

Science and Mathematics Education Centre

**Relative Effectiveness of Alternative Sequencing of Middle-School
Science Curriculum in Terms of Classroom Learning Environment
and Student Attitudes**

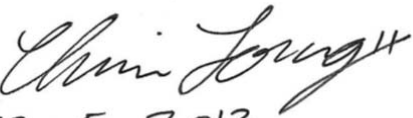
Christopher Sean Long

**This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University**

May 2013

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university. To the best of my knowledge and belief, this thesis contains no material previously published by any other person except where due acknowledgment has been made.

Signature: 
Date: 13-5-2013

ABSTRACT

This study compared the relative effectiveness of two approaches for sequencing middle-school science curriculum in use in the United States in terms of the learning environment and student attitudes towards science, as well as the differential effectiveness of the two curriculum sequences for students of two different ethnicities (Caucasians and Hispanics). Additionally, associations between learning environment and student attitudes were examined.

A new learning environment instrument, called the Outcomes-Related Learning Environments Survey (ORLES), was developed and validated for this study. The ORLES was administered to 367 students in 15 grade 8 science classes in Texas and California. Factor and item analysis supported a six-factor structure of the ORLES scales (Personal Relevance, Student Cohesiveness, Teacher Support, Investigation/Involvement, Task Orientation and Enjoyment of Science).

Various states' middle-school science standards were analyzed and classified as being either generalized (each grade level has topics from all fields of science) or topic-specific (each grade level covers topics from one field of science).

Comparisons of the two curriculum sequence approaches via MANOVA revealed a significant difference (effect size of 0.74 standard deviations) for Enjoyment of Science in favor of the topic-specific approach, and that the generalized approach was slightly more effective for Hispanic than Caucasian students in terms of Task Orientation (with an effect size of 0.34 standard deviations).

When data were analyzed for associations between learning environment and student attitudes towards science, significant associations between all five learning environment scales and attitudes were found.

ACKNOWLEDGEMENTS

First and foremost, I would like to acknowledge the most patient support offered by my supervisor, Professor Barry Fraser. His guidance and patience in reading and editing multiple revisions of my thesis were of immeasurable value. I also would like to recognize the assistance and statistical support of Dr Rekha Koul, who was instrumental in the analysis of my data as well as acting as my guide at the Curtin University campus. I would also like to thank Dr Christine Howitt for her teaching in the early days of my doctoral studies.

Special recognition goes to my colleagues at the University of Texas at Dallas. First, Dr Cynthia Ledbetter invited me into the Masters program at UT Dallas, bolstered my teaching career and introduced me to the world of educational research and to Dr Fraser. Dr Ledbetter has provided assistance above and beyond the call of duty on numerous occasions. Also in this group, is Dr Rebekah Nix who has supported me throughout my academic struggles as well as paving the way for me at Curtin. Others at UT Dallas, whose support I treasure dearly, include Dr Mary Urquhart, Barbara Curry, Kim Distin and Linda Cook. I would be remiss if I failed to mention the support of the cadre of the UT Dallas Collaborative for Excellence in Science Education in general.

Numerous people within the Garland Independent School District deserve recognition in support of my doctorate beginning with my principal, Jim Lewis, who, nearly 10 years ago, gave me the opportunity to teach middle-school science and begin a new career. I would also like to thank Jim Manley and Laurel Dingrando for their assistance in securing data necessary for my study.

Additionally, I would like to thank my countless friends and family for the support, especially my mother, Jo Long, for her support both in teaching and all of my studies throughout my entire life.

Of course, I owe a debt of gratitude to the well over 1000 students who have passed through my classrooms over the last decade and allowed me to have a hand in shaping their future. A special nod goes to my current set of students who ask me weekly if they get to call me Dr Long yet. Their encouragement means the world to me. It is, after all, for them that we teachers do what we do.

TABLE OF CONTENTS

DECLARATION	ii
ABSTRACT	iii
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	ix
LIST OF FIGURES	x
Chapter 1: INTRODUCTION.....	1
1.1 Introduction.....	1
1.2 Study Rationale.....	2
1.3 States Involved in Study	5
1.4 Field of Learning Environments.....	6
1.5 Middle-School Science: Issues and Standards	7
1.5.1 Middle-School Science Issues.....	7
1.5.2 Middle-School Science Standards.....	10
1.5.2.1 California Standards	13
1.5.2.2 Texas Standards.....	14
1.5.3 Classification of Standards.....	14
1.6 Research Questions.....	20
1.7 Overview of Thesis.....	21
Chapter 2: LITERATURE REVIEW.....	23
2.1 Introduction.....	23
2.2 History of Learning Environments Research	24
2.3 Learning Environment Instruments	26
2.3.1 Science Laboratory Environment Inventory (SLEI)	26
2.3.2 Constructivist Learning Environment Survey (CLES).....	29
2.3.3 What Is Happening In this Class? (WIHIC) Survey	31
2.3.4 Questionnaire on Teacher Interaction (QTI)	37
2.3.5 Other Instruments	38
2.4 Student Attitudes Towards Science.....	44
2.4.1 Test of Science Related Attitudes.....	46
2.4.2 Other Attitude Instruments	47
2.5 Applications of Learning Environment Instruments	48

2.5.1	Associations Between Classroom Environment and Student Outcomes	49
2.5.2	Evaluation of Educational Innovations.....	52
2.5.3	Ethnic Differences in Learning Environment Perceptions	54
2.5.4	Transition Between Levels of Schooling.....	57
2.6	Conclusions.....	58
Chapter 3: RESEARCH METHODS.....		61
3.1	Introduction.....	61
3.2	Sample Selection	61
3.3	Learning Environment and Attitude Instrument Used.....	64
3.4	Data Analysis.....	66
3.4.1	Validity and Reliability of the ORLES.....	67
3.4.2	Differences between ORLES Scales in Terms of Curriculum Sequence and Ethnicity	69
3.4.3	Associations between Learning Environment and Attitudes.....	70
3.5	Conclusion	71
Chapter 4: ANALYSES AND RESULTS.....		74
4.1	Introduction.....	74
4.2	Validity of Questionnaire	75
4.3	Curriculum-Sequence and Ethnicity Differences in ORLES Scales	79
4.3.1	Differences between Curriculum Sequences	81
4.3.2	Differences between Ethnicities	83
4.3.3	Differential Effectiveness of Different Curriculum Sequences for Different Ethnic Groups	85
4.4	Associations Between Learning Environment and Attitudes Towards Science.....	87
4.5	Conclusion	89
Chapter 5: CONCLUSIONS		92
5.1	Introduction.....	92
5.2	Introductory Chapter and Literature Review	92
5.3	Discussion of Methods and Results.....	93
5.3.1	Validity and Reliability of the ORLES Questionnaire	94
5.3.2	Relative Effectiveness of Alternative Curriculum Sequences.....	95

5.3.3	Differences Between Ethnicities and Differential Effectiveness of Curriculum Sequences for Two Ethnic Groups.....	97
5.3.4	Associations between Learning Environment and Attitudes.....	97
5.4	Limitations and Suggestions for Future Research.....	98
5.5	Significance and Implications.....	100
5.6	Concluding Remarks	101
REFERENCES		102
APPENDIX A: Outcomes-Related Learning Environment Survey (ORLES)		119

LIST OF TABLES

Table 1.1	Texas Essential Knowledge and Skills (TEKS) by Grade Level and Scientific Field.....	17
Table 1.2	California Middle School Science Standards by Grade Level and Scientific Field.....	18
Table 1.3	χ^2 Values for Distribution of Topics Across Three Fields of Scientific Study by Grade Level and State	20
Table 3.1	Gender and Ethnic Breakdown of Sample in Each State.....	64
Table 3.2	Scale Descriptions of Sample Items for Outcomes-Related Learning Environment Survey Scales (ORLES).....	66
Table 4.1	Factor Analysis Results for the ORLES	77
Table 4.2	Internal Consistency (Cronbach Alpha Reliability), Ability to Differentiate between Classrooms (ANOVA Results) and Discriminate Validity (Mean Correlation with Other Scales) for the ORLES	78
Table 4.3	ANOVA Results for Each ORLES Scale for Curriculum Sequence, Ethnicity and Sequence-by-Ethnicity Interaction.....	80
Table 4.4	Average Item Mean, Average Item Standard Deviation and Differences between Two Curriculum Approaches (ANOVA Results and Effect Size) for each Learning Environment and Attitude Scale	82
Table 4.5	Average Item Mean, Average Item Standard Deviation and Difference Between Ethnic Groupings for Learning Environment and Attitude Scales.....	84
Table 4.6	Average Item Mean and Average Item Standard Deviation for Each ORLES Scale for Two Ethnicities and Two Curriculum Sequences, and ANOVA Results for Sequence-by-Ethnicity Interaction.....	86
Table 4.7	Associations between Learning Environment Scales and Enjoyment of Science in Terms of Simple Correlations (r), Multiple Correlation (R) and Standardized Regression Coefficients (β).....	88

LIST OF FIGURES

Figure 4.1	Simplified Plot of Significant Differences Between Two Curriculum Sequences for ORLES Scales.....	83
Figure 4.2	Simplified Plot of Significant Differences Between Two Ethnicities for ORLES Scales	85
Figure 4.3	Average Item Means for Task Orientation for two Ethnicities and Two Curriculum Sequences	87

Chapter 1

INTRODUCTION

The scientist is not a person who gives the right answers, he's one who asks the right questions. – Claude Levi-Straus (1964)

Ever since Sputnik emitted its electronic beep from orbit, there has been urgency in the United States to improve science education. Over the years, there have been considerable efforts made to improve, modify or reinvent science instruction. Some of these efforts have fared better than others, according to various methods of evaluation. One frequent change made in science education is to the curriculum. Most states in the United States specify what topics should be taught in science classes at each grade level and each state has a different set of topics; but there appears to be little consensus as to how these topics should be arranged. My study, which is the subject of this thesis, evaluated the effectiveness of some alternative curriculum sequences at the middle-school level.

1.1 INTRODUCTION

This chapter introduces my study. Section 1.2 establishes the rationale for the study, including reasons for conducting the research. Section 1.3 provides brief background information and reasons for the selection of the two states for my research: California and Texas. Section 1.4 provides a brief introduction to the field of learning environments research on which my study drew and which is reviewed

comprehensively in Chapter 2. Section 1.5 reviews literature related to issues facing science education in middle schools, as well as standards for middle-school science and my method for classifying curricula as either general or specific. Section 1.6 delineates the four research questions that guided this study. Finally, Section 1.7 gives an overview of each chapter in the rest of the thesis.

1.2 STUDY RATIONALE

The impetus for this study comes from the change in the Texas public school science curriculum circa 1998. Before 1998, the middle-school science curriculum for students in grades 6–8 was arranged in topic-specific sequences. Sixth graders were taught physical science, seventh graders learned life science and eighth graders studied earth science. After 1998, sixth, seventh and eighth grade students all were taught with a spiraling curriculum of general science for three years. The main purpose of my study was to compare the sequence taught in the Texas public schools, with a sequence being implemented in public school systems in a state outside of Texas, namely, California.

The construction of a science curriculum is an important decision. There have been differences shown in the achievement of students based upon how and what they were taught. Walberg (1991) identifies nine factors that appear to increase learning. Among these are “the quality of the instructional experience including method (psychological) and curricular (content) aspects” (p. 42). Walberg points out that one of the reasons for the success of Japanese curricula is the “careful standardization and grading of the material to be learned” (p. 45) and that:

One ingredient of success in secondary school science is a uniform national science curriculum of high standards. Such a curriculum enables teachers to concentrate and cooperate on pedagogical means rather than inventing disparate goals. (Walberg, 1991 p. 63)

Certainly, course content is not the only factor. Walberg also lists factors under the grouping of student aptitude and psychological environment. Student aptitude includes ability or a measure of how the student has achieved in traditional tests, developmental level compared to age and motivation as measured by personality tests. Psychological environment factors in Walberg's model are the home environment, classroom morale, external peer groups and television viewing.

There have been numerous curriculum reform movements in the last couple of decades. One of these was the Iowa Scope, Sequence and Coordination (SS&C) that was conducted as part of the National Science Teachers Association's Scope, Sequence and Coordination project. The Iowa SS&C was implemented from 1990 to 1993. Liu and Yager (1997) conducted a quantitative study that showed that students who participated in the reforms implemented did demonstrate gains in achievement and understanding. They caution that the gains might have been attributable to a constructivist learning environment rather than curricular change, but still a positive impact was shown.

One major obstacle in comparing middle-school science curricula across state boundaries is that there is no standardized test of science achievement that covers multiple states. Each state has its own standardized test, such as the Texas Assessment of Knowledge and Skills (Texas Education Agency, n.d.) or the New York State Intermediate-Level Science Test (New York State Education Department, n.d.), but comparing scores on these different tests would be both impractical and invalid.

Therefore, my study followed a rich tradition in prior research of employing learning environment and student attitude as indicators of effectiveness in comparing different science curricula (Chionh & Fraser, 2009; Holding & Fraser, 2013; Lightburn & Fraser, 2007; Pickett & Fraser, 2009; Wolf & Fraser, 2008).

There is plentiful evidence in the literature that supports a link between learning environments and student outcomes (Fraser, 2012). Several examples of these studies are discussed in detail in Chapter 2 of this thesis. For example, when Dorman (2000) used 10 learning environment scales from the Constructivist Learning Environment Survey (CLES) (Taylor, Fraser & Fisher, 1997) and the What Is Happening In this Class? (WIHIC) (Aldridge, Fraser & Huang, 1999) and related them to student efficacy, the strongest association was found for the WIHIC scale of Task Orientation (the importance placed on completion of an assignment and sticking to the subject matter being studied). Task Orientation was shown to be associated with student achievement and attitudes by Allen and Fraser (2007). Nolen and Haldyna (1990) also found tentative connections between task orientation and student outcomes. Building upon this with a longitudinal study of classroom climates by using an instrument constructed from scales drawn from other instruments, Nolen (2003) found that the Ability–Meritocracy scale had a negative impact on achievement. When students felt that the focus was on doing well in class and had a belief that strong science aptitude is required, they scored lower on the end-of-course science test. When Chang, Hsiao and Barufaldi (2006) used an instrument called the Earth Science Classroom Learning Environment Instrument (ESCLEI), they found strong relationships with student attitudes as measured by the Attitudes Towards Earth Science Inventory (ATESI).

These four brief examples show that learning environments have been found to be connected to student outcomes such as achievement, efficacy and attitude. Many other studies that have reported similar connections (Fraser, 2007) are reviewed in Chapter 2 of this thesis.

1.3 STATES INVOLVED IN STUDY

All 50 U.S. states have some sort of standards for their middle-school science curriculum but, for the sake of economy in my study, the two states of California and Texas were chosen for comparison. There are several reasons why these two states were chosen. First, California and Texas represent a large proportion, about one-fifth, of the U.S. population. California is the most populous state with approximately 38 million residents (United States Census Bureau, n.d.). Texas has the second largest population in the U.S. with over 26 million inhabitants (United States Census Bureau, n.d.). Second, both states have a similar demographic makeup. Californian students are approximately 39% Caucasian, 7% African American, 38% Hispanic, and 16% other ethnicities. Texas is approximately 45% Caucasian, 13% African American, 38% Hispanic, and 4% other ethnicities (United States Census Bureau, n.d.). Third, both states' largest minority group, in this case Hispanics, speaks a language other than English. Finally, both states share similar histories as former Mexican colonies. The two states, while seemingly very different, have a number of striking similarities. However, these two states have very different yet well-defined middle-school science curriculum sequences that were ideal for comparison in my study.

1.4 FIELD OF LEARNING ENVIRONMENTS

My study drew upon, and attempted to make a modest contribution to, the field of learning environments, whose history covers half of a century (Fraser, 2012). The history of learning environments research is covered in more depth in Chapter 2, but is summarized here. Human environments research began with the work of Kurt Lewin and Henry Murray. Lewin (1936) introduced the 'person–environment paradigm' which states that behavior is a function of the person and the environment. Murray (1938) added the idea of *alpha* and *beta press* (i.e. external versus internal view of an environment). This idea was expanded upon by Stern, Stein and Bloom (1956) who further refined *beta press* into *personal* and *consensual* *beta press*.

Modern research specifically on classroom environment utilizing questionnaires began with Herbert Walberg, who developed the Learning Environment Inventory (LEI) to evaluate the effectiveness of Harvard Project Physics (Walberg & Anderson, 1968). Rudolf Moos pioneered the use of environment scales in institutional settings (e.g. hospitals and prisons), which led to the development of Classroom Environment Scale (CES) (Moos & Trickett, 1974). Beginning with this genesis, many learning environment instruments, including those used in my study, were developed, validated and utilized in a wide range of research applications and educational settings as reviewed in Chapter 2 of this thesis.

1.5 MIDDLE-SCHOOL SCIENCE: ISSUES AND STANDARDS

Because the sample in my study of general and specific science curriculum sequences involved middle-school students (grade 6–8), this section provides some background details of this important level of schooling in Section 1.5.1. Later, in Section 2.5.4, past learning environment research into the transition between levels of schooling is reviewed.

Because my study involved the two states of California and Texas, the middle-school science standards for each of these states are considered in Section 1.5.2. In particular, because my study compared general and specific curriculum sequences, Section 1.5.2 also outlines my methods of classifying curricula as either general or specific.

1.5.1 Middle-School Science Issues

Middle school is a difficult transitional time for students. Of particular relevance to my research is a study in Tasmania, Australia. Ferguson and Fraser (1998) studied the transition of 1,040 students in 47 primary schools and how their perceptions of their learning environment changed between the last year of their primary schooling (grade 6) and the first year of their secondary school (grade 7). One of the differences noted by the authors was the switch from a generalist learning environment into a series of specialized learning environments and the introduction of laboratory science classes. Ferguson and Fraser (1998) found that students entering

secondary school perceived many changes to their learning environment, particularly in terms of student cohesiveness.

Middle school is defined as grades 6–8 in most jurisdictions, but it also can encompass grades 7–8 or grades 7–9. In any case, because middle school covers the grade levels in between elementary and high school. As such, it can be described as a sort of educational grey area. In the states involved in my study (California and Texas), middle schooling encompasses grades 6 to 8 and typically involves students aged from 11 to 14 years. In the elementary grades, science is taught as a single generalized broad subject but, in high school, it is taught as specialized strand-specific courses. However, in middle school, science is taught in either the elementary mode or the high school mode depending on the preferences of the local or statewide authority.

The debate about the relative merits of a generalized or a specialized curriculum arrangement can be seen in the preparation of middle-school science teachers. German, Dumas and Barrow (1995) describe some of the concerns in preparing middle-school science teachers. Valentine and Mogar (1992) state that only 33 of 50 U.S. states have provisions for the certification of middle-school science teachers. Of those 33 states, only a few differentiate the middle school from elementary or high school. The National Science Teachers Association (1992) suggested that middle-school science teachers should be given a broad education (8–10 semester hours) in each of the three main science strands of biology, physical science and earth science. However, Germann, Dumas, and Barrow (1995) found that only 13 of 74 surveyed U.S. universities offered a specific middle-school science education degree, with 12 offering the suggested broad-based curriculum. The other university required specialization in one of the fields.

Although my study focused on middle-school science education in the U.S., concern for science education among this age group and in high school appears to be universal (Fraser, 1986; Lyons, 2006). For example, Turkish students exhibited a noticeable decrease in attitudes towards science between grades 5 and 11 (Baykul, 1990).

Learning environment has been shown to be an important factor in middle-school science. For example, Mason, Boscolo, Tornatora and Ronconi (2012) recently conducted a study of 696 students in Italy that included 202 middle-school students. They found that students who had a more constructivist view of science learning (e.g. the uncertainty of science) tended to have more scientific knowledge and demonstrate higher levels of outcomes such as self-efficacy and achievement.

Middle school provides a transition from elementary to high school. In this sense, it is neither fish nor fowl but both fish and fowl. Because of this transitional nature, there exists two fundamental modes of arranging science curriculum at the middle-school level: either the elementary general science approach; or the high school's course-based, specific approach. Because the debate over which mode of instruction is more effective is far from settled, my evaluated the relative effectiveness of these two different middle-school curriculum sequences.

The curriculum in most U.S. states is defined by its middle-school science standards. Therefore, these standards are considered in some detail in Section 1.5.3 below.

1.5.2 Middle-School Science Standards

Many states in the United States have attempted to codify exactly what should be taught in science at each grade level. At the time of this study, Texas used a general science approach that involves a smorgasbord of topics each year with increasing difficulty and detail as the student progresses from 6th to 8th grade (<http://ritter.tea.state.tx.us/rules/tac/ch112.html>). California, on the other hand concentrated on particular fields of science in each of the grades (<http://www.cde.ca.gov/be/st/ss/documents/sciencestnd.pdf>). The National Academy of Sciences has published a set of national science education standards and some states have utilized them in drafting their own standards. Private and parochial schools aren't bound by state directives and have their own ideas about course content.

National educational standards have their genesis in the efforts of the National Governors Association and the administrations of Presidents George Bush and William Clinton. In 1989, the National Council of Teachers of Mathematics published the Curriculum and Evaluation Standards for School Mathematics which represents the first set of nation-wide educational standards (National Academies Press, 1996).

The earliest call for a national science standard was issued in a document called *A Nation At Risk*, which outlined the need for standardized content and expectations (National Commission on Excellence in Education, 1983). From this document, many different efforts sprung forth. The latest of these was *The Content Core* published by the National Science Teachers Association (1992).

In 1996, the National Academy of Sciences published the *National Science Education Standards*, which were written over a five-year period and received input from thousands of individuals and groups. The National Science Education Standards address a whole spectrum of issues regarding the teaching of science. The National Science Education Standards contains chapters that cover standards for the teaching of science, professional development of teachers, assessment, design of school-wide and system-wide science programs and science content. My study was concerned with the arrangement of science content within different state curricula and the National Science Education Standards have had a very strong influence on most state education systems' decisions regarding their science education programs.

While virtually every state's science standards are based on the National Science Education Standards, the concept of local control still reigns and each individual state board of education has its own set of standards. The National Science Education Standards content standards list topics to be taught within fairly broad groupings of grade levels: K to 4, 5 to 8, and 9 to 12. Because this study was focused on grades 6 to 8, only the 5 to 8 grouping is pertinent to this study. Within this grouping, the National Science Education Standards identify topics to be covered but does not specify at what specific grade level the topic is to be covered. This allows states to make decisions as they see fit.

There appears to be a general consensus, at least informally, among supervisory-level educators about their own state's curriculum. Prior to the primary data collection, I sent e-mails to public school district science coordinators in the two states examined in this study. The participants were asked the following questions:

- 1) What is your professional position or title?
- 2) For how long have you held this or similar positions?
- 3) How would you describe your school's, district's, or state's arrangement of middle school science topics?
- 4) Why do you think the curriculum is arranged in the way that it is?
- 5) Do you like the current arrangement? Why or why not?
- 6) How would you change the arrangement, given the opportunity?

Three coordinators responded to the request for personal opinions regarding their state's set of standards. Most parties interviewed in this study supported their curriculum and felt that it is generally the best way to teach middle-school science. If they weren't enamored with their own arrangements, they liked them with some reservations.

For example, one example is a secondary science coordinator in a large suburban district in Texas. The coordinator, who has been in her position for four years, described the Texas arrangement as "integrated science concepts". She felt that the reasoning behind arranging the topics as they are is so that students understand the big picture. She thinks this is ideal because students need an understanding of how all the disciplines are connected.

A second Texas interviewee, a science coordinator from a smaller suburban district, described the middle-school curriculum as a spiraling set of concepts from all fields of science that build on strands that start in elementary schools. She feels that, while the concept is good, the number of topics could be narrowed. She also asserts that teaching is what ultimately makes the difference and she feels that the choice of arrangement makes little difference.

The third interviewed professional is a county science education specialist from a unified district in central California with six years in that position. She supported the course-based approach used in California and expressed concerns about the inclusion of a handful of standards at each grade level that fall outside the field of focus for that grade. For example, a physics topic is included in the grade 7 standards which primarily cover life sciences. The implication possibly is that she would prefer the curriculum to be more specialized than it is in its current form. She also feels that some of the standards are developmentally inappropriate for the grade level at which they are introduced. This is the same concern expressed by Liu (2006) in regards to the national standards which are the basis for most state standards.

In the subsection below, the focus is specifically on the middle-school science standards of the two states involved in my study, namely, California (Section 1.5.2.1) and Texas (section 1.5.2.2).

1.5.2.1 California Standards

In the early 1990s, officials in the state of California were increasingly concerned with their state's decline in National Assessment of Educational Progress (NAEP) scores which saw California drop to near the bottom of state rankings. This brought about a sequence of actions that began with a mandate for individual districts within the state to increase rigor and develop more challenging and engaging standards. In turn, this led to a set of state-wide assessments and accountability measures. However, officials felt that there was too much emphasis on testing and not enough emphasis on acquisition of science knowledge. Finally, after much political

debate, the current standards were written in 1997 and adopted in 1998. There has been no less than four attempts to review and change the standards since then but they have stalled in the legislative process (Bertrand, n.d.).

1.5.2.2 Texas Standards

Before 1984, Texas had no uniform science curriculum. Each individual school district was responsible for creating its own curriculum and some districts had none at all, leaving those decisions to the teachers. In the 1984/1985 school year, the state implemented the Essential Elements (EEs) which broadly outlined what every Texas student should learn. The EEs were the first uniform standards in Texas. Later, and after many different incarnations of accountability testing, the EEs proved to be inadequate (Nelson et al., 2007).

The EEs were replaced by the Texas Essential Knowledge and Skills (TEKS), which are a set of standards for Texas Publics schools from Kindergarten to 12th grade. The Science TEKS were first implemented in 1998. In the 2010/2011 school year, Texas revised the science TEKS into something more closely resembling the California TEKS. Because my study was conducted during the 2009/2010 school year, the discussion in this thesis is about the 1998 standards in use at the time (<http://ritter.tea.state.tx.us/rules/tac/ch112.html>).

1.5.3 Classification of Standards

Public School Curriculum in the United States is often controlled at the state level, which makes it relatively simple to collect and compare various sequences.

Although only two state middle-school science curricula (Texas and California) were analyzed in the development of my classification scheme, the techniques can be widely applied to any school system that has a grade-by-grade distinction of topics that are to be taught. This method does not work well for documents such as the National Science Standards which do not assign topics to individual grade levels.

The middle-school science TEKS in Texas are enumerated in Chapter 112, Subchapter B of the TEKS document. Subchapter B further is divided into sections by grade level: Section 112.22 for grade 6, Section 112.23 for Grade 7 and Section 112.24 for Grade 8. Within each of these sections, specific knowledge and skills, TEKS, are itemized. Each individual TEK is labeled as either 'scientific processes' or 'scientific concepts'.

Section 122.22 contains 14 TEKS. TEKS 6.1 to 6.4 involve scientific processes TEKS. For example, TEKS 6.1 states: “The student conducts field and laboratory investigations using safe, environmentally appropriate, and ethical practices. The student is expected to: (A) demonstrate safe practices during field and laboratory investigations; and (B) make wise choices in the use and conservation of resources and the disposal or recycling of materials” (Texas Education Agency, n.d.). TEKS 6.5 to 6.14 involve scientific concepts. For Example, TEX 6.14 states: "The student knows the structures and functions of Earth systems. The student is expected to: (A) summarize the rock cycle; (B) identify relationships between groundwater and surface water in a watershed; and (C) describe components of the atmosphere, including oxygen, nitrogen, and water vapor, and identify the role of atmospheric movement in weather change" (Texas Education Agency, n.d.). This pattern is repeated for Sections 112.23 and 112.24. The fifth TEKS in each section is a broad

scientific concept that applies equally to all fields of science. TEKS 6.5 and 7.5 require the student to know that systems are a part of science and that systems interact with other systems. TEKS 8.5 requires students to understand that science and technology are related.

The remaining nine TEKS for each grade were analyzed according to the following method. Each TEKS contains between 2 and 5 clauses. Each of these clauses was classified as Life Science, Earth/Space Science or Physical Science (Physics and/or Chemistry). For example, TEKS 6.7 contains two clauses: "(A) demonstrate that new substances can be made when two or more substances are chemically combined and compare the properties of the new substances to the original substances; and (B) classify substances by their physical and chemical properties" (Texas Education Agency, n.d.). Both of these clauses are classified as physical science topics. Table 1.1 shows the results of this classification for each grade level and field of science.

The same procedure as used above with the TEKS was applied to the standards in use in California. Table 1.2 shows the breakdown of California's content standards by grade level and by the three major scientific fields. Since 1998, the California State Board of Education has had in place science content standards for each grade level that are broken down into major areas of study each of which has several sub-topics (<http://www.cde.ca.gov/be/st/ss/documents/sciencstnd/pdf>). The last major area for each grade level is entitled Investigation and Experimentation. This section is comprised of process skills and was not included in Table 1.2.

Table 1.1

Texas Essential Knowledge and Skills (TEKS) by Grade Level and Scientific Field

Grade Level	Number of Topics		
	Life Science	Earth Science	Physical Science
6	11	8	8
7	14	11	8
8	6	11	12

In the sixth grade, 25 clauses were examined in TEKS 6.6 to 6.14. Of the 25 clauses classified, two applied to two fields. One of these is TEKS 6.8B which states “...explain and illustrate the interactions between matter and energy in the water cycle and in the decay of biomass such as in a compost bin” (Texas Education Agency, n.d.). This concept is classified as both life science and Earth science. With the two double-category TEKS, there were 27 items classified. Eleven of the items fell into Life Science, eight in Earth Science and eight in Physical Science (Table 1.1).

In Grade 7, TEKS 7.6–7.14 had 26 clauses considered with six doubly-classified clauses for a total of 32 items. Of these items, 14 were categorized as Life Sciences, 11 as Earth Science and eight as Physical Science. In the 8th grade, the analysis of TEKS 8.6 to 8.14 also involved 26 topics, three of which were classified into two fields, for a total of 29 items; six Life Science, 11 Earth Science and 12 Physical Science (Table 1.1).

Table 1.2

California Middle-school Science Standards by Grade Level and Scientific Field

Grade Level	Number of Topics		
	Life Science	Earth Science	Physics/Chemistry
6	6	21	7
7	28	7	9
8	3	6	34

Grade 6, as it is referred to in the California documents, encompasses 28 standards with four crossing over two or three categories for a total of 34 items. Analysis revealed six topics in Life Science, seven in the Physical Science and 21 in Earth Science. Grade 7's 40 standards, with four double-category topics, yielded 44 items comprised of seven Earth Science items, nine Physics/Chemistry and 28 Life Sciences items. Grade 8 had a similar composition but with physics/chemistry taking the lion's share of items with 34. Three standards had a basis in both Life Science and Physical Science. Six topics made up the Earth Science portion of the 43 items from the 43 topics and 39 standards.

It seems from a first cursory examination that either of these two state systems could be classified as having either generalized or specialized curriculum arrangements. Both states cover at least some topics from each of the three content fields of science at each grade level. In order to determine if a set of standards represents a generalized or specialized curriculum, a Chi Square or χ^2 test was used. The χ^2 test is used to determine if the proportions in frequency counts of the data fall within expectations or if they are different from what is expected. For this study, it was assumed that a perfectly generalized set of topics would have an even distribution of topics across all three scientific fields. For example, because the Texas 6th grade

data show 27 items, it is anticipated that each scientific field would hold 9 of the items. If the χ^2 test shows the distribution is not significantly different from expected, then it can be said that the curriculum is generalized. A specialized curriculum would show a high χ^2 value, which represents a distribution that is significantly different from that expected, meaning that the distribution is not even across the three scientific fields. A χ^2 value of less than 4.605 supports the null hypothesis (Gay et al., 2006). The data collected in this study show that, in each of the statewide curricula examined, the middle-school sequences can be characterized as specialized or generalized.

The χ^2 values in Table 1.3 confirm that the middle-school science curriculum in the Texas public schools is a generalized arrangement of science concepts taught over the three middle-school years. California is subscribing to a specialized grouping of standards in which a single scientific field is emphasized in each of the three years.

This technique can be used to classify any middle-school curriculum as either generalized or specialized. The one potential case that did not appear in these data is a hybrid. In a hybrid classification, one or two of the grade levels would be generalized while the other grade levels are specialized. This method cannot be applied to documents like the National Science Standards because they do not specify topics to single grade levels, but rather to grade-level groups. Some states also fall into this category, including Illinois and Florida. These states have broad concepts to be taught somewhere within a range of grades. This also holds true for Australian public schools. A framework is in place for each of the Australian states, but they are not so specific at each grade level. In order to analyze curricula in these places, the individual district or school needs to be examined. This is also the case for private and

parochial schools. The data can be obtained easily enough with interviews or site visits as required.

Table 1.3

χ² Values for Distribution of Topics Across Three Fields of Scientific Study by Grade Level and State

Grade Level	χ ² for Distribution of Topics Across Three Scientific Fields.	
	California	Texas
6	12.412*	0.667
7	18.318**	1.636
8	40.791**	2.138

Note. df=2, *p<0.01, **p<0.001

1.6 RESEARCH QUESTIONS

Keeping with the tradition of learning environments research, this study used a questionnaire to assess students' perceptions of their learning environment and attitudes. The following research questions were the foci for my study:

1. Is the Outcomes-Related Learning Environments Survey (ORLES), consisting of scales from the Constructivist Learning Environment Survey (CLES), What Is Happening In this Class? (WIHIC) and Test Of Science Related Attitudes (TOSRA), valid for use among middle-school science students in the United States?
2. Do scores on ORLES scales vary with (a) curriculum sequence (generalized versus topic-specific) and (b) student ethnicity and (c) the interaction of curriculum sequence with ethnicity?

3. Are there associations between classroom learning environments and student attitudes to science?

1.7 OVERVIEW OF THESIS

The current chapter (Chapter 1) introduced my study, including: the study's rationale; the two U.S. states involved; a brief introduction to the field of learning environments; issues in middle schooling, including the classification of curricula as generalized or specific; and my research questions.

Chapter 2 of the thesis reviews literature relevant to my study. Firstly, it offers a detailed history of learning environments research. Secondly, it discusses popular learning environment instruments currently in use in the field, especially those chosen for my study. Thirdly, there is a review of literature on student attitudes and the instruments used to measure attitudes. Lastly, the chapter explores various applications of learning environment research.

Chapter 3 covers the methodology used in my study. First, it provides details of the sample used in my study, including information about the schools, students and states that were involved. Next, the construction and nature of the ORLES instrument used in my study are discussed. Finally, the methods of data analysis used to answer the research questions are described.

Chapter 4 presents the analyses and results for the study. Each of the three research questions is covered along with the statistical analysis used to answer each.

The thesis concludes with Chapter 5. In addition to summarizing the previous chapters, this chapter includes a discussion of the implications of the findings, a review of the limitations of my study and suggestions for further research.

Chapter 2

LITERATURE REVIEW

2.1 INTRODUCTION

Educators, legislators and parents are constantly seeking ways to measure and improve education. The ways and means of improving education are often contentious and there are differences in viewpoints between politicians, researchers and educators as to the best ways to educate children. Rushton and Rushton (2008) explore the nature of this conflict in detail, pointing out the problems in reconciling United States federal initiatives, psychological research and efforts of teachers on the front lines.

The primary method of assessment of student, teacher and school performance in the United States is high-stakes testing. Perhaps no single educational policy change over the past 50 years has had as great an impact on the working lives of teachers in public schools in the United States and other developed nations as the imposition of high-stakes testing requirements. High-stakes testing refers to the use of standardized student achievement tests as a primary mechanism for evaluating the performance of students, their teachers, and their schools

High-stakes tests, while crucial to the current educational system, could not be used in this study to compare curricula between US states as each state has its own unique system of high-stakes testing and criteria. However, the field of learning environments provides a precedent for a method of comparing curricula in terms of the learning environment created (Lightburn & Fraser, 2007; Wolf & Fraser, 2008). Also, numerous studies have revealed connections between positive learning

environments and desired student outcomes such as achievement, efficacy, and positive attitudes towards learning (Fraser, 2012). The following literature review provides an overview of learning environments research and the connections often reported between the learning environment and student outcomes.

Section 2.2 offers a brief history of learning environment research. Section 2.3 reviews some popular instruments in current use in the field, particularly those that were used in my study. Section 2.4 presents information about student attitudes and the instruments used to measure attitudes. Section 2.5 discusses some applications of learning environment instruments, especially the evaluation of educational innovations (Section 2.5.1), relationships between environment on student outcomes (Section 2.5.2), ethnic differences in learning environment perceptions (Section 2.5.3) and studies of the transition between levels of schooling (Section 2.5.4). Finally, Section 2.6 provides a summary of the literature reviewed in this chapter.

2.2 HISTORY OF LEARNING ENVIRONMENTS RESEARCH

Discussion of human environments research begins with Lewin (1936) and Murray (1938). Lewin (1936) developed a 'person-environment interaction paradigm' which, stated simply, is $B = f(P,E)$, where B is Behavior, P is Person, E is Environment and f is the interactive function. Translated to the classroom, learning (B) would be a function of the student (P) and the classroom atmosphere (E). Murray (1938) introduced the concepts of 'alpha press' and 'beta press'. Alpha press is how an environment is observed and described by an external observer such as a researcher or administrator. Beta press is how the environment is perceived from within the

environment itself by students or teachers. The beta press concept was further refined by Stern, Stein and Bloom (1956) into 'private beta press' and 'consensual beta press'. Private beta press is the individual's perception of his or her environment. In a classroom setting, private beta press would be how an individual student sees his or her classroom environment from his or her own personal perspective. Consensual beta press is how an individual would judge the environment from a group perspective. In the classroom, this would be how an individual feels that the class as a whole perceives the classroom. This early research formed the foundation for more recent work in the learning environment field.

It can be said that modern research specifically focusing on learning environments in classrooms began with the work of Herbert Walberg and Rudolf Moos. Walberg developed the Learning Environment Inventory (LEI) as a part of research related to Harvard Project Physics (Walberg & Anderson, 1968). The LEI was used to evaluate the effectiveness of the project. Moos created psychosocial environment scales used in many settings such as hospitals and prisons (Moos, 1974). Eventually his studies resulted in the creation of the Classroom Environment Scale (CES) (Moos, 1979; Moos & Trickett, 1974). The final version of the CES contains 90 True or False questions divided into nine 10-item scales.

The Individualized Classroom Environment Questionnaire (ICEQ) (Fraser, 1990) was developed specifically to assess dimensions that differentiate an individualized classroom from a conventional classroom. Development of the ICEQ began in 1979 (Rentoul & Fraser, 1979) as a response to individualized, open and inquiry-based education. The ICEQ contains 50 items in five scales of 10 items each. Each item is scored with a five-point frequency response scale ranging from 'almost never' to 'very often'. Some of the items are negatively worded and scored in reverse.

2.3 LEARNING ENVIRONMENT INSTRUMENTS

Can scales from classroom environment instruments be used to predict student outcomes? If consistent relationships between learning environment and student outcomes such as efficacy, attitude and achievement can be found in the literature, then the scales that are found to be reliable indicators of these student outcomes can be used to make predictions as to the effectiveness and/or implied superiority of a particular curricular arrangement.

The subsections below review, in historical sequence, three important and frequently-used instruments in the field of classroom environment, namely, the Science Laboratory Environment Inventory (SLEI) (Section 2.3.1), Constructivist Learning Environment Survey (CLES) (Section 2.3.2) and What Is Happening In this Class? (WIHIC) (Section 2.3.3). The development of each of these instruments built upon historically important instruments introduced in Section 2.2 (LEI, CES, ICEQ). The Questionnaire on Teacher Interaction (QTI) is built on different theoretical foundations from other learning environment instruments, but is reviewed for completeness (Section 2.3.4.) Finally, Section 2.3.5 briefly discusses other less widely-used or more-recent instruments from the field of learning environments, again for completeness.

2.3.1 Science Laboratory Environment Inventory (SLEI)

The Science Laboratory Environment Inventory (SLEI) assesses the learning environment in science laboratory classrooms. The SLEI is an economical survey

that is appropriate for measuring learning environments in upper-secondary and tertiary educational settings (Fraser, Giddings & McRobbie, 1995). The SLEI is economical in that it contains only 35 items to measure five scales: Student Cohesiveness, Open-Endedness, Integration, Rule Clarity and Material Environment. While the SLEI is designed specifically for use in laboratory classes, its scales used were culled from several previous instruments used in non-laboratory settings. Fraser, Giddings and McRobbie (1995) interviewed numerous teachers and students at both the high-school and university levels to identify the most salient scales.

The SLEI is also noteworthy in that it was one of the first instruments mentioned in the literature that addresses the problem, identified by Fraser and Tobin (1993), that previous classroom environment instruments only measure a student's perception of the class and not his or her place within that class. Fraser, Giddings and McRobbie (1995) developed two versions of the SLEI, a Personal and a Class form, to accommodate this need in the field.

Fraser, Giddings and McRobbie (1995) validated the SLEI with a sample of 3,727 high school students in 198 classes and 1,720 university students in 71 classes in six different nations (Australia, United States, Canada, Israel and Nigeria). In Australia, Fraser and McRobbie (1995) cross-validated the SLEI with a second sample of 1594 high-school students in 92 classes, while Fisher, Henderson and Fraser (1997) cross-validated the SLEI with 489 high-school biology students. In all cases, the five SLEI scales showed reliabilities of 0.74 to 0.91, which suggests good reliability for scales consisting of just seven items.

Lightburn and Fraser (2007) used a modified version of the SLEI in their study of 761 high-school biology students in Florida. In that study, 186 students

participated in a set of new anthropometric (body measurement) activities, while 574 students who learned via more traditional laboratory activities served as a control group. Lightburn and Fraser (2007) used four of the SLEI scales (omitting Open-Endedness). In addition to the SLEI's learning environment scales, Lightburn and Fraser (2007) used scales from the TOSRA, which is discussed in detail later in this chapter, and the Modified Fennema-Sherman Science Attitude Scales (Doepken et al., 1993). In all, six attitude scales were selected in order to examine connections. As with the previous studies discussed in this section, all four of the SLEI scales, as well as the attitude scales, were found to be internally consistent with alpha reliabilities between 0.74 and 0.95. Also, factor analysis confirmed the *a priori* structure of the survey.

Quek, Wong and Fraser (2005) utilized the Chemistry Learning Environment Inventory (CLEI), which is a chemistry-specific version of the SLEI that was first used by Wong and Fraser (1995) and by Wong, Young and Fraser (1997). The CLEI scales are the same five scales used in the SLEI. Quek, Wong and Fraser (2005) also made use of scales from the Questionnaire on Teacher Interaction (QTI), which is discussed later in this chapter, and a variation of the TOSRA, which also is discussed later in this chapter. Quek, Wong and Fraser (2005) sampled 497 gifted and non-gifted students in Singapore and examined difference between streams (gifted vs. non-gifted) and genders in SLEI scores. For the most part, the CLEI scales were found to be factorially valid and internally consistent.

2.3.2 Constructivist Learning Environment Survey (CLES)

The Constructivist Learning Environment Survey (CLES) was originally developed in 1991 (Taylor & Fraser, 1991) to examine constructivist teaching practices. It was then revised in 1997 (Taylor, Fraser & Fisher, 1997) with several format and content changes to address problems identified in a series of studies. One refinement was reducing the number of negatively-worded questions which many respondents had found confusing. Another refinement was changing the arrangement of the items for each scale from a cyclic pattern to a block pattern that grouped the questions by scale. As it now stands, the CLES consists of 30 items arranged in five scales: Uncertainty of Science, Critical Voice, Personal Relevance, Shared Control, and Student Negotiation. Dorman (2000) used three of these scales in his study of efficacy.

When Johnson and McClure (2004) used a 30-item version of the CLES with 290 American teachers and pre-service teachers from the upper-elementary middle-school and high-school levels, they reported strong factorial validity and reliability. Johnson and McClure (2004) also created a 20-item version of the CLES that was field tested with a different sample of teachers within the same set of grade levels. This smaller and thus more efficient version of the CLES was also found to be valid and reliable. In Texas, Nix, Fraser and Ledbetter (2005) used a modified version of the CLES to study the effectiveness of a teacher education program. In their study of 1079 students from 59 classes in Texas, the CLES was found to be valid and reliable. Nix and Fraser (2011) again successfully used this relatively short 20-item version of the CLES in their study of 17 teachers and 845 middle-school students over three semesters. Nix and Fraser (2011) were able to demonstrate that changes in the

teachers' university classroom environment were associated with changes in their students' perceptions of the classroom environment, as well as to confirm the factorial validity and internal consistency of the CLES.

The CLES has been successfully validated in multiple locations and languages. Aldridge et al. (2000) conducted an international study utilizing the CLES with 1,081 students in 50 Australian classes and 1,879 students in 50 Taiwanese classes. Their study was able to validate the CLES in both English and Mandarin. Both the English and Mandarin versions of the CLES showed similar results for factor structure and internal reliability.

As part of an effort to improve teachers' self-reflection and constructivist orientation of mathematics classes, Aldridge, Fraser and Sebela (2004) validated a modified version of the CLES with 1864 students in 43 grade 4–9 classes in South Africa. Aldridge et al. found the CLES to be both valid and reliable for use in South Africa.

Peiro and Fraser (2009) translated the CLES into Spanish and modified it for use with 739 early-elementary (K–3) students in South Florida. Both the English and Spanish versions were found to be valid for use with these very young students.

In order to study the effectiveness of a new curriculum, Kim, Fisher, and Fraser (1999) cross-validated a Korean-language version of the CLES with 1,083 10th grade students in 24 different classrooms. Analysis confirmed the structural validity and reliability of the CLES and showed that the new curriculum led to more positive perceptions of the classroom environment among students who were taught using the curriculum.

2.3.3 What Is Happening In this Class? (WIHIC) Survey

The WIHIC evolved from a 90-item survey with nine scales in 1996 to the final form with 56 questions distributed among seven scales (Fraser, McRobbie, & Fisher, 1996). The WIHIC scales address a broader range of environment issues than the CLES which has a specific constructivist focus. When Dorman (2000) used all seven WIHIC scales (Student Cohesiveness, Teacher Support, Involvement, Investigation, Task Orientation, Cooperation, and Equity) together with three from the CLES, all scales positively correlated with student efficacy. While statistically significant, the correlations between efficacy and three CLES scales were small. The seven WIHIC scales had larger correlations with efficacy. The strongest correlation was found for the WIHIC scale of Task Orientation (the importance placed on completion of assignments and sticking to the subject matter being studied). The sample item offered by Dorman (2000) for Task Orientation is “I pay attention in class.” Classrooms that scored high on this scale were more than two points higher on a seven-point efficacy scale.

More recently, Task Orientation was shown to be associated with student achievement by Allen and Fraser (2007). One of the research questions posed by Allen and Fraser (2007) was whether there is an association between students’ or their parents’ perceptions of the learning environment in science classrooms. Using a modified version of the WIHIC, Allen and Fraser (2007) surveyed 520 fourth and fifth grade students from a large urban school district and their parents. Firstly, the language was simplified for use by the 9–11 year-old students. Secondly, 17 items were removed from the questionnaire, including the entire Cooperation scale, thus leaving only six scales (Student Cohesiveness, Teacher Support, Involvement,

Investigation, Task Orientation, and Equity). Thirdly, the items in the parent version of the questionnaires were modified from the first person (e.g. “I” and “my”) to the third person (e.g. “My child”). Similar to Dorman (2000), Allen and Fraser (2007) used the WIHIC along with the Test of Science Related Attitudes (TOSRA) and the science subtest from the Stanford Achievement Test (SAT) to investigate connections between learning environment and student outcomes.

Unlike Dorman (2000), Allen and Fraser (2007) used some qualitative research to triangulate with the quantitative data. They interviewed 10 parent-student pairs and made classroom observations of the 10 students. These interviews helped to clarify the responses and went some distance in explaining the quantitative results. Statistical analysis of the data revealed that there was a weak correlation between achievement and students’ perceptions of the learning environment, but that there was a much stronger correlation between Task Orientation and attitudes. Parents’ perceptions of the environment were more strongly correlated with achievement, especially Task Orientation and Investigation. It is interesting to note that students were generally more satisfied with their environment than their parents were. However, the interviews suggested that the parents had a limited view of the classroom and were making interpretations of the environment via student anecdotes, homework assignments and a science fair project.

Another very recent study involving associations between WIHIC scales and achievement was conducted in Southern California by Rita and Martin-Dunlop (2011). The researchers surveyed 261 high school biology students using the WIHIC and related students’ actual perceived learning environment to their scores on the California Standards Test (CST)–Biology. The test consists of 60 objective questions covering a broad spectrum of biology topics. Rita and Martin-Dunlop (2011) found

correlations between the WIHIC scales of Teacher Support, Involvement, Investigation, Task Orientation Cooperation and Equity and scores on the achievement test. Student Cohesiveness showed a negative correlation with the test scores. Not surprisingly, the researchers reported that students preferred a more favorable learning environment than the one that they actually perceived. This is a consistent finding in many learning environment studies (Fraser, 1982).

The WIHIC appears to be a very useful instrument for investigating links between learning environment and student outcomes. Dorman (2003) conducted a rather thorough statistical investigation of the validity of the instrument in a study in three countries (Australia, Canada and the UK) involving grade 8, 9 and 12 students of both genders. In addition to usual validity and reliability indices seen in most studies using questionnaires, Dorman (2003) used exploratory and confirmatory factor analyses. Exploratory factor analyses identifies the structure for the instrument, whereas confirmatory analyses verifies an a priori structure. Essentially Dorman (2003) used a weighted least squares method to confirm that each of the WIHIC items belongs to its scale and only its scale. Dorman's (2003) analysis confirmed that the WIHIC produced consistent results in all of these different settings.

Dorman (2003) mentioned that his study was constrained to Western, English-speaking countries and that therefore further research is needed in different cultures, etc. Such additional research has been performed and the WIHIC has been validated in other languages, countries and cultures. For example, Wolf and Fraser (2008) utilized the WIHIC in a study of New York schools to compare inquiry-based versus non-inquiry-based science instruction with 1,434 middle-school students in 71 classes. Wolf and Fraser (2008) found that the learning environment scales from the

WIHIC were valid and reliable, and they reported relationships between learning environment scales and student attitudes toward science.

Koul and Fisher (2005) conducted a study utilizing the WIHIC in seven schools in India. The 9th and 10th grade students surveyed were from four major cultural groups found in the city where the study took place. Hindi, Kashmiri, Dogri, and Punjabi students were all represented in the sample of 1021 respondents. Koul and Fisher (2005) demonstrated that the WIHIC was valid in India and in multicultural situations similar to those found in many places around the world.

Of significant note is Aldridge, Fraser and Huang's (1999) study of classroom environments across two countries (Australia and Taiwan) and involving two languages (English and Mandarin). They used a method of translation and back translation to ensure that the meaning of the survey items remained intact. Aldridge et al. (1999) were able to show that the WIHIC was valid for both countries, both languages and two very different cultures. Aldridge et al. (1999) also provided an example of the effective use of mixed research methods by adding in qualitative interviews to triangulate findings emerging from the quantitative study.

Kim, Fisher, and Fraser (1999) used versions of the WIHIC along with the QTI that had been translated into the Korean Language in their study of 542 students in 12 Korean schools. Kim et al. used a method of translation and back translation method that was similar to Aldridge et al. (1999) and obtained very similar validation and reliability results.

Khoo and Fraser (2008) used a modified version of the WIHIC to study learning environments among students in adult computer education classes in Singapore. Khoo and Fraser (2008) showed that the WIHIC was valid in adult

educational settings, not just in the primary and secondary classroom as in previous studies.

MacLeod and Fraser (2010) translated the WIHIC into Arabic and validated it through the use of parallel Arabic and English versions. Students chose in which language to take the test. Statistical analysis showed strong factorial validity and reliability for both the preferred and actual forms. MacLeod and Fraser (2010) used a modified form of the WIHIC that included three scales, Teacher Support, Involvement and Cooperation, as well as the two additional scales of Relevance of Learning and Use of Laptops.

Also, in the United Arab Emirates, Afari, Aldridge, Fraser and Khine (2013) used an Arabic version of five WIHIC scales (Student Cohesiveness, Teacher Support, Involvement, Cooperation and Equity) along with the Personal Relevance scale from the CLES. An Enjoyment of Mathematics Lessons scale was adapted from the Enjoyment of Science Lessons scale from the TOSRA and a Student Efficacy scale was adapted from Jinks and Morgan (1999) Student Efficacy Scale (MJSES). Afari et al. (2013) used this modified WIHIC questionnaire to evaluate the use of games in teaching mathematics in university classes. Like MacLeod and Fraser (2010), the questionnaire was translated into Arabic. Unlike MacLeod and Fraser (2010), the Arabic translation of each item was printed underneath the English-language item. Comparison of pretest and posttest administrations of the questionnaire showed that use of the games in the mathematics classes had a small but positive effect on Teacher Support, Involvement and Personal Relevance, Enjoyment and Academic Efficacy. Again, analyses confirmed the validity and reliability of the WIHIC in Arabic.

When Waldrup, Fisher and Dorman (2009) used the WIHIC as a tool to identify exemplary science teachers in Australia, they found difficulty in defining precisely what an exemplary teacher is. Because there are many factors that combine in multiple ways, widely-varied teaching practices, techniques and knowledge sets can contribute to or detract from a teacher's effectiveness. Direct observation of teachers is an expensive and time-consuming process and results can be suspect because of the subjective and anecdotal nature of occasional observation. Classroom environment is, in part, dependent upon the ability of the teacher but offers a broader view of students' classroom experiences than one-off observations. With tools such as the WIHIC, classroom environment can be easily measured. Waldrup, Fisher and Dorman (2009) used five WIHIC scales (Student cohesiveness, Teacher support, Involvement, Task orientation and Equity) to identify exemplary teachers as those who had scores of more than one standard deviation above the mean on three or more of the WIHIC scales.

WIHIC scales have also been used to define broad types of classroom environments. den Brok et al. (2010) examined 52 biology classrooms with 1,474 students in Turkey and found six distinct types of classroom environments based upon average scores on the various WIHIC scales. The typologies identified included low effective, high effective, mainstream, self-directed, task-oriented cooperative, and task-oriented individualized classrooms. Not surprisingly, mainstream classrooms, those with medium-high scores, were the most common.

2.3.4 Questionnaire on Teacher Interaction (QTI)

The previously discussed questionnaires (SLEI, CLES, WIHIC) have a common structure in that they have a number of items, usually scored on a five-point range, grouped into a number of scales. These are then reported linearly and independently of each other. These surveys form the basis for the questionnaire used in my study and receive the most attention in this review. However, for completeness, it is desirable to review another widely-used questionnaire.

The Questionnaire on Teacher Interaction (QTI) measures student perceptions along two axes or dimensions (Wubbels & Levy, 1993). The first is a *proximity* dimension which ranges from opposition (O) to cooperation (C). The second is an *influence* dimension which ranges from dominance (D) to submission (S). The eight scales of the QTI can be described as belonging to four quadrants. In the Dominance/Opposition quadrant are OD Admonishing Behavior and DO Strict Behavior. In the Dominance/Cooperation quadrant are found DC Leadership Behavior and CD Helping/Friendly Behavior. CS Understanding Behavior and SC Student Responsibility/Freedom Behavior are within the Cooperation/Submission quadrant. Finally the Submission/Opposition quadrant contains SO Uncertain Behavior and OS Dissatisfied Behavior.

In Turkey Telli, den Brok and Cakiroglu (2007) surveyed 7484 students in 55 schools from 13 cities representing all seven geographical regions of Turkey using a Turkish-language version of the QTI. This study supported the validity of the QTI. In India, Koul and Fisher (2005) reported support for the validity of the QTI for a sample consisting of 1021 students from 31 classes in seven private schools. The first use of the QTI in Australia by Rickards, den Brok and Fisher (2005) involved 6148

students and their 238 teachers from the four Australian states of New South Wales, Queensland, Victoria and Western Australia. Rickards et al. (2005) confirmed the factorial validity and reliability for the QTI for use in Australia.

The QTI was also found to be valid when used in China by Wei, den Brok, and Zhou (2009) for investigating connections between the interpersonal behavior of teachers of English as a Foreign Language and students' English fluency for a sample of 160 students from four classes in the southwest region. Goh and Fraser (1996) created a version of the QTI for use with Singaporean elementary students called the QTI-Primary (QTIP). Goh and Fraser (1998) validated the QTIP for use in Singapore with a sample of 39 classes with 1512 students aged from 10 to 11 years. Also, in Singapore, Quek, Wong and Fraser (2005) validated the QTI among 497 secondary school students and reported differences between gifted and non-gifted students and between genders. Kokkinos, Charalambous and Davazoglou (2009) translated the QTIP into Greek and validated it with 273 primary schools in 17 grade 5 and 6 classes on the island of Cyprus. Fraser, Aldridge, and Soerjaningsih (2010) validated an Indonesian-language version of the QTI with 422 university students in 12 classes at a private university in Indonesia.

2.3.5 Other Instruments

Although the above sections have reviewed literature pertaining to the most frequently-used learning environment instruments, for the sake of completeness, a range of other learning environment questionnaires is reviewed briefly in this section. Nolen and Haldyna (1990) also found possible connections between task orientation

and student outcomes. Nolen (2002) built upon this with a longitudinal study of classroom climates and examined the question “Are motivational orientations for science stable over the course of a school year, or are they shaped by elements of the learning environment?” (p. 352). Nolen (2002) surveyed 377 students in 22 9th grade introductory science classes with items that were drawn from previous studies and used a five-point Likert response format. The three scales of Monitoring, Elaboration, and Organization and Selection measured deep-processing strategies. These three scales were averaged into one overall score for analysis. Student motivations were assessed with Task Orientation, Work Avoidance and Ego Orientation scales. Three items each were included to form a Satisfaction with Learning scale, as well as a Cooperative climate scale. Finally 17 items formed two scales that measured a learning focus versus an ability focus. Nolen (2002) reported that classroom learning environment variables were significant predictors of both satisfaction and achievement in science.

Chang, Hsiao and Barufaldi (2006) conducted a study of 155 tenth grade students in four Taiwanese classrooms. Two classes were taught in a traditional teacher-centered learning environment and two were taught in a combination of a teacher-centered and student-centered environments. Chang et al. (2006) used an instrument called the Earth Science Classroom Learning Environment Instrument (ESCLEI), the Attitudes Toward Earth Science Inventory (ATESI), and the Earth Science Achievement Test (ESAT).

The ECSLEI is a survey made up of four scales each with 15 questions: Preferred Student Centered, Preferred Teacher Centered, Actual Student Centered and Actual Teacher Centered. In essence, the survey measures whether students prefer a student-centered or teacher-centered learning environment and whether they perceive

their environment as being teacher-centered or student-centered. Like all of the learning environment instruments examined, the ECSLEI uses a five-point response scale. The ATESI and ESAT were combined into the Earth Science Learning Outcomes Inventory. The combined ESLOI consists of 60 items, including 30 items for the ATESI portion with a five-point response format to gauge attitudes, and 30 multiple-choice items for the ESAT segment to judge students' mastery of the subject matter.

The goal of a study of Chang et al. (2006) was to determine if congruence between the students' preferred and perceived actual environments was related to better attitudes towards science and/or achievement. They called the distances between the preferred and actual environment data points on a graph the preferred-actual spaces. They found that distance made no difference to achievement but was correlated strongly with attitudes. Not surprisingly, the researchers found that reading skill was the strongest predictor of achievement. Variables such as reading skill should be controlled or at least accounted for in studies of connections between learning environment and achievement.

The Technology-Rich Outcomes-Focused Learning Environment Inventory (TROFLEI) (Aldridge, Dorman & Fraser, 2004) evolved from the WIHIC. It incorporates all seven WIHIC scales and includes the three additional scales of Differentiation from the ICEQ (which measures the extent to which a teacher accommodates students' differing interests and learning abilities), Computer Usage to gauge how students use computers to access information and collaborate with other students, and Young Adult Ethos to assess the extent to which student feel that they are treated as adults. The TROFLEI consists of 80 items which are presented in a side-by-side format so that respondents can answer an actual and preferred version of

an item at the same time. Aldridge, Dorman and Fraser (2004) first validated the TROFLEI in Western Australia and Tasmania with a sample of 1,249 students. Multitrait–multimethod modeling supported the questionnaire's construct validity and showed that both the preferred and actual forms shared a common structure.

The TROFLEI was used in the evaluation the success of a new post-secondary school that emphasized outcomes-focused teaching, rather than a set syllabus. Aldridge and Fraser (2008) studied changes in the perceptions of 4,146 grade 8–13 students over four years and were able to support the efficacy of the school's innovative programs. Aldridge and Fraser (2008) also reported differences between various groups within the sample, such as males and females or those taking college-entrance examinations as opposed to those who weren't.

Welch, Cakir, Peterson, and Ray (2012) investigated the cross-cultural validity of the TROFLEI with 985 Turkish and 131 American high-school students. The TROFLEI was translated and back translated three times to ensure that both the English-language and Turkish-language versions were equivalent. Analysis using Cronbach's alpha coefficient showed that all the scales in both versions were internally consistent. Confirmatory factor analysis was undertaken, with a chi-square test (χ^2), Comparative Fit Index (CFI) and root mean square error of approximation (RMSEA) being reported. These techniques confirmed that the TROFLEI has satisfactory structural validity.

Similar to the TROFLEI, the Constructivist-Oriented Learning Environment Survey (COLES) (Aldridge, Fraser, Bell & Dorman, 2012) is based on the WIHIC. Six WIHIC scales (Student Cohesiveness, Teacher Support, Involvement, Task Orientation, Cooperation and Equity) were combined with the Differentiation and

Young Adult Ethos scales found in the TROFLEI and the Personal Relevance scale from the CLES. In addition to these nine learning environment scales, Aldridge et al. (2012) included two new scales (Formative Assessment and Assessment Criteria) related to the assessment of student learning which had not been addressed in any previously-existing learning environment instrument.

With a sample of 2,043 grade 11 and 12 students from 147 classes from nine schools in Western Australia, Aldridge et al. (2012) reported that data analyses supported the factorial validity of internal consistency of both the preferred and actual forms of the COLES. It is worth noting that the researchers also applied the Rasch model to convert the data into interval data. The results from analysis using the Rasch model were virtually identical to those using raw score data. Aldridge et al. (2012) also reported the experiences of teachers using the COLES in their own action research using interviews and reflective journals.

An older but still widely-used instrument is the My Class Inventory (MCI), which was originally developed as a simplified version of the LEI to be used with younger children aged 8–12 years. The MCI was first created by Fraser, Anderson and Walberg (1982) and then further evolved and shortened to 25 items by Fraser and O'Brien (1985). The final form of the MCI contains both a long 38-item version and a short 25-item version. In order to be more accessible for elementary students, the MCI features four changes from the LEI: as an accommodation to shorter attention spans among younger children, the MCI utilizes five scales instead of the LEI's 15; the wording of each of the items was simplified for younger readers; the MCI uses a two-point response format (Yes or No) instead of the LEI's four-point response format; and the respondents can answer the questions right on the questionnaire page rather than using a separate answer sheet.

Goh and Fraser (1998) modified the MCI by using a three-point frequency response scale consisting of Seldom, Sometimes and Most of the Time by adding a Task Orientation scale in their study of 1512 primary students in mathematics classes in Singapore. The MCI was found to have good internal reliability. The discriminant validity of the scales was satisfactory, but some overlap between the scales was discovered. Nevertheless, factor analysis supported the independence of factors scores for this version of the MCI.

Majeed, Fraser and Aldridge (2002) used the MCI with a sample of 1,565 students in 81 mathematics classes in Brunei. With the exception of the Satisfaction scale, which was removed, Majeed et al. found a satisfactory factor structure and solid reliability for the three remaining scales (Cohesiveness, Difficulty and Competition).

Sink and Spencer (2005) validated the MCI with a sample of 2,835 upper-elementary students in an urban area of Washington State. Sink and Spencer advocate the use of the MCI in evaluating school counselors' programs and practices. All four MCI scales were found to be valid and internally reliable.

In Texas, Scott Houston, Fraser and Ledbetter (2008) evaluated the use of science kits by using the MCI with 588 grade 3–5 students. Three of the scales were found to have a satisfactory factorial structure (Cohesiveness, Competition and Friction). The Difficulty scale was removed because of poor factor loadings. Scott Houston, Fraser and Ledbetter reported that higher Cohesiveness among students who used the science kits.

2.4 STUDENT ATTITUDES TOWARDS SCIENCE

My study compared two middle-school science curriculum approaches (general and specific) in terms of both the classroom learning environment and students' attitudes. Whereas the previous sections in this chapter reviewed literature on conceptualizing, assessing, and investigating classroom environment, attention in this section moves to a review of literature on the conceptualization, assessment and investigation of attitudes to science.

It is safe to assume that most, if not all, science educators want their students to have positive attitudes towards science. Besides the personal satisfaction that an educator receives from students' sharing his/her own passion for science, there are practical reasons to promote positive attitudes. Mager (1968) mentions three reasons why a positive student attitude is desirable. First, there are associations between positive attitudes and student achievement. Second, positive attitudes on the part of the student result in his or her future interest in the subject. Finally, a student with positive attitudes has a positive influence on his or her peers.

Attitudes have been incorporated frequently into classroom environment research (Fraser & Butts, 1982; Fraser & Fisher, 1982; Kim et al., 2000). All of these studies have shown strong connections between student attitudes towards science and learning environment scales.

Fraser and Butts (1982) studied the relationship between classroom environment and student attitudes towards science among 712 middle-school students. Five classroom environment scales were selected from the Individualized Classroom Environment Questionnaire (ICEQ) and the Test of Science Related Attitudes

(TOSRA, Fraser 1981) was used to evaluate seven attitude scales. Multiple regression analysis showed significant connections between student attitudes towards science and the ICEQ scales for four of the seven TOSRA scales.

Fraser and Fisher's (1982) study was somewhat similar to Fraser and Butts (1982), but larger in scope and scale. Fraser and Fisher involved 1,083 middle-school students in 116 different classrooms in investigating connections between classroom environment and student outcomes. Learning environment was assessed with the ICEQ and the CES. Among the outcomes measured were student attitudes from the TOSRA and cognitive abilities measured with the Test of Enquiry Skills (Fraser, 1979). In total, Fraser and Fisher (1982) related 14 environment scales with nine outcome measures. Numerous connections between environment scales and some of the attitude scales were reported, such as positive attitudes towards the Social Implications of Science appear in classrooms with higher Participation and Order and Organization.

Kim, Fisher and Fraser (2000) studied 12 classrooms in South Korea encompassing 543 students. Seven scales from WIHIC (Student Cohesiveness, Teacher Support, Involvement, Investigation, Task Orientation, Cooperation, and Equity) were used in conjunction with all eight of the QTI scales and one scale from the TOSRA was used to assess student attitudes. The strongest relationships between WIHIC scales and attitudes were found for Equity, Teacher Support and Task Orientation.

2.4.1 Test of Science Related Attitudes

The Test of Science Related Attitudes (TOSRA) was developed and validated by Fraser (1978, 1981) in order to assess how students feel about science as a subject. The TOSRA consists of seven scales that measure distinct aspects of student attitudes based on categories defined by Klopfer (1971). Social Implications of Science and Normality of Scientists correspond to Klopfer's H.1 category (Manifestation of favorable attitudes to science and scientists). Attitude to Scientific Inquiry is analogous to the H.2 category (Acceptance of scientific inquiry as a way of thought). Adoption of Scientific Attitudes relates to category H.3 of the same name. Enjoyment of Science Lessons is derived from Category H.4 (Enjoyment of science learning experiences). Leisure Interest in Science is associated with H.5 (Development of interest in science and science related activities) and Career interest in Science is from H.6 (Development of interest in pursuing a career in science). TOSRA items are scored using a five-point Likert format with answer choices of Strongly Agree, Agree, Not Sure, Disagree and Strongly Disagree. The original form of TOSRA includes many negatively-worded and reverse-scored items.

The initial validation the TOSRA involved 1337 students in grades 7–10 from 44 classes, one at each grade level, from 11 different schools. The schools used in the study were selected to cover a wide range of geographic regions and socioeconomic situations. The internal consistency of the TOSRA scales was found to be fairly high, with Cronbach α coefficients ranging from 0.66 to 0.93, especially considering that each scale consists of only 10 items. Relatively low intercorrelations between the seven TOSRA scales indicated satisfactory discriminant validity. TOSRA has been

crossvalidated in Singapore by Wong and Fraser (1996) and in Indonesia and Australia by Fraser, Aldridge and Adolphe (2010).

2.4.2 Other Attitude Instruments

In addition to the TOSRA, there are other attitude instruments in use in the field. Seker (2011) created a questionnaire called the School Attitude Questionnaire (SAQ) to assess and investigate attitudes among students aged from 11 to 13 years old. The SAQ has a five-point Likert response scale, with three of the responses being positive (I strongly agree, I agree, and I partially agree) and two negative (I disagree and I strongly disagree). Factor analysis reduced the original 55 questions to 22, nine of which are reverse scored. The SAQ consists of the five scales of Teaching, School Image, Loneliness at School, Testing and Feedback, Reluctance, and Belongingness.

Van Rensburg, Ankiewicz and Mayburgh (1999) used the Pupils Attitudes Towards Technology (PATT), developed by Bame, Dugger, de Vries and McBee (1993), to explore attitudes among 1010 student in the Guateng province of South Africa. The PATT was validated in the United States and has been used in approximately 20 countries (de Vries, 1991). The PATT consists of 11 demographic items, 89 Likert type items and one open-ended question. The PATT is made of numerous scales: interest in technology (INT), the role of gender in technology (GEN), the importance and consequences of technology (CON), the difficulty of technology (DIF), technology as part of the school curriculum (CU), technology as human activity with societal aspects (SOC), technology as related to science (SCI), technology as a process of designing, making and using (PROC), and technology as

processing matter, energy and information, sometimes called the ‘pillars’ of technology (MEI).

2.5 APPLICATIONS OF LEARNING ENVIRONMENT INSTRUMENTS

Literature reviews (Fraser 1998a, 1998b, 2012) reveal that learning environment instruments have been used for many purposes. For example, they have been used in investigations of associations between the learning environment and outcomes (see Section 2.5.1), evaluations of educational innovations (see Section 2.5.2), and action research by teachers seeking to improve their classroom environments. Sinclair and Fraser (2002) tracked three teachers' use of the WIHIC as a feedback tool to guide improvements in classroom environments. Aldridge, Fraser and Ntuli (2009) studied 31 South African teachers who used an elementary-school version of the WIHIC in guiding improvements in their teaching and classroom environments. To a lesser extent, learning environment assessments have been used to examine differences between students' and teachers' perceptions of the same classroom environments. For example, when Fisher and Fraser (1983) conducted a study with a sample of 56 and 116 classes using the ICEQ, they found that students preferred a more positive learning environment than the one that they perceived, and that teachers perceived the learning environment more positively than their students. Learning environment research has also been undertaken to test the importance of person–environment fit (i.e. whether a respondent's preferred environment matches his/her actual environment) for achievement or attitudes. Fraser and Fisher (1983a, 1983b) used actual classroom environment and person-environment fit (defined as the

degree of match between actual and preferred environment) as predictors of achievement and attitudes.

One of my study's aims involved the use of the learning environment as an indicator of instructional effectiveness. There is convincing support for this approach within the literature (Lightburn & Fraser, 2007; Wolf & Fraser, 2008). The history of classroom environments research has an abounding tradition that is centered on exploring connections between students' perceptions of their classroom environment and student learning outcomes (Fraser, 2012). The following subsections explore those applications that relevant to my study. Section 2.5.1 reviews studies that have established associations between classroom environment and student outcomes. Section 2.5.2 reviews past uses of learning environment scales in evaluating classroom innovations. Because I also investigated ethnic differences in learning environment perceptions and attitudes (as well as the differential effectiveness of different curriculum sequences for students of different ethnicities), past studies of ethnic differences in learning environment perceptions are reviewed in Section 2.5.3. Because middle-school curricula and students were a major focus in my study, Section 2.5.4 is devoted to reviewing past studies of changes in the learning environment across the transition between different levels of education.

2.5.1 Associations Between Classroom Environment and Student Outcomes

There have been many studies that have demonstrated that classroom environment accounts for substantial amounts of variance in a wide variety of cognitive and affective outcomes. Fraser (1994) lists 40 separate studies that revealed

a relationship between perceived classroom environment and student outcomes. These studies used a variety of learning environment instruments, such as those detailed above and others, with a variety of learning outcomes and a broad assortment of sample types and sizes. In one of the earliest studies of the connection between classroom environment and outcomes, Walberg and Anderson (1968) showed strong relationships between learning environment scales and outcome measures and set the tone for almost half a century science education research. Walberg and Anderson (1968) used a preliminary version of the LEI, called the Classroom Climate Questionnaire, to survey 76 classes ranging in size from 76 to 96 students. When examining simple correlations between the learning environment dimensions and the outcome dimensions, Walberg and Anderson found four times as many significant correlations than would be expected by chance. Multiple correlation analysis revealed a significant influence of the Classroom Climate Questionnaire scales on achievement.

McRobbie and Fraser (1993) utilized the SLEI in a study of 1,594 high school chemistry students in 92 classes from 52 schools in Brisbane, Australia. All the students answered the SLEI but several different outcome instruments were used, with a subset of the student sample responding to different instruments. McRobbie and Fraser utilized three types of correlations (simple, multiple, and canonical) and consistently found associations between the SLEI scales and the six outcome measures of Conclusions and Generalizations, Design of Experiments, Attitude to Laboratory Learning, Nature of Chemistry Knowledge, Cooperative Learning and Adoption of Laboratory Attitudes. The results were consistent at the individual and class levels.

In a related study, Fraser and McRobbie (1995) cross-validated the SLEI with a sample consisting of 5,447 students in 269 classes across six nations (Australia, United States, Canada, England, Israel, and Nigeria) and related responses to the SLEI scales to four chemistry-related attitude scales: Attitude to Laboratory Learning, Nature of Chemistry Knowledge, Cooperative Learning and Adoption of Laboratory Attitudes. Fraser and McRobbie (1995) also reported numerous significant relationships between the learning environment scales and the outcome dimensions.

Telli, den Brok, and Cakiroglu (2010) conducted a study in Turkey with a Turkish-language version of the QTI and TOSRA with a sample of 7,484 high school students in 278 classes in 55 schools all across Turkey. Multilevel analyses established relationships between enjoyment of science and the Influence scale of the QTI, while the QTI Proximity scale was correlated with attitudes toward inquiry.

The WIHIC has had a strong and recent history of use in the study of associations between learning environment and outcomes, including many of the studies previously described in this chapter. Aldridge et al. (1999) found associations between WIHIC scales and Enjoyment of Science Lessons in Australia and Taiwan. In Singapore, Khoo and Fraser (2008) linked student satisfaction with classroom environment in their study of 250 adult computer education students in five classes. In the UAE, Afari et al. (2013) linked enjoyment and self-efficacy to positive learning environments. Wolf and Fraser (2008), in New York with a sample of 1,434 middle-school students in 71 classes, and Allen and Fraser (2007), with a sample of 520 grade 4–5 students in Miami, linked WIHIC scales with attitudes towards science and achievement..

Chionh and Fraser (2009) used the WIHIC with 2,310 high school students in 75 geography and mathematics classes in 38 schools in Singapore. They found that student cohesiveness was predictive of higher achievement, and that self-esteem and student attitudes were higher in classrooms with positive teacher support, task orientation and equity.

This handful of studies, while certainly not an exhaustive list of all such research, confirm the hypothesis that a positive classroom environment is associated with positive student outcomes, including achievement, attitudes and self-efficacy, among others. Therefore, often it is reasonable to use the learning environment as indirect indication of student outcomes when a direct outcome measure is not be readily accessible or even available at all.

2.5.2 Evaluation of Educational Innovations

An old adage states that perception is reality. That statement has been used to mean different things but, in the world of education, students' perceptions of their learning environment can give educators and researchers valuable insights into the reality of the classroom. Classroom environment instruments such as the WIHIC or CLES can be used in evaluating a wide range of educational innovations. These innovations can be something as simple as a new classroom activity or something as extensive as a prolonged professional development program. The versatility, well-established validity, and ready availability of the various classroom environment instruments make them invaluable tools in this application.

Nix, Fraser and Ledbetter's (2005) research, which was discussed above in Section 2.3.2, involved using a modified form of the CLES to show that teachers who participated in an extended, three-semester professional development program had created classrooms where students perceived an appreciably higher levels of Personal Relevance and Uncertainty of Science compared with the students' other classes.

In the south-eastern United States, Pickett and Fraser (2009) evaluated a two-year mentoring program for beginning elementary teachers using a version of the WIHIC that had been modified for use by elementary students. Pretests and posttests were given to the 573 students of seven teachers who participated in the mentoring program. Statistical analyses, specifically MANOVA and effect sizes, showed that the program was effective in promoting improvements in the students' perceptions of classroom environments, attitudes and achievement.

Martin-Dunlop and Fraser (2008) used learning environment scales from the SLEI and WIHIC to evaluate the effectiveness of an innovative science course for preservice elementary teachers in California. 525 female preservice teachers in 27 classes participated in the study, which showed very large differences between the new course and previous courses for all learning environment scales. Effect sizes ranged from 1.51 standard deviations for Student Cohesiveness to 6.74 standard deviations for Open-endedness.

As discussed in Section 2.3.1 of this chapter, Lightburn and Fraser (2007) provide another example of a classroom environment instrument being used in the evaluation of new lesson activities involving anthropometry with 761 high-school biology students in Florida. The new activities were found to have a positive effect in terms of SLEI scales as well as attitude scales. Also, as previously discussed in

Section 2.3.2, Wolf and Fraser's (2008) study in New York with 1,434 middle-school students in 71 classes showed that inquiry-based instruction promoted higher Student Cohesiveness and revealed that inquiry-based lessons were differentially effective based on gender.

In South Florida, Holding and Fraser (2013) used the WIHIC and TOSRA in evaluating the effectiveness of the National Board certification (NBC) for teachers. The perceptions and attitudes 442 8th and 10th grade students in 21 classes taught by National Board Certified teachers were compared with those of 484 students in 17 classes taught by non-National Board Certified teachers. Significantly higher scores were found for five WIHIC scales (Teacher Support, Involvement, Task Orientation, Investigation and Cooperation) and student attitudes in the classrooms taught by National Board Certified teachers. The effect sizes ranged from small to modest (0.07 for Student Cohesiveness to 0.35 for Attitudes).

The preceding studies illustrate the utility of learning environment instruments for assessing the effectiveness of classroom innovations, such as teacher training and certification programs, alternative courses, new classroom materials and novel learning activities. This application of learning environment questionnaires was important for my study which used learning environment scales to assess the relative effectiveness of different curriculum sequences.

2.5.3 Ethnic Differences in Learning Environment Perceptions

In Singapore, Lim and Fraser (2012) investigated ethnic differences in learning environment perceptions (using the WIHIC) and attitudes. The sample

consisted of 279 Chinese and 89 Malay students at the upper-primary level. Although Malay students reported more positive perceptions of their learning environment and attitudes than Chinese students for all scales, differences were statistically significant only for Teacher Support and Involvement. For these two scales, effect sizes for ethnic differences were 0.34 and 0.53 standard deviations, respectively, placing them in the modest to medium range according to Cohen (1988).

When Levy, den Brok, Wubbels and Brekelmans (2003) compared African-American students and Asian-American students, they found that the former held more favourable perceptions of their teachers in terms of leadership, helpfulness and friendliness. The Asian-American counterparts, however, had less favourable perceptions, indicating that their teachers were stricter and gave them significantly less responsibility and freedom. These findings were surprising to the researchers because earlier research had shown no differences between African-American students and their peers. The findings for Asian-American students also contradicted earlier studies in which Asians perceived less dominance and more submissive behaviour (den Brok, Levy, Rodriguez & Wubbels, 2002; Levy, Wubbels, Brekelmans & Morganfield, 1997).

Fisher and Rickards (1997) investigated associations between science students' perceptions of their interactions with their teachers and the cultural backgrounds of students. A sample of 3215 students in Australia completed the QTI and an attitude to class scale. Students from an Asian background tended to perceive their teacher more positively than those from the other cultural groups of Africa, America and Oceania. They also found that children from Asia had significantly higher mean scores on the QTI scales of Leadership, Helping/Friendly, Understanding, and Student Responsibility and Freedom.

Khine and Fisher (2004) investigated whether a sample of 1188 secondary science students in Brunei perceived their teachers differently depending on whether their teacher was Asian or Western. Use of the QTI revealed that students perceived that teacher–student interactions were more positive for Western teachers than Asian teachers for most QTI scales.

Hoang (2008) investigated different factors (grade level, sex and ethnicity) in the attitudes and learning environment perceptions of high school mathematics students in the US. The WIHIC and an attitude questionnaire based on the Test of Mathematics-Related Attitude (TOMRA) were administered to 600 grade 9 and 10 mathematics students in 30 classes in one high school. Anglo students consistently reported more positive perceptions of classroom environment and attitude than did their Hispanic counterparts.

Tulloch (2011) used the CLES and TOSRA with 544 students in 29 tertiary classes in the US to investigate sex, age and ethnicity as determinants of classroom environment. No significant differences emerged between African-Americans and students of other ethnicities for any learning environment scale or for enjoyment. This confirmed an earlier study by Moss (2003) who found no ethnic (black versus non-black) differences in classroom environment for a sample in the US.

Findings are generally less consistent in the area of ethnicity differences, partly because of the complexity of such a construct. It is difficult to draw conclusive results because different groups tend to exhibit different responses in different contexts. However, this does not diminish the importance of ethnicity as a construct for the learning environment, and therefore I investigated this aspect in my study.

2.5.4 Transition Between Levels of Schooling

In Section 1.5 in the previous chapter, the difficulty faced by students transitioning from one level of schooling to the next, such as moving from elementary school to middle school or from middle school to high school, was briefly discussed. Learning environment constructs have been used by several researchers in investigating the important transitions in a child's education.

Speering and Rennie (1996) conducted a set of longitudinal studies with a sample of 163 students in Western Australia. Student attitudes towards science were assessed using a short survey and interviews both before and after the transition from primary school (grade 7) to secondary school (grade 8). The students' attitudes towards science, especially among girls, declined during the transition to the secondary school setting.

In Tasmania, Australia, Ferguson and Fraser (1998) made use of the MCI and QTI with a sample of 1040 students, first at 47 primary schools at the end of primary schooling and later at 16 linked secondary schools during the following school year. Ferguson and Fraser found that students' perceptions of learning environment generally deteriorated across the primary-to-secondary transition. However some scales showed improved scores after transition, and the extent of change across transition varied with school size and gender.

Feldlaufer, Midgley and Eccles (1988) studied 1788 mathematics students from 143 primary classrooms in 12 school districts in the region around Detroit, Michigan both before and after they graduated to secondary school. Feldlaufer and colleagues developed an instrument called the Student Classroom Environment Measure (SCEM) which was based on the CES, LEI, ICEQ and other surveys. The

SCEM scales included Cooperation/Interaction, Competition, Social Comparison, Teacher - Unfair/Unfriendly, and Teacher - Valuing of Mathematics. Student perceptions of their teachers did change from primary to secondary school because their mathematics teachers were more specialized in the secondary school than in the primary school.

Pointon (2000) used personal interviews in a small study of 13 students in East Anglia, United Kingdom at the end of their first year of secondary school to assess their perceptions of the transition from primary school. This qualitative research uncovered some problems that students face in their first year on secondary school, such as moving from classroom to classroom several times a day, different styles of classroom, seating differences, and a loss of a sense of classroom ownership.

Bru, Stones, Munthe and Thuen (2010) compared student perceptions of their learning environments in primary and secondary school using a survey of 7,205 students from 98 schools across 30 municipalities in Norway. The survey measured teacher support with scales including academic, emotional, and managerial support. Another scale measured autonomy support. This study did not identify any changes in student perceptions of teacher support across the primary-to-secondary transition.

2.6 CONCLUSIONS

The history of learning environments research spans half of a century. The early beginnings of Lewin and Murray in non-educational settings evolved into the modern classroom research pioneered by Herbert Walberg and Rudolf Moos, and

stimulated the growing cohort of researchers currently working to assess and investigate learning environments.

Several classroom environments instruments have been used by researchers in the field. This chapter briefly reviewed several historically-significant learning environment instruments (namely, the LEI, CES and ICEQ) and then reviewed more comprehensively three frequently-used questionnaires (SLEI, CLES, WIHIC) that built upon the LEI, CES and ICEQ. All of these instruments were shown to have solid validity and reliability, and to be fairly easy to customize for use in specific applications. For completeness, some other learning environment instruments were briefly reviewed (QTI, TROFLEI, COLES).

Student attitudes towards science is a frequently-examined learning outcome by learning environment researchers. According to a literature review, the TOSRA and its variations seem to be the most popular choice among researchers, but other instruments have been used as well.

It is clear from this brief review of the literature in this chapter that there are connections between learning environment and student outcomes. Section 2.6.1 in this chapter considered numerous examples of past research that established consistent and significant associations between learning environment scales and student outcomes such as attitudes, achievement and efficacy.

It is also clear that learning environment assessments are useful in evaluating the effectiveness of instruction. Learning environment criteria have been used in past evaluations of university-level teacher preparation programs, graduate programs for working science teachers, teaching materials, novel lessons and innovative science courses.

This chapter also reviewed past research on ethnic differences in learning environment perceptions, as well as prior studies of changes in classroom learning environment across the transition between different levels of education.

There is an opportunity to construct a new instrument using elements of previously-established instruments that have been shown to be predictive of student outcomes. Such an instrument could be used to gather data that would make it possible to compare the effectiveness of educational systems, curricula or instructional techniques across a wide variety of locations, cultures and situations.

Chapter 3

RESEARCH METHODS

3.1 INTRODUCTION

This chapter discusses the methods used in the study, including details of the two schools used in the study and about the specific sample of students involved (Section 3.2), the survey instrument used in this study to assess students' perceptions of their learning environment and students' attitudes toward science (Section 3.3), and the methods of data analysis used in answering the research questions (Section 3.4).

Following validation of my learning environment and attitude scales (my first research question), my study focused on comparing the effectiveness of two alternative curriculum sequences (general and specific) in terms of students' perceptions of learning environment and attitudes. The design for this second research question was quasi-experimental and included comparison of two groups of students who were not randomly selected but still were fairly typical. My third research question (involving associations between classroom environment and student attitudes) involved a correlational design.

3.2 SAMPLE SELECTION

In California and Texas, middle schooling encompasses grades 6, 7 and 8 and caters for students aged between 11 and 14 years. An attempt was made to select two schools, one in each state, that were as similar as possible in terms of demographics in

order to make the fairest possible comparison of curriculum arrangements models. The school selected to represent the specific curriculum sequence was a grade 6–8 middle school located in a suburb in Alameda County, in the San Francisco Bay area of Northern California. For the purposes of this doctoral thesis, this school is referred to as Bayside Middle School or Simply Bayside. Bayside's student body of 489 students was approximately 2% African–American, 8% Asian, 47% Hispanic, 41% Caucasian and less than 1% Native American. As shown in Table 3.1, the ethnic comparison of my sample of 200 students in this school was a little different. 46% of Bayside students were classified as economically disadvantaged.

The representative Texas school was also a grade 6–8 middle school in a suburb of Dallas, which is located in North Central Texas. This school is referred to in this thesis as Riverdale Middle School or simply Riverdale. Riverdale's population is 18% African–American, 6% Asian, 27% Hispanic, 49% Caucasian and less than 1% Native American. As shown in Table 3.1, my sample of 167 students from this school has a quite similar ethnic mix. 35% of Riverdale Middle School is considered economically disadvantaged.

I surveyed as many of the eighth grade students from each school as possible. In both situations, informed consent was secured from the students, their parents and school administrations before the surveys were administered to any students.

At Bayside Middle School in California, the school principal and the science teachers secured informed consent under the direction of the researcher. Data from the Bayside sample were collected via an online survey using the Survey Monkey service. The participating students used their school's computer laboratory when answering the online version of the Outcomes-Related Learning Environment Survey (ORLES), which is the instrument that was used in my study.

Informed consent at Riverdale Middle School in Texas was collected by the science teachers for their respective classes. The Riverdale sample was collected via paper surveys administered by the researcher in each individual class over the course of a single school day.

The total sample consisted of 367 grade 8 students in 15 classes (eight in Texas, seven in California) taught by six teachers (4 in Texas, 2 in California). The sample was nearly evenly split between genders with 51.2% males, 48.2% females, and 0.6% declining to answer. The total sample represented the ethnic diversity of the communities in which the schools are located (49.3% were Caucasian, 27.8% were Hispanic, 11.6% were African–American, 8% were Asian, 3.3% identified themselves as Native American or Other and 0.6% declined to answer).

Table 3.1 shows the gender and ethnic make-up of the sample for each state. This table suggests a relatively similar composition for each state's samples. The largest discrepancy between the two samples is in the proportion of African–American students, which was 6% for Bayside Middle School and 18.3% for Riverdale Middle School.

The ethnic mix at Bayside and Riverdale, as shown in Table 3.1, is fairly representative of schools for their respective states. According to the 2010 Census, California was 39.7% Caucasian, 6.6% African–American and 38.1% Hispanic (United States Census Bureau. n.d.). Texas, in 2010, was 63.4% Caucasian, 13.1% Black and 38.1% Hispanic (United States Census Bureau. n.d.). While these two schools don't have exactly the same ethnic composition as their states, the ethnic mix of the sample in each state still is fairly similar to the state mix.

Table 3.1

Gender and Ethnic Breakdown of Sample in Each State

State	Gender		Ethnicity				
	Male	Female	Caucasian	Native American/Other	Hispanic	Asian	African– American
California	109	90	97	9	67	15	12
	54.8%	45.2%	48.3%	4.5%	33.3%	7.5%	6.0%
Texas	77	85	82	3	34	18	30
	47.5%	52.5%	50.0%	1.8%	27.8%	8.5%	18.3%

3.3 LEARNING ENVIRONMENT AND ATTITUDE INSTRUMENT USED

In my study, students' perceptions of their learning environment and attitudes towards science were measured with an instrument created specifically for this study. The Outcomes-Related Learning Environment Survey (ORLES) is based on scales from the Constructivist Learning Environment Survey (CLES) (Taylor & Fraser, 1991), What Is Happening In this Class? (WIHIC) questionnaire (Fraser, McRobbie, & Fisher, 1996), and Test of Science Related Attitudes (TOSRA) (Fraser, 1978, 1981).

Chapter 2, Section 2.3.2 discussed the CLES and its validation in past studies in a variety of geographical locations and languages. Aldridge et al. (2000) validated the CLES in English and Mandarin in Australia and Taiwan. Nix and Fraser (2011) and Nix et al. (2005) validated the CLES in the Texas in English. Peiro and Fraser (2009) validated the CLES in both Spanish and English for use in Florida. Kim, Fisher, and Fraser's (1999) study involved using a Korean version the CLES in South Korea.

As mentioned in Chapter 2, Section 2.3.3, the WIHIC has been utilized and validated in many countries and languages. The WIHIC has been validated in

English-speaking nations around the world including Australia, Canada, and the U.K (Dorman, 2003, 2008), the United States (Allen & Fraser, 2007; Martin-Dunlop & Fraser, 2008; Wolf & Fraser, 2008), India (Koul & Fraser, 2005), Singapore (Khoo & Fraser, 2008) and South Africa (Aldridge et al., 2009). The WIHIC has also been translated and found to be valid in several non-English languages including Arabic (Afari et al., 2013; MacLeod & Fraser, 2004), Mandarin (Aldridge et al., 1999), Korean (Kim et al., 2000), and Spanish (Allen & Fraser, 2007).

The TOSRA is described in Section 2.4.1. It was originally validated in multiple locations in Australia by Fraser (1978, 1981). The TOSRA has been used in conjunction with numerous learning environment instruments and therefore been found valid in the same broad range of countries and languages as the WIHIC and CLES. Telli, den Brok, and Cakiroglu (2010) translated and validated TOSRA scales in Turkey. The TOSRA has been validated in English in Florida (Helding & Fraser 2013) and Singapore (Quek, Wong & Fraser 2005). Afari et al. (2013) translated the TOSRA into Arabic and found it to be valid in the United Arab Emirates.

Because all three of these instruments have demonstrated validity in dozens of studies in numerous countries, languages, and applications, their scales were ideal for use in constructing the ORLES for my study.

Table 3.2 provides details for each ORLES scale, including a description and sample item. The initial form of the ORLES consisted of 40 questions divided among the seven scales. While each of the scales was derived from its parent questionnaire, the number of items was reduced in order to create a more-economical instrument. Each item is answered on the same five-point frequency scale with the response choices of Almost Always, Often, Sometimes, Seldom and Almost Never.

Table 3.2

Scale Descriptions of Sample Items for Outcomes-Related Learning Environment Survey Scales (ORLES)

Scale	Item Nos.*	Scale Description	Sample Item	Parent Survey
Personal Relevance	4-8	The extent to which teachers relate science to students out-of-school experiences	I get a better understanding of the world outside of school.	CLES
Student Cohesiveness	9-14	The extent to which students are friendly and supportive of one another	I work well with other class members.	WIHIC
Teacher Support	15-20	The extent to which the teacher helps, encourages and is interested in the students	The teacher considers my feelings.	WIHIC
Investigation	21-26	The extent to which that problem solving and inquiry skills are emphasized	I carry out investigations to answer the teacher's questions	WIHIC
Involvement	27-32	students participate in class and take part in activities	I discuss different answers to questions.	WIHIC
Task Orientation	33-38	The extent to which it is important for to complete activities and stay on the subject matter.	Class assignments are clear so I know what to do.	WIHIC
Enjoyment of Science	39-43	The extent to which students enjoy science lessons	School should have more science lessons each week.	TOSRA

*These item numbers correspond to those shown on the ORLES in Appendix A.

Like the original instruments, all the ORLES items have a positive scoring direction, thus avoiding confusing negative wording. Also, all items are grouped together in blocks with the other items in their own scale as opposed to a random or cyclical arrangement. A copy of the ORLES can be found in Appendix A.

3.4 DATA ANALYSIS

The data from the Bayside sample were converted into a spreadsheet by the Survey Monkey service that the students used when responding to the online version of the ORLES. The Riverdale paper surveys were scanned by an optical scanner and software which created a spreadsheet containing the data. The optical scanner results

were visually verified to ensure accuracy. Both spreadsheets were combined to form a master database containing the entire sample. The data were then analyzed using the Statistical Package for Social Scientists (SPSS) software package. These data were analyzed to answer my research questions. Section 3.4.1 discusses analyses for the first research question, namely, the validity and reliability of the ORLES, Section 3.4.2 covers analyses for differences between curriculum sequences, ethnicities and the sequence-by-ethnicity interaction in terms ORLES scales. Section 3.4.3 discusses the methods of analysis for investigating associations between learning environment and attitudes towards science.

3.4.1 Validity and Reliability of the ORLES

In determining if the ORLES was valid a reliable for use with middle-school students in Texas and California, factor analyses were applied to the data. Factor analysis is a method to reduce a large number of variables into a small number of clusters called factors (Gay, Mills & Airasian, 2006). The technique used in my study was principal axis factor analysis with a varimax rotation and Kaiser normalization (Kaiser, 1958). If an item had a factor loading of 0.40 or higher with its own scale and less than 0.40 with all other scales, then it was considered to have a satisfactory factor loading. In my study, the two ORLES scales, of Involvement and Investigation were merged into one scale because of similar factor loadings; thus the ORLES ended up with a six-scale structure. Full details of the factor analyses are found in Chapter 4 of this thesis.

In analyzing the internal consistency and discriminate reliability of the ORLES scales, two separate units of analysis were used, namely, the individual

student and the class. These two units of analysis represent the concepts of 'private beta press' (the individual's view of his environment) and 'consensual beta press' (the view of the environment shared by a group) (Stern, Stein & Bloom, 1956).

Internal consistency reliability is a measure of the extent to which each item in a particular scale measures the same concept as the other items in that scale or the homogeneity of the scale. Cronbach's alpha coefficient for both units of analysis (individual and class mean) was used to establish the internal consistency of the six ORLES scales.

Discriminant validity is the complement to internal reliability, in that it measures the extent to which items in a scale are only related to items within their own scale, and not to items in other conceptually-distinct scales. The mean of the correlations between a scale and each of the other scales was calculated in order to check further the discriminant validity that was established in the factor analysis.

In order to examine the ORLES' ability to distinguish between classes, a one-way analysis of variance (ANOVA), with class membership as the independent variable, was used to measure the extent to which members within a single class shared similar perceptions of a learning environment scale, and to which perceptions varied from class to class. The η^2 (η^2) statistic, which is the proportion of variance attributed to class membership, was used measure the magnitude of the variance between classes.

3.4.2 Differences between ORLES Scales in Terms of Curriculum Sequence and Ethnicity

Differences between ORLES scales in terms of curriculum sequence and student ethnicity were investigated with the use of a two-way Multivariate Analysis of Variance (MANOVA) with the ORLES scales as the dependent variables and with the curriculum sequence (general versus specific) and ethnicity (Caucasian versus Hispanic) as the two independent variables. Wilks' lambda (λ) criterion was used as a test of significance for the set of dependent variables as a whole. Significant findings in this multivariate analysis justified separate interpretation of each univariate two-way ANOVA individually for each ORLES scale.

Comparisons between curriculum approaches were further clarified using average item means and average item standard deviations for each ORLES scale. Whereas the statistical significance for the comparison of each scale was determined from the ANOVA results, Cohen's *d* (Cohen, 1988) was used to measure the magnitude of the difference between the curriculum approaches for each ORLES scale. Effect sizes were calculated by dividing the differences in two item means by the pooled standard deviation (Thompson, 1998). The same techniques were used to examine ethnic differences in scores on each ORLES scale.

The differential effectiveness of the two curriculum sequences for the two ethnic groups was explored by checking whether the sequence-by-ethnicity interaction was significant for each scale. The average item means and average item standard deviation for four groups of students was inspected for:

1. Caucasian students experiencing a generalized curriculum.
2. Caucasian students experiencing a specific curriculum.

3. Hispanic students experiencing a generalized curriculum.
4. Hispanic students experiencing a specific curriculum.

3.4.3 Associations between Learning Environment and Attitudes

Numerous previous studies have revealed association between learning environments and student attitudes towards science (Fraser, 1994, 2012). A more detailed discussion of past research is provided in Chapter 2, Section 2.5.1 of this thesis. My study investigated associations between five ORLES learning environment scales and Enjoyment of Science. Simple correlation analysis was used to examine the bivariate associations between Enjoyment of Science and each individual scale.

Multiple regression analysis (Gay, Mills, & Airaisian, 2006) was used to investigate the multivariate association between attitudes and the entire set of learning environment scales. The multiple correlation between Enjoyment of Science and the whole set of learning environment scales was calculated to quantify the strength of the multivariate association between Enjoyment of Science and environment. Then standardized regression coefficients were used to determine the contribution of individual learning environment scales to the multivariate association between Enjoyment of Science and learning environment. The standardized regression coefficients describe the influence of a particular learning environment scale on the attitude scale when all other learning environment scales are mutually controlled. The association between Enjoyment of Science and learning environment was investigated with the individual student as the unit of analysis.

3.5 CONCLUSION

This chapter provided a summary of the methods used in conducting my study. First, details were provided of the sample, including the rationale behind the selection of the schools used in the study, methods of data collection, and the ethnic composition of the sample (which consisted of 367 students from two schools). Second, the ORLES instrument was discussed in detail. Finally, the various methods of data analysis used to answer my research questions were identified.

The two states used in my study, California and Texas, have different curriculum arrangements. California's curriculum arrangement was determined to be a specific or course-based approach with each grade level concentrating on a particular field of science. Texas' set of standards was found to be a general curriculum approach with students at all three grade levels studying topics across the whole of science.

California and Texas were selected as the states for use in my study because of their similarity in area, population, and demographic make-up. Correspondingly, two representative schools were chosen, one from each state. Both schools were in suburban locations outside a major metropolitan area and had ethnically heterogeneous student populations that approximately reflected the demographics of the states. The total sample consisted of 367 grade 8 science students in 15 classes.

The instrument used in my study, the ORLES, was based on the CLES, WIHIC and TOSRA that have been well established in previous research (Fraser, 2012). I selected one scale, Personal Relevance came from the CLES, five scales (Student Cohesiveness, Teacher Support, Investigation, Involvement and Task

Orientation) from the WIHIC, and Enjoyment of Science Lessons from the TOSRA as an attitudinal outcome measure.

I collected data from the Texas school personally using paper surveys that were scanned into a database. Data from the California school were collected by the school principal and the science teachers via the online service, Survey Monkey. Informed consent from both the students and their parents was collected by the teachers of each class.

The data were analyzed with a variety of techniques in order to answer the three research questions. To address the first research question regarding the validity and reliability of the ORLES instrument, principal axis factor analysis with varimax rotation and Kaiser normalization was utilized. As well, I used Cronbach's alpha coefficient to check the internal consistency reliability of each ORLES scale and ANOVA to check the ability of each scale to differentiate between the perceptions of students in different classrooms.

To answer the second research question regarding differences in ORLES scale scores based on curriculum sequence and ethnicity, a two-way MANOVA was conducted. Effect sizes were calculated to express the magnitude of the differences in standard deviation units. When statistically significant results were revealed according to Wilks' lambda criterion, then the univariate two-way ANOVA was interpreted separately for each scale. In order to evaluate the differential effectiveness of each curriculum sequence for each ethnic group for each ORLES scale, the sequence-by-ethnicity interaction was examined.

The third research question, which dealt with associations between learning environment and student attitudes towards science, was answered with the use of

simple correlations for bivariate associations between attitude and individual scales and multiple regression analysis for multivariate associations between attitudes and the complete set of learning environment scales.

The results of these analyses, answers to the research questions and discussions of the findings are found in the next chapter of this thesis.

Chapter 4

ANALYSES AND RESULTS

4.1 INTRODUCTION

This chapter presents the results of analyses of the collected data in order to answer my study's three research questions:

1. Is the Outcomes-Related Learning Environments Survey (ORLES), consisting of scales from the Constructivist Learning Environment Survey (CLES), What Is Happening In this Class? (WIHIC) and Test Of Science Related Attitudes (TOSRA), valid for use among middle-school science students in the United States?
2. Do scores on ORLES scales vary with (a) curriculum sequence (generalized versus topic-specific) and (b) student ethnicity and (c) the interaction of curriculum sequence with ethnicity?
3. Are there associations between classroom learning environments and student attitudes to science?

Firstly, Section 4.2 addresses the first research question about the validation of the ORLES, which is an instrument based upon the CLES, WIHIC and TOSRA for measuring middle-school (grades 6–8) students' perceptions of their learning environment and attitudes towards science.

Secondly, Section 4.3 involves answering the second research question by focusing on differences in the ORLES scale scores according to curriculum sequence, student ethnicity, and the interaction of curriculum sequence with ethnicity.

Thirdly, Section 4.4 answers the third research question concerning links between the learning environment and students' attitudes towards science. For this question, the data were analyzed for the total sample of all students from both curriculum sequences.

Finally, Section 4.5 concludes the chapter with an overview of the preceding sections. The implications of the results are discussed in the Chapter 5.

4.2 VALIDITY OF QUESTIONNAIRE

As discussed in Chapter 3, the ORLES is composed of scales from the CLES, WIHIC and TOSRA. The Personal Relevance scale is from CLES, Student Cohesiveness, Teacher Support, Investigation, Involvement, and Task Orientation are from the WIHIC, and Enjoyment of Science is from TOSRA. In prior research, the three instruments used in the construction of the ORLES have been found consistently to be valid and reliable. Internal consistency and discriminate validity are measures of the correlations among items in an instrument. Items within the same scale should show a strong correlation to one another (internal consistency), but not to items in different scales (discriminate validity).

The CLES has been validated in multiple countries, languages and settings, including the USA by Nix, Fraser and Ledbetter (2005), South Africa by Aldridge, Fraser, and Sebela (2004), Australia and Taiwan by Aldridge et al. (2000), and Korea by Kim, Fisher, and Fraser (1999). The WIHIC has been well established as a valid

survey by, for example, Aldridge, Fraser and Huang (1999), Dorman (2003), Koul and Fisher (2005), Khoo and Fraser (2008), Wolf and Fraser (2008) and MacLeod and Fraser (2010). The TOSRA has been cross-validated and found to be valuable in studies of attitude–environment associations in numerous countries (Farenga & Joyce, 1998; Fraser, Aldridge, & Adolphe, 2010; Fraser & Butts, 1982; Fraser & Fisher, 1982; Wong & Fraser, 1996).

Factor analyses were applied to data collected from 367 students to validate the original 40-item ORLES. Principal axis factoring with varimax rotation and Kaiser normalization suggested a six-scale structure, with the original Involvement and Investigation scales coming together to form a new 10-item scale. An item was retained only if its factor loading was greater than or equal to 0.40 with its own scale and less than 0.40 with each of the other scales. Five items (Items 6, 12, 13, 25, and 28) were discarded because of their unsatisfactory factor loadings: one from the Personal Relevance scale (“I learn how science can be part of my out-of-school life”); two from the Student Cohesiveness scale (“I work well with other class members” and “I help other class members who are having trouble with their work”); and one each from Involvement (“I am asked to explain how I solve problems”) and Investigation (“I am asked to think about evidence behind statements”).

This resulted in an instrument with 35 items and six scales and with four items each in Personal Relevance and Student Cohesiveness, six items each in Teacher Support and Task Orientation, 10 items in Investigation and Involvement and five items in Enjoyment of Science Lessons. Table 4.1 shows the factor loadings for the 35 items retained in the ORLES.

Reliability is the degree to which an instrument consistently measures what it is intended to measure. Table 4.2 shows the reliability estimate (Cronbach alpha

coefficient) for each refined ORLES scale for two units of analysis (the individual student and the class mean). With the student as the unit of analysis, scale alpha reliabilities ranged from 0.63 to 0.91. With the class as the unit of analysis, scale alpha reliabilities ranged from 0.61 to 0.98. According to George and Mallery (2003), alpha reliabilities greater than 0.9 are considered excellent, reliabilities greater than 0.8 are considered good and reliabilities above 0.7 are considered acceptable. Table 4.2 shows that, for most ORLES scales, the reliability fell in the range from good to excellent according to these guidelines.

Table 4.1

Factor Analysis Results for the ORLES

Item	Factor Loadings					
	Personal Relevance	Student Cohesiveness	Teacher Support	Investigation/ Involvement	Task Orientation	Enjoyment of Science
4	0.66					
5	0.45					
7	0.75					
8	0.49					
9		0.68				
10		0.56				
11		0.45				
14		0.45				
15			0.48			
16			0.53			
17			0.56			
18			0.63			
19			0.63			
20			0.44			
21				0.45		
22				0.61		
23				0.71		
24				0.69		
26				0.51		
27				0.48		
29				0.46		
30				0.43		
31				0.56		
32				0.51		
33					0.58	
34					0.61	
35					0.61	
36					0.64	
37					0.68	
38					0.73	
39						0.65
40						0.79
41						0.76
42						0.85
43						0.77
%	28.54	9.19	5.99	5.12	4.46	4.23
Variance						
Eigenvalue	10.56	3.40	2.21	1.89	1.65	1.56

N = 365 students in 15 Classes. Factor loadings less than 0.40 have been omitted from the table. Principal axis factoring with varimax rotation and Kaiser normalization. Items 6, 12, 13, 25, and 28 were not retained.

A one-way ANOVA with class membership as the independent variable was used to determine the extent to which students within each of the 15 classes shared similar perceptions for each learning environment scale, while perceptions varied from class to class. Table 4.2 shows that five out of the six ORLES scales exhibited significant differentiation among classes ($p < 0.001$). Scores on the Personal Relevance scale, however, were not significantly different from class to class. The η^2 statistic (the amount of variance explained by class membership) ranged from 0.12 to 0.27 for the five scales.

Table 4.2

Internal Consistency (Cronbach Alpha Reliability), Ability to Differentiate between Classrooms (ANOVA Results) and Discriminate Validity (Mean Correlation with Other Scales) for the ORLES

Scale	No of Items	Sample Item	Unit of Analysis	Alpha Reliability	Eta ²	Mean Correlation
Personal Relevance	4	I learn about the world outside of school.	Individual Class	0.70 0.61	0.05	0.32 0.52
Student Cohesiveness	4	I work well with other class members.	Individual Class	0.63 0.86	0.13***	0.32 0.53
Teacher Support	6	The teacher talks with me.	Individual Class	0.81 0.90	0.15***	0.43 0.55
Investigation/ Involvement	10	I draw conclusions from investigations.	Individual Class	0.87 0.95	0.16***	0.44 0.60
Task Orientation	6	I know the goals of this class.	Individual Class	0.88 0.94	0.12***	0.42 0.61
Enjoyment of Science	5	I look forward to science lessons.	Individual Class	0.91 0.98	0.27***	0.38 0.51

$N=365$ students in 15 classes.

*** $p < 0.001$

Table 4.2 also reports the mean correlation of a scale with the other scales for each of the six ORLES scales at both the individual and class levels of analysis. The mean correlation is used as an index of discriminate validity or the extent to which a

scale is not related to the other scales. When using the individual as the unit of analysis, mean correlations ranged from 0.32 to 0.44. With the class as the unit of analysis, mean correlations ranged from 0.51 to 0.61. These results suggest that raw scores on the six ORLES scales overlap somewhat, but the factor analysis results support the independence of factor scores.

4.3 CURRICULUM-SEQUENCE AND ETHNICITY DIFFERENCES IN ORLES SCALES

This section reports analyses and results pertaining to my second research question which involves differences in ORLES scales according to:

- (a) curriculum sequence
- (b) student ethnicity
- (c) the interaction of curriculum sequence with ethnicity.

Multivariate Analysis of Variance (MANOVA) was conducted with the set of six ORLES scales as the dependent variable and with curriculum sequence (general versus specific) and ethnicity (Caucasian versus Hispanic) as the two independent variables.

Students were asked to identify themselves into one of five ethnic classifications: African American, Asian/Pacific Islander, Hispanic, Caucasian, or Native American/Other. For simplicity, convenience, validity and meaningfulness, only the two major ethnic groups (namely, Caucasian and Hispanic) were included in the analysis. This represented 77% of the sample. Students of other ethnicities were omitted for these analyses. The two-way MANOVA was restricted to the sample of 280 students whose ethnicity was either Caucasian or Hispanic.

Because the multivariate test using Wilks' lambda criterion revealed statistically significant results for the set of dependent variables as a whole, the individual ANOVA results for each of the ORLES scales were interpreted separately. Table 4.3 reports the ANOVA results for curriculum sequence, ethnicity and the sequence-by-ethnicity interaction for each ORLES scale. This table reveals the following statically significant findings.

- A significant difference between curriculum sequences for Enjoyment of Science.
- A significant difference between ethnicities for Task Orientation.
- A significant curriculum sequence-by-ethnicity interaction for Task Orientation.

The presence of a significant sequence-by-ethnicity interaction was used to identify the presence of differential effectiveness of different curriculum sequences for students of different ethnicities.

The discussion of these findings is provided in more detail below in separate

Table 4.3

ANOVA Results for Each ORLES Scale for Curriculum Sequence, Ethnicity and Sequence-by-Ethnicity Interaction

Scale	Curriculum Sequence	Ethnicity	Sequence-by-Ethnicity
Personal Relevance	3.05	0.30	0.35
Student Cohesiveness	0.59	0.63	1.60
Teacher Support	0.17	1.94	0.97
Investigation/ Involvement	0.53	1.12	0.39
Task Orientation	0.03	4.91*	4.39*
Enjoyment of Science	28.67***	0.25	3.37

N = 280

p*<0.05, **p*<0.001

subsections for differences between curriculum sequences (Section 4.3.1), differences

between ethnicities (Section 4.3.2) and differential effectiveness of curriculum sequences for student of different ethnicities (Section 4.3.3).

4.3.1 Differences Between Curriculum Sequences

Table 4.4 reports a comparison of the generalized and topic-specific curriculum approaches in terms of the average item means and average item standard deviations for student perceptions of their learning environments and attitudes toward science. To maximize statistical power, ANOVAs were conducted for the entire sample. Table 4.4 reports the results of ANOVA significance testing for the entire sample of 367 students of all ethnicities for whether or not scales for the two groups were significantly different. The average item mean is simply the scale mean divided by the number of items in the scale. The standard deviation is a measure of the variation of scores from the mean.

Differences between curriculum sequences on each ORLES scale were not only tested for statistical significance via ANOVA, but also effect sizes were used to describe the magnitude of these differences as recommended by Thompson (1998). Effect size shows the magnitude of the difference between the means for each scale. The larger the effect size, the more important is the difference between curriculum approaches for that specific scale. In my study, Cohen's d was used to measure effect size. Cohen's d is the difference between two means divided by the standard deviation. According to Cohen (1988), an effect size of 0.8 or more represents a large effect, 0.5 is a medium effect size and 0.2 is a small effect size. However, Cohen cautions against defining terms like small, medium or large.

Although the topic-specific approach overall was associated with slightly

higher scores for most of the learning environments scales relative the generalized approach (see Table 4.4), no statistically significant differences were found between the two approaches for any of the learning environment scales. The largest difference for any learning environment scale was for Personal Relevance (effect size of 0.15 deviations), which represents a small effect size (Cohen, 1988).

Table 4.4

Average Item Mean, Average Item Standard Deviation and Difference between Two Curriculum Approaches (ANOVA Results and Effect Size) for each Learning Environment and Attitude Scale

Scale	Average Item Mean		Average Item Standard Deviation		Difference	
	General	Specific	General	Specific	F	Effect Size
Personal Relevance	3.45	3.57	0.83	0.75	1.17	0.15
Student Cohesiveness	3.79	3.80	0.67	0.77	0.11	0.01
Teacher Support	3.45	3.49	0.83	0.83	0.29	0.04
Investigation/Involvement	2.97	3.03	0.77	0.83	0.42	0.07
Task Orientation	4.10	4.08	0.73	0.95	0.45	-0.02
Enjoyment of Science	2.61	3.42	1.01	1.17	43.43***	0.74

N = 362 (Generalized = 167, Specific =200)

***p<0.001

However, Table 4.4 shows that there was a significant and quite large difference with an effect size of 0.74 standard deviations for students' Enjoyment of Science Lessons. Enjoyment was higher among students experiencing a topic-specific curricular arrangement than among students following a generalized curriculum.

Figure 4.1 provides a plot of statistically significant curriculum-sequence differences for ORLES scales. In this graph, nonsignificant differences are depicted as zero differences. This figure illustrates the large differences in attitudes in favor of the specific curriculum approach.

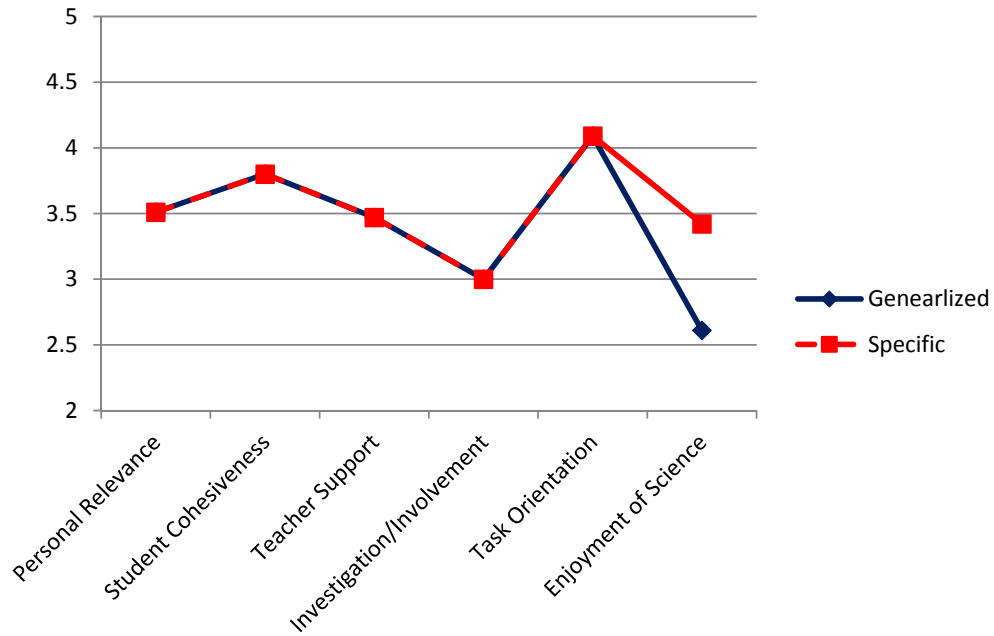


Figure 4.1 Simplified Plot of Significant Differences Between Two Curriculum Sequences for ORLES Scales

4.3.2 Differences between Ethnicities

This section reports my comparison of the two dominant ethnic subpopulations (namely, Caucasians and Hispanics) in terms of attitudes and perceptions of the learning environment. Table 4.5 reports the average item mean, average item standard deviation, and difference between Caucasian and Hispanic students for each learning environment scale and Enjoyment of Science. In the same manner as Table 4.4, the average item mean is the scale mean divided by the number of items in the scale. The effect size expresses the difference between the means in standard deviation units. The *F* values in Table 4.5 for ethnic differences are reported for ANOVAs that were conducted for the sample of 179 Caucasians and 101 Hispanics (total of 280 students).

Table 4.5 show that differences between Caucasian and Hispanic students were small and statistically nonsignificant for every scale except Task Orientation.

However ethnic differences were statistically significant ($p < 0.001$) for Task Orientation. Relative to Hispanic students, Caucasian students' perceptions of Task Orientation were more positive, with a small to moderate effect size of 0.34 standard deviations. However, because of the presence of a significant curriculum sequence-by-ethnicity interaction for Task Orientation, it is necessary to revisit this interaction in Section 4.3.3.

Table 4.5

Average Item Mean, Average Item Standard Deviation and Difference Between Ethnic Groupings for Learning Environment and Attitude Scales

Scale	Average Item Mean		Average Item Standard Deviation		Difference	
	Caucasian	Hispanic	Caucasian	Hispanic	<i>F</i>	Effect Size
Personal Relevance	3.56	3.52	0.73	0.48	1.47	0.06
Student Cohesiveness	3.83	3.79	0.69	0.64	2.59	0.06
Teacher Support	3.47	3.60	0.82	0.65	0.61	-0.18
Investigation/ Involvement	3.05	2.96	0.77	0.78	2.11	0.12
Task Orientation	4.25	3.99	0.81	0.73	20.06***	0.34
Enjoyment of Science	3.11	3.21	1.18	1.01	2.02	-0.09

N = 179 (Caucasian), 101 (Hispanic)

* $p < 0.05$, *** $p < 0.001$

Figure 4.2 illustrates the difference between the two ethnic groups' perceptions of classroom environment and attitudes. In this graphical format, statistically nonsignificant ethnic differences for five scales are depicted as zero differences. This figure also illustrates the larger between-ethnicities difference for Task Orientation.

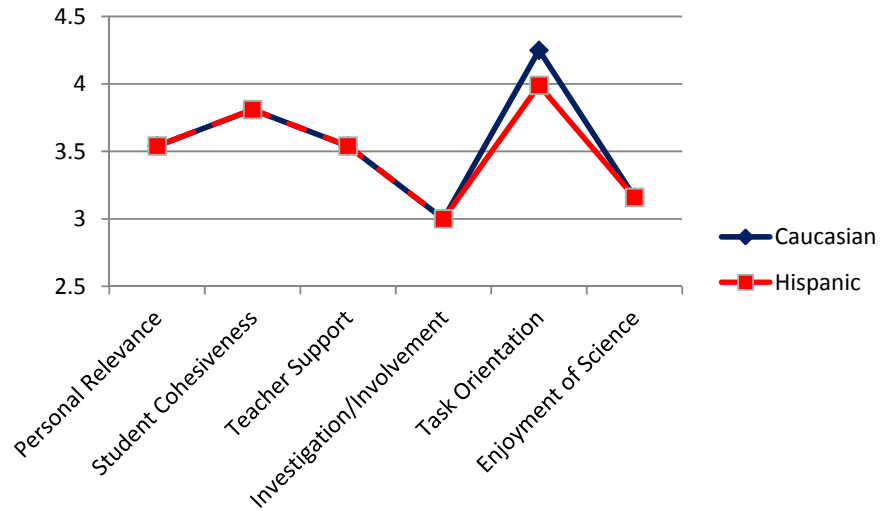


Figure 4.2 Simplified Plot of Significant Differences Between Two Ethnicities for ORLES Scales

4.3.3 Differential Effectiveness of Different Curriculum Sequences for Different Ethnic Groups

Table 4.6 shows the average item mean for each ORLES scale for the four groups of (1) Caucasian students experiencing a generalized curriculum, (2) Caucasian students experiencing a specific curriculum, (3) Hispanic students experiencing a generalized curriculum and (4) Hispanic students experiencing a specific curriculum. The ANOVA results for the sequence-by-curriculum interactions for each ORLES scale reported in Table 4.6 are repeated from Table 4.3. The sequence-by-curriculum interaction is an indication of the differential effectiveness of the two curriculum approach for the two different ethnic groups. Tables 4.3 and 4.6 shows that the only ORLES scale for which the sequence-by-ethnicity interaction was statistically significant was Task Orientation.

Table 4.6

Average Item Mean and Average Item Standard Deviation for each ORLES Scale for Two Ethnicities and Two Curriculum Sequences, and ANOVA Results for Sequence-by-Ethnicity Interaction

Scale	Model	Average Mean		Average Item Standard Deviation		F
		Caucasian	Hispanic	Caucasian	Hispanic	
Personal Relevance	Generalized	3.42	3.68	0.77	0.67	0.35
	Specific	3.43	3.57	1.00	0.64	
Student Cohesiveness	Generalized	3.85	3.81	0.68	0.70	1.60
	Specific	3.67	3.85	0.71	0.61	
Teacher Support	Generalized	3.40	3.54	0.81	0.82	0.97
	Specific	3.64	3.58	0.66	0.64	
Investigation Involvement	Generalized	3.04	3.05	0.78	0.76	0.39
	Specific	2.87	3.01	0.76	0.79	
Task Orientation	Generalized	4.15	4.34	0.80	0.81	4.39*
	Specific	4.11	3.92	0.59	0.79	
Enjoyment of Science	Generalized	2.58	3.57	0.91	1.19	3.37
	Specific	2.90	3.38	1.03	0.96	

N = 361 (Generalized Caucasian=82, Generalized Hispanic=34) and (Specific Caucasian=97, Specific Hispanic = 67)

* $p < 0.05$

Figure 4.3 provides a graph of these average item means for Task Orientation, which is the only ORLES scale for which the curriculum sequence-by-ethnicity interaction was statistically significant. Figure 4.3 shows the interpretation of the significant sequence-by-ethnicity interaction for Task Orientation in terms of the differential effectiveness of the curriculum sequences for the different ethnic groupings.

Hispanic students' Task Orientation was higher under the generalized curriculum approach than the specific approach. However, for Caucasians, Task Orientation scores were almost identical under the two alternative curriculum sequences.

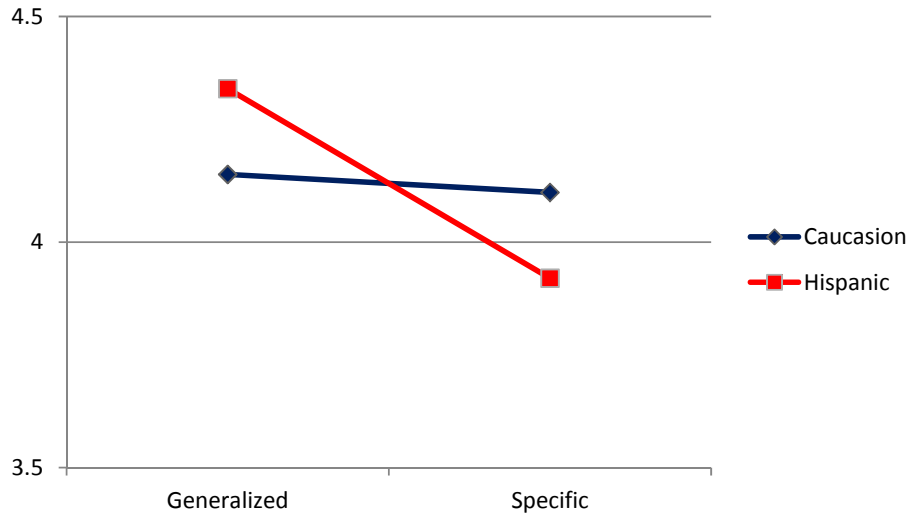


Figure 4.3 Average Item Means for Task Orientation for Two Ethnicities and Two Curriculum Sequences

4.4 ASSOCIATIONS BETWEEN LEARNING ENVIRONMENT AND ATTITUDES TOWARDS SCIENCE

The third research question was: "Are there connections between the students' perceptions of the learning environment and their attitudes towards science?" Table 4.7 reports the results of simple correlation and multiple regression analyses. The simple correlation analysis provided information about the bivariate association between Enjoyment of Science and each individual learning environment scale. On the other hand, the multiple regression analysis (Gay, Mills, & Airaisian, 2006) provided information about the multivariate association between Enjoyment of Science and the whole set of learning environment scales. In order to identify which individual learning environment scale contributed most to the multivariate association when the other environment scales were mutually controlled, the standardized regression coefficients were examined.

When analyzed using simple correlations, all five learning environment scales showed statistically significant associations with Enjoyment of Science (Table 4.7). Task Orientation had the strongest correlation of 0.52 with Enjoyment.

Table 4.7 shows the multiple correlation between Enjoyment and the set of all learning environment scales with 0.33 and statistically significant ($p < 0.01$).

When the standardized regression coefficients were used to identify how much influence each learning environment scale had upon the Enjoyment of Science while the other four learning environment scales were mutually controlled, four of the five learning environment scales were found to have a significant independent influence on Enjoyment of Science: Personal Relevance, Teacher Support, Investigation/Involvement and Task Orientation. The fifth scale, Student Cohesiveness, did not show a significant independent association with Enjoyment.

Table 4.7

Associations between Learning Environment Scales and Enjoyment of Science in Terms of Simple Correlations (r), Multiple Correlation (R) and Standardized Regression (β)

Scale	Associations with Enjoyment of Science	
	r	β
Personal Relevance	0.33**	0.11*
Student Cohesiveness	0.25**	0.03
Teacher Support	0.41**	0.12*
Investigation/Involvement	0.41**	0.14*
Task Orientation	0.52**	0.35***
Multiple Correlation, R		0.33**

$N=367$

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

It is noteworthy that every significant simple correlation and regression coefficient in Table 4.7 is positive, thus indicating that more emphasis on the environment dimensions was associated with greater student enjoyment. This result replicates considerable prior research around the world (Fraser, 2012).

4.5 CONCLUSION

This chapter reported results to answer three research questions. First, is the ORLES a valid and internally consistent instrument? Second, are the differences in students' perceptions of learning environment and attitudes towards science according to the curriculum sequence, ethnicity and the interaction of curriculum sequence and ethnicity. Third, are there associations between learning environment and attitudes towards science?

In answer to the first research question, the ORLES was found to be a valid questionnaire for assessing learning environment and attitudes among a sample of 367 middle-school students in California and Texas. The ORLES was constructed from three previously tried and tested questionnaires that have been successful in numerous studies across multiple age groups, geographic locations and fields of study. Principal axis factor analysis with varimax rotation and Kaiser normalization revealed that 35 of the original 40 items had satisfactory factor loadings. Two of the scales, Involvement and Investigation, were combined because their items loaded on the same factor. The percentage of variance for the scales were 28.54% for Personal Relevance, 9.19% for Student Cohesiveness, 5.99% for Teacher Support, 5.12% for Investigation/Involvement, 4.46% for Task Orientation and 4.23% for Enjoyment of Science, with the total percentage of variance being 57.53%. All six ORLES scales were found to have satisfactory internal consistency reliability using Cronbach's alpha coefficient. These results are not surprising given the history of the three parent instruments, namely, the CLES (Aldridge, Fraser, & Sebela, 2004; Aldridge et al., 2000; Kim, Fisher, & Fraser 1999; Nix, Fraser & Ledbetter, 2005), the WIHIC (Aldridge, Fraser & Huang, 1999; Dorman, 2003; Khoo & Fraser, 2008; Koul &

Fisher, 2005; MacLeod & Fraser, 2010; Wolf and Fraser, 2008) and the TOSRA (Farenga & Joyce, 1998; Fraser, Aldridge, & Adolphe, 2010; Fraser & Butts, 1982; Fraser & Fisher, 1982; Wong & Fraser, 1996). The history of the three parent instruments is reported in detail in Chapter 2. The statistical analyses reported in this chapter reinforce the validity and reliability of the CLES, WIHIC and TOSRA for the assessment of classroom learning environments and attitudes among middle-school students in Texas and California.

The second research question involved differences in learning environment and attitude scales in terms of curriculum sequences, student ethnicity and the interaction between curriculum sequence and ethnicity. Section 4.3 covers the analyses of this research question in detail. MANOVA, ANOVAs and effect sizes were used to answer this research question.

In answer to the first part of this research question, differences between curriculum sequences, a statistically significant difference between curriculum approaches emerged for Enjoyment of Science with an effect size of 0.74 standard deviations in favor of the specific curriculum approach. Some possible explanations for the findings are considered in Chapter 5.

Investigation of the second part of my second research question revealed that Caucasian students had higher scores than Hispanic students for all learning environment scales and enjoyment. However, these differences were statistically significant only for Task Orientation scale with a moderate effect size of 0.34 standard deviations.

However, because of the presence of a statistically significant sequence-by-ethnicity interaction for the Task Orientation scale (the third part of my second

research question), it was necessary to reconsider the interpretation of the main effect for Task Orientation. Whereas Hispanic students perceived higher Task Orientation under the generalized curriculum approach, Caucasian students perceived higher Task Orientation under the specific curriculum approach. Further research is needed to replicate and clarify this interesting result. Student interviews could be valuable in clarifying these patterns in future studies.

Regarding the third research question, there were significant associations between learning environment scales and attitudes towards science. Simple correlations showed significant positive associations between enjoyment and all five learning environment scales. Multiple regression analyses revealed that all learning environment scales with the exception of Student Cohesiveness showed a significant independent influence on Enjoyment of Science based on standardized regression coefficients. This is consistent with previous research (Fraser & Butts, 1982; Fraser & Fisher, 1982; Kim et al., 2000) reviewed in Section 2.5.1 of Chapter 2.

Chapter 5

CONCLUSIONS

5.1 INTRODUCTION

This chapter provides a summary of my study and discusses possible implications of findings, its limitations and suggestions for future research. Section 5.2 provides a summary of the introductory chapter and literature reviewed in Chapter 2. Section 5.3 summarizes the research methods and results. Section 5.4 discusses limitations of my study and details some suggestions for further research based on my study's findings and limitations. Section 5.5 considers some implications of the findings and the significance of my study.

5.2 INTRODUCTORY CHAPTER AND LITERATURE REVIEW

This section summarizes the introduction to my study in Chapter 1, as well as the literature reviewed in Chapter 2. The rationale for my study came from the change of the Texas middle-school science standards from a course-based specific curriculum arrangement to a generalized curriculum arrangement in the late 1990s. Curriculum is an important factor in student learning (Walberg, 1991) and is worthy of study. While each state has its own assessment methods for evaluating the effectiveness of its own curricula, there has not been much, if any, between-state comparisons of middle-school science standards. The field of learning environment research provided a promising method for comparing the effectiveness of different curriculum sequences

because of the links often reported between learning environments and student outcomes (Fraser, 2012).

Chapter 1 provided some of the background to my study, including the history of learning environments research, which built on the ideas of Lewin (1936) and Murray (1938) and led to development of the Learning Environment Inventory (LEI) Walberg (1968) and the Classroom Environment Scale (CES) (Moos & Trickett, 1974). Several studies relevant to middle-schooling and the transition between elementary and high school were considered. Middle-school science standards were considered, together with a description of my method for classifying curricula as either ‘general’ or ‘specific’.

Chapter 2 provided a review of literature covering published studies relevant to my study. Topics covered by the literature review included a history of learning environments research (Section 2.2), a range of learning environment instruments (Section 2.3), student attitudes towards science and associated instruments (Section 2.4), and past research applications of learning environment instruments (Section 2.5).

5.3 DISCUSSION OF METHODS AND RESULTS

This section provides a summary of the research methods and results for my study in terms of: the validity and reliability of the Outcomes-Related Learning Environments Survey (ORLES) (Section 5.3.1); the relative effectiveness of alternative curriculum sequences in terms of ORLES scales (Section 5.3.2); ethnic differences in perceptions of learning environment and attitudes and the differential effectiveness of two curriculum approaches for students of different ethnicities

(Section 5.3.3); and relationships between learning environment and attitudes (Section 5.3.4).

5.3.1 Validity and Reliability of the ORLES Questionnaire

The ORLES questionnaire used in my study had 40 items divided among seven scales. Six scales (Personal Relevance, Student Cohesiveness, Teacher Support, Investigation, Involvement and Task Orientation) measured students' perceptions of learning environment and one scale, Enjoyment of Science Lessons, measured students' attitudes towards science. The ORLES was constructed from scales selected from three previously well-established instruments, namely, the Constructivist Learning Environment Survey (CLES), What Is Happening In this Class? (WIHIC) and Test Of Science Related Attitudes (TOSRA). The ORLES was administered to 367 grade 8 students in 15 classes (eight in Texas, seven in California) taught by six teachers (four in Texas, two in California).

My first research question was concerned with the validity and reliability of the ORLES instrument. Principal axis factoring with varimax rotation and Kaiser normalization suggested a six-scale structure, with the Involvement and Investigation scales coming together to form a new, combined 10-item scale. Five of the original items were dropped because of unsatisfactory factor loadings. In order for an item to be retained, it had to have a factor loading greater than 0.40 within its own scale and less than a 0.40 with all other scales. The percentage of variance for the different scales ranged from 4.23% to 28.54%, with the total being 57.53%. The resulting ORLES instrument consisted of six scales and 35 items.

Cronbach's alpha coefficient was used to estimate the internal consistency of the ORLES scales. Using both the student and the class as units of analysis, most of the ORLES scales showed good to excellent reliability (George & Mallery, 2003). Discriminate validity of the six ORLES scales, using the mean correlation of each scale with the other scales, was found to be satisfactory with both the class and individual student as the unit of analysis (See Table 4.2). As well, ANOVA showed that each ORLES scale was capable of differentiating significantly between the perceptions of students in different classrooms.

The three instruments upon which ORLES was based have a strong history of reliability and validity. The CLES has been validated numerous times (Aldridge et al., 2000; Johnson & McClure, 2004; Nix & Fraser, 2011; Nix et al., 2005; Peiro & Fraser, 2009). The WIHIC, from which the bulk of the ORLES scales were derived, has been found valid and reliable in several different settings (Afari et al., 2013; Aldridge et al., 1990; Kim et al., 1990, Waldrip et al., 2009). The TOSRA has also been found valid and reliable in multiple studies in various settings (Afari et al., 2013; Fraser, 1978, 1981; Telli et al., 2010). The results of statistical analyses of the ORLES data for my sample are consistent with previous research undertaken with the CLES, WIHIC and TOSRA. A detailed review of literature related to the three parent instruments can be found in Chapter 2, Sections 2.3 and 2.4.

5.3.2 Relative Effectiveness of Alternative Curriculum Sequences

My second research question dealt with examining how scores on ORLES scales varied between curriculum sequences. Chapter 1, Section 1.5 discussed my

method of classification of middle-school curriculum sequences using two states' (California and Texas) middle-school science standards. California's middle-school science curriculum was classified as a topic-specific curriculum approach with students at each grade level (6th, 7th and 8th) studying a specific field of science such as physical science, life science or earth science. Texas's middle-school science standards were classified as a generalized curriculum approach, with students at each grade level being taught topics from all three fields of science.

A comparison of the two curriculum sequences was performed by examining the average item means and average item standard deviations for student perceptions of each learning environment and attitude scales. MANOVA was used to test the statistical significance of the differences between generalized and topic-specific curriculum approaches. Generally, the topic-specific approach was associated with slightly higher scores for the five learning environment scales, but no statistically significant differences were found. There was a large and significant difference (effect size of 0.74 standard deviations) between the two curriculum sequences in terms of student Enjoyment of science in favor of the topic-specific approach.

One possible explanation for this finding for Enjoyment is that, because science teachers often have an undergraduate degree in one particular field of science, such as biology, physics or geology, it is plausible that a teacher would have greater enthusiasm for his or her particular subject. That enthusiasm would be evident to the students and could lead to greater enjoyment of their science lessons when that teacher is teaching his or her favorite topics. In contrast, teachers trained in a particular field might not be as passionate about topics that are outside their field and thus students' enjoyment of some of the lessons could be somewhat diminished. Additional research could be undertaken to test this hypothesis.

5.3.3 Differences between Ethnicities and Differential Effectiveness of Curriculum Sequences for Two Ethnic Groups

Chapter 4, Section 4.3.2 reported the results of a comparison of ORLES scale scores for two ethnic groups, Caucasians versus Hispanics. Hispanic students had lower scores on most ORLES scales, but differences were statistically significant only for Task Orientation. For Task Orientation, a moderate-sized ethnic difference occurred with an effect size of 0.34 standard deviations.

Section 4.3.3 also reported the differential effectiveness of the two curriculum sequences for the two ethnic groups by examining the sequence-by-ethnicity interaction and inspecting average item means for each ORLES scale for the four groups (Caucasian/generalized curriculum, Caucasian/specific curriculum, Hispanic/generalized curriculum and Hispanic/specific curriculum). ANOVA revealed a statistically significant sequence-by-ethnicity interaction only for Task Orientation. For Task Orientation, the generalized curriculum approach was more effective for Hispanic students, while the specific approach was more effective for Caucasian students.

5.3.4 Associations between Learning Environment and Attitudes

My third research question involved associations between learning environment and student attitudes towards science. Simple correlation analysis and multiple regression analysis were used to investigate the statistical significance of connections between learning environment and student attitudes towards science. Simple correlation analysis showed that each of the five ORLES learning environment scales was statistically significantly associated with Enjoyment of

Science. The multiple correlation revealed that the learning environment scales, as a set, had a significant multivariate association with attitudes. The regression analysis showed that four of the five of the ORLES learning environment scales made a statistically significant and unique contribution to the variance in Enjoyment of Science when the other ORLES scales were mutually controlled. Student Cohesiveness was the only scale that didn't make a significant unique contribution. My study's results replicated numerous previous research that demonstrates positive associations between learning environment scales and student attitudes towards science (Fraser, 2012; Fraser, Aldridge & Adolphe, 2010; Wong, Young & Fraser, 1997).

5.4 LIMITATIONS AND SUGGESTIONS FOR FUTURE RESEARCH

Prudence should be used when generalizing the results of my study as it involved a relatively small sample of states, students, and schools. The United States is comprised of 50 states, each with its own set of science education standards and curriculum sequences. Because my study only included two of these states, the inclusion of more states in future research could allow for wider generalization of the results. Similarly, only two schools, one from each state, were included in my study. The selection of more schools in each state would have yielded a larger and broader sample of students. With only 367 students surveyed, the statistical power in my study could have been somewhat limited. In order to address these limitations, future research should be undertaken with larger samples encompassing more schools and more states.

My study revealed some interesting results that could be explored further in the future using qualitative research methods as recommended by Tobin and Fraser (1998). The most noteworthy finding reported in Chapter 4 was the higher Enjoyment of Science among students experiencing the specific curriculum sequence. Perhaps the higher attitude scores could be explained by the enthusiasm of teachers of the specific curriculum sequence who are operating within their own field of expertise, as opposed to the teachers of generalized curriculum sequences, who have to cover topics outside of their field of expertise. Some methods for testing this hypothesis would include conducting interviews with teachers and classroom observations. Expanding the sample to include students at all three middle-school grade levels could be another way to provide a broader perspective.

One limitation in my study was that ethnic differences were explored only for Caucasian and Hispanic students. A larger sample in future research would allow more ethnic categories and therefore a more in-depth analysis of ethnicity differences.

A significant sequence-by-ethnicity interaction was found for Task Orientation. Hispanic students perceived higher Task Orientation under the generalized curriculum sequence than under the specific sequence, while Caucasian students' Task Orientation scores were quite similar for the two different curriculum approaches. Again, qualitative data collection and analysis could be used in future studies in an attempt to explain this interesting pattern. Interviews with students could be fruitful in future research.

5.5 SIGNIFICANCE AND IMPLICATIONS

This study is significant because there has been little prior research into the effectiveness of alternative middle-school curriculum arrangements. State boards of education spend considerable time and effort in arranging middle-school curricula, but even a cursory glance at a handful of state standards shows that there is no universal agreement as to the best way to teach middle-school science. Studies such as mine potentially could help to shed some light on some of the advantages and disadvantages of different curriculum arrangements. This type of study also could be valuable as the debate continues about the merits of a common-core science curriculum.

My study is also significant in that it makes modest contributions to the field of learning environment research. First, my study is the first to apply learning environment ideas in the evaluation of middle-school science curriculum sequences. Second, the successful validation of the ORLES instrument in two U.S. states adds to the wealth of validity and reliability data for the CLES, WIHIC and TOSRA scales that make up the ORLES.

More research is needed in order to fully assess the implications of the third part of the second research question concerning the differential effectiveness alternative curriculum arrangements for different ethnic groups. My finding of a sequence-by-ethnicity interaction for Task Orientation suggests that some thought could be given to the idea of providing alternative curricular arrangements as options for students who could be more inclined to one approach than the other.

Another important implication of my study comes from the third research question, concerning associations between the learning environment and student

attitudes towards science. My study revealed a strong and positive association between ORLES learning environment scales and Enjoyment of Science (Chapter 4, Section 4.4), which is consistent with findings in other studies such as Aldridge and Fraser (2008), Chionh and Fraser (2009) and Afari et al. (2013). In light of these and other studies, teachers wishing to improve their students' attitudes towards science could seek ways to improve their classroom environments on the dimensions encompassed by the ORLES. This is especially salient in middle school when one considers the concerns about the changes in learning environment experienced by students transitioning from primary to secondary school, as discussed by Ferguson and Fraser (1998), and the decline in student attitudes between grades 5 and 11 as reported by Baykul (1990).

5.6 CONCLUDING REMARKS

From the insights provided in this thesis, what middle-schooling should be like is not clear. Should it be the pinnacle of primary education, the beginning of secondary education, or something else? That is a very complex question, but hopefully my study has contributed in a small way to this puzzle.

"The direction in which education starts a man will determine his future life." – Plato,
The Republic

REFERENCES

- Afari, E., Aldridge, J. M., Fraser, B. J., & Khine, M. S. (2013). Students' perceptions of the learning environment and attitudes in game-based mathematics classrooms. *Learning Environments Research, 16*, 131–150.
- Aldridge, J. M., Dorman, J. P., & Fraser, B. J. (2004). Use of multi-trait–multi-method modeling to validate actual and preferred forms of the Technology-Rich Outcomes-Focused Learning Inventory (TROFLEI). *Australian Journal of Educational and Developmental Psychology, 4*, 110–125.
- Aldridge, J. M., & Fraser, B. J. (2008). *Outcomes-focused learning environments: Determinants and effects* (Advances in Learning Environments Research series). Rotterdam, The Netherlands: Sense Publishers.
- Aldridge, J., Fraser, B., Bell, L., & Dorman, J. (2012). Using a new learning environment questionnaire for reflection in teacher action research. *Journal of Science Teacher Education, 23*, 259–290.
- Aldridge, J. M., Fraser, B. J., & Huang, I. T. -C. (1999). Investigating classroom environments in Taiwan and Australia with multiple research methods. *The Journal of Educational Research, 93*, 48–62.
- Aldridge, J. M., Fraser, B. J., & Ntuli, S. (2009). Utilising learning environment assessments to improve teaching practices among in-service teachers undertaking a distance education programme. *South African Journal of Education, 29*, 147–170.
- Aldridge, J. M., Fraser, B. J., & Sebela, M. P. (2004). Using teacher action research to promote constructivist learning environments in South Africa. *South African Journal of Education, 24*, 245–253.

- Aldridge, J. M., Fraser, B. J., Taylor, P. C., & Chen C. (2000). Constructivist learning environments in a cross-national study in Taiwan and Australia. *International Journal of Science Education*, 22, 37–55.
- Allen, D., & Fraser, B. J. (2007). Parent and student perceptions of classroom learning environment and its association with student outcomes. *Learning Environments Research*, 10, 67–82.
- Bame, E. A., Dugger, W. E., Jr., de Vries, M., & McBee, J. (1993). Pupils' attitudes towards technology–PATT-USA. *The Journal of Technology Studies*, 6, 40–48.
- Baykul, Y. (1990). *To what extent were students' attitudes towards mathematics and science subjects changed in Turkish schools in grades five through eleven: What factors might have significant relations with student performance in the Student Selection and Placement Exam?* [In Turkish]. Ankara, Turkey: ÖSYM Yayınları.
- Bertrand, C. (N.D.) *A brief history of the California science content standards and the outlook for their future*. Retrieved February 9, 21012 from <http://www.cascience.org/csta/pdf/standards%20history.pdf>.
- Bru, E., Stornes, T., Munthe, E., & Thuen, E. (2010). Students' perception of teacher support across the transition from primary to secondary school. *Scandinavian Journal of Educational Research*, 54, 519–533.
- Chang, C. Y., Hsiao, C. H., & Barufaldi, J (2006). Preferred-actual learning environment “spaces” and earth science outcomes in Taiwan. *Science Education*, 90, 420–433.

- Chionh, Y. H., & Fraser, B. J. (2009). Classroom environment, achievement, attitudes and self-esteem in geography and mathematics in Singapore. *International Research in Geographical and Environmental Education*, 18, 29–44.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- den Brok, P., Brekelmans, M., & Wubbels, Th. (2006). Multilevel issues in research using students' perceptions of learning environments: The case of the Questionnaire on Teacher Interaction. *Learning Environments Research*, 9, 199–213.
- den Brok, P. J., Levy, J., Rodriguez, R., & Wubbels, Th. (2002). Perceptions of Asian-American and Hispanic-American teachers and their students on interpersonal communication style. *Teaching and Teacher Education*, 18, 447–467.
- den Brok, P., Telli, S., Cakiroglu, J., Taconis, R., & Tekkaya, C. (2010). Learning environment profiles of Turkish secondary biology students. *Learning Environments Research*, 13, 187–204.
- De Vries, M. J. (1991). What do students in Dutch technology teacher programmes think of their subject? *Research in Science and Technological Education* 9, 173–179.
- Doepken, D., Lawsky, E., & Padwa, L. (1993, June). *Modified Fennema–Sherman attitude scales*. Paper presented at Woodrow Wilson Gender Equity in Mathematics and Science Congress (WWGEMS), Princeton, NJ.
- Dorman, J. P. (2000). Association between classroom environment and academic efficacy. *Learning Environments Research*, 4, 243–257.

- Dorman, J. P. (2003). Cross-national validation of the What Is Happening In this Class? (WIHIC) questionnaire using confirmatory factor analysis. *Learning Environments Research*, 6, 231–245.
- Dorman, J. P. (2008). Use of multitrait–multimethod modelling to validate actual and preferred forms of the What Is Happening In this Class? (WIHIC) questionnaire. *Learning Environments Research*, 11, 179–197
- Farenga, S. J., & Joyce, B. A. (1998). Science-related attitudes and science course selection: A study of high-ability boys and girls. *Roeper Review*, 20, 247–251.
- Feldlaufer, H., Midgley, C., & Eccles, J. S. (1988). Student, teacher, and observer perceptions of the classroom environment before and after the transition to junior high school. *The Journal of Early Adolescence*, 8, 133–156.
- Ferguson, P., & Fraser, B. (1998). Student gender, school size and changing perceptions of science learning environments during the transition from primary to secondary school. *Research in Science Education*, 28, 387–397.
- Fisher, D. L., & Fraser, B. J. (1983). A comparison of actual and preferred classroom environment as perceived by science teachers and students. *Journal of Research in Science Teaching*, 20, 55–61.
- Fisher, D., Henderson, D., & Fraser, B. (1997). Laboratory environments & student outcomes in senior high school biology. *American Biology Teacher*, 59, 214–219.
- Fisher, D. L., & Rickards, T. (1997). Cultural and gender differences in teacher–student interpersonal behaviour in science classrooms. In D.L. Fisher and T. Rickards (Eds.), *Science, mathematics and technology education and national development* (Vietnam conference proceedings) (pp. 1–9). Perth: Curtin University of Technology.

- Fraser, B. J. (1978). Development of a test of science-related attitudes. *Science Education*, 62, 509–515.
- Fraser, B. J. (1979). *Test of Enquiry Skills (TOES)*. Melbourne: Australian Council for Educational Research.
- Fraser, B. J. (1981). *Test of Science-Related Attitudes (TOSRA)*. Melbourne, Australia: Australian Council for Educational Research.
- Fraser, B. J. (1982). Differences between student and teacher perceptions of actual and preferred classroom learning environment. *Educational Evaluation and Policy Analysis*, 4, 511–519.
- Fraser, B. J. (1986). *Classroom environment*. London: Croom Helm
- Fraser, B. J. (1990). *Individualised Classroom Environment Questionnaire*. Melbourne, Australia: Australian Council for Educational Research.
- Fraser, B. J. (1994). Research on classroom and school climate. In D. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 527–564). New York: Macmillan.
- Fraser, B. J. (1998a). Science learning environments: Assessments, effects and determinants. In B.J. Fraser and K.G. Tobin (Eds.), *International handbook of science education* (pp. 527–564). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Fraser, B. J. (1998b) Classroom environment instruments: Development, validity and applications. *Learning Environments Research*, 1, 7–33.
- Fraser, B. J. (2007). Classroom learning environments. In S. K. Abell and N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 103–124). Mahwah, NJ: Lawrence Erlbaum.

- Fraser, B. (2012). Classroom learning environments: Retrospect, context and prospect. In B. J. Fraser, K. Tobin & C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 1191–1239). New York: Springer.
- Fraser, B. J., Aldridge J. M., & Adolphe, F. S. G. (2010). A cross-national study of secondary science classroom environments in Australia and Indonesia. *Research in Science Education, 40*, 551–571.
- Fraser, B. J., Aldridge, J. M., & Soerjaningsih, W. (2010). Instructor-student interpersonal interaction and student outcomes at the university level in Indonesia. *The Open Education Journal, 3*, 32–44.
- Fraser, B. J., Anderson, G. J., & Walberg, H. J. (1982). *Assessment of learning environments: Manual for Learning Environment Inventory (LEI) and My Class Inventory (MCI)* (third version). Perth, Australia: Western Australian Institute of Technology.
- Fraser, B. J., & Butts, W. L. (1982). Relationship between levels of perceived classroom individualization and student-related attitudes. *Journal of Research in Science Teaching, 19*, 143–154.
- Fraser, B. J., & Fisher, D. (1982). Predicting students' outcomes from their perceptions of classroom psychosocial environment. *American Educational Research Journal, 19*, 498–518.
- Fraser, B. J., & Fisher, D. L. (1983a). Student achievement as a function of person-environment fit: A regression surface analysis. *British Journal of Educational Psychology, 53*, 89–99.

- Fraser, B. J., & Fisher, D. L. (1983b). Use of actual and preferred classroom environment scales in person-environment fit research. *Journal of Educational Psychology, 75*, 303–313.
- Fraser, B. J., McRobbie, C.J., & Fisher D. L. (1996, April). *Development, validation and use of personal and class forms of a new classroom environment instrument*. Paper presented at the annual meeting of the American Educational Research Association, New York.
- Fraser, B. J., & McRobbie, C. J. (1995). Science laboratory classroom environments at schools and universities: A cross-national study. *Educational Research and Evaluation, 1*, 289–317.
- Fraser, B. J., Giddings, G. J., & McRobbie, C. J. (1995). Evolution and validation of a personal form of an instrument for assessing science laboratory classroom environments. *Journal of Research in Science Teaching, 32*, 399–422.
- Fraser, B. J., & O'Brien, P. (1985). Student and teacher perceptions of the environment of elementary-school classrooms. *Elementary School Journal, 85*, 567–580.
- Gay, L. R., Mills, G. E., Airasian, P. W., III. (2006). *Educational research competencies for analysis and applications*. Upper Saddle River, NJ: Pearson Education.
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Boston: Allyn & Bacon.
- Germann, P., Dumas, W., & Barrow, L. (1995). Preparation of middle school science teachers in major state research universities. *Journal of Science Teacher Education, 6*, 143–145.

- Goh, S. C., & Fraser, B. J. (1996). Validation of an elementary school version of the Questionnaire on Teacher Interaction. *Psychological Reports, 79*, 512–522.
- Goh, S. C., & Fraser, B. J. (1998). Teacher interpersonal behaviour, classroom environment and student outcomes in primary mathematics in Singapore. *Learning Environments Research, 1*, 199–229.
- Helding, K., & Fraser, B. (2013). Effectiveness of National Board Certified (NBC) teachers in terms of classroom environment, attitudes and achievement among secondary science students. *Learning Environments Research, 16*, 1–21.
- Hoang, T. N. (2008). The effect of grade level, gender, and ethnicity on attitude and learning environment in mathematics in high school. *International Electronic Journal of Mathematics, 3*, 48–59.
- Jinks, J. L., & Morgan, V. (1999). Children's perceived academic self-efficacy: An inventory scale. *Clearing House, 72*, 224–230.
- Johnson, B., & McClure, R. (2004). Validity and reliability of a shortened, revised version of the Constructivist Learning Environment Survey (CLES). *Learning Environments Research, 7*, 65–80.
- Kaiser, H. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika, 23*, 187–200.
- Khine, M. S., & Fisher, D. L. (2004). Teacher interaction in psychosocial learning environments: Cultural differences and their implications in science instruction. *Research in Science and Technological Education, 22*, 99–111.
- Khoo, H. S., & Fraser, B. J. (2008). Using classroom psychosocial environment in evaluation of adult computer application courses in Singapore. *Technology, Pedagogy and Education, 17*, 67–81.

- Kim, H. B., Fisher, D. L., & Fraser, B. J. (1999). Assessment and investigation of constructivist science learning environments in Korea. *Research in Science and Technological Education, 17*, 239–249.
- Kim, H., Fisher, D. L., & Fraser, B. J. (2000). Classroom environment and teacher interpersonal behavior in secondary classes in Korea. *Evaluation and Research in Education, 14*, 3–22.
- Klopfer, L. E. (1971). Evaluation of learning in science. In B. S. Bloom, J. T. Hastings, and G. F. Madaus (Eds.), *Handbook on summative and formative evaluation of student learning* (pp. 559–641). New York: McGraw-Hill.
- Kokkinos, C., Charalambous, K., & Davazoglou, A. (2009). Interpersonal teacher behaviour in primary school classrooms: A cross-cultural validation of a Greek translation of the Questionnaire on Teacher Interaction. *Learning Environments Research, 12*, 101–114.
- Koul, R. B., & Fisher, D. L. (2005). Cultural background and students' perceptions of science classroom learning environment and teacher interpersonal behavior in Jammu, India. *Learning Environments Research, 8*, 195–211.
- Levi-Strauss, C. (1964). *Le cru et le cuit*. Paris: Plon.
- Levy, J., den Brok, P., Wubbels, Th., & Brekelmans, M. (2003). Students' perceptions of interpersonal aspects of the learning environment. *Learning Environments Research, 6*, 5–36.
- Levy, J., Wubbels, Th., Brekelmans, M., & Morganfield, B. (1997). Language and cultural factors in students' perceptions of teacher communication style. *International Journal of Intercultural Relations, 21*, 29–56.
- Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw.

- Lightburn, M. E., & Fraser, B. J. (2007). Classroom environment and student outcomes among students using anthropometry activities in high-school science. *Research in Science & Technological Education*, 25, 153–166.
- Lim, D. C. T., & Fraser, B. J. (2012, April). *Learning environments in English classrooms in Singapore: Determinants and effects*. Paper presented at the annual meeting of the American Educational Research Association, Vancouver, Canada.
- Liu, X. (2006). Student competence in understanding the matter concept and its implications for science curriculum standards. *School Science & Mathematics*, 106, 220–227.
- Liu, C., & Yager, R. E. (1997). The Iowa scope, sequence, and coordination project: A middle school science reform program approved by the national diffusion network. *Research in Middle Level Education Quarterly*, 20, 77–105.
- Lyons, T. (2006). Different countries, same science classes: Students' experiences of school science in their own words. *International Journal of Science Education*, 28, 591–613.
- MacLeod, C., & Fraser, B. J. (2010). Development, validation and application of a modified Arabic translation of the What Is Happening In this Class? (WIHIC) questionnaire. *Learning Environments Research*, 13, 105–125.
- Mager, R. F. (1968). *Developing attitude toward learning*. Belmont, CA: Fearon Publishers.
- Majeed, A., Fraser, B. J., & Aldridge, J. M. (2002). Learning environment and its association with student satisfaction among mathematics students in Brunei Darussalam. *Learning Environments Research*, 5, 203–226.

- Martin-Dunlop, C., & Fraser, B. J. (2008). Learning environment and attitudes associated with an innovative course designed for prospective elementary teachers. *International Journal of Science and Mathematics Education*, *6*, 163–190.
- Mason, L., Boscolo, P., Tornatora, M., & Ronconi, L. (2013). Besides knowledge: A cross-sectional study on the relations between epistemic beliefs, achievement goals, self-beliefs, and achievement in science. *Instructional Science*, *41*, 49–79.
- McRobbie, C. J., & Fraser, B. J. (1993). Associations between student outcomes and psychosocial science environment. *Journal of Educational Research*, *87*, 78–85
- Moos, R. H. (1974). *The social climate scales: An overview*. Palo Alto, CA: Consulting Psychologists Press.
- Moos, R. H. (1979). *Evaluating educational environments: Procedures, measures, findings and policy implications*. San Francisco, CA: Jossey-Bass.
- Moos, R. H., & Trickett, E. J. (1974). *Classroom Environment Scale manual* (1st ed.). Palo Alto, CA: Consulting Psychologist Press.
- Moss, C.H. (2003). *Improving learning environment and student incomes in biology in North Carolina*. Unpublished doctoral thesis, Curtin University of Technology, Perth Western Australia.
- Murray, H. A. (1938). *Explorations in personality*. New York: Oxford University Press.
- National Academies Press. (1996). *National science education standards*. Retrieved February 8, 2012 from <http://www.nap.edu/openbook.php> .

- National Commission on Excellence in Education (1983). *A nation at risk: The imperative for educational reform*. Retrieved February 8, 2012 from <http://www2.ed.gov/pubs/NatAtRisk/title.html>.
- National Science Teachers Association. (1992). *Scope, sequence, and coordination: The content core: A guide to curriculum Designers*. Washington, DC: National Science Teachers Association.
- New York State Education Department. (n.d.). *Grade 8 intermediate-level science test manual for administrators and teachers*. Retrieved October 30, 2008 from <http://www.emsc.nysed.gov/osa/elintsci.html>
- Nelson S. W., McGhee M. W., Meno L. R., & Slater C. L. (2007). Fulfilling the promise of educational accountability. *Phi Delta Kappan*, 88, 702–709.
- Nix, R. K., & Fraser, B. J. (2011). Using computer-assisted teaching to promote constructivist practices in teacher education. In B. A. Morris & G. M. Ferguson (Eds.), *Computer-assisted teaching: New developments* (pp. 93–115). New York: Nova Science Publishers.
- Nix, R. K., Fraser B. J., & Ledbetter, C. E. (2005) Evaluating an integrated science learning environment using the Constructivist Learning Environment Survey. *Learning Environments Research*, 8, 109–133.
- Nolen, S. B. (2003). Learning environment and achievement in high school science. *Journal of Research in Science Teaching*, 40, 347–368.
- Nolen, S. B., & Haladyna, T. M. (1990). Personal and environmental influences on students' beliefs about effective study strategies. *Contemporary Educational Psychology*, 15, 116–130.
- Peiro, M. M., & Fraser, B. J. (2009). Assessment and investigation of science learning environments in the early childhood grades. In M. Ortiz & C. Rubio (Eds.),

Educational evaluation: 21st century issues and challenges (pp. 349–365).

New York: Nova Science Publishers.

Pickett, L. H., & Fraser, B. J. (2009). Evaluation of a mentoring program for beginning teachers in terms of the learning environment and student outcomes in participants' school classrooms. In A. Selkirk & M. Tichenor (Eds.), *Teacher education: Policy, practice and research* (pp. 1–15). New York: Nova Science Publishers.

Pointon, P. (2000). Students' views of environments for learning from the primary to the secondary school. *International Journal of Educational Research*, 33, 375–382.

Quek, C. L., Wong, A. F. L., & Fraser, B. J. (2005). Student perceptions of chemistry laboratory learning environments, student-teacher interactions and attitudes in secondary school gifted education classes in Singapore. *Research in Science Education*, 35, 399–321.

Rentoul, A. J., & Fraser B. J. (1979) Conceptualization of enquiry-based or open classroom learning environments. *Journal of Curriculum Studies*, 11, 233–245.

Rickards, T., den Brok, P., & Fisher, D. (2005). The Australian science teacher: A typology of teacher–student interpersonal behaviour in Australian science classes. *Learning Environments Research*, 8, 267–287.

Rita, R., & Martin-Dunlop, C. (2011). Perception of the learning environment and associations with cognitive achievement among gifted biology students. *Learning Environments Research*, 14, 25–38.

- Rushton, S., & Rushton, A. (2008). Classroom learning environment, brain research and the No Child Left Behind initiative: 6 years later. *Early Childhood Education Journal, 36*, 87–92.
- Scott Houston, L., Fraser, B. J., & Ledbetter, C. E. (2008). An evaluation of elementary school science kits in terms of classroom environment and student attitudes. *Journal of Elementary Science Education, 20*, 29–47.
- Seker, H. (2011). Developing a questionnaire on attitudes towards school. *Learning Environments Research, 14*, 241–261.
- Sink, C. A., & Spencer, L. R. (2005). My Class Inventory–Short Form as an accountability tool for elementary school counsellors to measure classroom climate. *Professional School Counseling, 9*, 37–48.
- Speering, W., & Rennie, L. (1996). Students' perceptions about science: The impact of transition from primary to secondary school. *Research in Science Education, 26*, 283–298.
- Stern, G. G., Stein, M. I., & Bloom, B. S. (1956). *Methods in personality assessment*. Glencoe, IL: The Free Press.
- Taylor, P. C. & Fraser, B. J. (1991, April). *Development of an instrument for assessing constructivist learning environments*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Taylor, P. C., Fraser, B. J., & Fisher, D. L. (1997). Monitoring constructivist classroom learning environments. *International Journal of Educational Research, 27*, 293–302.

- Telli, S., den Brok, P., & Cakiroglu, J. (2010). The importance of teacher–student interpersonal relationships for Turkish students’ attitudes towards science. *Research in Science and Technological Education*, 28, 261–276.
- Texas Education Agency. (n.d.). *Texas Administrative Code (TAC), Title 19, Part II Chapter 112. Texas Essential Knowledge and Skills for Science*. Retrieved October 18, 2006 from <http://www.tea.state.tx.us/rules/tac/chapter112/index.html>.
- Thompson, B. (1998). Review of ‘what if there were no significance tests?’ *Educational and Psychological Measurement*, 58, 334–346.
- Tobin, K., & Fraser, B. J. (1998). Qualitative and quantitative landscapes of classroom learning environments. In B. J. Fraser & K. G. Tobin (Eds.), *The international handbook of science education* (pp. 623–640). Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Tulloch, D. (2011). *Determinants and effects of the learning environment in college classes*. Unpublished thesis, Curtin University of Technology.
- United States Census Bureau. (n.d.). *State and county quick facts, California*. Retrieved Feb 2, 2013 from <http://quickfacts.census.gov/qfd/states/06000.html>.
- United States Census Bureau. (n.d.). *State and county quick facts, Texas*. Retrieved Feb 2, 2013 from <http://quickfacts.census.gov/qfd/states/48000.html>.
- Valentine, J. W., & Mogar, D. C. (1992). Middle level certification: An encouraging evaluation. *Middle School Journal*, 24, 36–43.
- van Rensburg, S., Ankiewicz, P., & Myburgh, C. (1999). Assessing South Africa learners’ attitudes towards technology by using the PATT (Pupils’ Attitudes

- Towards Technology) questionnaire. *International Journal of Technology and Design Education*, 9, 137–151.
- Walberg, H. J. (1991). Improving school science in advanced and developing countries. *Review of Educational Research*, 61, 25–69.
- Walberg, H. J., & Anderson, G. J. (1968). Classroom climate and individual learning. *Journal Educational Psychology*, 59, 414–419.
- Waldrip, B. G., Fisher, D. L., & Dorman, J. P. (2009). Identifying exemplary science teachers through students' perceptions of their learning environment. *Learning Environments Research*, 12, 1–13.
- Wei, M., den Brok, P., & Zhou, Y. (2009). Teacher interpersonal behaviour and student achievement in English as a foreign language classrooms in China. *Learning Environments Research*, 12, 157–174.
- Welch, A. G., Cakir, M., Peterson, C. M., & Ray, C. M. (2012). A cross-cultural validation of the Technology-Rich Outcomes-Focused Learning Environment Inventory (TROFLEI) in Turkey and the USA. *Research in Science & Technological Education*, 30, 49–63.
- Wolf, S. J., & Fraser, B. J. (2008). Learning environment, attitudes and achievement among middle-school science students using inquiry-based laboratory activities. *Research in Science Education*, 38, 321–341.
- Wong, A. F. L., & Fraser, B. J. (1995). Cross-validation in Singapore of the Science Laboratory Environment Inventory. *Psychological Reports*, 76, 907–911.
- Wong, A. F. L., & Fraser, B. J. (1996). Environment-attitude associations in the chemistry laboratory classroom. *Research in Science & Technological Education*, 14, 91–102.

Wong, A. L. F., Young, D. J., & Fraser, B. J. (1997). A multilevel analysis of learning environments and student attitudes. *Educational Psychology, 17*, 449–468

Wubbels, Th., & Levy, J. (Eds.). (1993). *Do you know what you look like: Interpersonal relationships in education*. London, UK: Falmer Press.

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.

APPENDIX A

Outcomes-Related Learning Environment Survey (ORLES)

Outcomes-Related Learning Environment Survey 3.6

1. Outcomes-Related Learning Environment Survey 3.6

This voluntary survey contains 43 statements about you and your science class. You will be asked how often each practice takes place. Your answer choices will include:

= Almost Always, = Often, = Sometimes, = Seldom, and = Almost Never

There are no 'right' or 'wrong' answers. We want to know how you feel about each statement. Please use a pencil and completely fill in the circle that shows your choice.

- First, read the statement. Don't worry if some of the statements sound the same.
- Next, think about how you feel right now and mark the best choice.
- Then read the next statement and so on for each statement on the back of this page.

If you change your mind, just erase your first mark and fill in another circle.

THIS IS NOT GRADED. YOUR TEACHER WON'T EVEN LOOK AT YOUR QUESTIONNAIRE RESPONSES!

1. Which School Do you attend?

2. Gender

Male

Female

3. What is your ethnicity?

African American

Asian/Pacific Islander

Hispanic

Native American/Other

White

4. I learn about the world outside of school.

Almost Always

Often

Sometimes

Seldom

Almost Never

Page 1

Numerous items in this survey are based on the What Is Happening In this Class? (WIHIC) survey (Fraser, McRobbie & Fisher, 1996), Constructivist Learning Environment Survey (CLES, Taylor, Fraser & Fisher, 1997) and the Test of Science-Related Attitudes (TOSRA, Fraser, 1981). These questionnaire items were used in my study and included in this thesis with the permission of the authors.

Outcomes-Related Learning Environment Survey 3.6

5. My new learning starts with problems about the world outside of school.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

6. I learn how science can be part of my out-of-school life.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

7. I get a better understanding of the world outside of school.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

8. I learn interesting things about the world outside of school.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

Outcomes-Related Learning Environment Survey 3.6

9. I make friendships among students in this class.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

10. I know other students in this class.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

11. I do favors for members of this class.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

12. I work well with other class members.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

Outcomes-Related Learning Environment Survey 3.6

13. I help other class members who are having trouble with their work.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

14. In this class, I get help from other students.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

15. The teacher goes out of his/her way to help me.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

16. The teacher's questions help me understand.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

Outcomes-Related Learning Environment Survey 3.6

17. It's alright for me to tell the teacher that I don't understand.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

18. The teacher talks with me

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

19. The teacher considers my feelings.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

20. It is alright with the teacher if I'm slower than other students in the class.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

Outcomes-Related Learning Environment Survey 3.6

21. I discuss ideas in class.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

22. The teacher asks me questions.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

23. My ideas and suggestions are used during discussions.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

24. When starting a new topic, I discuss what I already know.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

Outcomes-Related Learning Environment Survey 3.6

25. I am asked to explain how I solve problems.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

26. I discuss different answers to questions.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

27. I draw conclusions from investigations.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

28. I am asked to think about evidence for statements.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

Outcomes-Related Learning Environment Survey 3.6

29. I carry out investigations to answer questions which puzzle me.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

30. I carry out investigations to answer the teacher's questions.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

31. I solve problems by obtaining information from the library.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

32. I solve problems by using information obtained from my own investigation.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

Outcomes-Related Learning Environment Survey 3.6

33. I know what has to be done in this class.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

34. Class assignments are clear so I know what to do.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

35. I know the goals for this class.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

36. I pay attention during this class.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

Outcomes-Related Learning Environment Survey 3.6

37. I try to understand the work in this class.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

38. I know how much work I have to do.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

39. Science lessons are fun.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

40. I enjoy doing science lessons.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

Outcomes-Related Learning Environment Survey 3.6

41. School should have more science lessons each week.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

42. I look forward to science lessons.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never

43. Science is one of the most interesting school subjects.

- Almost Always
- Often
- Sometimes
- Seldom
- Almost Never