goal orientation than female classrooms. Stated somewhat differently, male classrooms tended to widen the gap between high and low efficacious students with regard to mastery goal orientation, whereas female classrooms seemed to narrow this gap. When compared to male classrooms, the feelings of more academic support and nurturance from the teacher in female classrooms (Covington & Dray, 2002) may be responsible for narrowing the gap in student mastery pursuits.

What was surprising in this study finding was the positive relationship between academic self-efficacy and performance-avoidance goal orientation at both levels of the hierarchy, student and class. Many studies reported negative relationships between academic self-efficacy and performance-avoidance goal orientations (e.g., Kaplan et al., 2002a; Middleton & Midgley, 1997). Given that performance-avoidance goal orientation is a form of avoidance motivation (Elliot, 1999) and that low academic competence perceptions lead to avoidance motivation (Elliot, 1999; Greene et al., 2004), why then high efficacious students in this study strive to avoid the demonstration of lack of academic competence? The findings from the current study suggest that teacher’s frequent use of traditional assessments may be one possible explanation for the positive relationship between academic self-efficacy and performance-avoidance goal orientation. The findings showed that teacher’s frequent use of traditional assessments tended to moderate the effect of student academic self-efficacy on performance-avoidance goal orientation. 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 academic self-
efficacy on performance-avoidance goal orientation 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, for the most part, convey to students that effort in class work should be in response to avoid the demonstration of incompetence to others. Therefore, results of the present 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 (Elliot, 1999) in classes where traditional assessments are more prevalent.

In addition, 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., Ablard & Lipschultz, 1998; Middleton & Midgley, 1997) and that teacher’s teaching experience is negatively related to student’s adoption of goal orientations 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 utilizing hierarchical linear modeling analyses that not only took the nested structure of the data into account, but also estimated the relations that crossed the class-level and the student-level characteristics. First, the present study findings revealed a similar pattern for gender differences in mastery goal orientations, but at the classroom-level, in that female classrooms tended to have higher levels of mastery goal orientations
than male classrooms. The present study findings, however, differed from previous research findings with regard to gender differences in performance goal orientations, in that female classrooms tended to be more performance-oriented than male classrooms. The difference between this study and previous studies (e.g., Ablard & Lipschultz, 1998; Deevers, 2005; 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 focus on competitive success. Classroom observations and interviews might shed more light on gender differences in achievement goals between segregated and desegregated classrooms.

Second, the current study findings revealed that the positive effect of student academic self-efficacy on performance-avoidance goal orientation depended jointly on class gender and teacher’s teaching experience. For example, male classes having a low experienced teacher tended to weaken the positive effect of academic self-efficacy on performance-avoidance goal orientation whereas male classes having a high experienced teacher tended to strengthen this effect. The opposite was found true in female classrooms. There is evidence from self-efficacy research (e.g., Daugherty, 2005; Deemer, 2004; Herman, 2001; Midgley, Feldlaufer, & Eccles, 1989) that teacher’s sense of efficacy could influence student motivation-related beliefs and behaviors. Therefore,
drawing upon self-efficacy research, the sampled low experienced male teachers and high experienced female teachers in this study might have high levels of teaching efficacy, which in turn might have been responsible for weakening the positive effect of student academic self-efficacy on performance-avoidance goal orientation. These findings demonstrate how important teacher’s experience and self-beliefs are for desirable student motivation-related outcomes.

The current study findings revealed that there was a trend for teacher’s frequent use of alternative assessments to correlate negatively with performance-approach goal orientation; suggesting that a strong emphasis on alternative assessments may less likely to orient students toward the adoption of performance-approach goals. In light of classroom assessment literature, alternative assessments tended to be more authentic, engaging, challenging, and emphasizing deep rather than superficial learning (Baker et al., 1993; Darling-Hammond et al., 1995; Hargreaves et al., 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 the matters more, the differentiating effect of academic self-efficacy on performance-approach goal orientation within a classroom was found in this study to be dependent jointly on class gender, teacher’s teaching experience, and frequent use of
alternative assessments. For example, in female classrooms where alternative assessments were more prevalent, a high experienced teacher tended to narrow the gap in performance-approach goal orientations between high and low efficacious 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, 2002), 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 teacher’s teaching experience and frequent use of certain assessment practices on student motivation to learn.

According to Brookhart’s (1997) theoretical model of classroom assessment and student motivation and achievement, students’ perceptions of the classroom assessment environment influence their achievement motivational beliefs. The current study findings showed that student perceived public-oriented classroom 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 classroom environment and achievement goals (e.g., Brookhart & DeVoge, 1999; Brookhart &
Durkin, 2003; Church et al., 2001; Wang, 2004). Ames (1992b) 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 emphasizing the importance of grades rather than learning and focusing on public rather than private evaluation and recognition practices on student achievement motivation.

Conclusion

Classroom assessment is a continual activity for teachers to improve the quality of instruction and motivate students to learn (Brookhart, 1999; Gronlund, 2006). Although there is a great deal of research on teachers’ classroom assessment practices, few empirical research attempts have been made to link teacher’s assessment practices to students’ perceptions of classroom assessment environment and motivation defined in terms of achievement goal orientations. When this issue was considered, little attention was devoted to the hierarchical structure of the data (e.g., Church et al., 2001; Wang, 2004). This study attempted to fill this gap by investigating the possible effects of teachers’ assessment practices on ninth grade students’ perceptions of classroom assessment environment and achievement goal orientations in Muscat science classrooms.
in Oman. Based upon the findings of this study, a number of recommendations for future research are suggested.

First, the results of this study tend to point to the conclusion that the effects of classroom assessment practices on student motivational beliefs are contextual and interact in unique ways with student and teacher characteristics thereby lending support to Rodriguez’s (2004) study. Therefore, the nested nature of students and classrooms needs to be continuously considered and extended in future research. Specifically, teachers’ classroom assessment practices are imbedded in a school, which is a formal organizational context (Natriello & Dornbusch, 1984), and as such future research may need to move beyond the classroom-level context and examine the potential influences of school-level policies, practices, and contextual features on teachers’ assessment practices and students’ motivational beliefs in a three-level hierarchical linear model.

Second, from a research perspective, the associations among student characteristics, class contextual features, and teacher characteristics and assessment practices revealed in this study are intriguing and call further exploration. For example, it may not be the assessment format whether being traditional or alternative per se that influences student perceptions of the classroom assessment environment and achievement goal orientations, but the assessment format in interaction with other factors such as teaching experience and efficacy. Future research needs to do more work to fully understand the effects of assessment format on student motivational beliefs. A qualitative approach may reveal a deeper understanding of the phenomenon from the perspectives of both students and teachers. Also, the findings underpin the importance for future research
to examine not only gender differences but also other group differences in patterns of achievement goals and perceived classroom assessment environment to identify which factors are facilitative for different groups of students. For example, certain classroom assessment practices may make different impacts to high versus low achieving students (Ames, 1992b; Black & Wiliam, 1998a), and as such a close attention may need to be given in future research to the differential effects of student’s prior academic achievement.

Moreover, the present study findings suggest that the collective (i.e., aggregate) perceived classroom assessment environment is operative in the assessment settings and as such it may be important to begin attending to this construct in the future empirical work on classroom assessment. In accordance with the social theory (Coleman, 1990), the findings from this study revealed that when the perceptions of an individual student about the assessment environment are incongruent with the shared perceptions of the class members, the student’s perceptions tended to be weakened. In addition, from a socio-cognitive perspective, students “are not social isolates of the influence of those around them” (Bandura, 1997, p. 469). In this particular regard, the present study findings showed that the shared perceptions of the class members about the assessment environment might influence student’s 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/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. These findings have the following implication for practice.

Specifically, the findings from this study 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 findings is that teachers might conduct periodic 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 the teacher make classroom assessment more motivating for students to learn.

However, a conceptual bias (Pedhazur, 1997) might have affected the results of this study in that the measurement of perceived classroom assessment environment at the classroom-level was based on a convenience basis. Specifically, the underlying dimensions of perceived classroom assessment environment may not be the same at both levels of the analysis, student and classroom. Future research may need to examine the underlying dimensions of perceived classroom assessment environment at the classroom-level using a larger sample size considered adequate for factor analytic techniques.

The results of this study might have also been affected by another systematic bias concerning students’ reports of their perceptions of the classroom assessment environment. When asking students about their classroom assessment environment, the
questionnaire items (e.g., the science tests in this class are difficult to students) did not specify an evaluative framework (i.e., a reference) that might have helped students make more accurate appraisals of their classroom assessment environment. A greater accuracy might have been obtained if students were given the opportunity to communicate their perceptions of the science classroom assessment environment with respect to a specific reference such as in relation to other academic subjects, teachers, or time of the academic year. Clearly framed questionnaire items with a suitable evaluative framework may help students make more accurate judgments and hence may provide more useful information.

In addition, the current study focused on the possible influences of the classroom assessment context on student achievement goal orientations. Two sources of information were used to examine the assessment context of the classroom. First, the perceived classroom assessment environment was measured by aggregating students’ perceptions of their science class assessment climate. Second, the science teachers were surveyed about their approaches to assessment. Although both student and teacher reports of the classroom assessment context are informative, they may not provide the complete picture yet. On one hand, what students report regarding their perceptions of the classroom assessment environment may not be the same as the perceptions and beliefs they make at the time they are actually engaging in an assigned classroom assessment task (Assor & Connell, 1992). On the other hand, teachers’ reports of their assessment practices may not always match their actual practices because they may answer survey questions in socially desirable ways (Ryan et al., 1998). Therefore, classroom observations and interviews
with students and teachers are critical in understanding how classroom assessment context influences student motivational beliefs and behaviors.

In this study, there was no evidence of teacher’s frequent use of recommended assessment practices’ effects on any dimension of the perceived classroom assessment environment and achievement goal orientations. One possible explanation for this is that advocate of educational assessment experts for certain classroom assessment practices may be to obtain a technical valid indication of student academic achievement rather than to have a direct impact on students’ perceptions of the classroom assessment environment and achievement goal orientations. Another possible explanation involves the assumption underlying the current investigation. Specifically, the present study was based on the assumption that aspects of the classroom assessment are interdependent, operating in an additive manner, and as such the study followed an integrative approach (Ames, 1992b) in the investigation of the effects of the recommended assessment practices on student perceived classroom assessment environment and achievement goal orientations. In other words, one limitation of this study was that it construed teacher’s frequent use of the recommended assessments as an omnibus combination of several factors including revision of assessments, 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 student perceived classroom assessment environment and achievement goal orientations. Future research may need to independently consider each aspect of the recommended assessment practices.
Compared with findings from past studies using conceptually similar scales, the present study findings might have been affected by the low reliability estimates for the scores from the students and teachers’ scales. In addition to the composition of the sample, the reduction in the number of response options from five and/or seven points to four points might have reduced the variance and as a result might have contributed to the lower score reliability for the students’ scales in this study (DeVellis, 2003). The difference between reliability data of this study and past studies using similar scales measuring teachers’ frequent uses of traditional and alternative assessments (e.g., Alsarimi, 2002; Bol et al., 1998) is that the present study sampled a much narrower population of teachers, who were Muscat ninth grade science teachers in Oman, whereas the other studies mixed grade levels, subject areas, and geographic regions. As a result, a restriction of the scores’ range might have occurred and reduced the reliability coefficients in this study (Crocker & Algina, 1986). Therefore, the homogeneous nature of the present study’s sample might need to be taken into account when interpreting the results. Future research increasing the number of scales’ items or response options, sampling a more heterogeneous group of individuals, or using a different procedure for estimating the reliability might enhance the reliability estimates and effect sizes. Another line of research might consider doing reliability generalization studies (Vacha-Haase, 1998) to empirically examine the possible factors that influence score reliability for the scales across diverse samples.

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 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. Also, 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.
APPENDICES
Student’s Questionnaire

First: To what extent each of the following items is true or not true for your ninth grade science class?

أولاً: ما مدى صحة أو عدم صحة العبارات الآتية بالنسبة لصفك في مادة العلوم؟

<table>
<thead>
<tr>
<th>Assessment task</th>
<th>Completely not true</th>
<th>Somewhat not true</th>
<th>Somewhat true</th>
<th>Completely true</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The science tests in this class are difficult to students.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2. In this class, teacher’s oral questions encourage thinking.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3. In this class, the homework is boring.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4. In this class, the assignments and activities are related to students’ every day lives.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5. In this class, the tests match what we learn in class.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6. In this class, the teacher uses more than one way to determine grades such as tests, projects…etc.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Assessment Feedback</td>
<td>Completely not true</td>
<td>Somewhat not true</td>
<td>Somewhat true</td>
<td>Completely true</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------------------</td>
<td>------------------</td>
<td>---------------</td>
<td>----------------</td>
</tr>
<tr>
<td>1. In this class, I can find out my strengths in science.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2. In this class, I can find out my weaknesses in science.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3. In this class, my teacher helps me identify the places where I need more effort in future.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4. In this class, my teacher encourages viewing mistakes as learning opportunities.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5. In this class, students are given a chance to correct their mistakes.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6. In this class, I receive continuous feedback about my performance from the teacher.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7. In this class, the teacher wants us to take responsibility for our learning.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Assessment Feedback (cont.)</td>
<td>Completely not true</td>
<td>Somewhat not true</td>
<td>Somewhat true</td>
<td>Completely true</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------------------</td>
<td>-------------------</td>
<td>--------------</td>
<td>-----------------</td>
</tr>
<tr>
<td><strong>9. In this class, assessment results are a fair representation of student learning.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. تعكس درجات مادة العلوم بصدق ما تعلم الطلاب بالفعل في المادة.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td><strong>10. In this class, assessment results fairly reflect the effort I have put in studying science.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. تعكس درجات مادة العلوم بصدق مقدار الجهد الذي بذلته في دراسة المادة بالفعل.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td><strong>11. In this class, my teacher compares my performance with performance of other students.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. يقارن معلم مادة العلوم مستوى أدائي في المادة بمستوى أداء الطلاب الآخرين في الصف.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td><strong>12. In this class, assignments and tests are returned in a way that keeps individual student scores private.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. يقوم معلم مادة العلوم بإعادة الواجبات والأمتحانات بعد التصحيح إلى الطلبة بطريقة تحافظ على سرية درجات كل طالب.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td><strong>13. In this class, students who do well are praised in front of the whole class.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. يمذج معلم مادة العلوم الطلاب المتفوقين في المادة أمام جميع طلبة الصف.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td><strong>14. In this class, students who do poorly are criticized in front of the whole class.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. ينتقد معلم مادة العلوم الطلاب الضعفاء في المادة أمام جميع طلبة الصف.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Assessment standards and criteria</td>
<td>Completely not true</td>
<td>Somewhat not true</td>
<td>Somewhat true</td>
<td>Completely true</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-----------------</td>
<td>--------------</td>
<td>---------------</td>
</tr>
<tr>
<td><strong>1. It is difficult to achieve high grades in this class.</strong></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>1. يصعب الحصول على درجات مرفعة في مادة العلوم.</td>
<td><strong>2. In this class, only a few students can get high grades.</strong></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2. طلاب قليلون فقط في هذا الصف يستطيعون الحصول على درجات مرفعة في مادة العلوم.</td>
<td><strong>3. In this class, my teacher’s grading is clear.</strong></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3. طريقة معلم مادة العلوم في تقدير الدرجات واضحة.</td>
<td><strong>4. In this class, the grades are used to penalize students who disturb the class.</strong></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4. يستخدم معلم مادة العلوم الدرجات كأداة لمعاقبة الطلبة المشاغبين.</td>
<td><strong>5. In this class, the tests and assignments are used to control students’ behavior.</strong></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5. يستخدم معلم مادة العلوم الأوراق والواجبات كأداة لضبط سلوك الطلبة داخل الصف.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Second: To what extent each of the following items is true or not true for you in this science class?

ثانياً: ما مدى صحة أو عدم صحة العبارات الآتية بالنسبة لك في مادة العلوم للصف التاسع؟

<table>
<thead>
<tr>
<th>Mastery-goal orientation</th>
<th>Completely not true (خطأ تمامًا)</th>
<th>Somewhat not true (خطأ إلى حد ما)</th>
<th>Somewhat true (صحيحة إلى حد ما)</th>
<th>Completely true (صحيحة تمامًا)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. It is important to me that I learn a lot of new science concepts this semester.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2. One of my goals in science class is to learn as much as I can.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3. One of my goals is to master a lot of new science skills this semester.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4. It is important to me that I thoroughly understand my science class work.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5. It is important to me that I improve my science skills this semester.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Performance-approach goal orientation</td>
<td>Completely not true</td>
<td>Somewhat not true</td>
<td>Somewhat true</td>
<td>Completely true</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>--------------------</td>
<td>------------------</td>
<td>---------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>1. It is important to me that other students in my class think I am good at my science class work.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2. One of my goals is to show others that I’m good at science class work.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3. One of my goals is to show others that science class work is easy for me.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4. One of my goals is to look smart in comparison to the other students in my class.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5. It is important to me that I look smart compared to others in my class.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Performance-avoidance goal orientation</td>
<td>Completely not true</td>
<td>Somewhat not true</td>
<td>Somewhat true</td>
<td>Completely true</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>--------------------</td>
<td>------------------</td>
<td>--------------</td>
<td>----------------</td>
</tr>
<tr>
<td>1. It is important to me that I don’t look stupid in science class.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2. One of my goals is to keep others from thinking I’m not smart in science class.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3. It is important to me that my teacher doesn’t think that I know less than others in science.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4. One of my goals in science class is to avoid looking like I have trouble doing the work.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Academic self-efficacy</td>
<td>Completely not true</td>
<td>Somewhat not true</td>
<td>Somewhat true</td>
<td>Completely true</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------------------</td>
<td>-------------------</td>
<td>---------------</td>
<td>----------------</td>
</tr>
<tr>
<td>1. I’m certain I can master the skills taught in science class this semester.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2. I’m certain I can figure out how to do the most difficult class work in science.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3. Even if the work is hard, I can learn it.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4. I am sure about my ability to do the assignments in this class.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5. I am sure I have the ability to understand the ideas and skills taught in this class.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6. I am certain I can understand the material presented in this class.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
Third: Please provide the appropriate information for each of the following questions.

ثالثاً: بيانات عامة يرجى الإجابة عنها بما هو مناسب

1. What was your age on your last birthday? ----------------------- years.
   ١. ما هو عمرك الآن بالسنوات؟ ............. سنة.

2. What is your gender?             --------- Male.                      --------- Female.
   ٢. ما هو جنسك؟ ...... ذكر. ...... أنثى.

3. What is your nationality?      --------- Omani.                  --------- Non-Omani.
   ٣. ما هي جنسيتك؟ ...... عمانى. ...... غير عمانى.

Thank you for your participation

شكرا على مشاركتك
APPENDIX B

TEACHER’S QUESTIONNAIRE
Teacher’s Questionnaire
استبيان المعلم

First: To what extent you use each of the following assessment practices in your ninth grade science class?
أولاً: إلى أي مدى تستخدم الجوانب الآتية في عملية قياس وتقويم تعلم الطلبة في مادة العلوم للصف التاسع؟

<table>
<thead>
<tr>
<th>Type of assessment</th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. True-false test items.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2. Multiple-choice test items.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3. Matching test items.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4. Completion test items (i.e., fill in the blanks).</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5. Short-answer questions (e.g., word, phrase, label, formula).</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6. Extended short answer questions (e.g., one sentence to three sentences).</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>7. Oral exams.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>8. Essay questions (i.e., one paragraph or more).</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Type of assessment (cont.)</td>
<td>Never</td>
<td>Rarely</td>
<td>Sometimes</td>
<td>Often</td>
<td>Always</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-------</td>
<td>--------</td>
<td>-----------</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td>9. Research paper (i.e., one full page or more that involves finding resources).</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>11. Models (e.g., inventions, applying theory to something tangible).</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>12. Unstructured performance assessment (rating students’ performance without preset criteria).</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>13. Structured performance assessment (rating students’ performance with preset criteria).</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Assessment revision</td>
<td>Never</td>
<td>Rarely</td>
<td>Sometimes</td>
<td>Often</td>
<td>Always</td>
</tr>
<tr>
<td>5. Using a table of specifications to plan assessments.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2. Calculating central tendency measures (e.g., mean, mode, median) to describe test scores.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3. Calculating variability measures (e.g., range, standard deviation, variance) to describe test scores.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
### Assessment revision (cont.)

<table>
<thead>
<tr>
<th>Activity Description</th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. Conducting item analysis (e.g., item difficulty, item discrimination) for the tests.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5. Calculating a reliability coefficient for test scores.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6. Verifying content validity of the test.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

### Communicating assessment

<table>
<thead>
<tr>
<th>Activity Description</th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. Informing students about the purpose of assessment prior to its administration.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2. Providing oral assessment feedback to each student.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3. Providing written assessment feedback to each student.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4. Informing every student about his or her strengths in the assessment.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5. Providing students with suggestions of ways to improve their performance in science.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
Communicating assessment (cont.)

<table>
<thead>
<tr>
<th></th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. Protecting students’ confidentiality with regard to assessment results.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Praising high achieving students in front of the whole class.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Criticizing low achieving students in front of the whole class.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Returning assignments and tests to students in a way that keeps individual student scores private.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessment standards and criteria</td>
<td>Mahkatat Wamayyir al-Taqim</td>
<td>Never</td>
<td>Rarely</td>
<td>Sometimes</td>
<td>Often</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>----------------------------</td>
<td>-------</td>
<td>--------</td>
<td>-----------</td>
<td>-------</td>
</tr>
<tr>
<td>1. Constructing a model answer for scoring essay questions.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2. Informing students in advance how grades are to be assigned.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3. Using zeros in calculating grades for work not completed.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4. Defining a rating scale for performance criteria in advance.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5. Communicating performance assessment criteria to students in advance.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Student-involved assessment</td>
<td>Never</td>
<td>Rarely</td>
<td>Sometimes</td>
<td>Often</td>
<td>Always</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------</td>
<td>--------</td>
<td>-----------</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td>1. Engaging students in using grading criteria to evaluate strong and weak samples of class work.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2. Providing students opportunities to write test questions based on their understanding of the instructional objectives.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3. Allowing students to choose assessment activities they want to work in the class.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4. Providing students with systematic ways to monitor their learning progress.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Nonachievement-based grading factors</td>
<td>Never</td>
<td>Rarely</td>
<td>Sometimes</td>
<td>Often</td>
<td>Always</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-------</td>
<td>--------</td>
<td>-----------</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td>1. Incorporating student’s behavior in the classroom in the calculation of grades.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2. Incorporating student’s class attendance in the calculation of grades.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3. Incorporating student’s interest in learning the subject in the calculation of grades.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4. Incorporating student’s class participation in the calculation of grades.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5. Comparing student’s performance with other students in determining student’s grade.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6. Incorporating student’s neatness of work in the calculation of grades.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
Second: Please provide the appropriate information for each of the following questions.

ثانياً: بيانات عامه برجى الإجابة عنها بما هو مناسب.

1. How many years of teaching experience do you have? ........................ years.

   1. كم عدد سنوات خبرتك في التدريس؟ ........................ سنة.

2. What is your gender?            ---------- Male.                      --------- Female.

   2. ما هو جنسيتك؟ ........................ ذكر. ........................ امرأة.

3. What is your nationality?       --------- Omani.                  --------- Non-Omani.

   3. ما هي جنسيتك؟ ........................ عماني. ........................ غير عماني.

4. What is your educational qualification?

     --------- A bachelor in education.                  --------- An educational diploma.

   4. ما نوع المؤهل التربوي لديك؟ ........................ بكالوريوس تربية. ........................ دبلوم تأهيل تربوي.

Thank you for your participation

شكرًا على مشاركتك
APPENDIX C

OMANI MINISTRY OF EDUCATION’S PERMISSION
LETTER FOR DATA COLLECTION
الموضوع: تطبيق دراسة ميدانية

بالإشارة إلى رسالة الفاضلة المذكورة سابقاً بتسمية البلولي سيدر المكتب الفني للدراسات والتطوير بوزارة التربية والتعليم بشأن موضوع البحث/البحث عملي بين طالب
الخوصفي في مدرستكم و ذلك من خلال التعرف على أثر القياس والتفاوت الشامل على أهداف
التحصيل في مادة العلوم العامة لدى طلاب وطالبات الصف التاسع بدءاً من محافظة مسقط وذلك
استكمالاً لمنطقتين برنامج درجة الدكتوراه في القياس والتفاوت جامعات كتكة الحكومية بالولايات
المتحدة الأمريكية في أوهايو.

يرجى النجوم تسهيل مهمة البحث في مدرستكم.

شكراً لكم حسن تعاملكم متنا
وكلم جزيل الشكر

[ลาย دائرية]

339
APPENDIX D

CONSENT FORMS
Parent’s English Consent Form

Consent Form: 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

I want to do research on the 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. I want to do this because it is part of my Doctor of Philosophy degree requirements in Evaluation and Measurement at Kent State University, Kent, Ohio. I would like you to let your child take part in this project. If you decide to do this, your child will be asked to complete one questionnaire regarding his or her perception of the classroom assessment environment, achievement goals, and academic self-efficacy in science. The requested information will take no more than twenty minutes to complete.

Risks and discomforts associated with the project: There are no risks involved with the participation in this project beyond the risks normally encountered in everyday life.

Anonymity/confidentiality: Your child’s responses obtained from the questionnaire will be anonymous and will only be used for and reported in the statistical analysis with no connections made to your child.

If your child takes part in this project, then information gathered from this project are hoped to improve instruction and classroom assessment of your child’s learning in the science class. Taking part in this project is entirely up to you, and no one will hold it against your child if you decide not to do it. If your child does take part, he or she may stop at any time.

If you want to know more about this research project, please call me at [Hussain Al Kharusi, Tel. 001.330.389.9871; & Dr. Rafa Kasim, Tel. 001.330.672.0601]. The project has been approved by Kent State University. If you have questions about Kent State University’s rules for research, please call Dr. John L. West, Vice President and Dean, Division of Research and Graduate Studies (Tel. 001.330.672.2704).

You will get a copy of this consent form.

Sincerely,
Principal Investigator: Hussain Al Kharusi  Academic Advisor: Dr. Rafa Kasim

B. CONSENT STATEMENT(S)

I agree to let my child take part in this project. I know what he or she will have to do and that he or she can stop at any time.

_________________________________  _______________________
Signature                                    Date

Educational Foundations and Special Services
P.O. Box 5190 • Kent, Ohio 44242-0001
330-672-2294 • Fax 330-672-2512 • http://efss.educ.kent.edu
Parent’s Arabic Consent Form

ẩuاءة مرفقات المعلمين في القياس والقياس المنظور على إدراك طفية الصن في جمعة مدرسية مباشرة مفتوحة مفتوحة في سلطة عمان لعبة اللعب والقياس المنظور وأهداف التحديق أنهم في مادة العلم.

أرغب في القيام ببحث حول أثر ممارسات المعلمين في القياس والقياس المنظم على إدراك طفية الصن.

الم Collapse 1/100

في القياس من القياس والعسسح، أراد التحديق أنهم في مادة العلم. أريد

التلبات المتحدثة الأمريكية، أجزأ أن تجعل الكلاب/إنوكيل في هذا المشروع إذا قررت قبل ذلك، فسيطلب من

كلاب/إنوكيل إكمال استبيان واحد حول إدراك/ها لبيئة اللعب والقياس المنظم، وأهداف التحديق، واقعية الذات الأكاديمية

شاك/ها في العالم. أن تكون المعلومات المنظم في الاستبيان لا تتجاوز 20 دقيقة.

المخطط والملاحظات المصاحبة مع المشروع: أن المشاركة في هذا المشروع غير مصحوبة بأي مخطط تتجاوز

حدود أدنى المخطط التي تحدد بالنص من خلال الأدبية.

السرية: إن إجابات كل/إنوكيل على الاستبيان ستكون سرية، واستخدام وسائل قطع في التحليل الإحصائي

للدراسة بدون أن يكون فيها أي رابط يحتوي بإمكان.

إذا شارك كل/إنوكيل في هذا المشروع، فإننا نأمل أن تساعد المعلومات التي سنحصل عليها من هذا

المشروع على تحسين التدريس والقياس والقياس المنظم لعلم الكلاب/إنوكيل في مادة العلم. أن كل/إنوكيل، كامل الحرية في مشاركة

كل/إنوكيل في هذا المشروع، فإننا نأمل أن كل/إنوكيل، إذا تم حذام مشاركة/ها. إذا شارك كل/إنوكيل، فسنتعل يثبب التوقف عن المشاركة في أي وقت.

إذا تفسر هذه الدراسة برجي الاستدلال بـ حمض الأوزومي على هاتف رقم

(012), 330.672.06001, (001). أما بالنسبة لهذه الدراسة فقد تم

التصنيف على أنها من قبل جامعة كلب الحربية، إذا كانت كلب أي أي أسئلة حولliv & والبحث على هاتف رقم

الإجابة بالكلام إنك أو ويجب تذكير السياسة والتعليم في المشاكل وご覧 على هاتف رقم

(001), 330.672.2704.

╕

مثنسل على نسخة من إستمارة الموافقة هذه.

المشرف الأكاديمي: د. رافص قاسم

بالحروف العربية.

التوقيع

التاريخ

Educational Foundations and Special Services
P.O. Box 5190 • Kent, Ohio 44242-0001
330-672-2294 • Fax 330-672-2512 • http://elss.educ.kent.edu
Teacher's English Consent Form

Consent Form: 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.

I want to do research on the 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. I want to do this because it is part of my Doctor of Philosophy degree requirements in Evaluation and Measurement at Kent State University, Kent, Ohio. I would like you to take part in this project. If you decide to do this, you will be asked to complete one questionnaire regarding your classroom assessment practices in the ninth grade science class. The requested information will take no more than ten minutes to complete.

Risks and discomforts associated with the project: There are no risks involved with the participation in this project beyond the risks normally encountered in everyday life.

Anonymity/confidentiality: Your responses obtained from the questionnaire will be anonymous and will only be used for and reported in the statistical analysis with no connections made to you.

If you take part in this project, then information gathered from this project are hoped to improve instruction and classroom assessment of student learning in the science class. Taking part in this project is entirely up to you, and no one will hold it against you if you decide not to do it. If you do take part, you may stop at any time.

If you want to know more about this research project, please call me at [Hussain Al Kharusi, Tel. 001.330.389.9871; & Dr. Rafa Kasim, Tel. 001.330.672.0601]. The project has been approved by Kent State University. If you have questions about Kent State University’s rules for research, please call Dr. John L. West, Vice President and Dean, Division of Research and Graduate Studies (Tel. 001.330.672.2704).

You will get a copy of this consent form.

Sincerely,
Principal Investigator: Hussain Al Kharusi    Academic Advisor: Dr. Rafa Kasim

B. CONSENT STATEMENT(S)
I agree to take part in this project. I know what I will have to do and that I can stop at any time.

Signature    Date

Educational Foundations and Special Services
P.O. Box 5190 • Kent, Ohio 44242-0519
330-672-2294 • Fax 330-672-2512 • http://efsx.educ.kent.edu
 البحث العلمي.

البحث الرئيسي: حسن الخروصي

المشرف الأكاديمي: د. رافق قاسم

(1) تسرير الموقفة

أوافق على المشاركة في هذا المشروع. أُ'aff * ما ينوي على القيام به. كما أُنين أستلم النتائج من المشاركة في أي وقت.

التاريخ

Educational Foundations and Special Services
P.O. Box 5190 • Kent, Oh. 44242-0001
330-672-2284 • Fax 330-672-2512 • http://ebs.os.edu.kent.edu
Student's English Consent Form

Consent Form: 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

I want to do research on the 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. I want to do this because it is part of my Doctor of Philosophy degree requirements in Evaluation and Measurement at Kent State University, Kent, Ohio. I would like you to take part in this project. If you decide to do this, you will be asked to complete one questionnaire regarding your perception of the classroom assessment environment, achievement goals, and academic self-efficacy in science. The requested information will take no more than twenty minutes to complete.

Risks and discomforts associated with the project: There are no risks involved with the participation in this project beyond the risks normally encountered in everyday life.

Anonymity/confidentiality: Your responses obtained from the questionnaire will be anonymous and will only be used for and reported in the statistical analysis with no connections made to you.

If you take part in this project, then information gathered from this project are hoped to improve instruction and classroom assessment of your learning in the science class. Taking part in this project is entirely up to you, and no one will hold it against you if you decide not to do it. If you do take part, you may stop at any time.

If you want to know more about this research project, please call me at [Hussain Al Kharusi, Tel. 001.330.389.9871; & Dr. Rafa Kasim, Tel. 001.330.672.0601]. The project has been approved by Kent State University. If you have questions about Kent State University's rules for research, please call Dr. John L. West, Vice President and Dean, Division of Research and Graduate Studies (Tel. 001.330.672.2704).

You will get a copy of this consent form.

Sincerely,
Principal Investigator: Hussain Al Kharusi Academic Advisor: Dr. Rafa Kasim

B. CONSENT STATEMENT(S)
I agree to take part in this project. I know what I will have to do and that I can stop at any time.

Signature

Date

Educational Foundations and Special Services
P.O. Box 5190 • Kent, Ohio 44242-0001
330-672-2294 • Fax 330-672-2512 • http://efss.educ.kent.edu
المشاركة في هذا المشروع غير مصحوبة بالمخاطر المعتادة.

إذا اشتركت في هذا المشروع، فإنك توافق على المشاركة في هذه الدراسة.

إذا أقررت عدم المشاركة في هذا المشروع، يرجى إخبارنا.

المشرف الأكاديمي: خالد عبد الجواد
تحسينات: (030330.672) 06/01
ال контакт: (030330.672) 06/01

البحث: اسم الشريحة
رقم التوثيق: الرقم

التوقيع:

التاريخ:

KENT STATE UNIVERSITY
Student's Arabic Consent Form

Educational Foundations and Special Services
P.O. Box 5190 • Kent, Ohio 44242-0001
330-672-2294 • Fax 330-672-2512 • http://efss.educ.kent.edu

346
Teacher’s English Cover Letter

Dear Ninth grade science teacher

The attached questionnaire is designed to identify your classroom assessment practices in the ninth grade science class. The questionnaire is part of my doctoral study in Evaluation and Measurement at Kent State University, Kent, Ohio, USA.

I would like you to participate in this study by completing the questionnaire. The time required to complete the questionnaire will take no more than 10 minutes. There are no risks involved with the participation in this study beyond the risks normally encountered in everyday life. You are not obligated to participate in the study.

If you decide to participate in the study, please sign the attached consent form. Then, please read each item of the questionnaire carefully and circle the number that best indicates the frequency of using the assessment practice described in the item in this science class as follows:
- Circle “1” if you “Never” use the assessment practice described in the item in this science class.
- Circle “2” if you “Rarely” use the assessment practice described in the item in this science class.
- Circle “3” if you “Sometimes” use the assessment practice described in the item in this science class.
- Circle “4” if you “Often” use the assessment practice described in the item in this science class.
- Circle “5” if you “Always” use the assessment practice described in the item in this science class.

Your responses to the questionnaire will remain anonymous and confidential. They will only be used for statistical analyses of the study with no connection made to you. Information gathered from this study are hoped to improve instruction, assessment, and learning in the ninth grade science class.

If you want to know more about this research project, please call me at [Hussain Al Kharusi, Tel. 001.330.389.9871; & Dr. Rafa Kasim, Tel. 001.330.672.0601]. The project has been approved by Kent State University. If you have questions about Kent State University’s rules for research, please call Dr. John L. West, Vice President and Dean, Division of Research and Graduate Studies (Tel. 001.330.672.2704).

Sincerely,
Principal Investigator: Hussain Al Kharusi
عزيزي معلم مادة العلوم للصف التاسع

بهدف الاستبيان المرفق إلى معرفة ممارسات معلم مادة العلوم للصف التاسع في القياس والتقييم الصف.

الاستبيان هو جزء من رسالتي للدكتوراه في التقويم والقياس بجامعة كنت الحكومية بمدينة كنت في ولاية أوهيا بالولايات المتحدة الأمريكية.

والمهم منك المشاركة في هذه الدراسة بالإضافة إلى الاستبيان المرفق، علماً بأن الإجابة على الاستبيان لا تغطي 10 دقائق. المشاركة في هذه الدراسة غير مصحوبة بأي مخاطر تتعدى حدود أدبي المخاطر التي تحدث بشكل طبيعي في الحياة اليومية، كما أنك تستلزم الرد على الاستبيان في هذه الدراسة.

إذا قررت المشاركة في الدراسة فأرجو التوقيع على إستمارة الموافقة المرفقة مع هذه الرسالة، ومن ثم قراءة كل عبارة من عبارات الاستبيان قراءة جيدة، وضع دائرة حول الرقم المناسب الذي يصف بدقة مدى استخدامك للجابن المذكور في الجواب في عملية قياس وتقييم تعلم الطلبة في مادة العلوم للصف التاسع على النحو التالي:

- ضع دائرة حول الرقم "1" إذا كنت "ابداً" لا تستخدم ذلك الجواب في عملية قياس وتقييم تعلم الطلبة في مادة العلوم للصف التاسع.

- ضع دائرة حول الرقم "2" إذا كنت "نادراً" ما تستخدم ذلك الجواب في عملية قياس وتقييم تعلم الطلبة في مادة العلوم للصف التاسع.

- ضع دائرة حول الرقم "3" إذا كنت "حياناً" تستخدم ذلك الجواب في عملية قياس وتقييم تعلم الطلبة في مادة العلوم للصف التاسع.

- ضع دائرة حول الرقم "4" إذا كنت "غالباً" تستخدم ذلك الجواب في عملية قياس وتقييم تعلم الطلبة في مادة العلوم للصف التاسع.

- ضع دائرة حول الرقم "5" إذا كنت "دائماً" تستخدم ذلك الجواب في عملية قياس وتقييم تعلم الطلبة في مادة العلوم للصف التاسع.

إن إجاباتك على الاستبيان ستظل سرية، ولن نستخدم إلا لغرض التحليل الأحصائي للدراسة دون أن يكون فيها أي رابط يفصل بك. تأمل أن تساعد المعلومات التي سوف نحصل عليها من هذه الدراسة على تحسين عملية التدريس، والقياس، والتعلم في مادة العلوم للصف التاسع.

إذا ترد أن تعرف المزيد عن هذه الدراسة برجي الاتصال بحسين الخروصي على هاتف رقم (01.330.389.9871) والدكتور راف قاسم على هاتف رقم (01.330.672.2704). علماً بأن هذه الدراسة قد تم التصديق عليها من قبل جامعة كنت الحكومية، فإذا كانت لديك أي أسئلة حول قواعد البحث بجامعة كنت الحكومية برجي الاتصال بالدكتور جون آل ويست مساعد الرئيس وعميد قسم الدراسات العليا والبحث على هاتف رقم (01.330.672.0601).

مع خالص التحية،

الباحث الرئيسي: حسین الخروصي
Dear Student

The attached questionnaire is designed to identify your perceptions of the classroom assessment environment, achievement goal orientations, and academic self-efficacy in your ninth grade science class. The questionnaire is part of my doctoral study in Evaluation and Measurement at Kent State University, Kent, Ohio, USA.

I would like you to participate in this study by completing the questionnaire. The time required to complete the questionnaire will take no more than 20 minutes. There are no risks involved with the participation in this study beyond the risks normally encountered in everyday life. You are not obligated to participate in the study. Participation in the study will not influence your science grade and your relation with the teacher.

If you decide to participate in the study, please sign the attached consent form. Then, please read each item of the questionnaire carefully and circle the number that best describes what you think about your science class or yourself as a student in this science class as follows:
- Circle “1” if you think that the item is “Completely not true” for your class or yourself.
- Circle “2” if you think that the item is “Somewhat not true” for your class or yourself.
- Circle “3” if you think that the item is “Somewhat true” for your class or yourself.
- Circle “4” if you think that the item is “Completely true” for your class or yourself.

Your responses to the questionnaire will remain anonymous and confidential. No one at home or at school will ever see your responses. Also, there is no right or wrong answer. Your responses will only be used for statistical analyses of the study with no connection made to you. Information gathered from this study are hoped to improve instruction, assessment, and learning in the ninth grade science class.

If you want to know more about this research project, please call me at [Hussain Al Kharusi, Tel. 001.330.389.9871; & Dr. Rafa Kasim, Tel. 001.330.672.0601]. The project has been approved by Kent State University. If you have questions about Kent State University’s rules for research, please call Dr. John L. West, Vice President and Dean, Division of Research and Graduate Studies (Tel. 001.330.672.2704).

Sincerely,
Principal Investigator: Hussain Al Kharusi
بهدف الاستفادة المفرفة إلى معرفة رأيك في بنية القياس والتقييم الصفي، وأهداف التحصيل لديك، وقرائك في مادة العلم للفصل التاسع. الاستفان هو جزء من رسالتي للدكتوراه في التقييم والقياس بجامعة كنت الحكومية بمدينة كنت في ولاية أو هاواي بالولايات المتحدة الأمريكية.

والمرجع ملكك المشاركة في هذه الدراسة بالإضافة إلى الاستفان المفرفة، علماً بأن الإجابة على الاستفان لا تنتهي 20 دقيقة. المشاركة في هذه الدراسة غير صحية بأي مخاطر تتعدى حوادث المخاطر التي تحدث بشكل طبيعي في الحياة اليومية، كما أنك تستمِّر ملزمًا بالمشاركة في هذه الدراسة، فإن تؤثر المشاركة في هذه الدراسة على معلدي في مادة العلم أو علاقتي مع معلم مادة العلم.

إذا قررت المشاركة في الدراسة فأرجو التوقيع على استمارة الموافقة المرفقة مع هذه الرسالة، ومن ثم قراءة كل عبارة من عبارات الاستفان قراءة جيدة، وأبدأ رأيك في الدراسة بوضع دانرة حول الرقم المناسب الذي يصف بدقّة ما تعتقد حول صفحك في مادة العلم أو حولك كطالب في هذا الصف على النحو التالي:

- ضع دائرة حول الرقم "1" إذا كنت تعتقد أن العبارة "صحبة تمامًا" بالنسبة لصفحك أو لك.
- ضع دائرة حول الرقم "2" إذا كنت تعتقد أن العبارة "صحبة إلى حد ما" بالنسبة لصفحك أو لك.
- ضع دائرة حول الرقم "3" إذا كنت تعتقد أن العبارة "خانطة إلى حد ما" بالنسبة لصفحك أو لك.
- ضع دائرة حول الرقم "4" إذا كنت تعتقد أن العبارة "خانطة تمامًا" بالنسبة لصفحك أو لك.

إن إجابتك على الاستفان ستظل سرية، ولن يطلع عليها أحد من البيت أو المدرسة، كما لا توجد إجابات صحبة وأخرى خاطئة، ولكن بإلتزام إجابتك فقط لغرض التحليل الأحصائي للدراسة بدون أن يكون فيها أي رابط يفصل بك. نأمل أن تساعد المعلومات التي سوف نحصل عليها من هذه الدراسة على تحسين عملية التدريس، والتقييم، والتعلم في مادة العلم للفصل التاسع.

إذا ترى أن تعرف المزيد عن هذه الدراسة برجي الاتصال بحسين الخروصي على هاتف رقم (001.330.389.9871) والدكتور راف قاسم على هاتف رقم (001.330.672.0601). أما بأن هذه الدراسة قد تم التصديق عليها من قبل جامعة كنت الحكومية، فإذا كانت لديك أي استفسار حول قواعد البحث بجامعة كنت الحكومية برجي الاتصال بالدكتور جون أل ويست مساعد الرئيس وعميد قسم الدراسات العليا والبحث على هاتف رقم (001.330.672.2704).

مع خالص التحية،

الباحث الرئيسي: حسين الخروصي
REFERENCES
REFERENCES


