**Science and Mathematics Education Centre**

# Effectiveness of National Board Certified Teachers in Terms of Classroom Environment, Attitudes and Achievement Among Secondary Science Students

**Karen A. Helding**

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

**April, 2006**

## 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 acknowledgement has been made.

Signature:

Date:

## ABSTRACT

A United States organization, called the National Board of Professional Teaching Standards (NBPTS), was initiated to strengthen the pedagogy of teaching and, subsequently, improve student achievement. The purpose of this study was to investigate the effectiveness of the NBPTS in terms of whether National Board Certified (NBC) teachers are effective in promoting positive classroom environments and student attitudes, and enhancing student achievement. The sample consisted of 927 Grade 8 and 10 science students from 12 secondary schools. Altogether, 443 students in 21 classes comprised the NBC teacher group and 484 students in 17 classes comprised the non-NBC teacher group. Students completed a learning environment questionnaire, the What Is Happening In this Class? (WIHIC), and an attitude scale based on the Test Of Science-Related Attitude (TOSRA). Scores from the science portion of the Florida Comprehensive Assessment Test FCAT, a state-mandated examination, were collected to measure achievement.

This research is unique in that it is the first time that a learning environments study has included a sample of National Board Certified (NBC) and non-NBC teachers in order to compare their effectiveness in terms of secondary students' perceptions of their science learning environment, attitudes toward science, and science achievement. The study revealed that the revised version of the WIHIC and the modified attitude scale are valid and reliable instruments for assessing perceptions of the classroom environment and attitudes toward science among secondary science students in Miami-Dade County, Florida. In addition, a much stronger association with learning environment was found for students' attitude than for students'

achievement. The contributions and significance of this study are not only that it adds to the area of research that pertains to the efficacy of NBC teachers, but it also adds to the field of learning environments research. This study is useful because it could be replicated to provide additional empirical evidence about the effect National Board teachers have on students in the classroom and add to the growth of educational data on the impact of National Teacher Certification and classroom learning environments research.

## DEDICATION

This thesis is largely dedicated to my Parents, Phil and Judy, for always believing in me and supporting me emotionally during many personal predicaments that I have created throughout my life. Additionally, I would like to especially thank my mother in helping me reach my educational goals. I can recall telling my mother in sixth grade I didn't think I would be able to make it in Junior High School. After completing Junior High school, I recalled informing her I didn't think that I would be able to make it in High School, and the same holds true for college. She never faltered, or got upset when I expressed these fears, she just always reminded me to do the best I can and to never quit. Without her support and her beliefs instilled in me, I would have never even considered this endeavor.

I am grateful to my loving husband, Paul, for initiating my pursuit in a Doctorate degree. Bless his heart; he always tried to provide quiet time for me to work on the thesis, but it hardly ever seemed to work out in my favor.

To Christian, my handsome little son, who played on the rug next to me while I worked on my thesis or slept on the bed beside me while I typed through out the night. You made sure we were always together whenever I 'did my homework'. I hope when you are older that you understand the importance of education and follow in my footsteps.

## ACKNOWLEDGEMENTS

There are several people without whose help this thesis would not of reached completion.

I wish to express my sincere appreciation to my supervisor, Barry Fraser, for his advice, time, effort and patience that he gave me throughout these past four years. It has been an honor to be directed by one of the greatest educational leaders of the time.

To Dr. Jill Aldridge, for her expeditious feedback to all of my questions and concerns. It is quite a feat to be able to successfully direct a person mathematically. But you do it with such style and ease that you make anything seem simplistic.

I would also like to acknowledge the help received by my colleague Dr. Maria Peiro who provided endless and invaluable support concerning every aspect of the study.

I would also like to thank the District Director of Assessment and Data Analysis, Gisela Field in facilitating in the collection of data at different schools. Without her, the study would have been sabotaged.

Lastly, I would like to thank those who participated in my study. Thank you to the teachers taking time out from their class activities in order for me to conduct the surveys and to the students who genuinely answered all the questions honestly.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# Chapter 1

# RATIONALE AND BACKGROUND

## 1.1 Introduction

At the national, state, and local levels in the U.S., much interest is placed on the quality of education that is provided to our nation's students. Consequently, a United States organization, called the National Board for Professional Teaching Standards (NBPTS), was initiated to strengthen the pedagogy of teaching and, subsequently, to lead to improvement in student achievement. NBPTS strongly believes that high-quality teachers are necessary for successful student learning. Therefore, its main goal is to advance the quality of teaching and learning by providing a national system of certification for teachers who meet high and rigorous standards for what accomplished teachers should know and be able to do. NBPTS is a non-profit organization that is heavily funded by the U.S. Department of Education and, as with any government-funded program, its effectiveness determines future funding of the program.

Considering that the U.S. government is increasingly seeking value for its money, this study investigated the efficacy of NBPTS by comparing the effectiveness of NBC (National Board Certified) and non-NBC science teachers in terms of their Grade 8 and 10 students' classroom environment perceptions, attitudes, and achievement. The What Is Happening In this Class? (WIHIC) questionnaire (Aldridge & Fraser, 2000) was used to assess classroom environment, an attitude scale derived from the Test of

Science-Related Attitudes (TOSRA) (Fraser, 1981) assessed student attitudes, and students' scores on the science component of Florida Comprehensive Achievement Test (FCAT) were utilized for assessing student achievement. What is unique about this study in that no previous research delved into whether NBC teachers are more effective in promoting positive student attitudes and/or classroom environments than non-NBC teachers.

This chapter describes the context of Miami-Dade County, Florida, where the study took place (Section 1.2), the background to my study (Section 1.3), the purpose of my study (Section 1.4) and the research questions (Section 1.5). Also presented in this chapter is an overview of the organization of the chapter in the remainder of the thesis (Section 1.6).

**1.2 About Miami-Dade County, Florida**

Greater Miami is regarded as a cultural melting pot due to its diverse population. The majority of Miami's population originates from Latin America and the Caribbean, and 59.5% of Miami's population are foreign born. The 2000 census reported that 65.76% of Miami's population is Latino. Additionally, the median income for a family in Miami is $27,225 U.S. Therefore, it is understandable why 23.5% of Miami families are below the poverty line (U.S. Census Bureau, 2000).

Miami's school district, where the study was carried out, is called Miami-Dade County Public Schools (M-DCPS). M-DCPS is the fourth largest school district in

the U.S. It's divided into six regions and has approximately 145,211 students. The ethnic make-up of the school district is 10.0% White, 29% Black, 59% Hispanic, and 2% other nationalities (Miami-Dade County Public Schools, 2004).

As for the set of academic courses, the Competency Based Curriculum (CBC) is the mandated curriculum in Miami-Dade County Public Schools. The science portion of the CBC is directly aligned with the National Science Educational Standards (standards that outline what America's students need to know, understand, and be able to do at different grade levels to reach scientific literacy) and Florida's Sunshine State Standards (standards created by the Florida's State Board of Education that mandate what each child is expected to achieve at each grade level).

In Miami-Dade County Public Schools, Grade 8 students have two basic science courses to choose from: General Science and Honors Earth Space Science. Some schools, due to the proximity of agriculture and its influence in the surrounding area, offer Agriscience as an elective. In addition to the two basic courses, classes could be divided into three academic levels: regular, honors and gifted honors. In this study, the students who participated were registered in one of the following courses: General Science, Earth Space Science (Regular), Honors Earth Space Science and Agriscience.

As for Grade 10 students, their courses are Earth Space Science, Biology and one science elective, respectively. A science elective can consist of Physical Science, Chemistry, Physics, Oceanography, Environmental Science or Agriscience. Again, these classes can be divided into four academic levels: regular, honors, gifted honors

and Advanced Placement (AP). Advanced Placement classes allow students to receive college credit for the class if they pass the AP state examination. In my study, the students who participated were registered in one of the following courses: Biology, Honors Biology, Gifted Honors Biology, Chemistry (Regular), Gifted Honors Chemistry, AP Chemistry and Agriscience.

Secondary schools in Miami-Dade County Public Schools (M-DCPS) are for the most part divided into two subgroups. Grades 6–8 are called *middle* schools and Grades 9–12 are called *senior* high schools. These two subgroups are not schooled at the same location; therefore, they report to different locations throughout the county. At the present time, M–DCPS has 54 middle schools and 40 senior high schools scattered throughout Miami. This secondary school count does not include choice schools such as K–8 Centers, Alternative Schools, Charter Schools, Specialized Centers and Vocational/Adult Education Centers that offer parents different options from the traditional schools.

Currently, in order for a child to be promoted to the next grade level in middle school, he or she needs to pass the required mathematics, science, language arts, and history classes and two elective classes with at least a letter grade of a D. In Grades 1–12, a common report card grading system is used. Academic letter grades for students are A, B, C, D, or F. Table 1.1 provides detailed information about the grading system used in Miami-Dade County Public Schools.

In order to receive a senior high school diploma, a student in M-DCPS must not only maintain a cumulative GPA (Grade Point Average) of 2.0 in all coursework, but also

pass the Grade 10 Florida Comprehensive Assessment Test (FCAT). Presently, reading, writing and mathematics are the three areas the students need to pass on the FCAT in order to graduate. However, science will be added as a graduation component on the FCAT for the school year 2006–2007.

The components of the FCAT assess a wide range of material. For example, the Reading test employs a wide variety of writing material to assess students' reading comprehension. Writing emphasizes the areas of focus, organization, support, and conventions. Mathematics assesses Number Sense, Concepts, and Operations, Measurement, Geometry and Spatial Sense, Algebraic Thinking, Data Analysis and Probability. While the Science test concentrates on Physical and Chemical Sciences, Earth and Space Sciences, Life and Environmental Sciences and Scientific Thinking (Florida Department of Education, 2004).

Table 1.1 Academic Grading System for Miami-Dade County Public Schools

| Letter Grade | Numerical Value | Verbal Interpretation | Grade Point Average |
|---|---|---|---|
| A | 90-100% | Outstanding progress | 4 |
| B | 80-89% | Above average progress | 3 |
| C | 70-79% | Average progress | 2 |
| D | 60-69% | Lowest acceptable progress | 1 |
| F | 0-59% | Failure | 0 |

Adapted from Miami-Dade County Public Schools (2003)

All subjects tested on the Florida Comprehensive Assessment Test (FCAT) is specifically designed for Florida and measures how well students are progressing toward meeting the benchmarks in the Florida Sunshine State Standards (Miami-Dade County Public Schools, 2004).

## 1.3 Background

This section provides background information relevant to the present study, including a brief introduction to National Board Professional Teaching Standards (Section 1.2.1) and the field of learning environments (Section 1.2.2).

### *1.3.1 Background to the National Board Professional Teaching Standards*

National Board for Professional Teaching Standards (NBPTS) certification is market driven because of the relatively large amount of money that becomes available to the educator after he or she becomes National Board Certified. Currently, the U.S. federal government has a big stake in the success of NBPTS, because it contributes more than 55% of its operating costs, while non-governmental sources fund the remaining amount (NBPTS, 2001). Consequently, federal and local governments are increasingly looking to justify the enormous expenditure for the operational costs and for the provision of monetary rewards given to individual teachers who become National Board Certified. Some research has been conducted to evaluate the impact of NBC (National Board Certified) teachers on students' achievement; however, few research studies have investigated other outcomes. Therefore, in addition to past research on evaluating the effectiveness of NBC (National Board Certified) teachers

in terms of their students' achievement, new research that focuses on the classroom environments and attitudes of students of these teachers could provide further justification for the enormous expenditure on NBPTS certification.

The National Board for Professional Teaching Standards (NBPTS) conducted an internal investigation of the value of NBPTS certification in order to further promote its credibility as a national teaching certificate. In 2000 and 2001, two studies compared NBPTS teachers to their counterparts: *A Distinction That Matters: Why National Teacher Certification Makes a Difference* (NBPTS, 2000) and *I Am a Better Teacher: What Candidates for National Board Certification Say about the Assessment Process* (NBPTS, 2001). These studies revealed that NBPTS teachers have stronger teaching characteristics than their counterparts including: an extensive knowledge of subject matter; adapting and improvising instruction; designing lessons that are challenging and engaging; promoting academic achievement by emphasizing both personal accomplishment and intellectual engagement and providing healthier teacher-student relationships. But controversy has loomed over these two reports due to inadequate research methods and for not investigating student achievement within the studies.

Because of this controversy about NBPTS, the organization has asked for the assistance of universities and research firms in conducting investigations of the impact of National Board Certification on teachers, students, public policies and educational reforms.

Stone (2002) conducted external research on National Board Certified teachers. In *Value-Added Achievement Gains of NBPTS-Certified Teachers in Tennessee: A Brief Report,* Stone reported that NBPTS-certified teachers in Tennessee were only average 'producers' of student achievement gains as compared with other teachers in their school districts. Most of the large body of research reported almost no link between certification and teacher effectiveness. Therefore, he stated that it made no sense to expect a relationship between being 'super-certified' and student performance.

However, much previous research reports positive associations between NBPTS certification and student achievement. For example, a multi-year study by the University of Washington, which was funded by the U.S. Department of Education, confirmed the effectiveness of National Board Certification by collecting more than 600,000 student records from students in North Carolina schools. It was found that children learned more from National Board Certified teachers (Goldhaber & Anthony, 2004). Also, Vandevoort, Amrein-Beardsley, and Berliner (2004), from Arizona State University, completed a study with students in the elementary school classrooms of 35 National Board Certified teachers and their non-certified colleagues in 14 Arizona school districts. It was found that students who attended classes taught by NBC (National Board Certified) teachers had greater academic gains than students in the classrooms of non-NBC teachers. Additionally, the CNA Corporation (CNAC) reported that Florida's Miami-Dade County Public Schools' high school mathematics teachers who had achieved National Board Certification helped their students to achieve larger testing gains than did colleagues who had not earned certification (Cavalluzzo, 2004).

Generally, current research on teaching reveals that teachers are influential in terms of students' academic achievement (NBPTS, 2001). However, the types of interrelationships between students and teacher that are associated with high-quality and effective teaching have yet to be sufficiently determined. Nonetheless, the National Board for Professional Teaching Standards (NBPTS) created a 'super certification', which outlines the components of a superior teacher. It did this without empirical evidence to support its claim that teachers who meet the standards set by the Board are superior to their counterparts in promoting academic achievement. In the 17 years since the founding of NBPTS, only a few empirical studies have addressed this important issue (Vandevoort, Amrein-Beardsley, & Berliner, 2004). The present study might serve as empirical evidence to support the effectiveness of National Board Certified (NBC) teachers because we compared the effectiveness of NBC and non-NBC teachers in terms of their science students' perceptions of the classroom learning environment, attitudes toward science, and science achievement.

### *1.3.2 Background to the Field of Learning Environments*

My study was grounded in and contributed to the field of learning environments. Two of the first formal studies relevant to the field of learning environments go back over 60 years and were most likely the first to recognize relationships between the environment and human behavior. Lewin's (1936) theory dealt with the relationship and interaction between the individual and his/her environment. Murray (1938) was the first person to use Lewin's scheme to propose a needs-press model, which refers to situational variables that are found in the environment and account for a degree of behavioral difference. The field of learning environments research built on Lewin

(1936) and Murray (1938), who probably created the foundation for the development of the first learning environment scales.

Following the work of Lewin and Murray, two research programs involved developing instruments that could be used to assess the learning environment. The modern era of learning environment research commenced when Rudolf Moos (1974) and Herbert Walberg (1968) began independent lines of research on the conceptualization and assessment of psychosocial environments. Herbert Walberg's Learning Environment Inventory (LEI) (Walberg & Anderson, 1968) and Rudolf Moos' Classroom Environment Scale (CES) (Moos, 1979; Moos & Trickett, 1987) were the first instruments developed to assess students' perceptions of their learning environment and paved the way for the subsequent development of other interpretations for learning environment research.

The impact of the learning environment on the education process has received a great deal of awareness, and there has been much development in terms of the conceptualization and assessment of learning environments (Fraser, 1994, 1998a). Both qualitative and quantitative methods have been used in conducting research in the field of learning environments, but the use of instruments to assess students' perceptions has been the predominant method. Learning environment questionnaires have been shown to be widely applicable in various classroom settings, in numerous countries around the world, and for a variety of purposes. The development of classroom environment instruments has facilitated learning environments research at the primary (Robinson, 2003), secondary (Fraser, 1989; Fraser, Dryden, & Taylor, 1998) and post-secondary (Yarrow, Millwater, & Fraser 1997) levels of education.

Learning environment research has become well established and internationally recognized over the past 35 years (Fraser, 1998a, 1999). For example, the American Educational Research Association has a Special Interest Group (SIG) specifically devoted to the study of learning environments.

As mentioned before, a great deal attention has been focused on studying student outcomes when conducting science education research. However, in the past 35 years, learning environments research has made remarkable progress by focusing on the assessment and investigation of the classroom environments worldwide. Much of the past research on learning environments has focused on students' perceptions of the classroom environment and its effects on student outcomes (Fraser, 1986, 1994, 1998a; Fraser & Walberg, 1991) and have provided compelling evidence that learning environments strongly influence student outcomes and play an important role in improving the effectiveness of learning (Fraser, 1998a, 2001). Furthermore, learning environment research has shown that teachers can improve classroom environments by introducing an intervention that addresses discrepancies between students's actual and preferred classroom learning environment (Fisher, Fraser, & Bassett, 1995; Fraser, 1998b).

My study is distinctive in that it compared National Board and non-National Board teachers in terms of their students' perceptions of the learning environment. The present study adds to the field of learning environments because it validated a widely-used and widely-applicable learning environment questionnaire among secondary science students. Also, it compared NBC (National Board Certified) and non-NBC science teachers in terms of their students' perceptions of their classroom learning

environment. In addition, it explored associations between the learning environment and student outcomes (attitudes and achievement) in science classrooms. A comprehensive review of literature for the field of learning environments can be found in Chapter 2.

## 1.4 Purpose of the Study

A main purpose of the present study was to provide important insights into the classroom learning environments of National Board Certified teachers and non-National Board Certified teachers at the secondary level in science classrooms in Miami-Dade County school district in Florida.

To date, there have been no studies of the learning environment that have been undertaken with National Board Certified Teachers. The study therefore has the potential to create a clearer picture of the differences between National Board teachers and non-National Board teachers in terms of the classroom learning environments that they create. In addition, this study could provide information to the directors of the National Board to guide them in amending the National Board standards to reflect their candidates' ability to modify their classroom environments and student interactions in order to cater more adequately for the needs of students.

A specific purpose of the present study was to examine similarities and differences between two distinct groups of students who were taught by two distinct groups of teachers, those who were National Board Certified and those who were not National Board Certified. By comparing the two different types of teaching licensure, the

study investigated which types of learning environments are most likely to enhance student outcomes in secondary science and identify ways in which the National Board can enhance the teaching and learning process.

The study also examined whether relationships exist between student cognitive and affective outcomes and the nature of the learning environment.

To measure learning environment, the What Is Happening In this Class (WIHIC)? questionnaire (Aldridge & Fraser, 2000) was selected. To measure students' achievement in science, their scores on the Florida Comprehensive Assessment Test (FCAT) were retrieved from the database of the Miami-Dade County school district in Florida. To measure students' attitudes toward science, a modified version of the *Test of Science-Related Attitudes* (TOSRA) was used (Fraser, 1981).

## 1.5 Research Questions

The present study addressed the following four main research questions:

1. Is a revised version of the What Is Happening In this Class? (WIHIC) questionnaire valid and reliable when used in secondary science classrooms in South Florida?

2. Is an attitude scale modeled on the Test Of Science-Related Attitudes (TOSRA) reliable when used in secondary science classrooms in South Florida?

3. Are National Board Certified (NBC) teachers more effective than non-NBC teachers in terms of classroom environment, student attitudes and student achievement in secondary science classrooms in South Florida?

4. Are there associations between student outcomes (attitudes and achievement) and classroom environment in secondary science classrooms in South Florida?

## 1.6 Organization of the Thesis

This thesis comprises five chapters. Chapter 1 discussed the rationale for the present study. It provided a brief background to the study including information about the National Board Professional Teaching Standards (NBPTS), the field of learning environments and the student achievement test, the Florida Comprehensive Assessment Test (FCAT). This chapter also discussed the purposes of the present study, gave an outline of research questions, and provided an overview of the organization of the thesis.

Chapter 2 reviews the background of National Board of Professional Teaching Standards and empirical evidence concerning National Board Certification. Also discussed in this chapter is the historical background of learning environment research and the numerous learning environment instruments available. In addition, the chapter reviews prior research on learning environments, including past studies involving the use of the What Is Happening In this Class? questionnaire and associations between learning environment and student outcomes.

Chapter 3 discusses methodology and provides insight into procedural aspects of the present study. This includes the research design used in different phases of the study, the choice of classes for the study, and the choice of the sample for the study. Also, discussed in this chapter are instrumentation of the study, field-testing of the instruments used, administration of the questionnaires, data collection, and the statistical procedures employed in the data analysis.

Chapter 4 reports the data analysis and findings for the present study, including: validation and reliability of the questionnaires; an investigation of whether if NBC (National Board Certified) teachers and non-NBC teachers are more effective in terms of learning environment and student outcomes; and an investigation of associations between the learning environment and students' achievement and attitudes.

Chapter 5 concludes the thesis with an overview of the whole thesis. Also, it discusses the findings of the study in terms of the validation of each assessment instrument, differences between NBC and non-NBC teachers' students in terms of learning environment perceptions, attitudes and achievement of students, and associations between the learning environment and students' achievement and attitudes. Furthermore, this chapter discusses the practical implications of the findings from this study, the significance of the study, limitations of the present study, and suggestions for further research on secondary science classrooms of National Board and non-National Board teachers in terms of learning environment, achievement and attitudes.

# Chapter 2

# LITERATURE REVIEW

## 2.1 Introduction

While there is almost unanimity that improving teacher quality is a top priority and a necessary precondition for boosting student achievement, there is less certainty about how to accomplish this. However, an educational reform movement in the United States stemming from the No Child Left Behind Act (NCLBA) of 2001 requires that all states ensure the availability of high-quality professional development for all teachers. The NCLBA has led educational scholars and policy makers in demanding professional development opportunities that will help teachers to enhance their knowledge and develop new instructional practices (Ball & Cohen, 1999). The National Board for Professional Teaching Standards (NBPTS) claims to have offered this high-quality professional development for the 40,200 teachers nationwide who have become National Board Certified (Keller, 2005). Furthermore, they claim that these teachers have passed their rigorous standards for certification and could be labeled as highly-qualified teachers.

Some research studies have been conducted to investigate the effectiveness of National Board Certification in terms of the quality of teachers' teaching practices (National Board for Professional Teaching Standards, 2000, 2001) and the improvement of student achievement (Cavalluzzo, 2004; Goldhaber & Anthony, 2004; Vandevoort, Amrein-Beardsley, & Berliner, 2004). However, none of the

documented research studies have focused on students' perceptions of their learning environments and their attitudes as outcome measures. My study is significant because it is the first documented research that applied a learning environment instrument and an attitudinal survey in assessing and investigating the differences in classroom learning environments and attitudes between students taught by National Board Certified (NBC) and non-NBC teachers.

In Chapter 1, I discussed the background of my study in terms of the NBPTS and the field of learning environments. I also discussed the context and purpose of my study. In this chapter, I review literature pertinent to the theoretical framework and methods of my research study. The literature review is organized into seven sections. Section 2.2 examines the background of the National Board for Professional Teaching Standards. Section 2.3 describes the empirical evidence, both positive and conflicting, concerning the effectiveness of National Board Certification. Section 2.4 discusses the historical background of the field of learning environments. Section 2.5 reviews eight learning environment instruments that are available to assess students' perceptions of their classroom learning environment. Section 2.6 discusses the development and characteristics of the widely-used What Is Happening In this Class? (WIHIC) learning environment questionnaire that was used in my study. A separate section is dedicated to the WIHIC because it was the learning environment questionnaire I used in my study to gather students' perceptions of their classroom learning environment. This section also discusses past validation studies using the WIHIC in Western and Asian countries. Section 2.7 discusses lines of past learning environments research with a specific focus on the two lines of research that particularly relevant to my study, namely, research on outcome-environment

associations and on determinants of classroom environment. Finally, Section 2.8 discusses literature about assessing attitudes to science, as well as the characteristics and validity of the Test Of Science-Related Attitudes (TOSRA), which I used in my study.

## 2.2 History of National Board for Professional Teaching Standards

The National Board for Professional Teaching Standards (NBPTS) was created in 1987 after the Carnegie Forum on Education and the Economy (1986) released *A Nation Prepared: Teachers for the 21st Century*. This piece of writing followed the landmark report *A Nation at Risk: The Imperative for Educational Reform* (U.S. Department of Education, 1983), which was developed by the President's Commission on Excellence in Education. Following the release of these important documents, educators, parents, business executives and legislators began worrying about the economic and social consequences of an educational system failing to keep pace with a changing America and global society. The report, *A Nation Prepared*, offered solutions such as urging the nation to set standards and to certify teachers who meet those standards. The members of the task force outlined a plan designed to retain, reward and advance accomplished teachers through a system of advanced certification. Hence, the National Board for Professional Teaching Standards (NBPTS) was created from the framework of these ideas.

The NBPTS is a nonprofit organization independently governed by a board of directors composed of individuals who are either directly or indirectly involved with the educational system. The NBPTS's basic task is to advance the quality of teaching

and learning by creating rigorous standards for what exemplary teachers should know and be able to do and to develop an assessment process to measure the teachers' ability to meet those standards. Teachers who go through this rigorous certification process and meet the standards set forth by the NBPTS become National Board Certified.

National Board Certification is valid for 10 years. This highly-developed system of National Board Certification complements, but does not replace, state licensing. Each state, school district and/or school decides how best to capitalize on the expertise of their National Board Certified teachers as it designs instructional arrangements to promote student learning and support professional practice. In addition, local governments and school districts offer local support and initiatives, such as payment of the certification fee, a means of meeting the requirements for teacher license renewal and licensure for out-of-state educator, and a salary increase for the life of the certificate.

Currently, a teacher can voluntarily become National Board Certified in one of the following areas: Generalist, Art, Career and Technical Education, English as a New Language, English Language Arts, Exceptional Needs, Library Media, Mathematics, Music, Physical Education, School Counseling, Science, Social Studies-History or World Languages other than English. In addition, each subject area has developmental levels: Early Childhood (ages 3–8 years), Early Childhood through Young Adulthood (ages 3–18+ years), Middle Childhood (ages 7–12 years), Early & Middle Childhood (ages 3–12 years), Early Adolescence (ages 11–15 years), Adolescence & Young Adulthood (ages 14–18+ years), and Early Adolescence

through Young Adulthood (ages 11–18+ years) and Exceptional Needs (ages 0–21+ years).

The National Board Certification process has evolved over the years into a more streamlined assessment. It is composed of the following five basic dimensions for what accomplished teachers should know and be able to do:

Deleted: ¶

1. Teachers are committed to students and their learning.
2. Teachers know the subjects that they teach and how to teach those subjects to students.
3. Teachers are responsible for managing and monitoring student learning.
4. Teachers think systematically about their practice and learn from experience.
5. Teachers are members of learning communities.

These five dimensions are the standards by which the teachers are evaluated.

Applicants who are interested in becoming National Board Certified must pay a one-time fee of US$2,300 to receive an instructional portfolio. After that, the candidates begin assembling their portfolio. Artifacts for the portfolio include videotapes of classroom interaction, student work samples and teacher reflective commentaries on evidence and documentation of educational/professional growth. Apart from the portfolio, teachers must demonstrate their knowledge of different pedagogical strategies within their subject area during a one-day written assessment.

A due date is established for each certificate and the portfolio is mailed to evaluators who are primarily teachers in the same field as the candidate. The assessors judge the components of the portfolio and the candidate's answers from the written assessment.

Applicants must receive a base score or higher from both the portfolio and the written assessment to achieve National Board Certification status.

The U.S. federal government contributes to more than 55% of the operating costs of the NBPTS, while non-governmental sources fund the remaining amount (National Board for Professional Teaching Standards, 2001). Interestingly enough, the NBPTS is mostly a public school phenomenon because private schools do not recognize it as an organization that grants national certification for exemplary teaching. As a result, less than one percent of the teachers who have achieved certification are private school teachers (Podgursky, 2001b).

## 2.3 Empirical Evidence for National Board Certification

Deleted:

Initially, empirical evidence about the effectiveness of National Board Certification was primarily generated by the National Board for Professional Teaching Standards (NBPTS). The organization completed two studies that compared National Board Certified teachers to their counterparts. The first study, *A Distinction That Matters: Why National Teacher Certification Makes a Difference* (National Board for Professional Teaching Standards, 2000), was conducted by a team of researchers from the University of North Carolina at Greensboro. It was the first comprehensive study to compare the teaching practices of National Board Certified (NBC) teachers with those of non-NBC teachers, as well as samples of student work from classrooms of the two groups of teachers. The sample consisted of 65 teachers from North Carolina, Ohio and the Washington D.C. metropolitan area. The study revealed that NBC

teachers scored higher on all 13 dimensions of teaching expertise than did teachers who sought, but did not achieve, National Board Certification.

The second study conducted by the NBPTS, *I Am a Better Teacher: What Candidates for National Board Certification Say about the Assessment Process* (National Board for Professional Teaching Standards, 2001), consisted of administering a 27-item survey to National Board Certification candidates. Five thousand, six hundred and forty-one (5,641) responses were received from candidates in 49 states, which represented a 53% response rate. The teachers in the study reported that the certification process made them better teachers, helped them to improve their teaching and their interactions with students and parents, helped them to create stronger curricula, allowed them to evaluate student learning, and helped them to develop a framework to use with state content standards for improving teaching.

The two studies reported by the NBPTS documented that National Board Certified teachers and/or National Board Certification candidates had stronger teaching practices and healthier teacher-student relationships than their counterparts (National Board for Professional Teaching Standards, 2000, 2001). However, student outcome data were never compiled. Controversy loomed over these two reports due to inadequate research methods, the lack of quantitative evidence in support of the National Board Certification process, and no proven method to establish that National Board Certified teachers were able to improve student achievement in an objectively-measurable manner. Thus, it was difficult to determine if the NBPTS was successfully accomplishing its stated mission: to advance the quality of teaching and learning by establishing standards for what accomplished teachers should know and be able to do.

Consequently, the NBPTS elicited research assistance to help to verify the worth of their national teaching certification. In this section, I discuss those studies conducted by independent researchers. Some provide positive support for the National Board Certification process (Section 2.3.1) and others provide conflicting evidence (Section 2.3.2).

### *2.3.1 Positive Support for National Board Certification*

One research study in favor of the National Board Certification process was a multi-year study conducted by the University of Washington and the Urban Institute. It was funded by the U.S. Department of Education. The researchers investigated the relationship between National Board Certification of teachers and elementary students' achievement. They measured the effectiveness of National Board Certified teachers by studying the annual test scores of North Carolina students in Grades 3–5 from three academic years: 1996–97, 1997–98 and 1998–99. Annual test scores from reading and mathematics were collected from more than 600,000 student records in North Carolina schools. Their findings indicated that National Board Certified (NBC) teachers appeared to be more effective than non-NBC teachers in terms of student achievement in reading and mathematics (Goldhaber & Anthony, 2004).

Another four-year study conducted by the University of Arizona (Vandevoort et al., 2004) compared students' academic performance on the Stanford Achievement Test-9$^{th}$ Edition (SAT-9) in order to investigate the effectiveness of National Board Certified (NBC) teachers. They compared students of NBC and non-NBC elementary teachers in Arizona's 14 school districts in Grades 3–6. In Grade 3, the sample consisted of 113 students in the NBC teachers' classes and 14,506 students from non-

NBC teachers' classes. In Grade 4, there were 184 students in the NBC teachers' classes and 15,487 students from the non-NBC teachers' classes. In Grade 5, there were 77 students in the NBC teachers' classes and 15,550 students from the non-NBC teachers' classes. In Grade 6, there were 79 students in the NBC teachers' classes and 11,752 students from the non-NBC teachers' classes. The results of the research showed that students in Grades 3–6 who were taught by NBC teachers in the 14 Arizona school districts outperformed their schoolmates on the nationwide SAT-9 in almost 75 % of reading, mathematics and language arts measures.

A third study was funded by the National Science Foundation and conducted by the CNA Corporation (Cavalluzzo, 2004). The sample for this study consisted of over 108,000 Grade 9 and 10 students in Miami-Dade County Public Schools in the state of Florida. Sixty-one (61) National Board Certified (NBC) teachers and 1,947 non-NBC teachers made their classes available for data collection. Cavalluzzo studied the associations between student gains in mathematics and certification status (NBC teachers versus non-NBC teachers). The data that were collected and statistically analyzed consisted of the students' end-of-grade examination scores for the school year. The results indicated strong evidence that National Board Certification is an effective indicator of teacher effectiveness in terms of students' mathematics achievement.

The results of these studies provide support for the NBPTS's success in accomplishing its stated mission: to advance the quality of teaching and learning by establishing standards for what accomplished teachers should know and be able to do. In addition, the results of these studies provide support for the policies in many states

that honor and provide extra remuneration for National Board Certified teachers. However, other studies have found conflicting evidence for the effectiveness of the National Board Certification process. These studies are discussed in the section that follows.

### *2.3.2 Conflicting Evidence of the Impact of National Board Certification*

In addition to research that supports the effectiveness of the National Board Certification process, there are research studies that provide conflicting results. For example, Stone (2002) used a value-added method of defining successful teaching. He claimed that the National Board for Professional Teaching Standards (NBPTS) had no studies to prove that National Board Certified teachers were able to improve student achievement using objectively-measurable means. Professor Stone collected *teacher-effect scores* (estimates of the impact a teacher has on a student's learning) for 16 National Board Certified (NBC) teachers and 40 non-NBC teachers from the state of Tennessee. He found that the NBC teachers were only average producers of student achievement gains as compared to the non-NBC teachers in their school districts. Stone concluded that most of the large body of research finds almost no correlation between certification and teacher effectiveness. Therefore, it made sense to expect no correlation between "super-certification" and student performance. It is risky, however, to generalize from this study because of the small sample size that was used and with the lack of student outcome data. While Stone further analyzed the value-added data with his colleague, Cunningham (Cunningham & Stone, 2005), it was found that the top 10% of non-NBC teachers in North Carolina produced student achievement gains 10 to 20 times larger than those produced by NBC teachers.

In addition to the conflicting evidence found by research studies, the educational community has criticized the research methods employed by researchers who showed positive results for the National Board Certification process. Podgursky, a professor of economics at the University of Missouri-Colombia, is highly critical of the research methods used in many of the National Board Certification studies, which were indirectly and/or directly funded by the NBPTS. Michael Podgursky (2001a) voiced his dissatisfaction with these issues in his article entitled *Defrocking the National Board*. Furthermore, Podgursky states that studies conducted to investigate the impact of National Board Certification on student outcomes will continually be flawed as long as the NBPTS pays for the study and all the researchers have had or continue to have professional ties to the Board (Wilcox, 2003). In addition, he argues that these studies offer only unconvincing evidence that National Board Certified teachers help students to learn more. He proposed that the research should link National Board Certification to direct measures of student achievement.

Furthermore, Finn and Wilcox (1999), who represented the Fordam Foundation (a conservative foundation that promotes a market-based approach to defining teacher quality), reported that, since its 12 years on inception, the NBPTS has been unable to prove that National Board Certified (NBC) teachers produced higher-achieving students than non-NBC teachers. One year later, Finn and Wilcox (2000) reiterated their stance on the ineptness of the NBC teachers' ability to produce higher-achieving students in their article written for the *Los Angeles Times*.

In conclusion, several studies have been conducted to investigate the effectiveness of the National Board Certification process. Some yielded results in favor of National

Board Certification and others contradicted those findings. The question still remains: Are National Board Certified (NBC) teachers more effective than non-NBC teachers? In my study, I sought the answer to that question. More specifically, I investigated if NBC teachers are more effective than non-NBC teachers in terms of secondary science students' learning environment perceptions, attitudes towards science, and achievement in science.

## 2.4 Historical Background of Learning Environments Research

Most science education research has focused on student academic achievement, while very little attention has been given to studying the learning environment as a determinant of learning outcomes. However, learning environments research has become well established and globally recognized over the last three decades as a means of assessing and investigating important aspects of education. Educational researchers have found connections between the learning environment and academic outcomes, and it is now well known that learning environments strongly influence student outcomes and play an important role in improving the effectiveness of learning (Fraser, 1998a, 2001). Therefore, it is important to discuss the historical beginnings of the study of learning environments.

Over 60 years ago, Lewin (1936) and Murray (1938) developed the first formal studies relevant to the field of learning environments and were most likely the first two individuals who recognized relationships between the environment and human behavior. Lewin's (1936) theory dealt with the relationship and interaction between the individual and his/her environment. He acknowledged that the environment is a

determinant of human behavior as stated in his formula $B = f(P, E)$. Behavior $B$ is a function of the person $P$ and the environment $E$. Murray (1938) was the first person to use Lewin's findings to propose a needs-press model, which refers to situational variables that are found in the environment and account for a degree of behavioral difference. As a result, Lewin and Murray are generally credited with providing a foundation for the later development of the field of learning environments research.

Murray's distinction between alpha press (the environment as observed by an external observer) and a beta press (the environment as perceived by milieu inhabitants) was extended by Stern, Stein, and Bloom (1956).

Stern et al. introduced the terms *private beta press* (the view that each person has of the environment) and *consensual beta press* (the view that members of a group hold about the environment). In general, learning environment researchers must decide whether their statistical analyses will involve the perception scores obtained by individual students (private press) or be combined to obtain the mean of the environment scores of all the students within the same class (consensual press) (Fraser, 1998b). Choosing the correct unit of analysis is critical. The choice of different units of analysis can lead to different interpretations if there are different levels of variability and differences (Fraser, 1994). Thus, relationships obtained using one unit of analysis could have different magnitudes and signs from relationships obtained using a different unit.

In my study, I used two units of analysis (the individual and class mean) when conducting most statistical analyses. For instance, I used two units of analysis

(individual and class mean) when estimating the internal consistency reliability of the revised version of the WIHIC (refer to Section 4.2.2 to view the results) and the modified attitude scale (refer to Section 4.2.5 to view the results), and discriminant validity of the revised version of the WIHIC (refer to Section 4.2.3 to view the results). Additionally when conducting simple correlation and multiple regression analyses, to investigate outcome environment association, I used two units of analysis (refer to Section 4.2.4 to view the results).

Deleted: ¶

The work of Lewin and Murray paved the way for the continuation of researchers to develop instruments to assess learning environments. Rudolf Moos (1974) and Herbert Walberg (1968) were two researchers who began independent studies on the conceptualization and assessment of psychosocial environments and developed the first two instruments to assess students' perceptions of their learning environment. Herbert Walberg designed the Learning Environment Inventory (LEI) as part of the research and evaluation activities of Harvard Project Physics (Walberg & Anderson, 1968). Another noteworthy accomplishment is Walberg's theory of educational productivity (Walberg, 1981, 1984, 1986) which states there are nine factors which contribute to students' variance in cognitive and affective outcomes: student ability, age and motivation; the quality and quantity of instruction; and the psychological environment of the home, the classroom social group, the peer group (pressure) outside of the home, as well as the amount of television viewing. Studies involving data collection from national samples have confirmed this model's validity in showing that student achievement and attitudes are influenced jointly by a number of factors rather than one prevailing factor (Walberg, 1986; Walberg, Fraser, & Welch, 1986).

Rudolf Moos (1974) was also a pioneer of the field of learning environments. He eventually developed the Classroom Environment Scale (CES) (Moos & Trickett, 1987) based on research involving perceptual measures of a variety of human environments including psychiatric hospitals, prisons, universities residences and work milieus (Moos, 1974). The final published version contains nine scales with 10 items of True-False response format in each scale.

As a result of the work of Walberg and Moos, classroom learning environments research has grown exponentially, including the continuing development of other learning environment instruments (Fraser, 1986, 1991, 1994, 1998a; Goh & Khine, 2002; Wubbels & Levy, 1993).

## 2.5 Classroom Learning Environment Instruments

The progression of the field of learning environments has seen the development of a plethora of instruments that can be used to assess classroom environments. Many of these classroom environment instruments have been modeled after Moos's (1974) initial work and, as a consequence, a common theme runs through those instruments. The common theme is Moos's three basic types of dimensions for classifying human environments: Relationship Dimensions (which identify the nature and intensity of personal relationships within the environment and assess the extent to which people are involved in the environment and support and help each other), Personal Development Dimensions (which assess basic directions along which personal growth and self-enhancement tend to occur) and System Change Dimensions (which involve

the extent to which the environment is orderly, clear expectations, maintain control and is responsive to change).

This section describes eight classroom learning environment instruments that follow Moos's three basic types of dimensions: Learning Environment Inventory, LEI (Section 2.5.1), Classroom Environment Scale, CES (Section 2.5.2), Individualized Classroom Environment Questionnaire, ICEQ (Section 2.5.3), My Class Inventory, MCI (Section 2.5.4), College and University Classroom Environment Inventory, CUCEI (Section 2.5.5), Questionnaire on Teacher Interaction, QTI (Section 2.5.6), Science Laboratory Environment Inventory, SLEI (Section 2.5.7), and Constructivist Learning Environment Survey, CLES (Section 2.5.8). A brief overview of these eight widely-applicable classroom learning environment instruments is given in Table 2.1. The table shows the name of each scale, the level for which each instrument is suited, the number of items contained in each scale, and the classification of each scale according to Moos's (1974) scheme for classifying human environments. The What Is Happening In this Class? (WIHIC) questionnaire is another widely-used classroom learning environments instrument; however, I chose not to discuss it in this section or include it in Table 2.1 because I would be discussing it in much more detail in a separate section (see Section 2.6 for detailed information about the development, characteristics, and validation of the WIHIC). I decided to dedicate a separate section to the WIHIC because it was the learning environment instrument I used in my study.

Deleted:

### *2.5.1 Learning Environment Inventory (LEI)*

The initial development of the LEI began in the late 1960s in conjunction with the evaluation and research related to Harvard Project Physics (Walberg & Anderson,

1968). The original version of the LEI evolved from Walberg's Social Climate Questionnaire (Walberg, 1968). The final version of the LEI contained 105 statements with seven descriptors of typical school classes. The individual expresses a degree of agreement or disagreement with each statement by choosing from four responses (Strongly Disagree, Disagree, Agree, and Strongly Agree). In addition, to reduce response bias on the part of the respondent, some items are negatively phrased. A typical question in the Cohesiveness scale is "All students know each other very well".

Deleted:

### *2.5.2 Classroom Environment Scale (CES)*

The CES, developed by Rudolf Moos at Stanford University, is based on social climate scales created as part of a comprehensive set of perceptual measures of a variety of human environments, including psychiatric hospitals, prisons, university residences and work environments (Fisher & Fraser 1983; Moos, 1979; Moos & Trickett, 1987). The original version of the CES consisted of 242 items representing 13 conceptual dimensions (Trickett & Moos, 1973). The final published version contains 90 items (9 scales with 10 items in each of the scales) with a True-False response format for each item. A typical question in the Teacher Support scale is "The teacher takes a personal interest in the students".

### *2.5.3 Individualized Classroom Environment Questionnaire (ICEQ)*

While the LEI and CES set the groundwork for the development of other classroom environment questionnaires, they did not include dimensions that distinguish individualized classrooms from conventional ones. Therefore, the ICEQ was developed to assess individualized classrooms as distinct from conventional ones. As

a result, Personalization and Participation dimensions were included as components of the ICEQ. The initial long version of the ICEQ (Rentoul & Fraser, 1979) was developed after interviewing teachers and secondary students. Afterwards, selected experts, teachers and junior high school students reviewed the questionnaire in draft form and modified it to a shorter version of 50 items (10 items in each of 5 scales) that are answered using a five-point frequency response format with the alternatives of Almost Never, Seldom, Sometimes, Often and Very Often, respectively. To avoid a biased response from the respondent, some items are negatively phrased. A typical question in the Personalization scale is "Different students use different books, equipment and materials".

### *2.5.4 My Class Inventory (MCI)*

The MCI is a simplified version of the LEI for use among children ages 8–12 years (Fisher & Fraser, 1981; Fraser, Anderson, & Walberg, 1982; Fraser & O'Brien, 1985; Majeed, Fraser & Aldridge, 2002; Mink & Fraser, 2005). The MCI was developed for the elementary level due to its simplicity in wording, but also it has found its niche at the junior high school level, especially with students who have limited reading skills in English. The MCI employs 38 items in five scales with a two-point (Yes-No) response format. However, Fraser and O'Brien (1985) developed an even shorter version with 25 items. A typical question in the Friction scale is "Children are always fighting with each other".

### *2.5.5 College and University Classroom Environment Inventory (CUCEI)*

The CUCEI was developed to gather information in higher education classrooms. It was not designed to assess lecture or laboratory settings, but rather to assess

perceptions in small class settings (Fraser & Treagust, 1986). The original version of the CUCEI used the more frequently-used features from the LEI, CES and ICEQ. The final version of the CUCEI has seven scales, each containing seven items. Each item has the four possible responses of Strongly Agree, Agree, Disagree, and Strongly Disagree (Fraser, 1998a). A typical question in the Task Orientation scale is "Activities in this class are clearly and carefully planned".

### *2.5.6 Questionnaire on Teacher Interaction (QTI)*

The QTI originated in the Netherlands and focuses on the nature and quality of interpersonal relationships between teachers and students (Wubbels & Levy, 1993). It assesses students' perceptions of eight behavioral aspects exhibited by teachers: Leadership, Helping/Friendly, Understanding, Student Responsibility and Freedom, Uncertain, Dissatisfied, Admonishing, and Strict. The QTI has 48 items with a five-point frequency response scale ranging from Never to Always. A typical item in the Student Responsibility and Freedom Behavior scale is "She/he gives us a lot of free time". In addition, the QTI has been cross-validated and found to be useful in various countries (Goh & Fraser, 1998; Lee & Fraser, 2003; Quek, Wong & Fraser, 2005a, 2005b; Scott & Fisher, 2004).

### *2.5.7 Science Laboratory Environment Inventory (SLEI)*

Laboratory settings are very important in science learning. Therefore, the SLEI was developed specifically for assessing the learning environment of science laboratory classes at the senior high and university levels (Fraser, Giddings, & McRobbie, 1995).

Table 2.1 Overview of Eight Classroom Environment Instruments (LEI, CES, ICEQ, MCI, CUCEI, QTI, SLEI, and CLES)

| Instrument | Level | Items per Scale | Scale Classified According to Moos's Scheme | | |
|---|---|---|---|---|---|
| | | | Relationship Dimension | Personal Development Dimension | System Maintenance and Change Dimension |
| Learning Environments Inventory (LEI) | Secondary | 7 | Cohesiveness<br>Friction<br>Favoritism<br>Cliqueness<br>Satisfaction<br>Apathy | Speed<br>Difficulty<br>Competitiveness | Diversity<br>Formality<br>Material Environment<br>Goal Direction<br>Disorganization |
| Classroom Environment Scale (CES) | Secondary | 10 | Involvement<br>Affiliation<br>Teacher Support | Task Orientation | Order and Organization<br>Rule Clarity<br>Teacher Control<br>Innovation |
| Individualized Classroom Environment Questionnaire (ICEQ) | Secondary | 10 | Personalization<br>Participation | Independence<br>Investigation | Differentiation |
| My Class Inventory (MCI) | Elementary | 6—9 | Cohesiveness<br>Friction<br>Satisfaction | Difficulty<br>Competitiveness | |
| College and University Classroom Environment Inventory (CUCEI) | Higher Education | 7 | Personalization<br>Involvement<br>Student Cohesiveness<br>Satisfaction | Task Orientation | Innovation<br>Individualization |
| Questionnaire on Teacher Interaction (QTI) | Secondary or Primary | 8—10 | Leadership<br>Helpful/Friendly<br>Understanding<br>Student Responsibility and Freedom<br>Uncertain<br>Dissatisfied<br>Admonishing<br>Strict | | |
| Science Laboratory Environment Inventory (SLEI) | Upper Secondary or Higher Education | 7 | Student Cohesiveness | Open-Endedness<br>Integration | Rule Clarity<br>Material Environment |
| Constructivist Learning Environment Survey (CLES) | Secondary | 7 | Personal Relevance<br>Uncertainty | Critical Voice<br>Shared Control<br>Investigation | Student Negotiation |

Fraser (1998a)

The SLEI has five scales with seven items in each. Each item has the five possible frequency responses of Almost Never, Seldom, Sometimes, Often and Very Often. Typical items are "I use the theory from regular science class sessions during laboratory activities" and "We know the results that we are supposed to get before we commence a laboratory activity". The SLEI has been found to be a valid and useful questionnaire in numerous studies (Henderson, Fisher & Fraser, 2000; Quek, Wong & Fraser, 2005a; Wong & Fraser, 1996).

### *2.5.8 Constructivist Learning Environment Survey (CLES)*

The CLES focuses on student-centered settings and was developed to assist researchers and teachers to assess the degree to which a particular classroom's environment is consistent with a constructivist epistemology (Taylor & Fraser, 1991). Additionally, the CLES assists teachers to alter their classroom learning environments in compliance with critical constructivist epistemology (Taylor, Dawson, & Fraser, 1995). The CLES is available in the actual and preferred form (Kim, Fisher, & Fraser, 1999). It has 30 items (5 scales with 6 items in each scale). Each item has a five-point response scale to choose from (Almost Never, Seldom, Sometimes, Often and Almost Always). A typical item in the Personal Relevance scale is "I learn how science can be part of my out-of-school life". The CLES has been cross-validated in the USA (Nix, Fraser & Ledbetter, 2005), Korea (Kim, Fisher & Fraser, 1999), South Africa (Aldridge, Fraser & Sebela, 2004) and Australia and Taiwan (Aldridge, Fraser, Taylor & Chen, 2000).

### *2.5.9 Preferred Forms of Classroom Environment Instruments*

As stated above, the CLES is available in the actual and preferred form. But, it should be noted that most of the learning environment instruments in Table 2.1 have the ability to measure perceptions of experienced classroom environment or 'actual' and the ability to measure perceptions of ideal classroom environment or 'preferred'. The preferred forms are concerned with goals and value orientation and measure perceptions of the classroom environment ideally liked or preferred. Although item wording is similar for actual and preferred forms, the statements are slightly different. For example, an item in the actual form states, 'There *is* a clear set of rules for students to follow', while in the preferred form it would state, 'There would be a clear set of rules for students to follow' (Fraser, 1998).

## 2.6 What Is Happening In this Class? (WIHIC) Questionnaire

In Section 2.5, I briefly discussed eight classroom learning environment questionnaires, which have facilitated the study of classroom learning environments in many parts of the world. However, I did not discuss the What Is Happening In this Class? (WIHIC), which is another widely-used and extensively-validated classroom learning environments instrument. The WIHIC questionnaire was the instrument I used in my study to measure students' perceptions of their science classroom environment.

In Section 2.6.1, I provide information about the development and characteristics of the WIHIC. In Section 2.6.2, validation studies using the WIHIC in the Western context are discussed. In Section 2.6.3, I discuss the validation of the original, translated, and/or modified versions of the WIHIC in the Asian context. Finally, in Section 2.6.4, I provide information about cross-validation studies using the WIHIC.

Table 2.2 Description and Sample Item for Each WIHIC Scale

| Scale | Description | Sample Item |
|---|---|---|
| Student Cohesiveness | Extent to which students know, help and are friendly toward each other. | I know other students in this class. |
| Teacher Support | Extent to which the teacher is interested in the students, while displaying characteristics of helpfulness, trustfulness, friendliness, etc. | The teacher takes a personal interest in me. |
| Involvement | Extent to which students are involved and participate in science class. | I explain my ideas to other students. |
| Investigation | Extent to which there is an emphasis on inquiry learning and problem solving. | I carry out investigations to test my ideas. |
| Task Orientation | Extent to which students complete activities and stay on subject matter. | I pay attention in this class. |
| Cooperation | Extent to which students cooperate rather than compete with one another on learning tasks. | I work with other students in this class. |
| Equity | Extent to which students are treated equally. | I am treated the same as other students in this class. |

Adapted from Koul and Fisher (2005)

### *2.6.1 Development and Characteristics of the WIHIC*

The WIHIC questionnaire, originally developed by Fraser, Fisher, and McRobbie (1996), brings parsimony to the field of learning environments by combining modified versions of the most leading environmental scales with additional scales that address contemporary concerns. In addition, the WIHIC has a class form and an individual form. Therefore, the instrument can be used to assess a student's perceptions of the class as a whole or assess a student's individual perceptions of his or her role in a classroom (Fraser, 1999). The original 90-item (10 statements in each of 9 scales) version of the WIHIC was modified to 54 items in seven scales after conducting statistical analyses of data collected from 355 junior high school science

students and extensive interviewing of students and teachers. The WIHIC was later expanded to 80 items and field-tested with 1,879 students in 50 classes in Australia (Fraser, 1998b; Huang, Aldridge, & Fraser, 1998). Based on the results of the field-testing, the WIHIC was modified to a 56-item version (8 items in each of 7 scales). In order to respond to the items in each of the seven scales, the WIHIC asks the respondent to indicate how often a practice occurs using the five-point scale of Almost Never, Seldom, Sometimes, Often, and Almost Always. The 56-item version of the WIHIC is the one that I chose to measure students' perceptions of their classroom learning environment in my study. Table 2.2 provides a description of the final 56-item seven-scale version of the WIHIC, along with a sample item from each scale.

The main reason why I chose the 56-item version of the WIHIC for my study is because it has been proven to be a valid and reliable tool in numerous countries around the world. In Section 2.6.2, I describe the validation of the WIHIC in Western countries. In Section 2.6.3, I provide more information about the validation of the WIHIC in Asian countries.

### *2.6.2 Validation of the WIHIC in Western Context*

The What Is Happening In this Class? (WIHIC) questionnaire has been extensively validated with students at various grade levels and in different subject areas within the Western context. For instance, Raaflaub and Fraser (2003) validated a modified version of the WIHIC with 1,173 mathematics and science high school students in Canada. After conducting principal components factor analysis with varimax rotation, the *a priori* eight-factor structure of the modified version of the WIHIC was

replicated with nearly all items having a factor loading of at least 0.40 on its *a priori* scale and no other scales. The internal consistency for each scale was established using Cronbach's alpha coefficient for two units of analysis. The alpha coefficient ranged from 0.76 to 0.92 (individual student) and from 0.78 to 0.95 (class mean) for each WIHIC scale, demonstrating satisfactory internal consistency. Additionally, the discriminant validity (mean correlation of one scale with the other scales) ranged from 0.10 to 0.38 using the individual student as the unit of analysis and from 0.18 to 0.45 for the class mean. These results indicated that each scale of the modified version of the WIHIC measures distinct aspects of the classroom learning environment, although some might overlap. The $eta^2$ statistic (the strength of association between class membership and the dependent variable) ranged from 0.01 to 0.11 for the WIHIC scales, thus indicating that each scale is capable of differentiating significantly ($p<0.05$) between the students' perceptions in the different classes.

Another study using the WIHIC questionnaire was validated using a sample of 3,980 high school mathematics students from Australia (1433 students) the United Kingdom (1596 students) and Canada (951 students) (Dorman, 2003). All scales had good internal consistency reliability using Cronbach's alpha coefficient. The $eta^2$ statistic showed that each WIHIC scale differentiated significantly ($p<0.01$) between the different classes in both countries. Confirmatory factor analysis supported the seven-scale *a priori* structure of the instrument, and all items loaded strongly on their *a priori* scale.

Zandvliet and Fraser (2004, 2005) provided another validation of the WIHIC using a sample of 1,404 students in 81 senior high school classes in Australia and Canada.

Principal components factor analysis followed by varimax rotation strongly supported the *a priori* seven-scale structure of the WIHIC. The alpha reliability coefficients suggested very good internal consistency. The discriminant validity (mean correlation of one scale with the other scales) demonstrates that these scales measure distinct, though somewhat overlapping, aspects of the psychosocial environment.

In North Carolina, United States, one small-scale study validated a modified version of the WIHIC with 364 biology students in Grades 9 and 10 (Moss & Fraser, 2001). Principal components factor analysis followed by varimax rotation resulted in the acceptance of the *a priori* seven-factor structure of the WIHIC with nearly all items loading on their *a priori* scale and no other scale. The alpha reliability estimates suggest that all scales of the WIHIC possess satisfactory internal consistency. Additionally, the discriminant validity for each WIHIC scale is low enough to support independence of the WIHIC scales. The use of ANOVA indicated that each scale is capable of differentiating significantly between students' perceptions in the different classes.

Original, modified, and/or translated versions of the WIHIC have been cross-validated in numerous studies that have been conducted in Miami-Dade County, Florida, including:

- MacDowell-Goggin's (2005) study with 860 Grade 4–6 students
- Allen's (2003) study with 520 Grade 4 and 5 students
- Pickett and Fraser's (2004) modified version with 573 Grade 3–5 students
- Castillo, Peiro and Fraser's (2004) study of 600 Grade 9–10 students

- Soto-Rodriguez and Fraser's (2004) use of English and Spanish versions with 1105 Limited English Proficient (LEP) and non-LEP students in Grades 2–5
- Robinson's (2003) use of English and Spanish versions with 172 kindergarten students and 72 parents
- Adamski, Peiro and Fraser's (2005) use of a modified Spanish version with 223 students in Grades 4–6.

The results from the studies replicate past research that has reported that the scales of the original, modified, and/or translated versions of the WIHIC exhibit sound factorial validity and internal consistency reliability and are able to differentiate significantly between the perceptions of the students in different classes.

In conclusion, the literature clearly shows that the WIHIC is a valid and reliable tool to assess students' perceptions of their classroom learning environment in the Western part of the world, especially in Miami-Dade County, Florida. For my study, it was important to select a learning environments instrument that was appropriate for secondary school students in Miami-Dade County, Florida because this is where my study took place. The fact that the WIHIC had been successfully validated in a variety of studies conducted in Miami-Dade County, Florida made this learning environments instrument the appropriate one to use in my study.

### *2.6.3 Validation of the WIHIC in Asian Context*

Although the WIHIC is a relatively recent instrument for assessing classroom environments, it has been translated into several Asian languages and modified for use in the Asian context. These translated and/or modified versions of the WIHIC

have been validated in studies conducted in countries such as Korea, Indonesia, Brunei and Singapore, just to name a few.

For example, in Korea, Kim, Fisher, and Fraser (2000) translated the WIHIC into Korean and back-translated into English to ensure that the English and Korean versions were the same. The sample consisted of 543 Grade 8 science students in 12 secondary schools who were administered the Korean version of the WIHIC. Principal components factor analysis followed by varimax rotation resulted in the acceptance of the *a priori* seven-factor structure of the Korean version of the WIHIC with nearly all items loading on their *a priori* scale and no other scale. The alpha reliability coefficient used as an index of scale internal consistency, ranged from 0.82 to 0.92 for each scale of the WIHIC suggesting that all scales possess satisfactory internal consistency. Additionally, the discriminant validity (mean correlation of one scale with the other scales) ranged from 0.32 to 0.49. These numbers suggest that each WIHIC scale measures distinct aspects of the classroom environment. The $eta^2$ statistic (the strength of association between class membership and the dependent variable) ranged from 0.06 to 0.20 for each WIHIC scale, and each scale was capable of differentiating significantly ($p<0.01$) between classes.

In Indonesia, a modified version of the WIHIC was validated by Margianti, Aldridge, and Fraser (2004). The study involved 2,498 university students in 50 computing classes. Results of principal components factor analysis showed that nearly all items of the Indonesian version of the WIHIC had factor loadings of at least 0.40 on their *a priori* scale and no other scale. Thus, the *a priori* seven-factor structure of the final version of the WIHIC was replicated. The alpha reliability coefficient ranged from

0.65 to 0.87 for each scale of the modified Indonesian version of the WIHIC, thus suggesting that all scales possess satisfactory internal consistency. Additionally, the discriminant validity (mean correlation of one scale with the other scales) suggested that each WIHIC scale measures distinct aspects of the classroom environment. The $eta^2$ statistic (the strength of association between class membership and the dependent variable) was calculated and each scale of the Indonesian version of the WIHIC was found to be capable of differentiating significantly ($p<0.01$) between classes.

A study conducted by Khine and Fisher (2001) in Brunei looked at the validity and reliability of the WIHIC. This study used a very large sample of 1,188 students from 54 science classes in 10 government secondary schools. Both the individual and the class mean were used as the unit of analysis to determine the internal consistency reliability and discriminant validity. ANOVA results suggested that each scale is capable of differentiating significantly ($p<0.01$) between the perceptions of students in the different classes.

An additional large-scale study conducted by Fraser and Chionh (2000) in Singapore looked at the validity and reliability of a modified English version of the WIHIC. This study involved 2,310 students in 75 Grade 10 geography and mathematics classes. The results suggest that all scales of the modified English version of the WIHIC exhibits adequate factorial validity, internal consistency and the ability to differentiate significantly between the perceptions of the students in the different classes.

Aldridge and Fraser (2000) examined classrooms in Australia and Taiwan using English and Mandarin versions of the WIHIC. The Mandarin version underwent a

back-translation process to achieve linguistic equivalence with the English version. The English and Mandarin versions of the WIHIC were administered to a sample of 1,081 junior high students in 50 classes in Australia and 1,879 junior high students in 50 classes in Taiwan. Principal components factor analysis followed by varimax rotation resulted in the acceptance of the *a priori* seven-factor structure of the WIHIC in both countries with nearly all items having a factor loading of at least 0.40 on their *a priori* scale and no other scale. The relatively high alpha reliability values for each scale of the English and Mandarin versions of the WIHIC suggest that the items in a scale assess a common concept. ANOVA statistic results indicated that each of the seven scales differentiated significantly ($p<0.01$) between the perceptions of the students in the different classes in both countries.

In addition to the studies I've just discussed, the literature clearly shows that the validity and reliability for original, modified, and/or translated versions of the WIHIC have been established in other Asian countries, such as in Indonesia with 422 students enrolled in 12 university level classes (Soerjaningsih, Fraser, & Aldridge, 2001), in Singapore with 250 adult learners in 23 computer classes (Khoo & Fraser, 1998), in Brunei with 644 Grade 10 Chemistry students (Riah & Fraser, 1998), and in Korea with 543 Grade 8 students in 12 schools (Kim et al., 2000).

In conclusion, the WIHIC is a learning environments questionnaire that could be used in the Asian context with positive results. The numerous studies conducted in Asian countries suggest that the WIHIC could be modified and translated into several Asian languages and still remain a valid and reliable instrument for gathering students' perceptions of their classroom learning environment.

## 2.7 Past Lines of Learning Environments Research

The numerous learning environment instruments available to teachers and researchers have facilitated research in the field of learning environments. Consequently, several lines of learning environments research have emerged in the past 30 years or so. The five basic lines of learning environments research are: associations between student outcomes and environment, determinants of classroom environment (including evaluations of educational innovations), differences between students' and teachers' perceptions of the same classrooms, use of qualitative research methods, and cross-national studies (Goh & Khine, 2002).

My study involved two past lines of learning environments research. First, it could be categorized under the line of research known as "determinants of classroom environment" (including evaluations of educational innovations) because I looked at the differences in students' classroom environment perceptions depending on the teacher's certification status (National Board Certified teacher versus non-National Board Certified teacher). This line of learning environments research looks at how the classroom environment varies depending on different factors (including using environment measures in evaluating educational evaluations). Second, my study could be categorized under the line of research known as "associations between student outcomes and environment" because I looked at associations between outcomes (achievement and attitudes) and the science learning environment. Section 2.7.1 will provide further information about studies that have looked at "determinants of classroom environment", and Section 2.7.2 will discuss studies that have investigated "associations between outcomes and the learning environment".

### *2.7.1 Past Research on Determinants of Classroom Environment*

Determinants of classroom environment dimensions have been used as dependent variables in research aimed at identifying how the classroom environment varies with different factors such as teacher personality, subject matter and grade level (Fraser, 1994). In my study, I looked at how the classroom environment varied depending on teacher certification status (National Board Certified teacher versus non-National Board Certified teacher).

Hirata and Sako (1998) found differences between the classroom environment perceptions of at-risk students (delinquents and non-attendees) and normal students in Japan with a population of 635 pupils from four junior high schools using a Japanese version of the Classroom Environment Scale (CES).

Quek, Wong and Fraser (2005a, 2005b) investigated the impact of the chemistry laboratory environment and teacher-student interaction on student attitudes towards chemistry for 200 gifted secondary school students in Singapore. The results reported interesting differences in the perceived learning environments of gifted and non-gifted students using the 35-item version of the Chemistry Laboratory Environment Inventory (CLEI), the 48-item version of the Questionnaire on Teacher Interaction (QTI) and the 30-item version of the Questionnaire on Chemistry-Related Attitudes (QOCRA).

In Brunei, Khine and Fisher (2001, 2002) reported cultural differences in students' classroom environment perceptions depending on whether the teacher was Asian or Western with a sample of 1,188 students from 54 science classes in ten secondary

schools. The students completed the What Is Happening In this Class? (WIHIC) questionnaire as a measure of learning environment perceptions and responded to two scales of the Test Of Science-Related Attitudes (TOSRA) that were used as attitudinal measures. The results showed that students perceived a more favorable learning environment in the classrooms of the Western teachers. For instance, students perceived that the science classrooms of Western teachers were more cohesive, they received more teacher support and they were more involved in the work of the class. Students also perceived that, in the science classes of Western teachers, there was more task orientation, cooperation among students and equity. The study also found that students in the classrooms of Western teachers enjoyed their science lessons more than those students in the other classes.

One of the most important determinants of the classroom learning environment is the instructional method. Consequently, numerous past studies have used the learning environment as one of the criteria in evaluating innovative teaching and learning methods. My study fits into this past line of learning environments research because I evaluated the effectiveness of National Board Certified (NBC) and non-NBC teachers in terms of the classroom learning environments that they created (Maor & Fraser, 1996; Mink & Fraser, 2005; Spinner & Fraser, 2005; Teh & Fraser, 1994).

### *2.7.2 Past Research on Outcomes-Environment Associations*

The most common past line of classroom learning environments research is the investigation of associations between students' outcomes (cognitive and affective) and their perceptions of their classroom environment (Fraser & Fisher, 1982; Haertel, Walberg, & Haertel, 1981; McRobbie & Fraser, 1993). This section reviews past

research on outcome-environment association because one of the aims of my study was to investigate associations between student outcomes (attitudes and achievement) and classroom environment.

Deleted: ,

For example, in Singapore, Khoo and Fraser (1998) established relationships between student attitudes and classroom environment using the What Is Happening In this Class? (WIHIC) for a sample of 250 adults attending 23 computing classes. More recently, Fraser and Chionh's (2000) comprehensive study in Singapore established association between WIHIC scales, and three student outcomes (attitudes, self-esteem and examination results) among 2310 mathematics and geography students in 75 classes. Furthermore, Henderson, Fisher, and Fraser (1998) reported that environmental science students' perceptions of cohesion, involvement, and task orientation were strongly associated with positive attitudinal outcomes when the WIHIC was used in Australia.

Using the Science Laboratory Inventory (SLEI), associations between students' cognitive and affective outcomes and the classroom environment have been established for a sample of 80 senior high school chemistry classes in Australia (Fraser & McRobbie, 1995; McRobbie & Fraser 1993), 489 senior high school biology students in Australia (Fisher, Henderson, & Fraser, 1995) and 1,592 grade 10 chemistry students in Singapore (Quek, Wong & Fraser, 2005a).

Using the Questionnaire on Teacher Interaction (QTI), associations between student outcomes and perceived patterns of teacher-student interaction were reported for samples of 489 senior high school biology students in Australia (Fisher et al., 1995),

3,994 high school science and mathematics students in Australia (Fisher, Fraser, & Rickards, 1997) and 39 primary school mathematics classes in Singapore (Goh & Fraser, 1998). Quek, Wong & Fraser (2005b) investigated associations between teacher-student interaction and students' attitudes towards chemistry among 497 tenth grade students from three independent schools in Singapore with a 48-item version of the Questionnaire on Teacher Interaction (QTI), while an elementary version of the QTI was translated into Standard Malay and carried out in 136 Brunei classrooms with 3,104 students (Scott & Fisher, 2004).

Some aspects of the learning environment in environmental science classrooms have been found to be associated with students' attitudinal outcomes and suggest that favorable student attitudes could be promoted in classes where the students perceive more cohesion amongst students, a greater degree of student involvement in classroom activities, and a higher level of task orientation (Henderson et al., 1998). Other studies that have established outcomes-environment associations include research in Indonesia by Soerjaningsih et al. (2001), in Brunei, by Riah and Fraser (1998), and in Korea by Kim et al. (2000).

## 2.8 Assessment of Students' Attitudes to Science

The inclusion of the measurement of students' attitudes to science was important in my study because I attempted to evaluate the effectiveness of NBC in terms of students' attitudes and to investigate the link between student satisfaction and aspects of the classroom learning environment. Open-ended questions, interviews, preference rankings, closed-item questionnaires (such as Likert scales), and projective techniques

are used to determine students' attitudes towards an academic subject (Laforgia, 1988; Schibeci, 1984). Closed-item questionnaires with Likert response scales (i.e. Strongly Disagree, Disagree, Not Sure, Agree, and Strongly Agree) are the ones most frequently used to determine students' attitudes toward science. One closed-item attitude questionnaire that is frequently used in science education is the Test Of Science-Related Attitudes (TOSRA) ( Fraser, 1981).

The original TOSRA has 70 statements (10 items in each of 7 scales), which are directly aligned with Klopfer's (1971) classification scheme for six affective aims for science education: attitude to science and scientists; attitude to inquiry; adoption of science attitudes like curiosity and open-mindedness; enjoyment of science learning experiences; interest in science apart from learning experiences and interest in a career in science. Table 2.3 shows a sample item and Klopfer's classification for each scale of the TOSRA.

The response format of the TOSRA is composed of Likert-type scale with students indicating their degree of agreement with each statement using Strongly Agree (SA), Agree (A), Not Sure (N), Disagree (D), and Strongly Disagree (SD). The TOSRA has both positively-phrased (i.e. "Science lessons are fun") and negatively-phrased (i.e. "Science lessons are boring") statements.

Several studies have cross-validated the TOSRA in Australia (Fraser & Butts, 1982; Lucas & Tulip, 1980; Schibeci & McGaw, 1981), the United States (Lightburn & Fraser, 2002; Pickett & Fraser, 2002), as well as Wong and Fraser (1996) in

Singapore and Adolphe, Fraser and Aldridge (2003) in Indonesia. These studies reported good validity and reliability for the TOSRA.

Table 2.3 Sample Item and Klopfer's Classification of Aims for each Scale of the TOSRA

| Scale | Sample Statement | Klopfer's (1971) Classification |
|---|---|---|
| Social Implications of Science | Money spent on science is well worth spending. | Manifestation of favorable attitudes towards science |
| Normality of Scientists | Scientists are about as fit and healthy as other people. | Manifestation of favorable attitudes towards scientists |
| Attitude to Scientific Inquiry | I would prefer to find out why something happens by doing an experiment than by being told. | Acceptance of scientific inquiry as a way of thought |
| Adoption of Scientific Attitudes | I enjoy reading about things which disagree with my previous ideas. | Adoption of 'scientific attitudes' |
| Enjoyment of Science Lessons | Science lessons are fun. | Enjoyment of science learning experiences |
| Leisure Interest in Science | I would like to belong to a science club. | Development of interest in science and science-related activities |
| Career Interest in Science | When I leave school, I would like to work with people who make discoveries in science. | Development of interest in pursuing a career in science |

Adapted from Fraser (1981)

For my study, I selected the TOSRA to assess attitudes. But, because of the overlap between some of the TOSRA scales, I followed the lead of previous researchers and

decided to use only some of the scales rather than all of them (Aldridge, Fraser, Taylor, & Chen, 2000; Lee, 2001; Quek, Wong & Fraser, 2005a).

Consequently, I created a 10-item attitude scale modeled on the TOSRA. I also modified the five-point response scale found in the original TOSRA to meet the needs of the students who were going to participate in my study and to minimize the loss of instructional time due to participating in the study. Thus, a three-point response scale was used with response alternatives of Disagree, Not Sure and Agree.

Initially, I conducted a pilot study to check if the students were able to understand the statements on the 10-item attitude scale. However, I had to make further modifications to it because some of the items proved troublesome for the students (refer to Section 3.4.2 for further information about how the attitude scale was modeled on the TOSRA and modified).

## 2.9 Summary of Literature Review

Because my research focused on the the effectiveness of National Board Certified (NBC) science teachers in terms of their secondary science students' perceptions of the classroom environment, attitudes, and achievement, the first section of this chapter discussed the history of the National Board for Professional Teaching Standards (Section 2.2) and the empirical evidence about the effectiveness of National Board Certification (Section 2.3).

In addition, this chapter reviewed literature about the field of learning environments (see Section 2.4) because my study provided the first to look at the effectiveness of National Board Certified teachers in terms of secondary students' perceptions of their classroom learning environment. Various instruments used to investigate and assess aspects of the learning environment were discussed (Section 2.5). In particular, the development, characteristics, and validation of the What Is Happening In this Class? (WIHIC) was covered in detail because it was the learning environment questionnaire used in my study (Section 2.6). Also, past lines of learning environments research were discussed, with especial emphasis on the two most relevant to my study, namely, determinants of classroom environment and outcome-environment associations (Section 2.7). In addition, literature on the study of students' attitudes to science, particularly the characteristics and validity of the Test Of Science-Related Attitudes (TOSRA), was reviewed (Section 2.8).

Chapter 3 provides a description of the methodology used to address my research questions in terms of the context where the study took place (Section 3.2), background and selection of the sample for data collection (Section 3.3), instruments used (Section 3.4), research design (Section 3.5), and procedures followed and data analysis methods used (Section 3.6).

# Chapter 3

# METHODOLOGY

## 3.1 Introduction and Overview

There is increasing concern that there needs to be a common core of knowledge for professional preparation to ensure that preservice and inservice teachers have the necessary skills. With schools in the U.S. striving to meet achievement standards set by the No Child Left Behind Act (2001), teacher scrutiny is higher than ever (Keller, 2005). The National Board for Professional Teaching Standards (NBPTS) was founded to improve the quality of teaching through the implementation of a national certification process. Such a national certification process requires teachers to demonstrate the effectiveness of their teaching via hand-written analyses of their teaching, videotaped lessons, and a rigorous assessment. The NBPTS is rooted in the belief that the single most important action that can be taken to improve schools and student learning is to strengthen teaching.

Miami-Dade County Public Schools (M-DCPS), where I conducted my study, encourages its teachers to become National Board Certified. Compared to many other counties, M-DCPS has an abundance of teachers who are currently National Board Certified. Thus, I thought it would be important to investigate the effectiveness of the National Board Certified (NBC) teachers in M-DCPS. The three criteria that I used in evaluating whether NBC teachers are more effective than non-NBC teachers were learning environment, student attitudes and student achievement.

Chapter 1 provided vital information about the background and theoretical framework of my study. Chapter 2 reviewed numerous past learning environment studies that focused on science classrooms. My study is significant because it is the first documented research that used learning environment instruments to assess and investigate students' perceptions of their science classroom environment when taught by NBC and non-NBC teachers. This is the first time that research has been conducted with a large American sample of National Board Certified (NBC) and non-NBC teachers while incorporating a contemporary educational tool such as the What Is Happening In this Class? (WIHIC). My study is also unique because it is the first time that the effectiveness of NBC teachers has been investigated in terms of students' attitudes toward science using an attitude scale modeled on the Test Of Science-Related Attitudes (TOSRA).

This study is consistent with past lines of learning environments research in that it also explored connections between students' perceptions of their science classroom environment and student outcomes (attitudes toward science and science achievement). Additionally, in my study, I validated the What Is Happening In this Class? (WIHIC) questionnaire and an attitude scale modeled on the Test Of Science-Related Attitudes (TOSRA) with students taught by National Board Certified (NBC) and non-NBC science teachers. The present study answered the following research questions:

*Research Question #1*

*Is a revised version of the What Is Happening In this Class? (WIHIC) questionnaire valid and reliable when used in secondary science classrooms in South Florida?*

*Research Question #2*

*Is an attitude scale modeled on the Test Of Science-Related Attitude (TOSRA) reliable when used in secondary science classrooms in South Florida?*

*Research Question #3*

*Are National Board Certified (NBC) teachers more effective than non-NBC teachers in terms of classroom environment, student attitudes and student achievement in secondary science classrooms in South Florida?*

*Research Question #4*

*Are there associations between student outcomes (attitudes and achievement) and classroom environment in secondary science classrooms in South Florida?*

This chapter provides a description of the methodology used to address the research questions in terms of the context where the study took place (Section 3.2), background and selection of the sample for data collection (Section 3.3), instruments used (Section 3.4), research design (Section 3.5), and procedures followed and data analyses methods used (Section 3.6).

## 3.2 Context of the Study

My study was conducted in the Miami-Dade County Public Schools (M-DCPS) system, which provides free public education to the children of Miami-Dade County, Florida. Miami-Dade County is one of the most transient and multicultural areas of the state of Florida. M-DCPS is considered one of the largest and most multicultural school systems in the state of Florida (refer to Section 1.2 for detailed information about the location of Miami-Dade County, Florida and the context of the Miami-Dade County Public Schools system). Currently, there are 54 middle schools (schools for students in Grades 6–8) and 40 senior high schools (schools for students in Grades 9–12) in M-DCPS.

The state of Florida mandates that each year public school students take the Florida Comprehensive Assessment Test (FCAT). The FCAT is an examination, which measures students' reading, writing, mathematics, and science achievement. The components of the FCAT assess a wide range of material. For example, the Reading test employs a wide variety of writing material to assess students' reading comprehension. Writing emphasizes the areas of focus, organization, support, and conventions. Mathematics assesses Number Sense, Concepts, and Operations, Measurement, Geometry and Spatial Sense, Algebraic Thinking, Data Analysis and Probability. While the Science test concentrates on Physical and Chemical Sciences, Earth and Space Sciences, Life and Environmental Sciences and Scientific Thinking (Florida Department of Education, 2004). M-DCPS's students in Grades 3–10 take

the reading and mathematics portions of the FCAT, students in Grades 4, 8, and 10 take the writing portion of the FCAT, and students in Grades 5, 8, and 10 take the science portion of the FCAT. The scores that students receive on the reading, writing, and mathematics portions of the FCAT are used at the state level to measure the success of each Florida school. The scores calculated are used to give each school a letter grade of an A, B, C, D, or F. A is the highest possible grade that represents excellence and F represents failure. The schools that receive an A or that show adequate yearly progress are given monetary compensation. Schools that do not make adequate progress must provide supplemental services for the students and make corrections. At the present time, the scores on the science portion of the FCAT are not used in the calculations of the school letter grade, but it is projected that, starting in the year 2006–2007, the science FCAT scores will become part of the school letter grade.

The present study took place in 12 secondary schools within the Miami-Dade County Public Schools (M-DCPS) system. More information about how the 12 schools were selected for the study is provided later in Section 3.3. Table 3.1 offers a profile for each school that was involved and for the entire M-DCPS district. The table outlines the following information for each participating school and the district as a whole: ethnic distribution, percentage of students on free or reduced-cost lunch (an index of coming from a low-income family), school achievement grade and the average of the science FCAT scores (Miami-Dade County Public Schools, 2004).

Table 3.1 Ethnic Distribution, School Achievement Grades, Percentage of Students from Low-Income Families, and FCAT Average for each Participating School and in the Entire Miami-Dade County Public Schools District

| Participating School | Ethnicities | | | | School Achievement Grade [a] | % of Students on Free or Reduced-Cost Lunch | Average Science FCAT Score[c] |
|---|---|---|---|---|---|---|---|
| | % White | % Black | % Hispanic | % Other | | | |
| A | 21 | 25 | 48 | 6 | A | 80 | 338 |
| B | 6 | 56 | 36 | 2 | F | 95 | 240 |
| C | 25 | 26 | 42 | 7 | A | 77 | 306 |
| D | 21 | 29 | 46 | 4 | C | 82 | 275 |
| E | 9 | 3 | 86 | 2 | B | 91 | 288 |
| F | 46 | 15 | 32 | 7 | A | 57 | 319 |
| G | 6 | 13 | 80 | 1 | C | 94 | 264 |
| H | 2 | 88 | 9 | 1 | C | 98 | 260 |
| I | 10 | 18 | 71 | 1 | C | 90 | 275 |
| J | 16 | 26 | 56 | 2 | C | 87 | 262 |
| K | 13 | 36 | 51 | 0 | D | 90 | 269 |
| L | 36 | 23 | 33 | 8 | A | 64 | 310 |
| District | 10.5 | 29 | 58.5 | 2 | X[b] | 63 | 273.5 |

[a] The achievement school grade is based on the percentage of students in the school who received a passing score and/or demonstrated adequate progress from one school year to the next on the reading, mathematics, and science portions of the Florida Comprehensive Assessment Test (FCAT). The achievement school grades could be A, B, C, D, or F where A is the highest grade possible and F is the lowest. The A represents excellence and F represents Failure.

[b] The district as a whole is not given an achievement grade; therefore, it was omitted.

[c] The possible science FCAT scale scores range from 100 to 500.

## 3.3 Background and Selection of Sample

The sample for the present study was selected by using the National Board for Professional Teaching Standards (NBPTS) website in order to locate Grade 8 (Early Adolescence in Science) and Grade 10 (Adolescence and Young Adulthood in Science) National Board Certified teachers in Miami-Dade County Public Schools (M-DCPS). The next process was to identify National Board Certified (NBC) science

teachers from among the 94 secondary schools available in M-DCPS. It is important to note that a researcher must get approval from the school district's Research Department prior to starting the research. When the district's Research Department approves the research, a profile number is issued to the researcher. This profile number must be given to the principal at the school site in order to allow any research to be conducted at his/her school. It is also fair to say that the profile number does not automatically allow the researcher to conduct research at any school. It is ultimately up to the school site principal if any research will be conducted.

For the current study, the technique used to obtain the principals' permission to carry out research at his/her school was by a letter that was either sent through the school district's internal-based email system or U.S. mail. If necessary, contact was eventually followed up with a telephone call. The letter gave a brief description of the research problem, tools used, amount of time needed to conduct the study, and the research profile number (refer to Appendix 4 to view a copy of the letter sent to the school principals). Eighty-seven percent (87%) of the principals who were contacted agreed to participate in the study.

After receiving the principals' consent, the NBC teachers were located via the school district's internal-based email system and by telephone. Once they responded to the email solicitation, I contacted them personally. Numerous telephone calls had to be made after not receiving responses to a second email attempt. This was necessary to ensure an adequate sample for my study. Finally, after the teachers consented and agreed to participate, they had to meet a set of criteria in order to be involved in the study.

The first necessary requirement was for the National Board Certified (NBC) teachers to be teaching Grade 8 or 10 science students. Only these two grade levels in secondary school are tested on the science component of the Florida Comprehensive Assessment Test (FCAT). This was important because the science FCAT scores were going to be used in my study to measure student achievement. Next, each NBC teacher was requested to recruit a non-NBC teacher who taught the same science course at his or her school, had been teaching for approximately the same number of years, and/or taught groups of students with similar characteristics. The NBC teachers were given the option of choosing a non-NBC teacher who met at least two out of those four requirements (see Section 3.5.1 where I explain the reasons why the NBC teachers and non-NBC teachers had to be matched based on these criteria).

The requirement for securing a non-NBC teacher was simple because many of the schools that participated in the study have only two or three science teachers per grade level. Thus, the sample of non-NBC teachers was chosen following a request made by the NBC participant, and based on the non-NBC teachers' willingness to participate in the study. Hence, not all schools had a non-NBC teacher for comparison purposes (see Table 3.2 for detailed information about the number of NBC and non-NBC participants in each school).

Forty-two percent (42%) of the NBC teachers who were contacted volunteered and met the criteria for participation in the study. Overall, 17 NBC teachers and 13 non-NBC teachers who taught science to Grade 8 or 10 students in 12 secondary schools participated.

Table 3.2 Descriptive Information about Each NBC and Non-NBC Participating Teacher

| Teacher | School Name | NBC | Teaching Assignment | Gender | Education | Years Taught |
|---|---|---|---|---|---|---|
| 1 | A | No | 8 | M | Masters | 4 |
| 2 | A | Yes | 8 | F | Specialist | 11 |
| 3 | B | Yes | 8 | F | Masters | 9 |
| 4 | B | Yes | 10 | F | Masters | 9 |
| 5 | C | Yes | 10 | F | Masters | 14 |
| 6 | C | Yes | 10 | F | Masters | 7 |
| 7 | D | Yes | 8 | F | Masters | 11 |
| 8 | D | No | 10 | F | Bachelors | 4 |
| 9 | E | Yes | 10 | M | Masters | 15 |
| 10 | E | No | 8 | F | Bachelors | 7 |
| 11 | F | No | 8 | M | Bachelors | 1 |
| 12 | F | Yes | 10 | M | Masters | 18 |
| 13 | F | Yes | 8 | F | Bachelors | 8 |
| 14 | F | Yes | 8 | F | Masters | 9 |
| 15 | G | Yes | 10 | F | Masters | 20 |
| 16 | G | No | 8 | F | Bachelors | 5 |
| 17 | G | Yes | 10 | F | Masters | 26 |
| 18 | H | No | 10 | F | Bachelors | 7 |
| 19 | H | Yes | 8 | F | Specialist | 17 |
| 20 | I | No | 8 | M | Masters | 11 |
| 21 | I | No | 10 | F | Doctorate | 4 |
| 22 | J | Yes | 8 | F | Specialist | 34 |
| 23 | J | No | 10 | F | Bachelors | 1 |
| 24 | J | No | 8 | M | Masters | 8 |
| 25 | J | Yes | 8 | M | Masters | 9 |
| 26 | K | Yes | 10 | F | Masters | 17 |
| 27 | L | No | 8 | F | Bachelors | 10 |
| 28 | L | No | 10 | F | Bachelors | 5 |
| 29 | L | No | 10 | F | Specialist | 28 |
| 30 | L | Yes | 10 | F | Specialist | 10 |

Table 3.2 gives a profile for each of the 30 participating teachers (NBC and non-NBC) in terms of the number of participants at each school site, National Board Certification status, teaching assignment during the time of the study, gender, amount of education received and number of years teaching. As you can see from Table 3.2, Schools G and K had an uneven number of participants which resulted in two less

non-NBC teachers for inclusion in the study. Seventy percent (70%) of the teacher sample had a Masters degree or higher as compared to 48% of the district-wide instructional staff. In addition, the average number of years teaching was 11.3 for the teacher sample as compared to 10 years at the district-wide level (Miami-Dade County Public Schools, 2004).

All 30 participating teachers made their classes available for the collection of data involving the classroom environment questionnaire (WIHIC), modified attitude scale, and science achievement scores. The students in this study came from the 38 science classes taught by those teachers.

Prior to gathering the data from the students, a parental permission letter was sent home with each student (refer to Appendix 3 to view a copy of this parental permission letter). This letter sought permission for the researcher to administer the questionnaires and to gain access to the students' science achievement scores in order to conduct the present study. After the participating teachers collected the majority of the parental permission letters from their classes, the WIHIC questionnaire and the attitude scale were administered to the students. The students who did not turn in the permission form were not allowed to participate in the study. It is also important to mention that many students returned the parental permission form, but the parents would not allow their child's science FCAT scores to be released. Therefore, these students were not part of the study group.

The total student sample in my study consisted of 927 science students (443 students in 21 classes taught by National Board Certified (NBC) teachers and 484 students in

17 classes taught by non-NBC teachers). The students were all enrolled in the third grading period of either Grade 8 or 10 science classes. It is important to note that, in the U.S., the school year is divided into four grading periods. Each grading period has about nine weeks of study. The secondary science students who participated in this study had been enrolled in Grade 8 or 10 science classes for approximately 27 weeks.

## 3.4 Instruments

As previously mentioned in Section 3.3, there were two questionnaires used in this study. A revised version of the What Is Happening In this Class? (WIHIC) questionnaire was the instrument utilized to gather students' perceptions of their learning environment and a modified attitude scale modeled on the Test Of Science-Related Attitudes (TOSRA) was used to assess their attitudes toward science. Additionally, the students' scores on the science portion of the Florida Comprehensive Assessment Test (FCAT) were used to assess students' science achievement.

In this section, information is provided in regards to the selection of the WIHIC and the revisions made to it for the present study (Section 3.4.1). Also, the creation and modifications of the attitude scale are discussed (Section 3.4.2). Finally, information about the achievement measure is provided (Section 3.4.3).

### *3.4.1 Selection and Revisions of the WIHIC*

The What Is Happening In this Class? (WIHIC) questionnaire was chosen for this study because it has been found to be valid and reliable when used with secondary school students in numerous studies around the world (Dorman, 2003; Fraser, Fisher, & McRobbie, 1996; Margianti, Aldridge, & Fraser, 2004; Riah & Fraser, 1998) and in Miami-Dade County, Florida (Allen, 2003; Pickett & Fraser, 2004; Robinson, 2003). It was important to select a learning environment instrument that was appropriate for secondary school students in a variety of contexts because in Miami-Dade County, Florida, where my study took place, the student population is ethnically diverse (see Section 3.2 for more information about the context of my study). The fact that the WIHIC had been validated in Miami-Dade County, Florida made this learning environment instrument even more appealing for my study.

The WIHIC questionnaire brings parsimony to the field of classroom learning environments by combining modified versions of the most salient scales from a wide range of existing questionnaires with additional scales that accommodate current educational concerns (Fraser, 1998a). The original 56-item WIHIC questionnaire (7 scales with 8 items in each scale) was developed by Fraser et al. (1996, March) to assess students' perceptions of seven dimensions: Student Cohesiveness, Teacher Support, Involvement, Investigation, Task Orientation, Cooperation, and Equity. The WIHIC questionnaire uses a frequency scale as the response format. The WIHIC presents numerous statements that ask the respondent to agree or disagree using a five-point frequency scale of Almost Never, Seldom, Sometimes, Often, and Almost Always (refer to Section 2.6.1 for more information about the characteristics and development of the original WIHIC).

To ensure that the original 56-item version of the WIHIC (7 scales with 8 items in each scale) was appropriate for the students who were going to participate in my study, I conducted a pilot study with 18 Grade 8 students who were enrolled in my own science class. I administered the instrument to the 18 science students in my class, and then interviewed each student independently after he/she had completed the survey. The questions that I asked during the interviews were aimed at finding out if the students were able to understand the statements on the questionnaire.

Based on the pilot study results, the WIHIC did not need any major modifications, and the 56 items were retained in the seven scales. However, one word in the WIHIC questionnaire was altered after discovering that the selected class found a phrase to be confusing. This class assisted me in rewording that phrase. Students did not see a relationship between the word 'investigations' in the original WIHIC and the laboratory work that they are used to in their science classrooms. Thus, they felt that changing the word 'investigations' to 'labs' in each of the statements in the Investigation scale of the WIHIC would improve the comprehensibility of the instrument. Hence, the word 'investigations' was replaced with the word 'labs' in six of the eight items of the Investigation scale. For example, a statement that read "I carry out investigations in class to answer questions which puzzle me" was changed to "I carry out labs in class to answer questions which puzzle me". A copy of the revised version of the WIHIC can be found in Appendix 1.

### *3.4.2 Creation and Modification of the Attitude Scale*

In addition to measuring students' perceptions of their learning environment, my study involved the assessment of students' attitudes toward science. Thus, I needed an

attitude questionnaire appropriate for measuring secondary students' attitudes toward science. After carefully reading literature pertinent to the topic of attitudes to science, I became interested in Fraser's (1981) Test Of Science-Related Attitudes (TOSRA). The original TOSRA measures seven distinct science-related attitudes among secondary science school students: Social Implications of Science, Normality of Scientists, Attitude to Scientific Inquiry, Adoption of Scientific Attitudes, Enjoyment of Science Lessons, Leisure Interest in Science, and Career Interest in Science. Each scale contains 10 items, making a total of 70 items for the whole instrument. The response scale is a five-point Likert-type scale and has response categories ranging from Strongly Agree to Strongly Disagree. Likert response scales are commonly used to measure attitudes, knowledge, perceptions, values, and behavior changes. The response scale is named after Rensis Likert who developed this format in 1932 (Anderson, 1998). The TOSRA, just like the WIHIC, has been carefully developed and extensively field-tested and has shown to be highly valid and reliable (Fraser, 1981). Section 2.8 provides more information about the conceptualization, characteristics, and validation of the original TOSRA.

A main advantage that the Test Of Science-Related Attitudes (TOSRA) has over other science attitude tests is that it yields a separate score for a number of distinct attitudinal aims instead of a single overall score like its counterparts. This makes it possible to obtain a 'profile' of attitude scores for groups of students (Fraser, 1981). In addition, teachers and researchers have found the TOSRA to be a useful and easy-to-use instrument for measuring and monitoring progress of science-related attitudes of individual students or whole class of students.

For my study, I created a 10-item attitude scale modeled on the original TOSRA's Enjoyment of Science Lessons and Adoption of Scientific Attitudes scales. The Enjoyment of Science Lessons scale measures contentment with science learning experiences at school or in the classroom and the Adoption of Scientific Attitudes scale measures scientific attitudes like eagerness to modify/change opinions after acquiring new information. Instead of using the five-point response scale found in the original TOSRA, which is the most practical for most common purposes, a three-point response scale was used with response alternatives of Disagree, Not Sure and Agree. Although three-point or even two-point response scales are usually employed with young children, I decided to use a three-point response scale with the secondary students as a tactic to deal with time constraints.

Following the creation of the attitude scale, I conducted a pilot study to check if the items were appropriate for the students who were going to participate in my study. The pilot study consisted of administering the attitude scale to the 18 Grade 8 students who were enrolled in my own science class, and then interviewing each of them independently. The questions that I asked during the interviews were aimed at investigating if the students were able to understand the statements on the attitude scale.

Based on the results of the pilot study, I had to make further modifications to the 10-item attitude scale because some of the items proved troublesome for the students. This was the case when students had a difficult time answering negatively-phrased items. Thus, all negatively-phrased items were reworded in a positive manner. For instance, the statement that read "Finding out about new things is unimportant" was

changed to “Finding out about new things is important”. Appendix 2 provides a copy of the modified 10-item attitude scale that was used in the present study.

### *3.4.3 Achievement Measure*

The students’ scores on the science portion of the Florida Comprehensive Assessment Test (FCAT) were used as a measure of science achievement (refer to Sections 1.2 and 3.2 where the FCAT is discussed in more detail). The main reason why I chose the FCAT to measure achievement is because it is a norm-referenced test, which means that the students’ scores are compared to the scores of an original sample of students who were given an early version of the same test.

The purpose of the science portion of the FCAT is to assess students’ achievement of the science standards mandated by the Florida’s Sunshine State Standards and to compare the performance of Florida’s students to the rest of the nation. The science FCAT for Grade 8 and 10 students is given during the second half of the school year. The FCAT combines multiple-choice items with gridded-response items to measure students’ understanding of science concepts. The science FCAT at each grade level includes approximately the same number of questions from each of four clusters: physical and chemical science, earth space science, life and environmental science and scientific thinking. The science FCAT scores used to measure achievement in my study were obtained from the school district’s Division of Data Quality Management.

## 3.5 Research Design

Two distinct types of research designs were used in the present study: causal-comparative and correlational. The research questions dictated the choice of research designs. For instance, one research question dealt with investigating the differences between the effectiveness of National Board Certified (NBC) teachers and non-NBC teachers in terms of classroom environment, student attitudes and achievement in secondary science classes. This research question was best investigated by utilizing a causal-comparative research design. However, another research question reflected the need to understand the relationship between secondary science students' outcomes (attitudes toward science and science achievement) and their perceptions of the classroom learning environment. A correlational research design was more appropriate to investigate this particular research question. Further discussion about how I used the causal-comparative (Section 3.5.1) and correlational (Section 3.5.2) research designs in conducting my study is provided below.

The other two research questions concentrated on validating a revised version What Is Happening In this Class? (WIHIC) and an attitude scale modeled on the Test Of Science-Related Attitudes (TOSRA). The results from all of the statistical analyses demonstrated that the revised version of the WIHIC and the modified attitude scale modeled on the TOSRA were valid and reliable instruments that could be used with confidence in my study. Further detail on the validation of the two instruments can be found in Section 3.6.

### *3.5.1 Causal-Comparative Research Design*

To answer the following research question, a causal-comparative research design was found to be most appropriate:

*Research Question #3*

*Are National Board Certified (NBC) teachers more effective than non-NBC teachers in terms of classroom environment, student attitudes and student achievement in secondary science classrooms in South Florida?*

The main reason why I decided to use the causal-comparative research design to answer this research question was because the students were already divided into two groups that differed on the independent variable (one group was being taught science by National Board Certified (NBC) teachers and the other group by non-NBC teachers) and that were going to be compared on three dependent variables (perceptions of the science classroom environment, attitudes toward science, and science achievement). According to Gay and Airasian (1996), causal-comparative research is similar to experimental research because they both attempt to establish cause-effect relationships and involve group comparisons. However, in causal-comparative research, the researcher does not manipulate the independent variable. Although an experimental design is the most valid approach to investigating cause-and-effect relationships, it is not always possible to utilize such an approach, as it was the case in my study. When comparing the group of students taught by the National Board Certified (NBC) teachers with the group taught by the non-NBC teachers, the students had already been assigned to and taught by either a National Board Certified (NBC) or non-NBC teacher. Thus, the independent variable (teacher certification status) could not be manipulated making it impossible to utilize an experimental design.

It is important to mention the internal validity here. The term ‘internal validity’ means whether relationships between variables in the study have been correctly interpreted, especially whether the independent-dependent variable relationships are correctly interpreted (Punch, 1998). In other words, internal validity is commonly used in educational research when judging if the differences on the dependent variable are a direct result of the independent variable and not some other variable. In the case of my study, it would be difficult to determine if the difference on the three dependent variables (students’ learning environment perceptions, attitudes towards science, and academic achievement in science) were a direct result of students having been taught by a National Board Certified teacher. I was aware that other independent variables, such as the type of school that students attend, type of science course in which students are enrolled, the number of years that the teacher had been teaching, and the characteristics of the students in the class, also could have been the partial causes of higher or lower scores on the dependent variables.

Because I was aware of some of those extraneous variables before conducting the study, I was able to particularly control for some of them. The control technique that I used consisted of matching each National Board Certified (NBC) teacher with a non-NBC teacher whose teaching experience was similar. For instance, after I selected the NBC teachers who were willing and able to participate in my study, I requested that each NBC teacher recruit one non-NBC teacher who met at least two of four criteria. First, the non-NBC teacher must have been teaching at the same school as the NBC teacher. Second, the non-NBC teacher must have been teaching the same science course (i.e. biology, physics) as the NBC teacher. Third, the non-NBC teacher must have been teaching for approximately the same number of years as the NBC teacher.

Fourth, the non-NBC teacher must have been teaching a group of students with similar characteristics (i.e. science achievement levels, level of English proficiency) as the one taught by the NBC teacher. Although I controlled several extraneous variables by matching the NBC to non-NBC teachers, it can't be guaranteed that all extraneous variables were controlled for. Therefore, the results should be interpreted with caution (also see Section 5.5 where I address this limitation in more detail).

Mathison (1988) investigated methods of good research practice and proposed that it was essential for the researcher to triangulate. Triangulation means to employ "multiple methods, data sources, and researchers to enhance the validity of research findings in order to withstand critique by colleagues" (Mathison, 1988, p. 13). To increase the credibility of my findings, I employed researcher triangulation in my study. Two other researchers assisted me in the collection and analyses of the data. Denzin (1978) proposed data triangulation, which refers not only to using several data sources but to include multiple uses of time and space. I employed data triangulation in my study by collecting the data during different times of the school day and in different settings.

### *3.5.2 Correlational Research Design*

To answer the following research question, a correlational research design was found to be most appropriate:

*Research Question #4*

*Are there associations between student outcomes (attitudes and achievement) and classroom environment in secondary science classrooms in South Florida?*

Correlational research attempts to describe the relationships among naturally occurring variables or phenomena without attempting to change them (Shulman, 1997). In my study, I collected measures of the students' learning environment perceptions, attitudes toward science, and science achievement to assist me in accurate predictions. Learning environment perceptions and attitudes were measured by administering questionnaires to the students and achievement was measured by examining students' scores on a statewide science achievement test.

Shulman (1997) explains that the investigator should use techniques of correlation and regression to investigate the relationships between predictor variables and sets of useful outcomes measures. For this reason, correlation as a statistical procedure could be used to determine how two variables are related or how much they are related. Anderson (1998) states that correlational research relies on pairs of observations, each related to an individual or other unit of analysis. This does not mean that one variable causes another but only that there is a relationship. The cause of the relationship could be something else entirely. I chose the correlational method to answer the above research question because I was not able to manipulate the predictor variables.

### 3.6 Procedures for Data Analysis

Several steps were taken to collect and analyze the data in my study. First, the revised version of the WIHIC questionnaire and the modified attitude scale were administered to 927 Grade 8 and 10 science students in 38 classes from 13 secondary schools in

Miami-Dade County, Florida. Two trained researchers and I administered the instruments during different periods of the day on different days of the week.

The data gathered from the 927 students were statistically analyzed to investigate if the revised WIHIC was valid and reliable when used with secondary science students. First, principal axis factor analysis with oblique rotation was conducted to check the *a priori* seven-factor structure of the revised version of the WIHIC. Second, internal consistency reliability at two units of analysis (individual and class mean) was calculated for each of the seven scales of the revised version of the WIHIC. The Cronbach alpha reliability coefficient was used as an index of scale internal consistency. Next, discriminant validity (using the mean correlation of a scale with the other scales as a convenient index) was used to indicate the extent to which each scale measured a distinct construct. Finally, one-way ANOVA for each WIHIC scale was conducted to check if the WIHIC was able to differentiate between the perceptions of students in the different classrooms. Results of these analyses are reported in Section 4.2.

To investigate if the modified attitude scale modeled on the TOSRA was reliable when used with science students in secondary school, the data collected from the 927 secondary science students were subjected to internal consistency reliability analysis at two units of analysis (individual and class mean). The Cronbach alpha reliability coefficient was used as an index of scale internal consistency to indicate whether each item in the 10-item attitude scale assesses a similar construct. The results of the internal consistency reliability analysis for the modified attitude scale are discussed in Section 4.2.5.

After the results from all of the statistical analyses demonstrated that the revised version of the WIHIC and the modified attitude scale modeled on the TOSRA were valid and reliable instruments that could be used with confidence in my study, the data that had been gathered with these two instruments were used to answer the remaining research questions. The science FCAT scores from all of the participants also had to be collected to answer the rest of the research questions. To assist in the tedious process of collecting science FCAT scores from the 927 students, the students recorded their identification numbers on their surveys. Using the Excel program, I transferred the student identification numbers along with their responses from the WIHIC and the attitude scale. To ensure anonymity, no names were attached to the student identification numbers. Then, with the assistance of the Miami-Dade County Public Schools' Department of Assessment and Data Analysis, the science FCAT scores were downloaded. After I received the downloaded science FCAT scores from all students in the district, I had to research and locate each student's science FCAT score from the data sheets that had been provided to me. After the science FCAT scores were obtained, the student identification numbers were replaced by a sequential code.

The data gathered with the revised WIHIC, modified attitude scale modeled on the TOSRA, and the science FCAT scores were all statistically analyzed. First, to investigate the differences between National Board Certified (NBC) and non-NBC teachers in terms of secondary science students' learning environment perceptions, attitudes toward science, and science achievement, MANOVA was conducted and effect sizes were calculated. The set of dependent variable consisted of seven learning environment scales assessed by the WIHIC, one attitude scale, and science FCAT

scores. The independent variable was the two-level variable (NBC and non-NBC teacher). The MANOVA results and the effect sizes are reported in Section 4.3. Next, to investigate the relationships between secondary students' outcomes (attitudes toward science and science achievement) and their perceptions of the science classroom environment, simple correlation and multiple regression analyses were conducted at two units of analysis (individual and class mean). The results of these combined analyses are reported in Section 4.4.

## 3.7 Summary of Methodology

This chapter described the methods used in my study. It also discussed the context in which the study was conducted, data sources used and how they were selected, instruments, research design, procedures followed, and data analyses conducted. The main objective of my study was to investigate differences between National Board Certified (NBC) and non-NBC teachers in terms of their students' perceptions of the learning environment and outcomes (attitudes toward science and science achievement). The study involved a sample of 927 students (443 students in 21 classes taught by NBC teachers and 484 students in 17 classes taught by non-NBC teachers) in Miami-Dade County Public Schools, Florida.

Because I was going to assess students' perceptions of their science classroom environment and their attitudes toward science, I needed a learning environment and attitude questionnaire that would serve this purpose. First, I selected the What Is Happening In this Class? (WIHIC) questionnaire to assess students' perceptions of their classroom learning environment. The WIHIC was pilot-tested with a small

sample of students in Grade 8 to check if modifications were needed. Based on the pilot study, only minor revisions were necessary. Thus, a revised version of the WIHIC was used in my study. Next, I created a 10-item attitude scale modeled on the Test Of Science-Related Attitudes (TOSRA) to measure students' attitudes to science. The 10-item attitude scale was also pilot-tested with a small sample of Grade 8 students, and the results of the pilot study helped me to realize that modifications were needed in order to make the instrument more comprehensible for the students. Thus, I made modifications such as changing phrases which were negatively-worded. Therefore, a modified 10-item attitude scale modeled on the TOSRA was used in my study.

Because I was also interested in measuring students' science achievement, I used students' scores on the science portion of the Florida Comprehensive Assessment Test (FCAT). The science FCAT is the state-mandated examination that is used to measure science achievement in Grades 8 and 10. The district's Department of Assessment and Data Analysis provided these scores.

The data gathered with the revised WIHIC were statistically analyzed in terms of its factor structure, internal consistency reliability and ability to differentiate between classrooms to determine the validity and reliability of this learning environment instrument. The modified attitude scale modeled on the TOSRA was also administered to the student sample, and the data were statistically analysed in terms of internal consistency reliability to determine if it was reliable when used with secondary school students.

Because the revised WIHIC and modified attitude scale were found to be valid and reliable tools, they were used to investigate the main research questions of my study. The students' science FCAT scores were also used. Thus, the data gathered with the revised version of the WIHIC and the modified attitude scale and students' scores on the science FCAT were statistically analyzed using MANOVA and effect sizes to investigate any differences between the effectiveness of National Board Certified (NBC) and non-NBC teachers in terms of their students' perceptions of the science learning environment, attitudes toward science and science achievement.

The data gathered with the revised WIHIC and modified attitude scale, as well as the students' science FCAT scores, were also statistically analysed using simple correlation and multiple regression analyses to determine whether associations exist between students' perceptions of the learning environment and students' outcomes (attitudes toward science and science achievement).

Chapter 4 reports in detail the results of all the analyses that were conducted to answer the research questions of my study.

# Chapter 4

# DATA ANALYSES AND RESULTS

## 4.1 Introduction

The main focus of the present study, as discussed in Chapter 1, was to investigate the effectiveness of National Board Certified (NBC) science teachers in terms of their secondary science students' perceptions of the classroom environment, attitudes, and achievement. In addition, I explored associations between student outcomes (attitudes and achievement) and the classroom learning environment among secondary science students in Miami-Dade County, Florida schools. Because valid and reliable tools were needed to assess the participating students' perceptions of their science classroom environment and their attitudes toward science, another aim of my study was to validate a learning environment and an attitude questionnaire that could be used with confidence when gathering the data to answer the main research questions.

This chapter reports the analyses of the data collected for my study and the results for each research question. The data were collected utilizing a revised version of the What Is Happening In this Class? (WIHIC) questionnaire, a modified attitude scale modeled on the Test Of Science-Related Attitudes (TOSRA), and students' achievement scores on the science component of the Florida Comprehensive Assessment Test (FCAT). The study sample consisted of 927 Grade 8 and 10 students in 38 classes from 13 secondary schools (see Section 3.3 for a more detailed description of the sample used).

The sections that follow report the results of the data analyses used to answer the research questions delineated in Section 1.5. Section 4.2 reports the validity and reliability of the revised version of the What Is Happening In this Class? (WIHIC) questionnaire and the reliability of a modified attitude scale derived from the TOSRA. Section 4.3 reports the results for the differences between the effectiveness of National Board Certified (NBC) teachers and non-NBC teachers in terms of classroom environment, student attitudes and achievement in their secondary science classes. Section 4.4 reports the results of analyses for associations between student outcomes (attitudes and achievement) and the classroom environment.

## 4.2 Validation of the Revised Version of the WIHIC and the Modified Attitude Scale

The What Is Happening In this Class (WIHIC) questionnaire was developed by Fraser, Fisher, and McRobbie (1996) to gather secondary students' perceptions of their classroom environment. It was designed to bring economy to the field of learning environments by combining the most relevant scales from existing learning environment questionnaires with new scales that measure cooperative learning and equity, which are two of the most important aspects of today's classrooms (Aldridge, Fraser, & Huang, 1999). While the WIHIC is a relatively new instrument, it has already been used in Australia and Taiwan (Aldridge & Fraser, 2000), Brunei (Riah & Fraser, 1998), Korea (Kim, Fisher, & Fraser, 1999), Singapore (Chionh & Fraser, 1998; Khoo & Fraser, 1998) and the United States (Moss & Fraser, 2001; Sinclair & Fraser, 2003). More detailed information about the development, characteristics, and validity of the original WIHIC can be found in Section 2.6.

The original WIHIC is made up of 56 items (7 scales with 8 items in each scale). The seven scales are Student Cohesiveness, Teacher Support, Involvement, Investigation, Task Orientation, Cooperation and Equity (Aldridge & Fraser, 2000; Fraser et al., 1996). All 56 items in the seven WIHIC scales were used in the present study. However, before it could be field-tested with a large sample, it was important to conduct a pilot study to ensure that young adolescent students were able to understand the statements on the questionnaire. After the pilot study, minor revisions were made to the original WIHIC questionnaire because students had difficulty with some of the items (see Section 3.4.1 for details about the revisions that were made to the WIHIC based on the pilot study and Appendix 1 to view a copy of the revised version of the WIHIC used in my study).

The Test Of Science-Related Attitudes (TOSRA) (Fraser, 1981), just like the WIHIC, has been carefully developed and extensively field-tested and has shown to be highly valid and reliable when used with secondary students. The original TOSRA has 70 statements (7 scales with 10 statements in each). The seven scales are Social Implications of Science, Normality of Scientists, Attitude to Scientific Inquiry, Adoption of Scientific Attitudes, Enjoyment of Science Lessons, Leisure Interest in Science, and Career Interest in Science. Section 2.8 provides more information about the conceptualization, characteristics, and validation of the original TOSRA.

For the present study, I created a 10-item attitude scale modeled on the original TOSRA. Items were modeled on selected items in TOSRA's Enjoyment of Science Lessons and Adoption of Scientific Attitudes scales. The attitude scale derived from the original TOSRA was pilot tested with adolescent students to ensure that the items

were comprehensible to them before field-testing with a larger sample. Further modifications were made to the 10-item attitude scale when some of the items proved troublesome for the students. Section 3.6 provides details about the revisions that were made to the 10-item attitude scale based on results of the pilot study. Appendix 2 provides a copy of the modified 10-item attitude scale that was used in the present study.

This section presents a discussion of analysis and results pertaining to the first and second research questions:

*Research Question #1*

*Is a revised version of the What Is Happening In this Class? (WIHIC) questionnaire valid and reliable when used in secondary science classrooms in South Florida?*

*Research Question #2*

*Is an attitude scale modeled on the Test Of Science-Related Attitude (TOSRA) reliable when used in secondary science classrooms in South Florida?*

In order to check the validity and reliability of the revised version of the WIHIC questionnaire, the following statistical measures were used: factor structure (Section 4.2.1), internal consistency reliability (Section 4.2.2), discriminant validity (Section 4.2.3), and the ability to differentiate between classes using one-way ANOVA (Section 4.2.4). In order to check the reliability of the single modified attitude scale derived from the Test Of Science-Related Attitudes (TOSRA), the Cronbach alpha

coefficient was calculated to establish the internal consistency reliability using the individual student and class mean as the units of analysis (Section 4.2.5).

### *4.2.1. Factor Structure of the Revised Version of the WIHIC*

The data collected from 927 students were analyzed to provide statistical validation for the revised 56-item seven-scale version of the WIHIC questionnaire. Factor and item analyses were conducted to identify questionnaire items whose removal would improve the internal consistency reliability and factorial validity of each WIHIC scale. Item analysis of the 56 WIHIC items showed that all items had sizeable item-remainder correlations.

To check the structure of the revised version of the WIHIC, principal axis factor analysis with oblique rotation was conducted. The factor loadings obtained are reported in Table 4.1. Only loadings of 0.30 or higher are reported in Table 4.1. For an item to be retained, two criteria needed to be satisfied: the loading on an item's *a priori* scale needs to be at least 0.30; and the loading on each of the other six WIHIC scales needs to be less than 0.30. Though it needs to be noted, 0.40 is a more acceptable value with factor loadings (Dryden & Fraser, 1998; Quek, Wong, & Fraser. 2005b; Soerjaningsih, Fraser, & Aldridge, 2001).

There are 392 possible loadings in Table 4.1 (56 items x 7 scales = 392). Only two of the possible 392 cases in the original seven-factor structure were not replicated, leaving a total 54 of the 56 items. Items 6 and 45 each had a loading of less than 0.30 on its own scale and therefore were omitted from subsequent analyses.

Table 4.1 Factor Loadings for the Revised Version of the WIHIC (Principal Axis Factoring with Oblique Rotation)

| Item No | Factor Loadings | | | | | | |
|---|---|---|---|---|---|---|---|
| | Student Cohesiveness | Teacher Support | Involvement | Investigation | Task Orientation | Cooperation | Equity |
| 1 | 0.68 | | | | | | |
| 2 | 0.69 | | | | | | |
| 3 | 0.48 | | | | | | |
| 4 | 0.81 | | | | | | |
| 5 | 0.39 | | | | | | |
| 7 | 0.57 | | | | | | |
| 8 | 0.31 | | | | | | |
| 9 | | 0.71 | | | | | |
| 10 | | 0.74 | | | | | |
| 11 | | 0.68 | | | | | |
| 12 | | 0.51 | | | | | |
| 13 | | 0.74 | | | | | |
| 14 | | 0.79 | | | | | |
| 15 | | 0.65 | | | | | |
| 16 | | 0.48 | | | | | |
| 17 | | | 0.69 | | | | |
| 18 | | | 0.86 | | | | |
| 19 | | | 0.56 | | | | |
| 20 | | | 0.69 | | | | |
| 21 | | | 0.49 | | | | |
| 22 | | | 0.56 | | | | |
| 23 | | | 0.41 | | | | |
| 24 | | | 0.47 | | | | |
| 25 | | | | 0.72 | | | |
| 26 | | | | 0.63 | | | |
| 27 | | | | 0.80 | | | |
| 28 | | | | 0.64 | | | |
| 29 | | | | 0.77 | | | |
| 30 | | | | 0.77 | | | |
| 31 | | | | 0.67 | | | |
| 32 | | | | 0.58 | | | |
| 33 | | | | | 0.59 | | |
| 34 | | | | | 0.67 | | |
| 35 | | | | | 0.37 | | |
| 36 | | | | | 0.59 | | |
| 37 | | | | | 0.62 | | |
| 38 | | | | | 0.63 | | |
| 39 | | | | | 0.66 | | |
| 40 | | | | | 0.62 | | |
| 41 | | | | | | 0.54 | |
| 42 | | | | | | 0.51 | |
| 43 | | | | | | 0.63 | |
| 44 | | | | | | 0.62 | |
| 46 | | | | | | 0.69 | |
| 47 | | | | | | 0.68 | |
| 48 | | | | | | 0.52 | |
| 49 | | | | | | | 0.64 |
| 50 | | | | | | | 0.66 |
| 51 | | | | | | | 0.78 |
| 52 | | | | | | | 0.80 |
| 53 | | | | | | | 0.83 |
| 54 | | | | | | | 0.78 |
| 55 | | | | | | | 0.71 |
| 56 | | | | | | | 0.79 |
| % Variance | 2.76 | 2.93 | 4.77 | 7.00 | 3.69 | 31.83 | 6.36 |
| Eigenvalue | 1.49 | 1.58 | 2.58 | 3.78 | 2.00 | 17.19 | 3.43 |

The sample consisted of 927 students.

Factor loadings smaller than 0.30 have been omitted.
Items 6 and 45 were omitted.

Table 4.1 shows that all of the remaining 54 items had a factor loading of 0.30 or greater on their *a priori* scale and less than 0.30 on each of the other six scales. Previous validation studies conducted in Singapore (Fraser & Chionh, 2000), Australia and Taiwan (Aldridge & Fraser, 2000), Canada (Raaflaub & Fraser, 2003; Zandvliet & Fraser, 2004) and the U.S. (Allen, 2003) have reported similar factor structures for the original and/or modified versions of the WIHIC questionnaire, thus further supporting this factor structure.

The bottom of Table 4.1 shows that the percentage of variance accounted for by the different factors ranges from 2.76% to 31.83%. The total proportion of variance accounted for by the 54 WIHIC items in seven scales is 59.34%. The bottom of Table 4.1 also shows that the eigenvalue for each of the seven WIHIC scales ranges from 1.49 to 17.19.

### *4.2.2 Internal Consistency Reliability of the Revised Version of the WIHIC*

The internal consistency reliability indicates whether each item in a scale assesses a similar construct. The Cronbach alpha reliability coefficient was used as an index of scale internal consistency. Table 4.2 reports the Cronbach alpha coefficient for each of the seven scales of the revised version of the WIHIC using two units of analysis (individual and class mean) for the whole sample of 927 students in 38 classes. The alpha reliability estimates for different scales of the revised version of the WIHIC range from 0.82 to 0.92 for the individual as the unit of analysis and from 0.69 to 0.97 for the class mean as the unit of analysis. These internal consistency indices are comparable to those obtained when the WIHIC was used with an Australian sample

(Fraser et al., 1996), which ranged from 0.67 to 0.88, and with an Indonesian sample (Margianti, Fraser, & Aldridge, 2002), which ranged from 0.65 to 0.87. Therefore, the results suggest reasonable reliability for the revised version of the WIHIC.

### *4.2.3 Discriminant Validity of the Revised Version of the WIHIC*

The discriminant validity (using the mean correlation of a scale with the other scales as a convenient index) indicates whether each scale is distinct in what it measures.

Table 4.2 Internal Consistency Reliability (Cronbach Alpha Coefficient) and Discriminant Validity(Mean Correlation with Other Scales) for Two Units of Analysis and Ability to Differentiate between Classrooms (ANOVA Results) for the Revised Version of the WIHIC and a Modified Attitude Scale

| Scale | Unit of Analysis | No of Items | Alpha Reliability | Mean Correlation with other Scales | ANOVA $Eta^2$ |
|---|---|---|---|---|---|
| Student Cohesiveness | Individual | 7 | 0.84 | 0.44 | 0.12** |
| | Class Mean | | 0.91 | 0.61 | |
| Teacher Support | Individual | 8 | 0.90 | 0.48 | 0.14** |
| | Class Mean | | 0.96 | 0.59 | |
| Involvement | Individual | 8 | 0.88 | 0.49 | 0.11** |
| | Class Mean | | 0.94 | 0.64 | |
| Investigation | Individual | 8 | 0.90 | 0.41 | 0.15** |
| | Class Mean | | 0.96 | 0.58 | |
| Task Orientation | Individual | 8 | 0.82 | 0.45 | 0.12** |
| | Class Mean | | 0.69 | 0.56 | |
| Cooperation | Individual | 7 | 0.89 | 0.50 | 0.13** |
| | Class Mean | | 0.95 | 0.68 | |
| Equity | Individual | 8 | 0.92 | 0.48 | 0.11** |
| | Class Mean | | 0.97 | 0.71 | |
| Attitude | Individual | 10 | 0.81 | | |
| | Class Mean | | 0.93 | | |

** $p<0.01$
The sample consisted of 927 students in 38 classes.
The $eta^2$ statistic (which is the ratio of 'between' to 'total' sums of squares) represents the proportion of variance explained by class membership.

Table 4.2 shows that the mean correlation with other scales for the individual student as the unit of analysis ranges from 0.41 to 0.50 for different WIHIC scales. The mean correlation with other scales for classes ranges from 0.56 to 0.71 for different WIHIC scales. Even though some discriminant validity values (see Table 4.2) are quite high for raw scores on the revised version of the WIHIC, the factor analysis results (see Table 4.1) attest to the independence of factor scores.

### *4.2.4 Ability of the Revised Version of the WIHIC to Differentiate between Classrooms*

As further evidence of the validity of the revised version of the WIHIC, a one-way ANOVA was used to indicate whether each scale of the questionnaire was able to differentiate significantly between the perceptions of students in different classes. The results reported in Table 4.2 suggest that all scales were able to do so. Scores on a WIHIC scale were used as the dependent variable and class membership was used as the independent variable. The $eta^2$ statistic (an estimate of the strength of association between class membership and the dependent variable) ranges from 0.11 to 0.15 for different scales, and was statistically significant ($p<0.01$) for all seven WIHIC scales (see Table 4.2). On the whole, these figures are similar to those obtained in validation studies conducted in Australia with 1,081 students in 50 classes (Huang & Fraser, 1997), Brunei with 1,188 students in 54 classes (Khine & Fisher, 2001), Singapore with 2,310 students in 75 classes (Chionh & Fraser, 1998), and Miami-Dade County, Florida with 1,105 students in 54 classes (Soto-Rodriguez & Fraser, 2004) for the original, modified, and/or translated versions of the WIHIC.

### *4.2.5 Internal Consistency Reliability of the Modified Attitude Scale*

The internal consistency reliability was used to indicate whether each item in the 10-item attitude scale assesses a similar construct. The Cronbach alpha reliability coefficient was used as an index of scale internal consistency. Table 4.2 reports the Cronbach alpha coefficient for the modified attitude scale using two units of analysis (individual and class mean) for the whole sample of 927 students in 38 classes. For the single modified attitude scale, the alpha reliability coefficient is 0.81 with the individual student as the unit of analysis and 0.93 for the class mean (see Table 4.2). These results suggest reasonable reliability for the modified attitude scale.

## 4.3 Differences between National Board Certified (NBC) and Non-NBC Teachers in Terms of Classroom Environment and Student Attitudes and Achievement

This section reports differences between secondary science students who were taught by National Board Certified (NBC) teachers and those who were taught by non-NBC teachers in terms of their perceptions of the classroom learning environment, attitudes toward science, and science achievement. The revised version of the What Is Happening In this Class? (WIHIC) questionnaire was used to measure students' perceptions of their science classroom environment. A modified attitude scale derived from the Test Of Science-Related Attitudes (TOSRA) was used to measure students' attitudes to science. To assess achievement, the students' state science examination scores from the Florida Comprehensive Assessment Test (FCAT) were obtained from the school district's Division of Data Quality Management. This test is given during the Spring each year and was used as a measure of achievement.

The sample for this portion of my study consisted of 927 Grade 8 and 10 students in 38 science classes in 13 secondary schools in Miami-Dade County, Florida. NBC teachers taught 443 of those students in 21 classes and non-NBC teachers taught 484 students in 17 classes. Because the revised version of the WIHIC and the modified attitude scale were found to be valid and reliable instruments, the same data gathered during the validation stage of my study (previously discussed in Section 4.2) were statistically analyzed along with the FCAT scores obtained from the school district to answer the third research question:

*Research Question #3*

*Are National Board Certified (NBC) teachers more effective than non-NBC teachers in terms of classroom environment, student attitudes and student achievement in secondary science classrooms in South Florida?*

MANOVA was used to determine whether differences existed between students in classes that are taught by NBC teachers and those taught by non-NBC teachers in terms of scores on each WIHIC scale, the modified attitude scale and the science FCAT. Because the MANOVA produced statistically significant results using Wilks' lambda criterion, the univariate ANOVA results were interpreted for each of the nine dependent variables. The results of the nine ANOVA's are shown in the last column of Table 4.3.

The MANOVA results in Table 4.3 indicate statistically significant ($p<0.05$) differences between students taught by National Board Certified (NBC) teachers and

those taught by non-NBC teachers for five of the seven WIHIC scales (namely, Teacher Support, Involvement, Task Orientation, Investigation and Cooperation) and the modified attitude scale. Differences between the two groups were nonsignificant for the WIHIC scales of Student Cohesiveness and Equity and for achievement.

Table 4.3 Average Item Mean, Average Item Standard Deviation and Difference between NBC and non-NBC Teachers (Effect Size and MANOVA Results) for Student Scores on WIHIC, Modified TOSRA and Achievement Scales Using the Individual as the Unit of Analysis

| Scale | Average Item Mean | | Average Item Standard Deviation | | Difference | |
|---|---|---|---|---|---|---|
| | NBC | Non-NBC | NBC | Non-NBC | Effect Size | *F* |
| Student Cohesiveness | 4.22 | 4.17 | 0.72 | 0.66 | 0.07 | 1.11 |
| Teacher Support | 3.75 | 3.57 | 0.93 | 0 .92 | 0.19 | 1.73** |
| Involvement | 3.42 | 3.16 | 0.91 | 0.89 | 0.29 | 2.09** |
| Task Orientation | 4.36 | 4.25 | 0.74 | 0.73 | 0.15 | 1.53* |
| Investigation | 3.25 | 2.99 | 1.00 | 1.02 | 0.25 | 1.96** |
| Cooperation | 4.16 | 4.05 | 0.78 | 0.84 | 0.14 | 1.46* |
| Equity | 4.29 | 4.19 | 0.88 | 0.88 | 0.11 | 1.34 |
| Attitudes | 2.44 | 2.30 | 0.41 | 0.40 | 0.35 | 2.26** |
| Achievement | 31.15 | 30.47 | 6.11 | 5.67 | 0.12 | 1.29 |

*$p<0.05$, **$p<0.01$
The sample consisted of 927 students in 38 science classes (443 students in NBC teachers' classes and 484 students in non-NBC teachers' classes).

Table 4.3 reports the average item mean and average item standard deviation for each environment, attitude and achievement scale for NBC and non-NBC teachers. The average item mean is the scale mean divided by the number of items in that scale and

was used to enable easy comparison of the average scores on scales containing different numbers of items. Figure 4.1 graphically depicts a simplified plot of significant differences between NBC and non-NBC teachers in terms of mean scores on each WIHIC scale, the modified attitude scale and science achievement.

In Figure 4.1 the points plotted for Student Cohesiveness, Equity and Achievement was the average of the two statistically non-significantly different item means, rather than plotting both values for each scale (see Table 4.3). This was done so that non-significant differences in this graph appear as zero differences.

The graphical representation of the differences between the two groups (NBC vs. non-NBC teachers) in Figure 4.1 indicates that students in NBC classes had somewhat higher scores on learning environment, attitude and achievement scales than did the students of the non-NBC teachers. However, the graph suggests that the differences between NBC and non-NBC teachers are relatively small for all scales of the revised WIHIC, modified attitude scale, and science FCAT.

The effect size, or the difference between means expressed in standard deviation units (Thompson, 1998a, 1998b), provides an index of the magnitude of effect and its educational importance. Table 4.3 shows that the effect sizes for the five scales of the WIHIC questionnaire for which the differences between students taught by NBC and non-NBC teachers were statistically significant range from 0.14 to 0.29 standard deviations. The effect size for differences between groups on the attitude scale was approximately one third of a standard deviation. Overall, effect sizes range from small

to modest. It is noteworthy that the direction of differences for all the nine scales in Table 4.3 is consistently in favor of the group of students taught by NBC teachers.

In summary, the difference between groups is statistically significant for six of the nine measures. The effect size is of small to modest magnitude, with a range from 0.14 to 0.35 standard deviations.

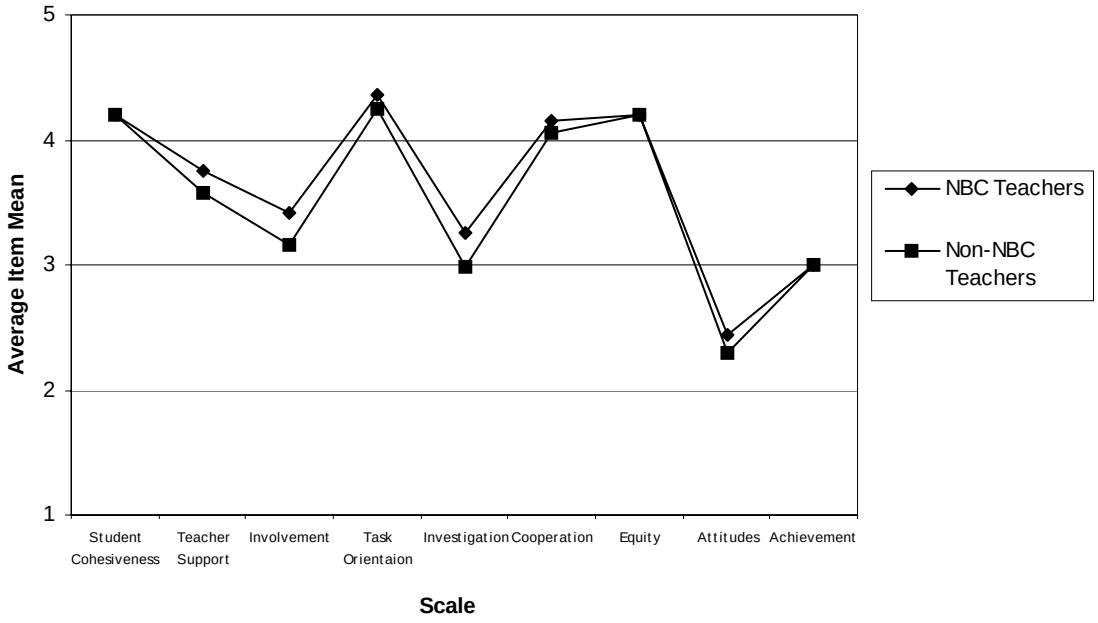


Figure 4.1 Simplified Plot of Significant Differences of National Board Certified (NBC) and Non-NBC Teachers' Scores on the Revised WIHIC Modified TOSRA, and Science Achievement Scales (*N*=927)

Overall, these results tentatively suggest that National Board Certified (NBC) science teachers are more effective than non-NBC science teachers in providing more teacher support to their students and creating a learning environment in which the students have more opportunities to become involved, carry out investigations, and work

cooperatively. Furthermore, the students of NBC science teachers report having more favorable attitudes toward science than students of non-NBC science teachers.

It is important to point out that generalizability of these findings are limited because my study was conducted in only one geographical area of the United States (namely, Miami-Dade County, Florida), and therefore cannot be generalized to the rest of the nation. I suggest that further research be conducted in other states of the U.S. to investigate if NBC science teachers are more effective that non-NBC science teachers in terms of their students' perceptions of the learning environment, attitudes toward science, and achievement. This limitation is further addressed in Section 5.5.

## 4.4 Associations between Students' Outcomes (Attitudes and Achievement) and Classroom Environment

This section reports the results of an investigation of whether differences exist between secondary science students' outcomes (attitudes and achievement) and their classroom environment perceptions. The revised version of the WIHIC, a modified attitude scale derived from the TOSRA, and the science FCAT were used to measure, respectively, students' perceptions of their science classroom environment, their attitudes to science, and science achievement. The same data gathered from the 927 Grade 8 and 10 students to answer Research Question #3 (previously discussed in Section 4.3) were statistically analyzed to answer the fourth research question:

*Research Question #4*

*Are there associations between student outcomes (attitudes and achievement) and classroom environment in secondary science classrooms in South Florida?*

Simple correlation ($r$) and multiple regression analyses for two units of analysis (individual and class mean) were used to determine whether associations exist between students' perceptions of the learning environment and each student outcome measure (attitudes and achievement).

The simple correlation analysis provides information about the bivariate association between each student outcome and each learning environment scale. The results of the simple correlation analysis (reported in Table 4.4) suggest a positive and statistically significant ($p<0.05$) correlation between student attitudes and all seven learning environment scales, with either the individual or the class mean as the unit of analysis.

The simple correlation analysis results for associations between students' science achievement and the learning environment, however, show that positive and statistically significant ($p<0.05$) correlations occurred only for Involvement, Investigation and Equity for the student level of analysis. For the class mean as the unit of analysis, Table 4.4 shows no significant correlation between student achievement and any of the learning environment scales.

Multiple regression analysis was undertaken using the set of seven scales of the WIHIC questionnaire as independent variables and either science attitudes or achievement as the dependent variable. This analysis provided a more parsimonious picture of the joint influence of a set of correlated environment scales on each outcome and reduced the Type I error rate. A multiple regression analysis was

performed separately using the individual student and the class mean as the units of analysis.

Table 4.4 Simple Correlation and Multiple Regression Analyses for Associations Between Two Student Outcomes and Scores on the WIHIC for Two Units of Analysis

| Scale | Unit of Analysis | Association with Learning Environment | | | |
|---|---|---|---|---|---|
| | | Attitudes | | Achievement | |
| | | *r* | *β* | *r* | *β* |
| Student Cohesiveness | Individual | 0.28** | 0.01 | -0.01 | -0.09 |
| | Class | 0.49* | -0.24 | 0.10 | 0.18 |
| Teacher Support | Individual | 0.47* | 0.23** | 0.05 | -0.06 |
| | Class | 0.70** | 0.54 | -0.01 | 0.46 |
| Involvement | Individual | 0.46** | 0.23** | 0.07* | 0.05 |
| | Class | 0.63** | 0.34 | -0.11 | -0.20 |
| Investigation | Individual | 0.34** | 0.03 | 0.10** | 0.08 |
| | Class | 0.52** | -0.02 | 0.12 | 0.19 |
| Task Orientation | Individual | 0.40** | 0.17** | 0.04 | -0.03 |
| | Class | 0.60** | 0.19 | -0.29 | -0.31 |
| Cooperation | Individual | 0.32** | -0.01 | 0.04 | 0.00 |
| | Class | 0.66** | 0.34 | 0.14 | 0.45 |
| Equity | Individual | 0.38** | 0.05 | 0.11** | 0.14** |
| | Class | 0.73** | -0.21 | -0.11 | -0.75 |
| Multiple Correlation (*R*) | Individual | | 0.56** | | 0.15** |
| | Class | | 0.81** | | 0.54 |

*$p<0.05$ **$p<0.01$
*N*=927 students in 38 classes.

The bottom of Table 4.4 shows that the multiple correlation (*R*) between the set of seven learning environment scales and the attitude scale is 0.56 (individual student) and 0.81 (class mean), and it is statistically significant ($p<0.01$) with either unit of analysis. For achievement, the multiple correlations are 0.15 and 0.54 with the student and the class mean as the unit of analysis, respectively. However, the multiple

correlation is statistically significant ($p<0.01$) only with the individual student as the unit of analysis.

In order to interpret which individual WIHIC scales were responsible for explaining the significant multiple correlations found, the standardized regression weights reported in Table 4.4 were examined. The regression weight (β) describes the association between an outcome and a particular WIHIC scale while controlling for all other WIHIC scales. For the attitude outcome and with the student as the unit of analysis, Table 4.4 shows that Teacher Support, Involvement and Task Orientation were all statistically significant independent predictors of attitudes when the other WIHIC scales were mutually controlled. For the achievement outcome and with the student as the unit of analysis, Table 4.4 shows that only the WIHIC scale of Equity was a statistically significant and independent predictor of achievement.

It is worth mentioning that every statistically significant association between an environment scale and an outcome is positive in Table 4.4. These results replicate studies in numerous countries (Aldridge & Fraser, 2000; Fraser, 1998a; Khine & Fisher, 2001, 2002; Kim, Fisher, & Fraser, 2000; Riah & Fraser, 1998; Wong & Fraser, 1996) and generally suggest that a more positive learning environment on all WIHIC scales is linked to better student outcomes, especially their attitudes.

## 4.5 Summary of Analyses and Results

This chapter reported the analyses and results for a comparison of National Board Certified (NBC) and non-NBC science teachers in terms of their students' perceptions

of the science learning environment, attitudes toward science and science achievement at the secondary school level. As well, this chapter reported analyses and results for an investigation of environment-outcomes associations, and the validation of a learning environment questionnaire and a scale to measure attitudes toward science.

The What Is Happening In this Class? (WIHIC) questionnaire was pilot-tested, revised to ensure its suitability for American secondary science students, validated, and used to assess the students' perceptions of their science classroom learning environment. A scale modeled on the Test of Science-Related Attitudes (TOSRA) was created and later modified after pilot-testing it with a small group of students. The attitude scale was later validated and used to measure students' attitudes to science. The science Florida Comprehensive Assessment Test (FCAT) scores were collected to provide a measure of student achievement.

The data collected from 927 students (443 students in 21 classes taught by NBC teachers and 484 students in 17 classes taught by non-NBC teachers) were statistically analysed to answer the research questions of the present study, which were delineated in Section 1.5. At the beginning, the data were statistically analysed to determine the validity and reliability of the revised version of the WIHIC in terms of its factor structure, internal consistency reliability and ability to differentiate between classrooms using one-way ANOVA.

Factor analysis of the revised version of the WIHIC revealed that 54 of the 56 items had a factor loading of at least 0.30 on their a priori scale and no other scale. The

original seven-factor structure was replicated. The total proportion of variance accounted for was over 59% for the revised 54-item version of the WIHIC. Generally, alpha reliability coefficients exceeded 0.69 for the different WIHIC scales using either the individual or class mean as the unit of analysis. The one-way ANOVA results demonstrated that all seven WIHIC scales are able to differentiate significantly ($p<0.01$) between the perceptions of the students in the different classrooms.

Overall, the results supported the WIHIC's validity and reliability for assessing perceptions of the classroom environment among secondary science students in Miami-Dade County, Florida. The same data were analyzed to determine the internal consistency reliability of the modified attitude scale using two units of analysis (individual and class mean). The alpha reliability coefficient was 0.81 (individuals) and 0.93 (class means), suggesting that the modified attitude scale is reasonably reliable when used with secondary science students in Miami-Dade County, Florida.

Once the revised version of the WIHIC questionnaire was found to be valid and reliable and the modified attitude scale was found to be a reliable tool to use in the main study, the data gathered during the validation stage were statistically analyzed using MANOVA and effect sizes to investigate any differences between the effectiveness of National Board Certified (NBC) and non-NBC science teachers in terms of their students' perceptions of the science learning environment, attitudes toward science and science achievement. The results indicated a consistent direction of differences in favor of NBC teachers in terms of learning environment perceptions and attitudes. In particular, MANOVA showed a statistically significant difference

between the two groups for five of the seven learning environment scales (namely, Teacher Support, Involvement, Task Orientation, Investigation, and Cooperation) and the modified attitude scale. The effect sizes for the scales for which differences were statistically significant ranged from 0.14 to 0.29 standard deviations for the WIHIC scales and 0.35 standard deviation for the attitude scale. Overall, results suggest that the students taught by NBC teachers perceive a somewhat better learning environment and report having better attitudes than those taught by non-NBC teachers.

The data gathered during the validation of the WIHIC and attitude scale were once again analyzed using simple correlation and multiple regression analyses at two units of analysis (individual and class mean) to determine whether associations exist between students' perceptions of the learning environment and their outcomes (attitudes and achievement). A positive and statistically significant ($p<0.05$) correlation was found between student attitudes and all seven learning environment scales with the two units of analysis (individual and class mean). Also, it was found that Teacher Support, Involvement and Task Orientation are significant independent predictors of attitudes when using the student as the unit of analysis. However, a much stronger association between the learning environment and students' attitudes was found than for students' achievement. For instance, positive and statistically significant ($p<0.05$) correlations were found between student achievement and only the Involvement, Investigation and Equity scales when using the student as the unit of analysis. Furthermore, only Equity was found to be a statistically significant and independent predictor of achievement when using the student as the unit of analysis.

Chapter 5 provides a discussion of the findings. Additionally, it identifies the distinctive contributions made by the present study, possible limitations of the study and future directions for research.

# Chapter 5

# DISCUSSION AND CONCLUSION

## 5.1 Introduction

In an era of results-driven school reform, little information is available on the outcome of student achievement among National Board Certified (NBC) teachers. Furthermore, no research is available about either the student outcome of attitudes on students' perceptions of their learning environment among NBC teachers. The Federal government in the USA is taking steps to ensure highly-qualified teachers become nationally certified in order to fulfill the No Child Left Behind Act (2001) by providing millions of dollars to the National Board for Professional Teaching Standards (NBPTS). Therefore, the main goal of the NBPTS is to assist the Federal government in advancing the quality of teaching and learning by providing a national system of certification for teachers who meet high and rigorous standards based on what accomplished teachers should know and be able to do. But the question remains: Do National Board Certified teachers provide a higher quality of education to their students than non-National Board Certified teachers? My study sought to answer this question. I investigated the effectiveness of National Board Certified (NBC) science teachers in terms of secondary students' perceptions of the science classroom environment, attitudes toward science, and science achievement.

My study not only adds to the area of research that pertains to the effectiveness of NBC teachers, but it also contributes to the field of learning environments research for

numerous reasons. First, in my study, I selected, revised, and validated the widely-used learning environment questionnaire, What Is Happening In this Class? (WIHIC), with secondary students in Miami-Dade County, Florida. Second, this is the first time that a learning environments study has included a sample of National Board Certified (NBC) and non-NBC teachers in order to compare their effectiveness in terms of secondary students' perceptions of their science learning environment, attitudes toward science, and science achievement. Third, my study provides further information about outcomes-environment associations because I explored associations between the science classroom environment and secondary students' attitudes toward science and science achievement.

The purpose of this chapter is to provide an overview of the thesis in Section 5.2 by summarizing each of the chapters. Section 5.3 specifically provides a summary of the major findings of the study. Next, the significance of the study is discussed in Section 5.4. Section 5.5 discusses the constraints and limitations of my study. Last, Section 5.6 provides recommendations and suggestions for future research in two areas: the efficacy of National Board Certified teachers and classroom learning environments.

## 5.2 Overview of the Thesis

Chapter 1 introduced the National Board for Professional Teaching Standards (NBPTS) organization and its need to provide empirical evidence in favor of National Board Certification. Therefore, this study sought to compare the effectiveness of National Board Certified teachers and non-NBC teachers in terms of their students' perceptions of the science classroom environment, attitudes toward science, and

science achievement. Chapter 1 also included a brief discussion on the background of the field of learning environments and the rationale for my study. Section 1.4 established the aims of the study to be:

1. To provide information about the validity and reliability of a revised version of the What Is Happening In this Class? (WIHIC) questionnaire when used in secondary science classrooms in South Florida.
2. To provide information about the reliability of an attitude scale modeled on the Test Of Science-Related Attitude (TOSRA) when used in secondary science classrooms in South Florida.
3. To investigate whether NBC (National Board Certified) teachers are more effective than non-NBC teachers in terms of classroom environment, student attitudes and student achievement in secondary science classrooms in South Florida.
4. To investigate associations between student outcomes (attitudes and achievement) and classroom environment in secondary science classrooms in South Florida.

Chapter 2 reviewed literature associated with my study. Section 2.2 discussed the emergence and background of the National Board for Professional Teaching Standards (NBPTS). The NBPTS was created in response to *A Nation Prepared: Teachers for the 21st Century* (Carnegie Forum on Education and the Economy, 1986) and *A Nation at Risk: The Imperative for Educational Reform* (U.S. Department of Education, 1983). These landmark reports criticized the United State's Department of Education for the lack of highly-qualified educators who could better prepare

American students for the demands of the 21st century. Consequently, the NBPTS was created as a solution. The NBPTS is a nonprofit organization, which sets rigorous standards for what exemplary teachers should know and be able to do and to develop an assessment process to measure the teachers' ability to meet those standards. Teachers who go through this rigorous certification process and meet the standards set forth by the NBPTS become National Board Certified.

Section 2.3 described the positive and negative empirical evidence concerning National Board Certification. The earliest empirical evidence in regards to the effectiveness of National Board Certified (NBC) teachers was generated by the NBPTS themselves (National Board for Professional Teaching Standards, 2000, 2001). These early studies, which compared NBC to non-NBC teachers, showed favorable results for the effectiveness of the NBC teachers. Later, other independent studies were carried out to further investigate the effectiveness of NBC teachers (Cavalluzzo, 2004; Goldhaber & Anthony, 2004; Vandevoort, Amrein-Beardsley, & Berliner, 2004). The results of these studies indicated strong evidence that National Board Certification is an effective indicator of teacher effectiveness in terms of student achievement. However, other studies found contradictory evidence for the effectiveness of the National Board Certification process (Cunningham & Stone, 2005; Finn & Wilcox, 1999; Podgursky, 2001a, 2001b; Stone, 2002). Thus, in my study, I attempted to shed some light into this controversial issue by investigating whether NBC teachers are more effective than non-NBC teachers in terms of classroom environment, student attitudes, and student achievement in secondary science classrooms in South Florida.

Section 2.4 discussed the historical background of the field of learning environments, especially the work of Lewin (1936), Murray (1938), Stern, Stein, and Bloom (1956), Moos (1974), and Walberg (1968) who began research in the field of learning environments more than 30 years ago. Since then, learning environments research has progressed immensely thanks to the development of a variety of instruments, modeled after Moos's (1974) three basic dimensions for human environments classification. Therefore, Section 2.5 overviewed eight classroom environment instruments (Learning Environment Inventory, Classroom Environment Scale, Individualized Classroom Environment Questionnaire, My Class Inventory, College and University Classroom Environment Inventory, Questionnaire on Teacher Interaction, Science Laboratory Environment Inventory, and Constructivist Learning Environment Survey).

Section 2.6 reviewed literature about one other widely-used classroom environment instrument, the What Is Happening In this Class? (WIHIC). It was important to review literature about the WIHIC in a separate section and in much more detail because it was the learning environment instrument that I used in my study to assess the students' perceptions of their classroom learning environment. This section concentrated on discussing the characteristics and development of the WIHIC and reviewing past studies involving its use. The WIHIC was originally created by Fraser, Fisher, and McRobbie (1996) to bring parsimony to the field of learning environments by combining modified versions of existing environment scales with additional scales that address contemporary classroom dimensions. The WIHIC has been validated in various parts of the world, translated into different languages, and used in studies conducted at various grade levels (Adamski, Peiro, & Fraser, 2005; Aldridge &

Fraser, 2000; Castillo, Peiro, & Fraser, 2005; Kim, Fisher, & Fraser, 2000; MacDowell-Goggin, 2005; Margianti, Fraser, & Aldridge, 2001; Moss & Fraser, 2001; Raaflaub & Fraser, 2003; Soto-Rodriguez & Fraser, 2004).

Section 2.7 discussed the past lines of learning environments research with a specific focus on the two lines of research involved in my study, namely, research on determinants of classroom environment and research on outcome-environment associations. Finally, Section 2.8 reviewed literature about assessing attitudes to science and the characteristics and validity of the Test of Science-Related Attitudes (TOSRA) questionnaire. Providing literature about the TOSRA was important because, in my study, I created an attitude scale modeled on the TOSRA.

Chapter 3 thoroughly discussed the methodology used in the present study. Section 3.2 provided detailed information about the context of Miami-Dade County, Florida, which is the location where my study took place. In Section 3.3, I discussed in detail the background of the sample and explained how students were selected for the study. The students in this study came from 38 science classes to form a sample of 927 science students from 12 secondary schools. Altogether, 443 students in 21 classes comprised the NBC (National Board Certified) teacher group and 484 students in 17 classes comprised the non-NBC teacher group. The students were all enrolled in the third grading period of either Grade 8 or 10 science classes. The participating students were chosen based on the NBC and non-NBC teachers' willingness to make their classes available and on whether or not those teachers met certain criteria. First, the NBC teachers who were teaching Grade 8 and 10 science students were contacted and asked if they wanted to participate in the study. Second, the NBC teachers who agreed

to participate were asked to recruit a non-NBC teacher based on at least two of three criteria. The three criteria were that the non-NBC teacher had to be teaching the same science course at the same school as the NBC teacher, had been teaching for approximately the same amount of years as the NBC teacher, and/or was teaching students with similar characteristics as the students of the NBC teacher. The matching of the NBC and non-NBC teachers was important to increase the internal validity of my study.

Section 3.4 reviewed the instruments that I used in my study. The What Is Happening In this Class? (WIHIC) questionnaire was the instrument I chose to assess students' perceptions of their science classroom environment. The original 56-item WIHIC questionnaire (7 scales with 8 items in each scale) measures students' perceptions of seven dimensions: Student Cohesiveness, Teacher Support, Involvement, Investigation, Task Orientation, Cooperation, and Equity. In my study, I administered all seven scales of the WIHIC to a small group of students to ensure that they comprehended the items on the questionnaire. The results of the pilot study were favorable in that students were able to comprehend the items and answer in a short amount of time. However, minor revisions had to be made to the WIHIC.

To measure students' attitudes toward science, I created a 10-item attitude scale modeled on the original Test Of Science-Related Attitudes (TOSRA). The items were selected from two TOSRA scales (namely, the Enjoyment of Science Lessons and Adoption of Scientific Attitudes scales). I pilot-tested the attitude scale with a small group of students to check for comprehensibility. Because some of the negatively-

phrased items proved to be troublesome to most of the students, I modified the attitude scale by rewording several items.

The students' scores on the science portion of the Florida Comprehensive Assessment Test (FCAT) were used to measure science achievement. The science FCAT assesses achievement of science standards mandated by Florida's Department of Education. It combines multiple-choice items with gridded-response items from each of four clusters: physical and chemical science, earth space science, life and environmental science, and scientific thinking.

In Section 3.5, I explained the two research designs used to answer the two main research questions. Research Question #3 pertained to determining if National Board Certified (NBC) teachers are more effective than non-NBC teachers in terms of classroom environment, student attitudes and student achievement in secondary science classrooms in South Florida. To answer Research Question #3 a causal-comparative research design was used. On the other hand, to answer Research Question #4, a correlational research design was more appropriate. Research Question #4 asked if there are associations between student outcomes (attitudes and achievement) and classroom environment in secondary science classrooms in South Florida. Section 3.6 addressed the procedures followed in conducting my study and the statistical analyses I used to answer my research questions.

Chapter 4 presented the analyses and results of the quantitative data. The data gathered from the 927 secondary science students were statistically analyzed to provide information about the validity and reliability of the revised version of the

WIHIC and the modified attitude scale modeled on the TOSRA. First, to check the *a priori* seven-factor structure of the revised version of the WIHIC, principal axis factor analysis with oblique rotation was conducted. Next, to check whether each item in each scale of the revised version of the WIHIC and in the attitude scale assesses a similar construct, internal consistency reliability for two units of analysis (individual and class mean) was calculated for each scale. The Cronbach alpha reliability coefficient was used as an index of scale internal consistency. Additionally, discriminant validity (using the mean correlation of a scale with the other scales as a convenient index) was conducted to indicate to what extent each scale of the revised version of the WIHIC measured a distinct construct. Discriminant validity was conducted at two units of analysis (individual and class mean). Finally, to check if the revised version of the WIHIC was able to differentiate between the perceptions of students in the different classrooms, a one-way ANOVA for each WIHIC scale was conducted.

Once the results of these statistical analyses indicated that the revised version of the WIHIC and attitude scale were valid and reliable, the same data gathered from the 927 science students, along with these students' scores on the science FCAT, were statistically analyzed using multivariate analysis of variance (MANOVA) to determine the differences between the effectiveness of National Board Certified (NBC) and non-NBC teachers in terms of classroom environment, student attitudes and achievement. These sets of data were also statistically analyzed using simple correlation and multiple regression analyses for two units of analysis (individual and class mean) to determine if associations exist between secondary science students' outcomes (attitudes toward science and science achievement) and their perceptions of

the science classroom environment. The results of all of the statistical analyses mentioned in Chapter 4 are summarized in Section 5.3 below.

## 5.3 Major Findings of the Study

The major findings of the present study are discussed under the following headings:

- Findings for the Validity and Reliability of the Revised Version of the WIHIC and a Modified Attitude Scale (Section 5.3.1);
- Findings for the Differences between National Board Certified (NBC) and Non-NBC Teachers in Terms of Classroom Environment, Attitudes and Achievement (Section 5.3.2);
- Findings for Associations between Student Outcomes and Classroom Environment (Section 5.3.3).

### *5.3.1 Findings for the Validity of the Revised Version of the WIHIC and a Modified Attitude Scale*

The data gathered from 927 secondary science students in 38 science classes in 12 secondary schools were statistically analyzed to determine the validity and reliability of the revised version of the WIHIC in terms of its factor structure, internal consistency reliability, discriminant validity and ability to differentiate between classrooms using one-way ANOVA. In addition, in order to check the reliability of the single modified attitude scale modeled on the TOSRA, the Cronbach alpha coefficient was calculated to establish the internal consistency reliability using the individual student and class mean as the units of analysis.

The findings based on the results of these statistical analyses are reported below:

*Finding 1: The* a priori *factor structure for the revised version of the WIHIC was replicated for all items except only two of the original 56 items. All of the remaining 54 items had a factor loading of 0.30 or greater on their* a priori *scale and less than 0.30 on each of the other six scales. Therefore, the revised version of the WIHIC displayed strong factorial validity.*

*Finding 2: The internal consistency reliability estimate (Cronbach alpha coefficient) for each of the seven scales of the revised version of the WIHIC, using both the individual and the class mean as the unit of analysis, was above 0.69 for all scales. Thus, the revised version of the WIHIC displayed strong internal consistency reliability.*

*Finding 3: Some discriminant validity values (See Chapter 4, Table 4.2) are quite high for raw scores on the revised version of the WIHIC, thus indicating that some scales overlap in what they measure. However, the factor analysis results (see Chapter 4, Table 4.1) attest to the independence of factor scores on the revised version of the WIHIC.*

*Finding 4: All scales of the revised version of the WIHIC were able to differentiate significantly between the perceptions of students in different classes.*

*Finding 5: The modified attitude scale modeled on the TOSRA displayed reasonable internal consistency reliability.*

Overall, it was found that the revised version of the WIHIC and the modified attitude scale are valid and reliable instruments for assessing perceptions of the classroom environment and attitudes toward science among secondary science students in Miami-Dade County, Florida. These findings are consistent with past studies (Aldridge & Fraser, 2000; Dorman, 2003; Fraser, Fisher, & McRobbie, 1996; Margianti, Aldridge, & Fraser, 2004).

### *5.3.2 Findings for the Differences between National Board Certified (NBC) and Non-NBC Teachers in Terms of Classroom Environment, Attitudes and Achievement*

The data gathered from 927 secondary science students in 38 science classes in 12 secondary schools were statistically analyzed using MANOVA to determine whether differences exist between students in classes that are taught by National Board Certified (NBC) teachers and those taught by non-NBC teachers in terms of scores on each WIHIC scale, the modified attitude scale and science FCAT achievement. Because the MANOVA produced statistically significant results using Wilks' lambda criterion, the univariate ANOVA results were interpreted for each of the nine dependent variables. The findings based on these results are listed below:

*Finding 6: A statistically significant difference was found between the two groups (NBC and non-NBC teachers) for five of the seven learning environment scales (namely, Teacher Support, Involvement, Task Orientation, Investigation, and Cooperation) and the attitude scale.*

*Finding 7: A consistent direction was found in favor of NBC teachers in terms of learning environment perceptions and attitudes.*

*Finding 8: Differences between NBC and non-NBC teachers were nonsignificant for two WIHIC scales (namely, Student Cohesiveness and Equity) and for achievement.*

### *5.3.3 Findings for Associations between Student Outcomes and Classroom Environment*

The data gathered from 927 secondary science students in 38 science classes in 12 secondary schools were statistically analyzed using simple correlation and multiple regression analyses, calculated at two units of analysis (individual and class means), to determine whether associations exist between students' outcomes (attitudes toward science and science achievement) and their perceptions of the classroom learning environment). The findings based on the results of these statistical analyses are reported below:

*Finding 9: A positive and statistically significant correlation existed between students' attitudes toward science and all seven learning environment scales, with either the individual or the class mean as the unit of analysis.*

*Finding 10: For associations between students' science achievement and the learning environment, positive and statistically significant correlations occurred only for Involvement, Investigation and Equity for the student level of analysis. For the class mean as the unit of analysis, there were no significant correlations between student achievement and any of the learning environment scales.*

*Finding 11: The multiple correlation for the set of seven learning environment scales was statistically significant with either the individual student or the class mean as the units of analysis for the attitude scale, but only with the individual student as the unit of analysis for science achievement.*

*Finding 12: For the student as the unit of analysis, Teacher Support, Involvement and Task Orientation were all statistically significant independent predictors of students' attitudes toward science, but only Equity was a statistically significant and independent predictor of science achievement.*

*Finding 13: All statistically significant associations between an environment scale and an outcome were positive.*

Overall, it was found that a much stronger association with learning environment existed for students' attitude than for students' achievement. My findings replicate past research on outcome-environment associations (Adolphe, Fraser, & Aldridge, 2003; Allen, 2003; Dorman, 2003; Margianti et al., 2004).

## 5.4 Unique Contributions of the Present Study

The present research is significant to the field of learning environments because it is the only learning environments study to be carried out with U.S. National Board Certified teachers. Also, it is noteworthy because it provided further statistical validation of a learning environment questionnaire, the WIHIC, among secondary school students in South Florida. Furthermore, results of my study provide a baseline

for other researchers to investigate differences between National Board Certified (NBC) and non-NBC teachers in terms of students' perceptions of the classroom environment and affective/cognitive outcomes, as well as associations between the classroom environment and students' affective/cognitive outcomes in the U.S.

For the National Board for Professional Teaching Standards (NBPTS) organization, the present study is of great potential significance. First, it provides additional results supporting the effectiveness of National Board Certified (NBC) teachers in promoting positive classroom environments and attitudes to science, which supports the United States Department of Education's continued efforts in recognizing the importance of National Board Certification. Furthermore, because the NBPTS's objective is to engage teachers in reflecting on and improving their own teaching practices, our study could encourage NBPTS to make the WIHIC questionnaire available to teachers who are undergoing the National Board Certification process to help them to monitor and improve their own classroom environments. Also, our study could encourage the NBPTS and/or sponsors of the organization to conduct further research in establishing if NBC teachers are more effective than non-NBC teachers in terms of students' perceptions of the classroom learning environment, attitudes, and achievement for other student samples in Florida and/or other regions of the U.S. Finally, this study could help to support the NBPTS in providing empirical evidence of NBC teachers' impact on classrooms and students.

**5.4 5.5 Limitations of the Study**

Before the results from the present study can be interpreted, there are several limitations that need to be taken into account. First, the lack of a pretest-posttest design to answer the third research question listed below could be considered a limitation of the present study:

*Research Question #3*

*Are National Board Certified (NBC) teachers more effective than non-NBC teachers in terms of classroom environment, student attitudes and student achievement in secondary science classrooms in South Florida?*

Students of participating teachers were not pretested with the learning environment questionnaire and the attitude scale at the beginning of the study. Additionally, pretest science achievement scores were not gathered at the beginning of the study. Therefore, it wasn't possible to determine students' perceptions of the learning environment, attitudes toward science, and science achievement prior to the teachers becoming National Board Certified. Consequently, there is no way to evaluate if the NBC process improved their teaching methods in a way that would improve the science classroom environment and students' outcomes (attitudes toward science and science achievement). How do we know if the NBC teachers were already exceptional educators prior to the certification process, or if some improved after the certification process due to reflection of standards? Although I attempted to gather pretest and posttest scores in my study, I was not granted permission by the district to gather data twice during the school year. To address this limitation in my study, I used a control group (students taught by non-NBC teachers) for comparison with those in the experimental group (students taught by NBC teachers).

The students who did not have their permission forms filled out completely, did not wish to participate due to their own or parents' preference, or did not have a complete data for all dependent variables (i.e. WIHIC, TOSRA and FCAT scores) were eliminated from the study. Overall, approximately 20 subjects chose not to participate in the study. An additional 285 participants were dropped from the study due to incomplete data (mainly the absence of FCAT scores). Consequently, the sample size was reduced from 1230 to 927 students, thus reducing the statistical power of some analyses.

Another limitation involves the representativeness of the student sample that I used in my study. The generalizability of the findings from the present study could be limited by the uniqueness of the student sample involved. The student sample in my study consisted of science students in 12 high schools where more than 60% of the population is Hispanic and African-American and more than 70% of students are on free or reduced cost school lunch (an indication of low socioeconomic status). Therefore, the findings of my study should be applied with caution to other groups. I recommend that further research be conducted with a variety of other large samples.

In the case of my study, it would be difficult to determine if the difference on the three dependent variables (students' learning environment perceptions, attitudes towards science, and academic achievement in science) were a direct result of students having been taught by a National Board Certified teacher. For instance, other independent variables such as the type of school that students attend, type of science course in which students are enrolled, the number of years that teachers had been teaching and the characteristics of the students in the class could have been partial

causes of higher or lower scores on the dependent variables. Because the comparability of NBC and non-NBC teachers was a potential limitation in my study, I particularly controlled for some of those extraneous variables by matching each National Board Certified (NBC) teacher with a non-NBC teacher whose teaching experience was similar. For instance, after I selected the NBC teachers who were willing and able to participate in my study, I requested that each NBC teacher recruit one non-NBC teacher who met at least two of four criteria. First, the non-NBC teacher must have been teaching at the same school as the NBC teacher. Second, the non-NBC teacher must have been teaching the same science course (i.e. biology, physics) as the NBC teacher. Third, the non-NBC teacher must have been teaching for approximately the same number of years as the NBC teacher. Fourth, the non-NBC teacher must have been teaching a group of students with similar characteristics (i.e. science achievement levels, level of English proficiency) as the one taught by the NBC teacher. Although I controlled several extraneous variables by matching the NBC to non-NBC teachers, it can't be guaranteed that all extraneous variables were controlled for. Therefore, the results should be interpreted with caution. I recommend that future researchers attempt to control for more extraneous variables when comparing the effectiveness of NBC and non-NBC teachers.

## 5.6 Recommendations and Suggestions for Future Research

I recommend that future researchers implement a pretest-posttest design to help to create a better degree of accuracy in determining environment, attitudinal, and achievement scores prior to the teachers' pedagogical influence. Using a pretest to measure students' learning environment perceptions, attitudes, and achievement prior

to being taught by National Board Certified teachers could help the researcher determine if the students in the two groups (taught by National Board Certified teacher and taught by non-National Board Certified teacher) were initially comparable on the independent variables.

The sample size was small because the data came from a variety of classes and schools. Also, each class consisted of a different science discipline as well as two different grade levels ($8^{th}$ and $10^{th}$). To enhance future findings concerning this study, it is recommended that the study be conducted with one grade level and to concentrate on one science discipline in order to control as many variables as possible.

A very important limitation in my study was the possible lack of comparability between NBC and non-NBC teachers. As explained in Section 5.5, I addressed this limitation by matching the NBC and non-NBC teachers who participated in my study using certain criteria to ascertain that they were comparable to a certain extent. I recommend that future researchers use a pretest-posttest control group design. The experimental group would include the classes of the NBC teachers, and the control group would include the classes of the non-NBC teachers. Students in both the experimental and control groups would be pretested and posttested to determine which group, NBC or non-NBC teachers, is more effective in increasing students' attitude, classroom environment, and achievement scores.

I also recommend future researchers to use National Board Certified teachers who are certified in other disciplines such as reading and mathematics as samples in future studies. The effectiveness of these National Board Certified teachers could be

compared to that of non-NBC teachers in terms of classroom environment, attitudes, and achievement.

I also recommend that future studies include qualitative data-collection methods to enhance the quantitative findings. According to Tobin and Fraser (1998), adding qualitative research methods can help to provide a clearer picture of the learning environments, which is something that cannot be accomplished by just administering questionnaires. Future researchers should gather qualitative data in the form of teacher and student interviews after administering the questionnaires. During interviews, the researcher could ask questions related to the items presented in the questionnaires. Future researchers could also conduct observations of NBC and non-NBC teachers to gather a more personal perspective of the classroom environment created by each group of teachers. Finally, in future studies, the researcher could ask the NBC and non-NBC teachers to keep a reflective journal (Leitch, 1993) about their daily science lessons to gain a deeper understanding of their teaching practices, pragmatic decisions, student-teacher interactions, etc. in order to make a more conclusive comparison between NBC and non-NBC teachers.

## 5.7 Concluding Comments

This is the first study conducted to investigate the effectiveness of National Board Certified teachers in terms of secondary students' outcomes (attitudes toward science and science achievement) and their perceptions of their science classroom environment in Miami-Dade County, Florida. Additionally, it is the first in providing comprehensive validation data for the revised version of the WIHIC and a modified

attitude scale modeled on the TOSRA among students of National Board Certified science teachers in Miami-Dade County, Florida.

The findings of my study have implications for science teachers who are interested in improving the learning environments for their students. In addition, the study suggests that National Board Certified (NBC) teachers could be more effective in developing a more positive and productive learning environment for their students when compared with non-NBC teachers. Conclusions of my study are designed to be distinctively applicable to the science learning environment and, by and large, to the learning environments for other disciplines. The findings from this study add to the accumulation of educational data concerning the impact of National Teacher Certification and classroom learning environments research.

Perhaps, as suggested by Fisher and Fraser (1992), teacher training programs might benefit from incorporating and modeling learning environments research methodologies as a means of ensuring that teachers become aware of the alternatives that exist in evaluating and modifying the classroom environment for the benefit of students.

# REFERENCES

Adamski, A., Peiro, M.M., & Fraser, B.J. (2005, August). *Relationships between parental involvement in schooling, classroom environment, and students' attitudes and achievement*. Paper presented at the Fourth International Conference on Science, Mathematics, and Technology Education, Victoria, BC, Canada.

Adolphe, G.F., Fraser, B.J., & Aldridge, J.M. (2003, January). *A cross-national study of classroom environment and attitudes among junior secondary science students in Australia and Indonesia.* Paper presented at the Third International Science, Mathematics and Technology Education Conference, London, South Africa.

Aldridge, J.M., & Fraser, B.J. (2000). A cross-national study of classroom environments in Taiwan and Australia. *Learning Environments Research: An International Journal, 3,* 101-134.

Aldridge, J.M., Fraser, B.J., & Huang, I.T.C. (1999). Investigating classroom environments in Taiwan and Australia with multiple research methods. *Journal of Educational Research, 93,* 48-62.

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.C. (2000). Constructivist learning environments in a cross-national study in Taiwan and Australia. *International Journal of Science Education, 22,* 37-55.

Allen, D. (2003). *Parent and student perceptions of the science learning environment and its influence on student outcomes.* Unpublished doctoral thesis, Curtin University of Technology.

Anderson, G. (Ed.). (1998). *Fundamentals of educational research* (2nd ed.). Bristol, PA: Falmer Press.

Ball, D.L., & Cohen, D.K. (1999). *Instruction, capacity, and improvement*. (CPRE Research Report No. RR-43). Philadelphia, PA: University of Pennsylvania, Consortium for Policy Research in Education. Retrieved July 30, 2003, from http://www.cpre.org/Publications/rr43.pdf

Carnegie Forum on Education and the Economy. (1986). *A nation prepared: Teachers for the 21st century*. New York: Author.

Castillo, G.E., Peiro, M.M., & Fraser, B.J. (2005, August). *Grade-level, gender and ethnic differences in attitudes and learning environment in high school mathematics*. Paper presented at the Fourth International Conference on Science, Mathematics, and Technology Education, Victoria, BC, Canada.

Cavalluzzo, L.C. (2004, November). *Is National Board Certification an effective signal of teacher quality?* Paper presented at the Consortium for Policy Research in Education's National Conference on Teacher Compensation and Evaluation, Chicago, IL.

Chionh, Y.H., & Fraser, B.J. (1998, April). *Validation and use of the 'What Is Happening In this Class' (WIHIC) questionnaire in Singapore*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Cunningham, G.K., & Stone, J.E. (2005). Value-added assessment of teacher quality as an alternative to the National Board for Professional Teaching Standards:

What recent studies say. In Robert Lissitz (Ed.), *Value added models in education: Theory and applications*. Maple Grove, MN: JAM Press.

Denzin, N.K. (1978). *The research act: A theoretical introduction to sociological methods*. New York: MacGraw-Hill.

Dorman, J.P. (2002). Classroom environment research: Progress and possibilities. *Queensland Journal of Educational Research, 18* (12), 112-140.

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.

Dryden, M., & Fraser, B.J. (1998, April). *The impact of systemic reform efforts in promoting constructivist approaches in high school science.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA,

Finn, C.E., & Wilcox, D.D. (1999). *Board games: Failure of the National Board for Professional Teaching Standards to accomplish objective of improving quality of teaching in the US: Business backs a losing strategy.* Retrieved January 1, 2006 from http://www.edexcellence.net/foundation/publication/publication.cfm?id=161

Finn, C.E., & Wilcox, D.D. (2000). *Teachers should be graded on how well their students are learning.* Retrieved January 1, 2006 from http://www.edexcellence.net/foundation/publication/publication.cfm?id=102

Fisher, D.L., & Fraser, B.J. (1981). Validity and use of My Class Inventory. *Science Education, 65,* 145-156.

Fisher, D.L., & Fraser, B.J. (1983). A comparison of actual and preferred classroom environments as perceived by science teachers and students. *Journal of Research in Science Teaching, 20,* 55-61.

Fisher, D.L., & Fraser, B.J. (1992). Validity and use of school environment instruments. *Journal of Classroom Interaction, 26*(2), 13-18.

Fisher, D.L., Fraser, B.J., & Bassett, J. (1995). Using a classroom environment instrument in an early childhood classroom. *Australian Journal of Early Childhood, 20*(3), 10-15.

Fisher, D.L., Fraser, B.J., & Rickards, T. (1997, April). *Gender and cultural differences in teacher-student interpersonal behavior*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Fisher, D.L., Henderson, D., & Fraser, B.J., (1995). Interpersonal behaviour in senior high school biology classes. *Research in Science Education, 25,* 125-133.

Florida Department of Education. (2004). *FCAT Florida Comprehensive Assessment Test.* Retrieved July 1, 2005, from http://www.fldoe.org/default.asp?bhcp=1

Fraser, B.J. (1981). *Test of Science-Related Attitudes*. Melbourne, Australia: The Australian Council for Educational Research.

Fraser, B.J. (1986). *Classroom environment.* London: Croom Helm.

Fraser, B.J. (1989). *Assessing and improving classroom environment* (What Research Says, No. 2). Perth, Australia: Curtin University of Technology.

Fraser, B.J. (1991). Two decades of classroom learning environment research. In B.J. Fraser & H.J. Walberg (Eds.), *Educational environments: Evaluation, antecedents and consequences* (pp. 3-27). London, England: Pergamon Press.

Fraser, B.J. (1994). Research on classroom and school climate. In D. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 493-541). New York: Macmillan.

Fraser, B.J. (1996, March). *NARST's expansion, internationalization, and cross nationalization: History in the making.* Presidential address at the annual meeting of the National Association for Research in Science Teaching, St Louis, MO.

Fraser, B.J. (1998a). Classroom environment instruments: Development, validity and applications. *Learning Environments Research: An International Journal*, *1*, 7-33.

Fraser, B.J. (1998b). Science learning environments: Assessment, effects, and determinants. In B.J. Fraser & K.G. Tobin (Eds.), *International handbook of science education* (pp. 527-564). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Fraser, B.J. (1999). "Grain sizes" in learning environment research: Combining qualitative and quantitative methods. In H.C. Waxman & H.J. Walberg (Eds.), *New directions for teaching practice and research* (pp. 285-296). Berkeley, CA: McCutchan.

Fraser, B.J. (2001). Twenty thousand hours: Editor's introduction. *Learning Environments Research: An International Journal*, *4*, 1-5.

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).* Perth, Australia: Western Australian Institute of Technology.

Fraser, B.J., & Butts, W.L. (1982). Relationship between perceived levels of classroom individualization and science-related attitudes. *Journal of Research in Science Teaching, 19*, 143-154.

Fraser, B.J., & Chionh, Y.H. (2000, April). *Classroom environment, self-esteem, achievement, and attitudes in geography and mathematics in Singapore.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Fraser, B.J., Dryden, M., & Taylor, P. (1998, April). *The impact of systemic reform efforts on instruction in high school science classes*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Diego, CA.

Fraser, B.J., & Fisher, D.L. (1982). Predicting students' outcomes from their perceptions of classroom psychosocial environment. *American Educational Research Journal, 19*, 498-518.

Fraser, B.J., Fisher, D.L., & McRobbie, C.J. (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., 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., & McRobbie, C.J. (1995). Science laboratory classroom environments at school and universities: A cross-national study. *Educational Research and Evaluation, 1*, 289-317.

Fraser, B.J., & O'Brien, P. (1985). Student and teacher perceptions of the environment of elementary-school classrooms. *Elementary School Journal, 85*, 567-580.

Fraser, B.J., & Treagust, D.F. (1986). Validity and use of an instrument for assessing classroom psychosocial environment in higher education. *Higher Education, 15*, 37-57.

Fraser, B.J., & Walberg, H.J. (1991). *Educational environments: Evaluation, antecedents and consequences*. Oxford, England: Pergamon Press.

Gay, L.R., & Airasian, P. (Eds.). (1996). *Educational research: Competencies for analysis and application* (6th ed.). Upper Saddle River, NJ: Merrill Publishing Company.

Goh, S.C., & Fraser, B.J. (1998). Teacher interpersonal behaviour, classroom environment and student outcomes in primary mathematics in Singapore. *Learning Environment Research, 1*, 199-229.

Goh, S.C., & Khine, M.S. (Eds.). (2002). *Studies in educational learning environments: An international perspective*. Singapore: World Scientific Publishers.

Goldhaber, D., & Anthony, E. (2004). *National Board Certification: Who applies and what factors are associated with success?* The Urban Institute, Education Policy Center. [On-line], Available: http://www.urban.org/EmilyAnthony

Haertel, G.D., Walberg, H.J., & Haertel, E.H. (1981). Socio-psychological environments and learning: A quantitative synthesis. *British Educational Research Journal, 7*, 27-36.

Henderson, D.G., Fisher, D.L., & Fraser, B.J. (1998, April). *Learning environment, student attitudes and effects of students' sex and other science study in environmental science classes.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Henderson, D.G., Fisher, D.L., & Fraser, B.J. (2000). Interpersonal behavior, learning environments and student outcomes in senior biology classes. *Journal of Research in Science Teaching, 37,* 26-43.

Hirata, S., & Sako, T. (1998). Perceptions of school environment among Japanese junior high school, non-attendant, and juvenile delinquent students. *Learning Environments Research, 1,* 321-331.

Huang, T.-C.I., Aldridge, J.M., & Fraser, B.J. (1998) A collaborative study of science classroom environments between Taiwan and Australia: Combining qualitative and quantitative research methods. *Chinese Journal Science Education, 6,* 343-362.

Huang, T.-C.I., & Fraser, B.J. (1997, March). *The development of a questionnaire for assessing student perceptions of classroom climate in Taiwan and Australia.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Chicago, IL.

Keller, B. (2005). Panel urges new testing for teachers. *Education Week, 24*(38), 1-16.

Khine, M.S., & Fisher, D.L. (2001, December). *Classroom environment and teachers' cultural background in secondary science classes in an Asian context.* Paper presented at the annual meeting of the Australian Association for Research in Education, Perth, Australia.

Khine, M.S., & Fisher, D.L. (2002, April). *Analysing interpersonal behaviour in science classrooms: Associations between students' perceptions and teachers' cultural background.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, New Orleans, LA.

Khoo, H.S., & Fraser, B.J. (1998, April). *Using classroom environment dimensions in the evaluation of adult computer courses.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Kim, H., Fisher, D.L., & Fraser, B.J. (1999). Assessment and investigation of constructivist science learning environments in Korea. *Research in Science & Technological Education, 17*, 239-249.

Kim, H.B., Fisher, D.L., & Fraser, B.J. (2000). Classroom environment and teacher interpersonal behaviour in secondary school 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, & G.F. Madaus (Eds.). *Handbook on summative and formative evaluation of student learning* (pp. 559-641). New York: McGraw-Hill.

Koul, R.B., & Fisher, D.L. (May 2005) Cultural background and students' Perceptions of science classroom learning environment and teacher interpersonal behaviour in Jammu, India. *Learning Environments Research, 8*, 195-211**.**

Laforgia, J. (1988). The affective domain related to science education and its evaluation. *Science Education, 72*, 407-421.

Lee, S.S.U. (2001). *Assessment, description and effects of science classroom environments in Korea.* Unpublished doctoral dissertation, Curtin University of Technology, Perth, Australia.

Lee, S.S.U., & Fraser, B.J. (2003). Teacher-student interactions in Korean high school science classrooms. *International Journal of Science and Mathematics Education, 1*, 67-85.

Leitch, R. (1993, July) *The Implications of Intra-psychic Reflection on the Professional and Moral Development of Teacher*. Paper presented to University of Exeter Conference on Moral and Spiritual Development, RIMSCUE Centre, Exeter, United Kingdom.

Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw Hill.

Lightburn, M.E., & Fraser, B.J. (2002, April). *Classroom environment and student outcomes associated with using anthropometry activities in high school science.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Lucas, K.B., & Tulip, D.F. (1980, September). *Scientific literacy of high school students.* Paper presented at the annual conference of the Australian Science Teachers Association, Canberra.

MacDowell-Goggin, M. (2005). *Effect of using graphic organizers on the attitudes and classroom environment of elementary science students*. Unpublished doctoral thesis, Curtin University of Technology.

Majeed, A., Fraser, B.J., & Aldridge, J.M. (2002). Learning environment and its associations with student satisfaction among mathematics students in Brunei Darussalam. *Learning Environments Research, 5*, 203-226.

Maor, D., & Fraser, B.J. (1996). Use of classroom environment perceptions in evaluating inquiry-based computer assisted learning. *International Journal of Science Education, 18*, 401-421.

Margianti, E.S., Fraser, B.J., & Aldridge, J.M. (2001, December). *Investigating the learning environment and students' outcomes in university computing courses in Indonesia.* Paper presented at the annual conference of the Australian Association for Research in Education, Fremantle, Australia.

Margianti, E.S., Fraser, B.J., & Aldridge, J. (2002, April). *Learning environment, attitudes and achievement: Assessing the perceptions of Indonesian university students.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Margianti, E.S., Aldridge, J.M., & Fraser, B.J. (2004). Learning environment perceptions, attitudes and achievement among private Indonesian university students. *International Journal of Private Higher Education.* Retrieved January 18, 2005, from http://www.xaiu.com/xaiujournal/

McRobbie, C.J., & Fraser, B.J., (1993). Associations between student outcomes and psychosocial science environments. *Journal of Educational Research, 87*, 78-85.

Mathison, S. (1988). Why triangulate? *Educational Researcher, 17*, 13-17.

Miami-Dade County Public Schools. (2003). *Miami-Dade County Public Schools student progression plan 2003-2004.* Retrieved July 1, 2005, from http://www.dadeschools.net/ehandbook/student03/Section%20I/2003%20General%20Procedures%20for%20Student%20Progression,%20K%20-%201.pdf

Miami-Dade County Public Schools. (2004). *District demographic*. Retrieved July 1, 2005, from Miami-Dade County Public Schools Web site: http://oada.dadeschools.net/04Profiles.htm

Mink, D.V., & Fraser, B.J. (2005). Evaluation of a K-5 mathematics program which integrates children's literature: Classroom environment and attitudes. *International Journal of Science and Mathematics Education, 3*, 59-85.

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. (1987). *Classroom Environment Scale manual* (2nd ed). Palo Alto, CA: Consulting Psychologists Press.

Moss, C.H., & Fraser, B.J. (2001, April). *Using environment assessments in improving teaching and learning in high school biology classrooms.* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Murray, H.A. (1938). *Explorations in personality*. New York: Oxford University Press.

National Board for Professional Teaching Standards. (2000). *A distinction that matters: Why national teacher certification makes a difference.* Retrieved February 8, 2003, from http://www.mprinc.com/pubs/pdf/distinction_that_matters.pdf

National Board for Professional Teaching Standards. (2001). *I am a better teacher: What candidates for National Board Certification say about the assessment process.* Retrieved February 8, 2003, from http://www.nbpts.org/pdf/better_teacher.pdf

Nix, R., Fraser, B.J., & Ledbetter, C. (2005, May). Evaluating an Integrated Science Learning Environment Using the Constructivist Learning Environment Survey. *Learning Environments Research: An International Journal, 8*, 109-133.

No Child Left Behind Act. (2001). Retrieved October 15, 2004, from U.S. Department of Education Web site: http://www.ed.gov/nclb/landing.jhtml

Pickett, L.H., & Fraser, B.J. (2002, April). *The role of learning environment, achievement, and student and teacher attitudes in a science mentoring program for beginning elementary school teachers*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Pickett, L.H., & Fraser, B.J. (2004, April). *An evaluation of a science mentoring program for beginning elementary school teachers in terms of learning environment, student achievement and attitudes and teacher attitudes*. Paper presented at the annual meeting of the National Association of Research in Science Teaching, Vancouver, BC, Canada.

Podgursky, M. (2001a). Defrocking the National Board. *Education Next, 1* (2), 79-82.

Podgursky, M. (2001b). Should state subsidize National Certification? Retrieved February 11, 2003, from www.edweek.org

Punch, K.F. (1998). *Introduction to social research*. London, England: Sage Publications.

Quek, C.L., Wong, A.F.L., & Fraser, B. (2005a). 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*, 299-321.

Quek, C.L., Wong, A.F.L., & Fraser, B. (2005b). Teacher-student interaction and gifted students' attitudes toward chemistry in laboratory classrooms in Singapore. *Journal of Classroom Interaction, 40*(1), 18-28.

Raaflaub, C., & Fraser, B.J. (2003, March). *The learning environment associated with the use of laptop computers in Canadian science classes.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Philadelphia, PA.

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

Riah, H., & Fraser, B.J. (1998, April). *Chemistry learning environment and its association with students' achievement in chemistry*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Robinson, E. (2003). *Kindergarten students' and their parents' perceptions of science environments: Achievement and attitudes.* Unpublished doctoral thesis, Curtin University of Technology. Retrieved December 27, 2005, from Curtin University of Technology Library and Information Services Web site: http://adt.curtin.edu.au/theses/available/adt-WCU20031117.155353/

Schibeci, R.A. (1984). Attitudes to science: An update. *Studies in Science Education, 11*, 26-59.

Schibeci, R.A., & McGaw, B. (1981). Empirical validation of the conceptual structure of a test of science-related attitudes. *Educational and Pyschological Measurement, 41,* 1195-1201.

Scott, R.H., & Fisher, D.L. (2004). Development, validation and application of a Malay translation of an elementary version of the Questionnaire on Teacher Interaction. *Research in Science Education. 34,* 173-194.

Shulman, L.S. (1997). Disciplines of inquiry in education: A new overview. In R.M. Jaeger (Ed.), *Complementary methods for research in education* (2nd ed., pp. 3-29). Washington, DC: American Educational Research Association.

Sinclair, B.B., & Fraser, B.J. (2003). Changing classroom environments in urban middle schools. *Learning Environments Research: An International Journal, 5,* 201-328.

Soerjaningsih, W., Fraser, B.J., & Aldridge, J.M. (2001, April). *Achievement, satisfaction and learning environment among Indonesian computing students at the university level.* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Soto-Rodriguez, M., & Fraser, B.J. (2004, April). *A comparison of attitudes, achievement, and classroom environment perceptions of LEP (Limited English Proficient) and non-LEP students in integrated science classrooms*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Spinner, H., & Fraser, B.J. (2005). Evaluation of an innovative mathematics program in terms of classroom environment, student attitudes, and conceptual development. *International Journal of Science and Mathematics Education, 3,* 267-293.

Stern, G.G., Stein, M.I., & Bloom, B.S. (1956). *Methods in personality assessment.* Glencoe, IL: Free Press.

Stone, J.E. (2002, May). *Value-added achievement gains of NBPTS-certified teachers in Tennessee: A brief report.* Retrieved July 21, 2005, from www.education-consumers.com/briefs/may2002.asp

Taylor, P.C., Dawson, V., & Fraser, B.J. (1995, April). *Classroom learning environments under transformation: A constructivist perspective.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

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, San Francisco, CA.

Teh, G., & Fraser, B. (1994). An evaluation of computer-assisted learning in terms of achievement, attitudes and classroom environment. *Evaluation and Research in Education, 8*, 147-161.

Thompson, B. (1998a). Review of 'What if there were no significance tests?' *Educational and Psychological Measurement, 58*, 334-346.

Thompson, B. (1998b, April). *Five methodology errors in educational research: The pantheon of statistical significance and other faux pas*. Invited address presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Tobin, K., & Fraser, B.J. (1998). Qualitative and quantitative landscapes of classroom learning environments. In B.J. Fraser & K.G. Tobin (Eds.),

*International handbook of science education* (pp. 623-640). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Trickett, E.J., & Moos, R.H. (1973). Social environment of junior high and high school classrooms. *Journal of Educational Psychology, 65*, 93-102.

U.S. Census Bureau. (2000). *Florida quick facts*. Retrieved June 25, 2005, from http://quickfacts.census.gov/qfd/states/12000.html

U.S. Department of Education. (1983). *A nation at risk: The imperative for education reform* (Publication No. 065-000-00177-2). Washington, DC: Author. Retrieved August 20, 2004, from http://www.ed.gov/pubs/NatAtRisk/risk.html

Vandevoort, L.G., Amrein-Beardsley, A., & Berliner, D.C. (2004). National Board Certified teachers and their students' achievement. *Education Policy Analysis Archives*. Retrieved January 20, 2005, from http://epaa.asu.edu/epaa/v12n46/

Walberg, H.J. (1968). Teacher personality and classroom climate. *Psychology in the Schools, 5*, 163-169.

Walberg, H.J. (1981). A psychological theory of educational productivity. In F. Farley & N. Gordon (Eds.), *Psychology and education: The state of the union* (pp. 81-108). Berkley, CA: McCutchan.

Walberg, H.J. (1984**).** Improving the productivity of America's schools. *Educational Leadership*, *41*, 19-27.

Walberg, H.J. (1986). Synthesis of research on teaching. In M.C. Wittrock (Eds.), *Handbook of research on teaching* (3rd ed., pp. 214-229). Washington, DC: American Educational Research Association.

Walberg, H.J., & Anderson, G.J. (1968). Classroom climate and individual learning. *Journal of Educational Psychology*, *59*, 414-419.

Walberg, H.J., Fraser, B.J., & Welch, W.W. (1986). A test of a model of educational productivity among senior high school students. *Journal of Educational Research, 39,* 133-139.

Wilcox, D.D. (2003). The National Board for Professional Teaching Standards: Can it live up to its promise? Retrieved February 11, 2003, from http://www.edexcellence.net/better/tchrs/13.htm

Wong, A.F.L., & Fraser, B.J., (1996). Environment-attitude associations in the chemistry laboratory classroom. *Research in Science and Technological Education, 14,* 91-102.

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

Yarrow, A., Millwater, J., & Fraser, B.J. (1997). Improving university and primary school classroom environments through pre-service teachers' action research. *International Journal of Practical Experiences in Professional Education, 1* (1), 68-93.

Zandvliet, D.B., & Fraser, B.J. (2004). Learning environments in information and communication technology classrooms. *Technology, Pedagogy and Education, 13* (1), 97-123.

Zandvliet, D.B., & Fraser, B.J. (2005). Physical and psychosocial environments associated with networked classrooms. *Learning Environments Research, 8,* 1-17.

# REFERENCES

Adamski, A., Peiro, M.M., & Fraser, B.J. (2005, August). *Relationships between parental involvement in schooling, classroom environment, and students' attitudes and achievement*. Paper presented at the Fourth International Conference on Science, Mathematics, and Technology Education, Victoria, BC, Canada.

Adolphe, G.F., Fraser, B.J., & Aldridge, J.M. (2003, January). *A cross-national study of classroom environment and attitudes among junior secondary science students in Australia and Indonesia.* Paper presented at the Third International Science, Mathematics and Technology Education Conference, London, South Africa.

Aldridge, J.M., & Fraser, B.J. (2000). A cross-national study of classroom environments in Taiwan and Australia. *Learning Environments Research: An International Journal, 3,* 101-134.

Aldridge, J.M., Fraser, B.J., & Huang, I.T.C. (1999). Investigating classroom environments in Taiwan and Australia with multiple research methods. *Journal of Educational Research, 93,* 48-62.

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.C. (2000). Constructivist learning environments in a cross-national study in Taiwan and Australia. *International Journal of Science Education, 22,* 37-55.

Allen, D. (2003). *Parent and student perceptions of the science learning environment and its influence on student outcomes.* Unpublished doctoral thesis, Curtin University of Technology.

Anderson, G. (Ed.). (1998). *Fundamentals of educational research* (2nd ed.). Bristol, PA: Falmer Press.

Ball, D.L., & Cohen, D.K. (1999). *Instruction, capacity, and improvement*. (CPRE Research Report No. RR-43). Philadelphia, PA: University of Pennsylvania, Consortium for Policy Research in Education. Retrieved July 30, 2003, from http://www.cpre.org/Publications/rr43.pdf

Carnegie Forum on Education and the Economy. (1986). *A nation prepared: Teachers for the 21st century*. New York: Author.

Castillo, G.E., Peiro, M.M., & Fraser, B.J. (2005, August). *Grade-level, gender and ethnic differences in attitudes and learning environment in high school mathematics*. Paper presented at the Fourth International Conference on Science, Mathematics, and Technology Education, Victoria, BC, Canada.

Cavalluzzo, L.C. (2004, November). *Is National Board Certification an effective signal of teacher quality?* Paper presented at the Consortium for Policy Research in Education's National Conference on Teacher Compensation and Evaluation, Chicago, IL.

Chionh, Y.H., & Fraser, B.J. (1998, April). *Validation and use of the 'What Is Happening In this Class' (WIHIC) questionnaire in Singapore*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Cunningham, G.K., & Stone, J.E. (2005). Value-added assessment of teacher quality as an alternative to the National Board for Professional Teaching Standards:

What recent studies say. In Robert Lissitz (Ed.), *Value added models in education: Theory and applications*. Maple Grove, MN: JAM Press.

Denzin, N.K. (1978). *The research act: A theoretical introduction to sociological methods*. New York: MacGraw-Hill.

Dorman, J.P. (2002). Classroom environment research: Progress and possibilities. *Queensland Journal of Educational Research, 18* (12), 112-140.

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.

Dryden, M., & Fraser, B.J. (1998, April). *The impact of systemic reform efforts in promoting constructivist approaches in high school science.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA,

Finn, C.E., & Wilcox, D.D. (1999). *Board games: Failure of the National Board for Professional Teaching Standards to accomplish objective of improving quality of teaching in the US: Business backs a losing strategy.* Retrieved January 1, 2006 from http://www.edexcellence.net/foundation/publication/publication.cfm?id=161

Finn, C.E., & Wilcox, D.D. (2000). *Teachers should be graded on how well their students are learning.* Retrieved January 1, 2006 from http://www.edexcellence.net/foundation/publication/publication.cfm?id=102

Fisher, D.L., & Fraser, B.J. (1981). Validity and use of My Class Inventory. *Science Education, 65,* 145-156.

Fisher, D.L., & Fraser, B.J. (1983). A comparison of actual and preferred classroom environments as perceived by science teachers and students. *Journal of Research in Science Teaching, 20,* 55-61.

Fisher, D.L., & Fraser, B.J. (1992). Validity and use of school environment instruments. *Journal of Classroom Interaction, 26*(2), 13-18.

Fisher, D.L., Fraser, B.J., & Bassett, J. (1995). Using a classroom environment instrument in an early childhood classroom. *Australian Journal of Early Childhood, 20*(3), 10-15.

Fisher, D.L., Fraser, B.J., & Rickards, T. (1997, April). *Gender and cultural differences in teacher-student interpersonal behavior*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Fisher, D.L., Henderson, D., & Fraser, B.J., (1995). Interpersonal behaviour in senior high school biology classes. *Research in Science Education, 25,* 125-133.

Florida Department of Education. (2004). *FCAT Florida Comprehensive Assessment Test.* Retrieved July 1, 2005, from http://www.fldoe.org/default.asp?bhcp=1

Fraser, B.J. (1981). *Test of Science-Related Attitudes*. Melbourne, Australia: The Australian Council for Educational Research.

Fraser, B.J. (1986). *Classroom environment.* London: Croom Helm.

Fraser, B.J. (1989). *Assessing and improving classroom environment* (What Research Says, No. 2). Perth, Australia: Curtin University of Technology.

Fraser, B.J. (1991). Two decades of classroom learning environment research. In B.J. Fraser & H.J. Walberg (Eds.), *Educational environments: Evaluation, antecedents and consequences* (pp. 3-27). London, England: Pergamon Press.

Fraser, B.J. (1994). Research on classroom and school climate. In D. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 493-541). New York: Macmillan.

Fraser, B.J. (1996, March). *NARST's expansion, internationalization, and cross nationalization: History in the making.* Presidential address at the annual meeting of the National Association for Research in Science Teaching, St Louis, MO.

Fraser, B.J. (1998a). Classroom environment instruments: Development, validity and applications. *Learning Environments Research: An International Journal*, *1*, 7-33.

Fraser, B.J. (1998b). Science learning environments: Assessment, effects, and determinants. In B.J. Fraser & K.G. Tobin (Eds.), *International handbook of science education* (pp. 527-564). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Fraser, B.J. (1999). "Grain sizes" in learning environment research: Combining qualitative and quantitative methods. In H.C. Waxman & H.J. Walberg (Eds.), *New directions for teaching practice and research* (pp. 285-296). Berkeley, CA: McCutchan.

Fraser, B.J. (2001). Twenty thousand hours: Editor's introduction. *Learning Environments Research: An International Journal*, *4*, 1-5.

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).* Perth, Australia: Western Australian Institute of Technology.

Fraser, B.J., & Butts, W.L. (1982). Relationship between perceived levels of classroom individualization and science-related attitudes. *Journal of Research in Science Teaching, 19*, 143-154.

Fraser, B.J., & Chionh, Y.H. (2000, April). *Classroom environment, self-esteem, achievement, and attitudes in geography and mathematics in Singapore.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Fraser, B.J., Dryden, M., & Taylor, P. (1998, April). *The impact of systemic reform efforts on instruction in high school science classes*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Diego, CA.

Fraser, B.J., & Fisher, D.L. (1982). Predicting students' outcomes from their perceptions of classroom psychosocial environment. *American Educational Research Journal, 19*, 498-518.

Fraser, B.J., Fisher, D.L., & McRobbie, C.J. (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., 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., & McRobbie, C.J. (1995). Science laboratory classroom environments at school and universities: A cross-national study. *Educational Research and Evaluation, 1*, 289-317.

Fraser, B.J., & O'Brien, P. (1985). Student and teacher perceptions of the environment of elementary-school classrooms. *Elementary School Journal, 85*, 567-580.

Fraser, B.J., & Treagust, D.F. (1986). Validity and use of an instrument for assessing classroom psychosocial environment in higher education. *Higher Education, 15*, 37-57.

Fraser, B.J., & Walberg, H.J. (1991). *Educational environments: Evaluation, antecedents and consequences*. Oxford, England: Pergamon Press.

Gay, L.R., & Airasian, P. (Eds.). (1996). *Educational research: Competencies for analysis and application* (6th ed.). Upper Saddle River, NJ: Merrill Publishing Company.

Goh, S.C., & Fraser, B.J. (1998). Teacher interpersonal behaviour, classroom environment and student outcomes in primary mathematics in Singapore. *Learning Environment Research, 1*, 199-229.

Goh, S.C., & Khine, M.S. (Eds.). (2002). *Studies in educational learning environments: An international perspective*. Singapore: World Scientific Publishers.

Goldhaber, D., & Anthony, E. (2004). *National Board Certification: Who applies and what factors are associated with success?* The Urban Institute, Education Policy Center. [On-line], Available: http://www.urban.org/EmilyAnthony

Haertel, G.D., Walberg, H.J., & Haertel, E.H. (1981). Socio-psychological environments and learning: A quantitative synthesis. *British Educational Research Journal, 7*, 27-36.

Henderson, D.G., Fisher, D.L., & Fraser, B.J. (1998, April). *Learning environment, student attitudes and effects of students' sex and other science study in environmental science classes.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Henderson, D.G., Fisher, D.L., & Fraser, B.J. (2000). Interpersonal behavior, learning environments and student outcomes in senior biology classes. *Journal of Research in Science Teaching, 37,* 26-43.

Hirata, S., & Sako, T. (1998). Perceptions of school environment among Japanese junior high school, non-attendant, and juvenile delinquent students. *Learning Environments Research, 1,* 321-331.

Huang, T.-C.I., Aldridge, J.M., & Fraser, B.J. (1998) A collaborative study of science classroom environments between Taiwan and Australia: Combining qualitative and quantitative research methods. *Chinese Journal Science Education, 6,* 343-362.

Huang, T.-C.I., & Fraser, B.J. (1997, March). *The development of a questionnaire for assessing student perceptions of classroom climate in Taiwan and Australia.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Chicago, IL.

Keller, B. (2005). Panel urges new testing for teachers. *Education Week, 24*(38), 1-16.

Khine, M.S., & Fisher, D.L. (2001, December). *Classroom environment and teachers' cultural background in secondary science classes in an Asian context.* Paper presented at the annual meeting of the Australian Association for Research in Education, Perth, Australia.

Khine, M.S., & Fisher, D.L. (2002, April). *Analysing interpersonal behaviour in science classrooms: Associations between students' perceptions and teachers' cultural background.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, New Orleans, LA.

Khoo, H.S., & Fraser, B.J. (1998, April). *Using classroom environment dimensions in the evaluation of adult computer courses.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Kim, H., Fisher, D.L., & Fraser, B.J. (1999). Assessment and investigation of constructivist science learning environments in Korea. *Research in Science & Technological Education, 17*, 239-249.

Kim, H.B., Fisher, D.L., & Fraser, B.J. (2000). Classroom environment and teacher interpersonal behaviour in secondary school 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, & G.F. Madaus (Eds.). *Handbook on summative and formative evaluation of student learning* (pp. 559-641). New York: McGraw-Hill.

Koul, R.B., & Fisher, D.L. (May 2005) Cultural background and students' Perceptions of science classroom learning environment and teacher interpersonal behaviour in Jammu, India. *Learning Environments Research, 8*, 195-211**.**

Laforgia, J. (1988). The affective domain related to science education and its evaluation. *Science Education, 72*, 407-421.

Lee, S.S.U. (2001). *Assessment, description and effects of science classroom environments in Korea.* Unpublished doctoral dissertation, Curtin University of Technology, Perth, Australia.

Lee, S.S.U., & Fraser, B.J. (2003). Teacher-student interactions in Korean high school science classrooms. *International Journal of Science and Mathematics Education, 1*, 67-85.

Leitch, R. (1993, July) *The Implications of Intra-psychic Reflection on the Professional and Moral Development of Teacher*. Paper presented to University of Exeter Conference on Moral and Spiritual Development, RIMSCUE Centre, Exeter, United Kingdom.

Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw Hill.

Lightburn, M.E., & Fraser, B.J. (2002, April). *Classroom environment and student outcomes associated with using anthropometry activities in high school science.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Lucas, K.B., & Tulip, D.F. (1980, September). *Scientific literacy of high school students.* Paper presented at the annual conference of the Australian Science Teachers Association, Canberra.

MacDowell-Goggin, M. (2005). *Effect of using graphic organizers on the attitudes and classroom environment of elementary science students*. Unpublished doctoral thesis, Curtin University of Technology.

Majeed, A., Fraser, B.J., & Aldridge, J.M. (2002). Learning environment and its associations with student satisfaction among mathematics students in Brunei Darussalam. *Learning Environments Research, 5*, 203-226.

Maor, D., & Fraser, B.J. (1996). Use of classroom environment perceptions in evaluating inquiry-based computer assisted learning. *International Journal of Science Education, 18*, 401-421.

Margianti, E.S., Fraser, B.J., & Aldridge, J.M. (2001, December). *Investigating the learning environment and students' outcomes in university computing courses in Indonesia.* Paper presented at the annual conference of the Australian Association for Research in Education, Fremantle, Australia.

Margianti, E.S., Fraser, B.J., & Aldridge, J. (2002, April). *Learning environment, attitudes and achievement: Assessing the perceptions of Indonesian university students.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Margianti, E.S., Aldridge, J.M., & Fraser, B.J. (2004). Learning environment perceptions, attitudes and achievement among private Indonesian university students. *International Journal of Private Higher Education.* Retrieved January 18, 2005, from http://www.xaiu.com/xaiujournal/

McRobbie, C.J., & Fraser, B.J., (1993). Associations between student outcomes and psychosocial science environments. *Journal of Educational Research, 87*, 78-85.

Mathison, S. (1988). Why triangulate? *Educational Researcher, 17*, 13-17.

Miami-Dade County Public Schools. (2003). *Miami-Dade County Public Schools student progression plan 2003-2004.* Retrieved July 1, 2005, from http://www.dadeschools.net/ehandbook/student03/Section%20I/2003%20General%20Procedures%20for%20Student%20Progression,%20K%20-%201.pdf

Miami-Dade County Public Schools. (2004). *District demographic*. Retrieved July 1, 2005, from Miami-Dade County Public Schools Web site: http://oada.dadeschools.net/04Profiles.htm

Mink, D.V., & Fraser, B.J. (2005). Evaluation of a K-5 mathematics program which integrates children's literature: Classroom environment and attitudes. *International Journal of Science and Mathematics Education, 3*, 59-85.

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. (1987). *Classroom Environment Scale manual* (2nd ed). Palo Alto, CA: Consulting Psychologists Press.

Moss, C.H., & Fraser, B.J. (2001, April). *Using environment assessments in improving teaching and learning in high school biology classrooms.* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Murray, H.A. (1938). *Explorations in personality*. New York: Oxford University Press.

National Board for Professional Teaching Standards. (2000). *A distinction that matters: Why national teacher certification makes a difference.* Retrieved February 8, 2003, from http://www.mprinc.com/pubs/pdf/distinction_that_matters.pdf

National Board for Professional Teaching Standards. (2001). *I am a better teacher: What candidates for National Board Certification say about the assessment process.* Retrieved February 8, 2003, from http://www.nbpts.org/pdf/better_teacher.pdf

Nix, R., Fraser, B.J., & Ledbetter, C. (2005, May). Evaluating an Integrated Science Learning Environment Using the Constructivist Learning Environment Survey. *Learning Environments Research: An International Journal, 8*, 109-133.

No Child Left Behind Act. (2001). Retrieved October 15, 2004, from U.S. Department of Education Web site: http://www.ed.gov/nclb/landing.jhtml

Pickett, L.H., & Fraser, B.J. (2002, April). *The role of learning environment, achievement, and student and teacher attitudes in a science mentoring program for beginning elementary school teachers*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Pickett, L.H., & Fraser, B.J. (2004, April). *An evaluation of a science mentoring program for beginning elementary school teachers in terms of learning environment, student achievement and attitudes and teacher attitudes*. Paper presented at the annual meeting of the National Association of Research in Science Teaching, Vancouver, BC, Canada.

Podgursky, M. (2001a). Defrocking the National Board. *Education Next, 1* (2), 79-82.

Podgursky, M. (2001b). Should state subsidize National Certification? Retrieved February 11, 2003, from www.edweek.org

Punch, K.F. (1998). *Introduction to social research*. London, England: Sage Publications.

Quek, C.L., Wong, A.F.L., & Fraser, B. (2005a). 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*, 299-321.

Quek, C.L., Wong, A.F.L., & Fraser, B. (2005b). Teacher-student interaction and gifted students' attitudes toward chemistry in laboratory classrooms in Singapore. *Journal of Classroom Interaction, 40*(1), 18-28.

Raaflaub, C., & Fraser, B.J. (2003, March). *The learning environment associated with the use of laptop computers in Canadian science classes.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Philadelphia, PA.

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

Riah, H., & Fraser, B.J. (1998, April). *Chemistry learning environment and its association with students' achievement in chemistry*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Robinson, E. (2003). *Kindergarten students' and their parents' perceptions of science environments: Achievement and attitudes.* Unpublished doctoral thesis, Curtin University of Technology. Retrieved December 27, 2005, from Curtin University of Technology Library and Information Services Web site: http://adt.curtin.edu.au/theses/available/adt-WCU20031117.155353/

Schibeci, R.A. (1984). Attitudes to science: An update. *Studies in Science Education, 11*, 26-59.

Schibeci, R.A., & McGaw, B. (1981). Empirical validation of the conceptual structure of a test of science-related attitudes. *Educational and Pyschological Measurement, 41,* 1195-1201.

Scott, R.H., & Fisher, D.L. (2004). Development, validation and application of a Malay translation of an elementary version of the Questionnaire on Teacher Interaction. *Research in Science Education. 34,* 173-194.

Shulman, L.S. (1997). Disciplines of inquiry in education: A new overview. In R.M. Jaeger (Ed.), *Complementary methods for research in education* (2nd ed., pp. 3-29). Washington, DC: American Educational Research Association.

Sinclair, B.B., & Fraser, B.J. (2003). Changing classroom environments in urban middle schools. *Learning Environments Research: An International Journal, 5,* 201-328.

Soerjaningsih, W., Fraser, B.J., & Aldridge, J.M. (2001, April). *Achievement, satisfaction and learning environment among Indonesian computing students at the university level.* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Soto-Rodriguez, M., & Fraser, B.J. (2004, April). *A comparison of attitudes, achievement, and classroom environment perceptions of LEP (Limited English Proficient) and non-LEP students in integrated science classrooms*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Spinner, H., & Fraser, B.J. (2005). Evaluation of an innovative mathematics program in terms of classroom environment, student attitudes, and conceptual development. *International Journal of Science and Mathematics Education, 3,* 267-293.

Stern, G.G., Stein, M.I., & Bloom, B.S. (1956). *Methods in personality assessment.* Glencoe, IL: Free Press.

Stone, J.E. (2002, May). *Value-added achievement gains of NBPTS-certified teachers in Tennessee: A brief report.* Retrieved July 21, 2005, from www.education-consumers.com/briefs/may2002.asp

Taylor, P.C., Dawson, V., & Fraser, B.J. (1995, April). *Classroom learning environments under transformation: A constructivist perspective.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

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, San Francisco, CA.

Teh, G., & Fraser, B. (1994). An evaluation of computer-assisted learning in terms of achievement, attitudes and classroom environment. *Evaluation and Research in Education, 8*, 147-161.

Thompson, B. (1998a). Review of 'What if there were no significance tests?' *Educational and Psychological Measurement, 58*, 334-346.

Thompson, B. (1998b, April). *Five methodology errors in educational research: The pantheon of statistical significance and other faux pas*. Invited address presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Tobin, K., & Fraser, B.J. (1998). Qualitative and quantitative landscapes of classroom learning environments. In B.J. Fraser & K.G. Tobin (Eds.),

*International handbook of science education* (pp. 623-640). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Trickett, E.J., & Moos, R.H. (1973). Social environment of junior high and high school classrooms. *Journal of Educational Psychology, 65*, 93-102.

U.S. Census Bureau. (2000). *Florida quick facts*. Retrieved June 25, 2005, from http://quickfacts.census.gov/qfd/states/12000.html

U.S. Department of Education. (1983). *A nation at risk: The imperative for education reform* (Publication No. 065-000-00177-2). Washington, DC: Author. Retrieved August 20, 2004, from http://www.ed.gov/pubs/NatAtRisk/risk.html

Vandevoort, L.G., Amrein-Beardsley, A., & Berliner, D.C. (2004). National Board Certified teachers and their students' achievement. *Education Policy Analysis Archives*. Retrieved January 20, 2005, from http://epaa.asu.edu/epaa/v12n46/

Walberg, H.J. (1968). Teacher personality and classroom climate. *Psychology in the Schools, 5*, 163-169.

Walberg, H.J. (1981). A psychological theory of educational productivity. In F. Farley & N. Gordon (Eds.), *Psychology and education: The state of the union* (pp. 81-108). Berkley, CA: McCutchan.

Walberg, H.J. (1984**).** Improving the productivity of America's schools. *Educational Leadership*, *41*, 19-27.

Walberg, H.J. (1986). Synthesis of research on teaching. In M.C. Wittrock (Eds.), *Handbook of research on teaching* (3rd ed., pp. 214-229). Washington, DC: American Educational Research Association.

Walberg, H.J., & Anderson, G.J. (1968). Classroom climate and individual learning. *Journal of Educational Psychology*, *59*, 414-419.

Walberg, H.J., Fraser, B.J., & Welch, W.W. (1986). A test of a model of educational productivity among senior high school students. *Journal of Educational Research, 39,* 133-139.

Wilcox, D.D. (2003). The National Board for Professional Teaching Standards: Can it live up to its promise? Retrieved February 11, 2003, from http://www.edexcellence.net/better/tchrs/13.htm

Wong, A.F.L., & Fraser, B.J., (1996). Environment-attitude associations in the chemistry laboratory classroom. *Research in Science and Technological Education, 14,* 91-102.

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

Yarrow, A., Millwater, J., & Fraser, B.J. (1997). Improving university and primary school classroom environments through pre-service teachers' action research. *International Journal of Practical Experiences in Professional Education, 1* (1), 68-93.

Zandvliet, D.B., & Fraser, B.J. (2004). Learning environments in information and communication technology classrooms. *Technology, Pedagogy and Education, 13* (1), 97-123.

Zandvliet, D.B., & Fraser, B.J. (2005). Physical and psychosocial environments associated with networked classrooms. *Learning Environments Research, 8,* 1-17.