Science and Mathematics Education Centre

Sustained, Job-Embedded Professional Development and the Learning Environment of Middle-Level Mathematics Classrooms

Craig Thomas Gabler

This thesis is presented as part of the requirements for the award of the Degree of Doctor of Philosophy of the Curtin University of Technology

September 2007

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: …March 12, 2008…….

Abstract

As the need for educational reform is increasingly recognized, so too is the need for effective professional development (Guskey, 2000). Historically the evaluation of professional development experiences has been limited to exit surveys, noticeably failing to examine the long-term impact of the effort. This study assessed the impact on the classroom learning environment of a yearlong, job-embedded professional development opportunity for middle-school mathematics teachers. The application of learning environment instruments to the evaluation of professional development is a unique feature of this study. The research employed the Questionnaire on Teacher Interactions (QTI) and a modified version of the What Is Happening In this Class? (WIHIC) survey with over 1000 middle-school mathematics students in 57 classrooms in the state of Washington. Both instruments were administered at the beginning and end of the school year. Teacher interviews were conducted with a sample of participants in order to further illuminate the impact of the professional development. Data from the study were examined for changes in the learning environment and to cross-validate the QTI and WIHIC with this specific population. Results indicate that the QTI and WIHIC are valid and reliable with the middle-school population is this study. Statistical analyses of learning environment data indicate that any pretest-posttest changes that were observed are mostly likely too small to be of educational significance. This study contributes to a better general understanding of the impact of this professional development, and its findings could be utilized in the preparation of future professional development opportunities.

Acknowledgements

No man who has ever wrestled with a self-adjusting card table
can ever be quite the man he once was.

James Thurber

Mr. Thurber's comment was never truer than when describing the journey to completion of my doctoral studies. Coming to this point, however, would not have been possible without the supportive interaction of many folks, and to them I would like to extend my acknowledgement and heartfelt thanks.

First and foremost, I would like to thank Dr. Barry Fraser, Professor of Education and Director, Science and Mathematics Education Centre, Curtin University of Technology, for serving as my supervisor and ever-patient mentor. Without his leadership at SMEC and tireless support of my study and thesis preparation I would not have come to this point.

The entire staff of SMEC has been extraordinarily supportive, seemingly reducing to almost nothing the nearly 10,000 miles separating us. My sincere thank you to Dr. Jill Aldridge for her assistance with data analysis. And a special thank you to Dr. Tony Rickards, for those motivational quips along the way.

To Margaret Flink, former student and doctoral program scout, thank you for thinking of me and for sharing our common doctoral travails.

Most importantly to my family - Thank you, Lori, Andrea & Geoff, for bearing with the impact of my advanced studies. Your continuing support made this all possible. It is to you, Andrea and Geoff, that I dedicate this thesis - May you carry on the value of life-long learning to the next generation.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction and Overview

> A new president, governor, state legislator, superintendent, or board member might come along who wants to know the payoff from the district's or school's investment in professional development. If the evidence isn't there, heads may roll and programs may get axed. To provide evidence in this age of accountability, we must get serious about evaluation. (Guskey, 2005, p. 11)

The spark for my study was precisely what Guskey predicted in the statement above; the need for a systematic evaluation of professional development programs. In my work as a mathematics instructional specialist for an educational service agency in the state of Washington, I was faced with trying to provide evidence that the professional development work that we were doing with teachers was 'paying dividends'. My study undertook to evaluate the impact of sustained, job-embedded professional development from the perspective of the learning environment in the participating teachers' middle-school mathematics classrooms.

This chapter introduces my study in five sections. Section 1.1 presents the motivation for the study and the context within which the study was conducted; Section 1.2 briefly introduces the field of learning environments research; Section 1.3 presents the research aims which informed the evaluation of the professional development and hence laid the foundation for the study design; Section 1.4 discusses the potential contributions and significance of the study; Section 1.5 introduces the study design and potential limitations; and Section 1.6 provides a general overview of the entire study.

1.1 Context for the Study

Around the world, particularly in the United States, and very prominently in my work as a professional development provider in Washington State classrooms, more and more attention is being focused on educational reform. While changes in mathematics curricula have been made, it has been strongly suggested (Bybee, 1997; Sparks & Hirsch, 1997) that the proper professional development of teachers of mathematics is an even more significant factor than curricula in bringing about the needed changes. In the State of Washington, significant financial resources have been brought to bear on the issue of reform, including development of methods for assessing the new student learning standards, and most recently on assisting teachers to be better prepared to help students to achieve these new standards. Now the question has to be asked as to whether those teacher development efforts are having an impact. It is in this context that the motivation for this study was found, and that same impetus will be sustained by the ongoing nature of the professional development of teachers in Washington and around the world.

In recognizing that teaching is a very critical component for the enhancement of student achievement (Glenn, 2000; McCoy, 2005), the State of Washington set out to address the need to improve the pedagogical skills of middle-school mathematics teachers (U.S. Department of Education, 2005). In an attempt to meet that need, the *New Outcomes: Learning Improvement in Mathematics Integrating Technology* (NO LIMIT) teacher professional development project was initiated in 2001. The NO LIMIT project's focus is to "develop classroom models in which students experience standards-based instruction, with the goal of improving student achievement in mathematics. Its implementation includes working with teams of middle-school teachers to improve teaching practices in mathematics through the integration of technology and other research-based methodologies" (Popejoy, Myhre, & Carney, 2005, p. 1). While the intent of improved mathematics achievement was for all students, a priority was placed on those classrooms with demonstrated need based on student mathematics achievement data, high poverty levels and/or high technology needs.

At the state level, the decision was made to regionalize the work of the NO LIMIT professional development project. The existing structure of the Educational Service Districts (ESD) was chosen. In the State of Washington, there are nine ESDs which provide services to up to as many as 50 surrounding school districts within the service region. NO LIMIT funding was provided to each ESD for staff, based on the number of selected schools in the region. Collectively these staff were referred to as Mathematics Integration Specialists (MIS) because their role included infusing the use of technology in the improvement of instruction.

Each of the nine ESDs in Washington was charged with developing its own professional development plan to address the focus of NO LIMIT. My ESD chose to approach the NO LIMIT project goals with a focus on the following six standards for the teaching of mathematics, as presented by the National Council of Teachers of Mathematics (1991):

- worthwhile mathematical tasks
- the teacher's role in discourse
- the student's role in discourse
- tools for enhancing discourse
- the learning environment
- the analysis of teaching and learning.

This focus was chosen because we believed that, because the teaching standards were nationally recognized and were clear and concise statements, the teachers would more readily incorporate them into the structure of their routine.

From a careful review of the six standards for teaching mathematics, and drawing from past experiences in providing mathematics professional development, a classroom observation tool was crafted. A copy of this observation tool can be found in Appendix D. A set of ‘look fors’ was created for each of the six standards. This added detail assisted the other mathematical instructional staff from the ESD and myself with classroom observations. In addition to aiding my observations, the ‘look fors’ were used as discussion points with the participating teachers throughout the entire course of the NO LIMIT project. A sample alignment of these ‘look fors’ with the scales of the WIHIC are discussed in greater detail in Chapter 3.

The professional development experience for the participating teachers began with an intensive five-day summer institute. During this institute the teachers engaged deeply with the Washington State mathematics standards and how they are addressed in their districts' curricular materials. From the state standards for students, the teachers moved onto the teaching standards as presented by NCTM. Video case studies of standards-based classrooms were analyzed and linked to the selected classroom observation tool. The summer institute concluded with a series of concurrent sessions in which the teachers learned about the new technology being made available to them, questioning protocols to encourage student discourse, and strategies for enhancing traditional curricular materials to include more worthwhile mathematical tasks.

The Mathematics Integration Specialists from the ESD were assigned to a set of schools within the service region. In a limited number of cases, two specialists were assigned to the same building. The on-site professional development process began with a great deal of time being spent on building relationships with the participating teachers. Typically, the process initially involved observing in the classrooms, using the observation tool, followed by a conversation with the teachers about which area of the observation tool they would like to have as a focus. This conversation led to decisions about how the teacher wanted to proceed. For some teachers, it included additional observations and follow-up discussions. For other teachers who exhibited a greater readiness to benefit, I began by modeling lessons or team teaching lessons with the teacher. When the teachers were ready to try something on their own, I was there to observe and then guide the reflective conversation afterwards.

Mid-way through the school year and again at the end of the first year a two-day large group institute was conducted. During these two follow-up sessions teacher were given the opportunity to share successes and challenges from their classroom work. Sessions were offered to provide teachers with additional strategies for strengthening student problem-solving skills, writing in mathematics and greater familiarity with the state high stakes assessment system.

This level of intensity of support was a characteristic of the NO LIMIT teacher professional development project. As suggested by the research (Loucks-Horsley, Love, Stiles, Mundry, & Hewson, 2003; Love, 2002; Sparks, 2002; Sparks & Hirsch, 1997), the project had adopted both a job-embedded and a sustained approach. Unlike other professional development projects in the past, we devoted the majority of our time to the teachers in their classrooms. In fact, the outside evaluators found that the "MISs attempt[ed] to spend approximately 80% of their time in the schools with teachers" (Popejoy et al., 2005, p. 3). This feature of the NO LIMIT professional development project alone warrants the additional attention to its evaluation as provided by my study.

The evaluation plan for the NO LIMIT teacher professional development project, as developed by the contractor hired by the State of Washington, had an overarching goal of examining the extent to which the grant was being implemented (Popejoy et al., 2005). This approach was chosen rather than trying to make value judgments about participants or materials because the professional development plans from ESDs were so different.

To assess the implementation of the NO LIMIT project, the evaluator selected a multi-prong strategy. The evaluation included surveys of teachers' pedagogical approaches, knowledge of mathematical content, and perceptions of implementation and administrative issues. And in order to be able to describe the implementation in greater detail, two case studies of selected ESDs were conducted. Results of the examination of student performance on the assessments are discussed in Chapter 6.

In an attempt to examine the impact of NO LIMIT on student achievement in mathematics, the evaluators used data from the Washington Assessment of Student Learning (WASL) for mathematics and the mathematics component of the Iowa Test of Basic Skills (ITBS). The ITBS was administered by the evaluators to NO LIMIT and non-NO LIMIT classrooms for comparison purposes.

Because I was looking for additional evidence of the impact of the NO LIMIT project, beyond what the hired evaluators could provide, and knowing that the connections between classroom learning environment and student cognitive and affective outcomes have been investigated for many years, I chose to pursue an evaluation from a learning environments perspective.

1.2 Learning Environments

Almost 70 years ago, Lewin (1936) published his influential work on role of environment on behavior. Out of his theory has grown the field of educational research known as learning environments. This field significantly matured as a result of the independent research of Herbert Walberg and Rudolf Moos with the development of instruments suitable for capturing students' perceptions of the classroom learning environment. In conjunction with his evaluation of the Harvard Project Physics, Walberg created the Learning Environment Inventory (LEI) (Walberg & Anderson, 1968). And coming out of his work in a variety of institutional settings, Moos developed the Classroom Environment Scale (CES) (Moos, 1979). From these early learning environments instruments have emerged a wealth of other instruments particularly suited to assessing different aspects of the learning environment.

The first of the two learning environments instruments used in my study was the What is Happening In this Class? (WIHIC) questionnaire (Fraser, Fisher, & McRobbie, 1996), which represents a consolidation of previously-used learning environment scales into a very efficient instrument. The second instrument used in my study was the Questionnaire on Teacher Interaction (QTI), which was developed in the Netherlands (Wubbels, Levy, & Brekelmans, 1997) for examining the learning environment from the perspective of teacher-student interpersonal behaviors.

A variety of learning environment instruments have frequently been used in investigating associations between student perceptions of the classroom learning environment with affective and cognitive outcomes (Brekelmans, van den Eeden, Terwel, & Wubbels, 1997; Brekelmans, Wubbels, & Levy, 1993; Fraser, 1994; Fraser & Walberg, 1991; Goh & Khine, 2002; Goh, Young, & Fraser, 1995; Pickett & Fraser, 2002). These studies consistently revealed a positive association between the learning environment and student outcomes.

If positive changes in the learning environment in the NO LIMIT classrooms in my study could be demonstrated, the evidence in the literature would suggest that student achievement outcomes could be positively impacted as well. Using the learning environment perceptions as an indicator therefore could provide some evidence relevant to evaluation of professional development at Guskey's Level 5 – impact on student learning outcomes (Guskey, 2005).

The development, validation, and uses of a wide array of learning environments instruments available to educational researchers and teachers are considered in Chapter 2.

1.3 Research Aims

The purpose of this study, then, was to examine middle-school mathematics classroom learning environments in the context of the ongoing NO LIMIT teacher professional development project. To achieve this purpose, the study utilized the following specific research aims:

1. To investigate whether the WIHIC and QTI are valid and reliable in the context of Washington State middle-school mathematics classrooms.

2. To investigate whether there are differences between how students and teachers perceive the learning environment.

3. To investigate whether changes occur in the learning environment over the course of sustained, job-embedded professional development.

4. To investigate whether boys and girls differ in their perceptions of the learning environment.

1.4 Significance and Contributions

One of the key contributions of my study was the application of a pretest/posttest design within the same group, separated in time by nearly a year's worth of sustained, job-embedded professional development. While other studies have utilized instruments to evaluate the impact of professional development or educational innovations, my study is unique in its use of learning environment criteria in a pretest/posttest design with middle-school mathematics learning students in the United States.

Another contribution of my study is the application of two learning environments instruments as professional development evaluation tools. An external evaluator contracted to evaluate the NO LIMIT project based on its implementation (Popejoy et al., 2005) included interviews, observations and student achievement data. The feature brought to the evaluation by my study is, as Love (2002) suggests, the critical component of student voice.

Middle-school mathematics classrooms have been studied extensively both in the United States and internationally. The significance of my study in the field of middle-school mathematics education is the focus on changes in students' perceptions of the learning environment in relationship to the professional development experience of their teacher. As Bransford et al. (2000, p. 247) point out, "supportive learning environments … need to focus on the characteristics of classroom environments that affect learning". My study could contribute to a greater understanding of those very characteristics.

While other studies utilizing learning environments instruments can claim large sample sizes, my study is the first of its kind and magnitude in the State of Washington. It is also one of few of its size to examine the learning environment in the context of an evaluation of teacher professional development. With the statistical power afforded by the large sample size, the results of my study are likely to contribute to greater understanding of the learning environment of middle-school mathematics classrooms in general, and in Washington classrooms in particular.

Taken together, these features of my study contribute to both the field of learning environments research and the evaluation of professional development. The field of learning environments research is enhanced by an additional study of middle-school mathematics classroom learning environments in the United States. The examination of a large scale professional development project adds to the field of evaluation of professional development through the relatively new application of the QTI and WIHIC learning environments instruments in that context.

1.5 Study Design and Limitations

A contemporary approach to research, as suggested by the current literature (L. Cohen, Manion, & Morrison, 2000; Fraser & Tobin, 1991; Tobin & Fraser, 1998a), was adopted as the methodology for my study. My study included a mixed-methods approach to data collection and analysis. Quantitative data were obtained from both students and teachers through the use of the QTI and WIHIC learning environments instruments. Having a sample size of approximately 1200 students contributes to the uniqueness of the study. The qualitative component of my study involved the collection of data through classroom observations, conversations, and interviews using a convenience sampling approach. Both quantitative and qualitative data were examined in different-sized aggregations. The variation in 'grain size' allows for development of a more complete study.

The need for a more complete picture of what takes place in classrooms has been pointed out by researchers (Guskey, 2000) from both the learning environments and professional development evaluation fields. This more complete picture of the classroom can be obtained through collection of both quantitative and qualitative data, which are then woven together to provide the richer image. In addition, examining the data at varying grain sizes allowed my study the potential of picking up details which might otherwise be lost. As was introduced in Section 1.1 and is expanded on in Chapter 2, the links between classroom environment and student outcomes (cognitive and affective) have been shown to be strong. Because Guskey (2000) has called for the inclusion of 'student voice' in the evaluation of teacher professional development, it was included in my study.

While a more complete discussion of my study's limitations can be found in Chapter 6, a brief overview will be provided here. The NO LIMIT project was a statewide project but was designed by each Educational Service District (ESD) and then delivered in its own service region. While each ESD was designing professional development to achieve the same set of program goals, each ESD approached the task slightly differently. This diversity of professional development plans limits the degree to which my results can be generalized statewide.

In the original design for my study, I had intended to examine the relationship between the psychosocial factors measured by the QTI and WIHIC with the results of the statewide, high-stakes mathematics assessment. That component had to be dropped as I was unable to obtain access to student scores on the assessment. The Washington State Office of Superintendent of Public Instruction (OSPI) was implementing a program in which all students were being assigned unique identifier codes and, by providing a list of the names of participating students to OSPI, I could obtain the achievement scores and preserve the privacy of the students. However the program for assigning the unique identifier codes took much longer than anticipated and thus restricted my program evaluation to the psychosocial domain.

Another limitation of my study is the fact that, during the year in which the teachers were participating in NO LIMIT, it was not their only professional development experience. While the majority of the teachers' professional development time and energy was directed toward the NO LIMIT project, not being able to isolate that professional development as the single variable limits the generalizability of the results.

Finally, as many in the professional development field suggest, the timeline for measurable change in teacher behavior as the result of professional development is more in the order of years than months. While unique in its pretest-posttest design, my study was limited by the fact that the pretest and posttest were only separated by about the length of one school year.

1.6 Overview of Thesis

The design and findings of my study are presented in six chapters. Chapter 1 describes the motivational and contextual background for this study, presents the four research questions, and describes the potential significance and limitations of the study.

Chapter 2 provides a thorough review of the literature pertaining to the research behind current practices in professional development, and current thinking about the importance and role of evaluating professional development. The chapter reviews pertinent literature from the field of learning environments, as well as literature relevant to the research methods which were selected for the research design.

Chapter 3 discusses the methodology used in my study to obtain and analyze the data, including the rationale for combining quantitative and qualitative methods. Included in the discussion is the nature and size of the sample, data-collection instruments, and approaches to analysis of the quantitative data. Finally, the chapter presents the conceptual framework behind the collection and analysis of the qualitative data.

Chapter 4 describes the statistical analysis and results obtained from using the QTI and WIHIC used in my study, as quantitative measures of the classroom learning environments. The reliability and validity results for the QTI and WIHIC are presented.

Chapter 5 communicates the results from both the qualitative and quantitative components of this study. The quantitative results are presented as comparisons between subgroups within the sample or between the group before and after the professional development took place. The qualitative results (classroom observations and interviews) are presented as case studies of four individual classrooms.

Chapter 6 provides discussion and conclusions related to the research questions. The chapter also includes further discussion of the limitations of my study, suggestions for further study, and the potential implications.

1.7 Summary

This first chapter began with the context within which the study was originated and was carried out. The remaining sections outlined the aims of the research, the potential significance and contributions of the study, and an introduction to the design and limitations of the study. Both the contributions and limitations are discussed more fully in Chapter 6. Finally, this chapter has presented an overview of the contents of each of the chapters contained in this thesis.

Chapter 2

Review of the Literature

2.1 Introduction

My study set out to examine the school classroom learning environment in the classrooms of teachers who were participating in an ongoing, job-embedded teacher professional development project. The previous chapter provided an introduction to and background information about this study. In order to lay the foundation for my study, this chapter reviews literature connected with the two main areas of focus within my study: professional development, especially research into its evaluation (Section 2.2); and the field of learning environments research (Section 2.3).

The literature consistently shows that the learning environment in a classroom is strongly linked to student outcomes, including achievement. And, as Fraser points out, there is "compelling evidence that the classroom environment so strongly influences student outcomes that it should not be ignored by those wishing to improve the effectiveness of schools and universities" (Fraser, 2001, p. 2). What is also clear is that there is a need to study what types of professional development best enable teachers to create learning environments suited to enhancing student achievement (P. C. Taylor, Fraser, & Fisher, 1997). The literature clearly points toward a need to include the assessment of learning environments in any evaluation of educational reform efforts (Guskey, 2000). This study set out to do just that.

A review of the research indicates that few studies have been conducted with the intent of examining the linkage between sustained, job-embedded professional development of middle-school mathematics teachers and changes in their classrooms' learning environments. While the research base is rich with evidence of the effect of learning environments on student achievement and attitudes, and with studies involving various approaches to evaluating professional development, few past evaluations of professional development have been conducted from a learning environments perspective. The dearth of research into associations between professional development and learning environments is not due to the need not being clearly communicated. Several years ago, Taylor, Fisher and Fraser (1997) suggested that such research could deal with questions about relationships between professional development and the learning environments created in classrooms.

This chapter considers the literature relevant to the two main areas of focus in my study: the field of professional development, especially evaluation of its activities and programs; and the field of learning environments research. Both fields have an extensive history and have employed a number of tools to further their goals. This review outlines pertinent historical perspectives on professional development (Section 2.2), principles of professional development design (Section 2.3), and methods of evaluating professional development and suggestions for future work (Section 2.4). The review goes on to examine the historical background of learning environments work (Section 2.5), the variety of instruments used to assess learning environments (Section 2.6), and the types of learning environments research undertaken over the years (Section 2.7). Finally, this chapter ends with a summary of the literature in the two areas, with the intent of seeking out potential linkages between evaluating professional development and assessing learning environments.

2.2 Professional Development

2.2.1 History of Professional Development

Professional development, also often referred to as inservice education, teacher training and staff development, has historically been thought of as the special events that occur at discrete times periodically throughout a school year. This view of professional development has contributed to the predominate set of beliefs by teachers and administrators that professional development is distinct and separate from the daily work done in the school. As this chapter develops we will see that part of the work to be done by designers of professional development is to shift that perception to the more current view that professional development is “those processes and activities designed to enhance the professional knowledge, skills, and attitudes of educators so that they might, in turn, improve the learning of students” (Guskey, 2000, p. 16).

While teaching has been going on for centuries, the formal preparation of teachers in the United States has a relatively short history, and inservice education an even shorter history. In the early 1800s, the requirements to teach could be as scant as convincing a local board of your moral character and a rudimentary knowledge of the subjects. By the 1850s, states in the United States were beginning to adopt certification requirements, including completion of a course of studies at a state-regulated institution (Conant, 1963). The focus on a broad preparation to teach continued for nearly the next 100 years.

In mathematics and science education, significant secondary school teaching reform began shortly after the close of World War II (Conant, 1963) and was intensified by the Russian launch of Sputnik in 1957. Rather than placing a focus on significant teacher education, both preservice and inservice, the ‘Golden Age’ (Bybee, 1997) of mathematics and science education took the approach that improvements could be driven out through curriculum development and preparation of better instructional materials. This era is known for the development of such programs as PSSC (Physical Science Study Committee, 1960) for physics; the chemistry curriculum known as CHEMStudy (Chemical Education Materials Study) (Pimentel, 1963), and the School Mathematics Study Group (SMSG) materials first published in 1960. As a consequence of the emergence of these new materials, the professional development in mathematics and science during this era was focused on preparation to use these new materials.

The mid-1970s found the United States’ public dissatisfied with the overall performance of schools and thus focusing greater accountability on schools (Adey, 2004). During this period of social unrest and dissatisfaction with schools, Hopkins and Lagerweij (1996) portray school improvement efforts as consisting of predominately ‘hand-wringing’. Sashkin and Egermeier (1993) characterized this same period, from the late 1970s to the early 1980s, as “distinguished by state level mandates for change, often legislated; [and] generally, it was no more effective” (p. 17) than the ‘Golden Age’. The changes being brought about in schools today rest in part on the components of the effective schools research (Coleman et al., 1966) that found a foothold in the late 1980s and the 1990s. In addition to taking a systems approach to school improvement, today we recognize the need for not only a commitment to professional development (Hopkins, 2001), but an understanding of how to design and evaluate professional development.

2.2.2 Professional Development Design

This section discusses the field of professional development research from the perspective of the current thinking on designing effective professional development. Included in this discussion also is the potential role that professional development can play in the making the latest reforms efforts more successful than the reform efforts mentioned in Section 2.2.1.

Over the past 100 years, professional development can best be described as a ‘sit and git’ situation. In the earliest days, the professional development for teachers consisted of traditional college courses attended after school or during the summer. More recently, this same approach to professional development has moved into schools where the teachers attend ‘workshops’. This model of professional development has been described by Lieberman (1995) as ‘direct teaching’, with the teacher as the passive recipient of new knowledge. In light of this ‘training’ mode of professional development, it should not be a surprise to find that, in a 2003 survey, teachers responded in equal numbers as to whether their recent experiences with professional development improved their performance or had little impact (Public Agenda, 2003).

In addition to teacher responses such as in the previously-mentioned survey, educational leaders have been indicating for many years a need to change how professional development is planned and conducted (Elmore, 2002; Fullan & Stiegelbauer, 1991; Loucks-Horsley et al., 2003; National Staff Development Council, 2001; Sparks & Hirsch, 1997). Stigler and Hiebert point out that teaching is “the new frontier in the struggle to improve schools” (Stigler & Hiebert, 1999, p. 2), while Sparks and Hirsch (1997) suggest that what is required to address that struggle is a paradigm shift in professional development .

The literature is abundant with research and suggestions on how to improve the planning and delivery of professional development. Sparks and Hirsch (1997) suggest that there are three transformative ideas that are key to designing effective professional development. These ideas are results-driven education, systems thinking, and constructivism. In results-driven professional development, the success of the effort is not judged by mere attendance at a professional development activity, but by how much change is observable in teaching practices that are beneficial to students (e.g. the learning environment). The systems thinking approach to professional development incorporates Senge’s (1990) belief that we have to move from a view of life that predominately focuses on discrete events to one that capitalizes on the interconnectedness of all things. And, finally, professional development designed from a constructivist basis integrates the general concepts of what recent brain research tells us about learners and learning (Bransford et al., 2000).

While complete agreement about what constitutes effective professional development is still elusive (Guskey, 2003), a review of the current literature reveals significant support, and application of, the three-pronged approach suggested by Sparks and Hirsch (Sparks & Hirsch, 1997).

In the widely-used *Designing Professional Development for Teachers of Science and Mathematics* (Loucks-Horsley et al., 2003) the authors identified the values underlying the framework as:

1. Professional development experiences need to have students and their learning at their core.
2. Excellent science and mathematics teachers have a very special and unique kind of knowledge that needs to be developed through their professional learning experiences.
3. Principles that guide the reform of student learning should also guide professional learning for educators.

4. The content of professional learning must come from both inside and outside the learner, and from both research and practice.
5. Professional development must both align with and support systems-based changes that promote student learning. (p. xxv)

From their list, it can be seen that their values encompass a results-driven schema (Item 1), a systemic perspective (Items 2 and 5) and a reliance on a constructivist approach to learning (Items 2 and 4). The framework for designing teacher professional development promoted by the authors reinforces these three ideas throughout their suggested process.

Others have identified and incorporated these same three key ideas into their suggestions not only for improving professional development, but also for school reform in general. Love (2002) identifies eight guiding principles which are based in part or in whole on getting results and a reliance on system-wide support. Similarly, in describing his vision for transforming professional development, Guskey (2000) identifies the need not only to have a student learning focus and a systemic approach, but to utilize what research says about how all people learn. Guskey (1994; 1996) also advocates a balanced approach to evaluating professional development and the need to have effective evaluation.

More recently, a new emphasis is appearing in the literature concerning well-designed professional development (Darling-Hammond, 1998; Loucks-Horsley et al., 2003; Loucks-Horsley & Matsumoto, 1999; Love, 2002). This new emphasis focuses on the learning environment with the professional development itself and, as such, is very much in concert with the focus of my research. Coming out of their work with data from the Third International Mathematics and Science Study (TIMSS), Stigler and Hiebert (2004) suggest that, in designing a professional development program:

> A focus on teaching must avoid the temptation to consider only the superficial aspects of teaching: the organization, tools, curriculum content, and textbooks. The cultural activity of teaching – the ways in which the teacher and students interact about the subject – can be more powerful than the curriculum materials that teachers use. (p. 16)

With a similar emphasis on the importance of the learning environment for the teacher, Weiss and Pasley (2004) indicate that for the continuing professional growth of teacher the "professional development activities should reflect the elements of high-quality instruction with clear, explicit learning goals, a supportive but challenging learning environment, and ways that ensure that teachers are developing their understanding" (p. 28).

The professional growth opportunities for teachers have not only expanded over the last century, but also have changed significantly in their content and approach. The research is not clear on precisely which approach is best, and it is likely no one model will work in all settings. What is clear is that, if professional development is going to meet the needs of teachers, and ultimately students, further refinements are necessary. Possibly most significant of those changes at this time is the way in which professional development is evaluated, which is discussed in the next section.

2.2.3 Evaluating Professional Development

Although the current literature makes clear the connection between the classroom learning environment and student outcomes (Brekelmans, Wubbels et al., 1993; Fraser, 2001; Marzano & Marzano, 2003; Stockard & Mayberry, 1992), the effective evaluation of the efforts of professional development providers, including those targeting the learning environment, is noticeably lacking (P. C. Taylor et al., 1997).

As recently as the mid-1990s, Clarke (1994) had identified what he deemed were the ten key principles of professional development of mathematics teachers, but no mention was made of the need to include evaluation of the effort. It was also at about this same time when substantial interest in evaluating professional development was emerging (Guskey, 2000). Guskey describes the four reasons for the growing interest:

1. A better understanding of the "dynamic nature" of professional development
2. Recognition of professional development as "an intentional process"

3. The need for better information to guide reform efforts
4. Increased pressure for “accountability”. (Guskey, 2000, p. 8)

Despite over ten years of awareness of the need for, and interest in, evaluating professional development little progress has been made, as evidenced by Mizell’s statement that “[t]he field of staff development needs better evaluation both to improve the effectiveness of teachers’ learning experiences and to produce credible evidence that will garner more support for professional development” (2003, p. 12).

In the past, what was misnamed as the evaluation of professional development was in fact more about documentation of the event (Killion, 2002). To a large extent, this is due to the presence of the predominate model for professional development described in Section 2.2, namely, the ‘sit and git’ approach. The documentation typically focused on the participants’ reactions to whether they liked the training, the quality of the presenter, whether their own comfort needs were met, and how useful they thought that the information might be to them (Guskey, 2000). Sparks & Hirsch summarize this sort of documentation as a measure of the “happiness quotient” (1997, p. 1). When the professional development included teacher acquisition of content knowledge, the documentation might have included use of an instrument to measure the extent to which the information was learned, often administered at the beginning and end of a project. While most teacher professional development has had a goal of enhancing student outcomes, rarely have the attempts at evaluation of such programs included direct measures of the impact of the professional development on students.

Only recently has the prevailing school of thought around evaluation of teacher professional development begun to change. As a result of the evaluation efforts of numerous National Science Foundation funded projects in the United States, several key characteristics have surfaced. These areas include not only the quality of the professional development activities themselves but also changes in the curriculum, instruction and assessment, as well as the classroom culture or context (Loucks-Horsley et al., 2003). As Mizell (2003) points out, “while it will continue to be important to evaluate the delivery of staff development, it will be essential to understand what participants actually learn, when and how they begin to apply their learning, and when and how it benefits students" (p. 12).

In 1995, Guskey and Roy (1995) suggested that a minimum set of guidelines be established for evaluating professional development programs. These guidelines might include such elements as:

- Evaluation should be ongoing.
- Evaluation expectations and procedures should be explicit and public.
- Evaluation should be informed by multiple sources of data.
- Evaluation should use both quantitative and qualitative data.
- Evaluation should focus on all levels of the organization.
- Evaluation should be considerate of participants' time and energy.
- Evaluation results should be presented in forms that can be understood by all program participants and patrons.

Building from these basic ideas, Guskey (2000) developed a robust model for evaluating teacher professional development. As seen in the model presented in Table 2.1, Guskey has identified five critical levels: (1) Participants' reaction, (2) Participants' learning, (3) Organization support and change, (4) Participants' use of new knowledge and skills, and (5) Student learning outcomes. It is in Level 5 that Guskey has acknowledged that a desired outcome of teacher professional development should include an impact on the students' affective outcomes, and thus this impact could be used as a measure of success. This component of the Guskey model for teacher professional development is the one that is most closely related to my research.

Evidence of inclusion of these five levels of professional development evaluation can be found in many of the current guidelines for professional development, including: the National Staff Development Council standards (National Staff Development Council, 2001), Florida Department of Education (Florida Department of Education, 2006), and in the language of the Washington State professional development standards, which states that effective professional development

Table 2.1
Five Critical Levels of Professional Development Evaluation

Evaluation Level	Crucial Questions What Questions are Addressed?	How Will Information be Gathered?	What is Measured or Assessed?	How Will Information be Used?
1. Participants' reactions	• Did they like it? • Was their time well spent? • Did the material make sense? • Will it be useful? • Was the leader knowledgeable and helpful? • Were the refreshments fresh and tasty? • Was the room the right temperature? • Were the chairs comfortable?	• Questionnaires administered at the end of the session	• Initial satisfaction with the experience	• To improve program design and delivery
2. Participants' learning	• Did participants acquire the intended knowledge and skills?	• Paper-and-pencil instruments • Simulations • Demonstrations • Participant reflections (oral and/or written) • Participant portfolios • Case study analyses	• New knowledge and skills of participants	• To improve program content, format, and organization
3. Organization support and change	• What was the impact on the organization? • Did it affect organizational climate and procedures? • Was implementation advocated, facilitated, and supported? • Was the support public and overt? • Were problems addressed quickly and efficiently? • Were sufficient resources made available? • Were successes recognized and shared?	• District and school records • Minutes from follow-up meetings • Questionnaires • Focus groups • Structured interviews with participants and school or district administrators • Participant portfolios	• The organization's advocacy, support, accommodation, facilitation, and recognition	• To document and improve organizational support • To inform future change efforts
4. Participants' use of new knowledge and skills	• Did participants effectively apply the new knowledge and skills?	• Questionnaires • Structured interviews with participants and their supervisors • Participant reflections (oral and/or written) • Participant portfolios • Direct observations • Video- or audiotapes	• Degree and quality of implementation • To document and improve the implementation of program content	
5. Student learning outcomes	• What was the impact on students? • Did it affect students performance or achievement? • Did it influence students' physical or emotional wellbeing? • Are students more confident as learners? • Is student attendance improving? • Are dropouts decreasing?	• Student records • School records • Questionnaires • Structured interviews with students, parents, teachers, and/or administrators • Participant portfolios	• Student learning outcomes: o Cognitive (performance & achievement) o Affective (attitudes and dispositions) o Psychomotor (skills and behaviors)	• To focus and improve all aspects of program design, implementation, and follow-up • To demonstrate the overall impact of professional development

Source (Guskey, 2000)

"[i]ncludes a strong program evaluation component based on evidence of improvements in student learning and teacher practice" (Office of Superintendent of Public Instruction, 2005, p. 12). While these plans for evaluation of teacher professional development have incorporated a much richer approach, few studies are currently available (C. C. Johnson, Kahle, & Fargo, 2007) which describe the benefits that might accrue to this new approach.

2.3 Learning Environments

This section describes the field of learning environments research, beginning with its history and foundations in Section 2.3.1. The establishment and application of a range of the instruments used to measure the various constructs of learning environments within the classroom are discussed next in Section 2.3.2. Following those sections is a detailed discussion of the two instruments used in my study, namely, the WIHIC (Section 2.3.3) and the QTI (Section 2.3.4). This section concludes with a discussion of past studies of learning environments and their possible connection to my study (Section 2.3.5).

2.3.1 Learning Environments History

As far back as the early 1930s (Waller, 1932), the social nature of the classroom was recognized. However the actual field of learning environments had its roots in the work of Lewin (1936) and Murray (1938). Lewin put forth that the interaction between an individual and the environment could be a significant determinant of human behavior. As a formula, Lewin's theory describes the behavior of the individual (B) as a function of both the personality of the individual (P) and the individual's environment (E):

$$B = f(P, E)$$

Murray (1938) incorporated this construct into his approach to studying the educational environment. He coined the terms *alpha press* to describe the observation of the educational environment by a detached observer and *beta press* to describe the observations made by an observer (e.g. a student in the classroom) intimately embedded in the environment. The combination of these perspectives in studying the classroom was used in my study to provide a much richer representation of the classroom environment.

The development of instruments for assessing the learning environment began over 35 years ago through the work of Herbert Walberg of Harvard University and Rudolf Moos, a Stanford University psychology researcher. As part of a study of the Harvard Project Physics program, Walberg developed the *Learning Environment Inventory* (LEI) (Walberg & Anderson, 1968). Through his work in clinical and institutional settings (Moos, 1974), Moos created social climate scales that eventually developed into the *Classroom Environment Scale* (CES) (Moos, 1979; Moos & Trickett, 1974, 1987; Trickett & Moos, 1973). In the ensuing 35 years, more than a dozen other classroom questionnaires have been developed and validated for use in assessing the learning environment (Fraser, 2002). These instruments range from the *My Class Inventory* (MCI) for use at the elementary level to the *Science Laboratory Environment Inventory* (SLEI) and *Constructivist Learning Environment Survey* (CLES) which are useful at the secondary and post-secondary level. Of particular note, because they form the core of learning environment instruments in my study of professional development effectiveness, are the *Questionnaire on Teacher Interaction* (QTI) and the *What Is Happening In this Class* (WIHIC).

2.3.2 Learning Environments Instruments

Classroom learning environments have most commonly been studied through the use of survey instruments. Beginning with Walberg's LEI (Walberg & Anderson, 1968), and spanning nearly four decades of using survey instruments to collect data on classroom learning environments, a variety of instruments have been created. Moos' (1974) developed a classification scheme for sorting human environments into three dimensions: *relationship*, *personal development*, and *system maintenance and change*. Recognition of the three dimensions led Moos (Moos, 1974; Moos & Trickett, 1974) to develop the Classroom Environment Scale (CES). Moos' significantly influenced the development of instruments for assessing the learning environment from the student's perspective. All of the instruments have this underlying structure inherent in them, and their scales can be classified into one of the dimensions of Moos' scheme. The instruments described below are those considered to be most relevant to my study.

The learning environment instruments briefly discussed in the subsections below are:

- Learning Environment Inventory (LEI)
- My Class Inventory (MCI)
- College and University Classroom Environment Inventory (CUCEI)
- Classroom Environment Scale (CES)
- Individualized Classroom Environment Questionnaire (ICEQ)
- Science Laboratory Environment Inventory (SLEI)
- Constructivist Learning Environment Survey (CLES)
- Computer Classroom Environment Inventory (CCEI)
- Cultural Learning Environment Questionnaire (CLEQ)
- Distance and Open Learning Environment Scale (DOLES)
- Socio-Cultural Environment Scale (SCES)
- What Is Happening In this Class? (WIHIC)
- Questionnaire on Teacher Interaction (QTI)

2.3.2.1 Learning Environment Inventory (LEI)

As mentioned in the previous section, Walberg utilized the *Learning Environment Inventory* (LEI) in a pioneering study of the impact of the Harvard Project physics in the late 1960s (Walberg & Anderson, 1968). The LEI had its origins in the 18-scale *Classroom Climate Questionnaire* developed a by Walberg (1968). The LEI assesses 15 different learning environment dimensions of secondary science classrooms. Each of the 15 scales has seven items, for a total of 115 items. The items are scored on a four-point Likert scale (Likert, 1932), with responses ranging from Strongly Disagree to Strongly Agree, with some of the items utilizing reverse scoring. The two dimensions introduced by this early learning environment instrument that most closely relate to my research are the scales for Cohesiveness and Democracy. The Democracy scale has evolved into the scale now known as Equity, which appears in the WIHIC. Over the years, the LEI has been analyzed for internal consistency reliability and discriminate validity (Fraser, Anderson, & Walberg, 1982) and found to be both reliable and valid.

2.3.2.2 My Class Inventory (MCI)

The *My Class Inventory* (MCI) (G. J. Anderson & Walberg, 1968; Fraser et al., 1982; Fraser & O'Brien, 1985; Majeed, Fraser, & Aldridge, 2002; Mink & Fraser, 2005; Walberg & Anderson, 1968) is an adaptation of the original LEI. The MCI is intended for younger audiences, and so the adaptations to the LEI were made with this in mind. The MCI has five scales compared to the 15 of the LEI and now has 38 items compared to the 105 items of the LEI. The original MCI had 45 items but, through analysis of scale reliabilities (Fisher & Fraser, 1981), this was reduced to the present 38 items. In addition to reducing student fatigue during administration (Fraser, 1991), the MCI has been found to be useful for those students who struggle with reading the LEI (Fraser & Wubbels, 1995). Additional enhancements to the MCI include the response format being changed to a simple Yes or No format and being answered directly on the questionnaire.

2.3.2.3 College and University Classroom Environment Inventory (CUCEI)

To address the need for an instrument to assess the learning environment at the College and University level, the *College and University Classroom Environment Inventory* (CUCEI) was created by Fraser and Treagust in 1986 (Fraser & Treagust, 1986; Fraser, Treagust, Williamson, & Tobin, 1987). While the CUCEI was created for the small classroom environment typical at a university, it has found applicability at the secondary level as well (Logan, Crump, & Rennie, 2006; Nair & Fisher, 2001). The CUCEI consists of seven scales with seven items per scale. The items are scored on a Likert scale, with approximately half of the items utilizing reverse scoring. The CUCEI contains the Task Orientation scale in the personal development dimension as well as the Involvement and Student Cohesiveness scales under Moos' (Moos, 1974) relationship dimension.

2.3.2.4 Classroom Environment Scale (CES)

Out of Rudolf Moos'(1974) pioneering investigations of human interactions in a wide variety of situations, including some educational settings, the *Classroom Environment Scale* (CES) was developed (Fisher & Fraser, 1983b). The CES measures the psychosocial environment of secondary school classrooms from the perspective of interactions, including teacher-student, student-student and teacher behaviors (Moos & Trickett, 1974, 1987). The CES includes such scales as Involvement and Teacher Support from the relationship dimension, Task Orientation from the personal development dimension, and Teacher Control from the system maintenance and change dimension. The CES instrument consists of nine scales utilizing 10 items per scale. The 90 items utilize a True/False response format with approximately half of the items being reverse scored.

2.3.2.5 Individualized Classroom Environment Questionnaire (ICEQ)

The *Individualized Classroom Environment Questionnaire* (ICEQ) was developed by Rentoul and Fraser (1979) and published by Fraser (1990) to assess the learning environment in classrooms where individualized and inquiry-based learning was taking place. In its original version, the ICEQ consisted of 5 scales with 15 items per scale. The current version retains the five scales, but only contains 10 items per scale for a total of 50 items. The items are scored based on a five-point frequency response scale, ranging from Almost Never to Very Often. Some of the items are reverse scored.

Unique features of the ICEQ are focus on the individualized (as distinct from teacher-centered) setting, and its introduction of the Investigation scale. With many classrooms shifting toward more student-centered teaching during the 1970s, the inclusion of a scale focusing on student investigation made possible a richer examination of the learning environment. The ICEQ has been utilized in a number of studies and been found to be valid and reliable; this includes a study in Brunei examining the relationship between student classroom environment perceptions and their attitudes (Asghar & Fraser, 1995).

2.3.2.6 Science Laboratory Environment Inventory (SLEI)

To address the need for an instrument which would assess the learning environment in science laboratory classrooms at the secondary and postsecondary levels, the Science Laboratory Environment Inventory (SLEI) was developed by Fraser, Giddings, and McRobbie (1995). The SLEI consists of 35 items equally divided among five scales of Cohesiveness, Open-Endedness, Integration, Rule Clarity and Material Environment. The 35 items are scored on a five-point frequency response scale with approximately half of the items being reverse scored. In addition, the SLEI was the first instrument to utilize a 'personal' version to supplement the traditionally-used 'class' version. The personal version allows greater differentiation within class subgroups, such as boys and girls, by collecting student perceptions of their own position in the classroom rather than just their perception of the classroom as a whole as is done using the class form (Fraser, Giddings, & McRobbie, 1992).

The SLEI has been utilized in a number of international studies and its cross-validation has been demonstrated. These studies include Singapore (Quek, Wong, & Fraser, 2005a; Wong & Waldrip, 1996; Wong, Young, & Fraser, 1997), Brunei (Riah & Fraser, 1998), Tasmania (Fisher, Harrison, Hofstein, & Henderson, 1998), United States (Lightburn & Fraser, 2007), and Korea (Kim & Kim, 1995, 1996; Lee & Fraser, 2002; Lee & Kim, 2002). Additionally, the SLEI has been used to study chemistry, biology, and physics classrooms (Fisher et al., 1998; Henderson & Fisher, 1998).

2.3.2.7 Constructivist Learning Environment Survey (CLES)

The *Constructivist Learning Environment Survey* (CLES) was constructed (P. C. Taylor, Dawson, & Fraser, 1995; P. C. Taylor & Fraser, 1991; P. C. Taylor et al., 1997) about the same time that the SLEI was being developed. The CLES, originally designed for use with secondary students, aids teachers and researchers as they examine classroom learning environments from a constructivist perspective. The constructivist theory of learning is based on the premise that learners construct their own understanding through interactions with their existing knowledge and recent experiences. The CLES is intended to assess the degree to which a classroom conforms to the constructivist view of learning. The original version of the CLES consisted of four scales, containing from nine to twenty items. In its final validated form, the CLES has five scales with six items each. Students respond on a five-point frequency response scale ranging from Almost Always to Almost Never.

The CLES has found wide application internationally, and in a variety of school settings. The CLES has been utilized in studies of science classrooms in the United States (Dryden & Fraser, 1998; B. Johnson & McClure, 2004; Nix, Fraser, & Ledbetter, 2005; Spinner & Fraser, 2005) and Korea (Kim, Fisher, & Fraser, 1999; Lee & Fraser, 2001). Additionally, the CLES has been used as a tool in action research projects in South Africa (Aldridge, Fraser, & Sebela, 2004) and in a cross national study between Taiwan and Australia (Aldridge, Fraser, Taylor, & Chen, 2000). The extensive application of the CLES in researching classrooms from a constructivist learning perspective has created a learning environments instrument trusted among researchers.

2.3.2.8 Computer Classroom Environment Inventory (CCEI)

With the introduction of technology into classrooms, both the style of teaching and the materials used for teaching changed. To assist teachers and researchers to assess these classrooms, the Computer Classroom Environment Inventory (CCEI) was developed (Maor & Fraser, 1993, 1996). Similar to the CLES, the CCEI investigates the computer classroom environment with a focus on the perceptions of inquiry and investigation. The CCEI consists of 30 items equally divided over the five scales. Responses are scored using a five-point frequency scale (ranging from Almost Never to Almost Always) with some items being reverse scored.

2.3.2.9 Cultural Learning Environment Questionnaire (CLEQ)

To better address the increasing diversity in classrooms, the *Cultural Learning Environment Questionnaire* (CLEQ) was developed (Fisher & Waldrip, 2002). Drawing from the research on the dimensions of culture by Hofstede (1984), and Moos' (1979) dimensions, Waldrip and Fisher constructed the CLEQ with the seven scales of Equity, Collaboration, Deference, Competition, Teacher Authority, Modeling and Congruence. Each of the scales consists of five items which are scored on a five-point Likert scale with responses ranging from Disagree to Agree. The

CLEQ has been administered to nearly 4000 students in a range of school settings, including a teacher training program (Dhindsa & Fraser, 2004). Analyses of those results (Fisher & Waldrip, 1997, 1999) indicate that the CLEQ is a valid and reliable instrument and a valuable tool for investigating the learning environment from a cultural perspective.

2.3.2.10 Distance and Open Learning Environment Scale (DOLES)

As the use of computers in education has grown, and distance learning become more prevalent, the need for a research tool in this area was evident. To address this need Jegede at al. (1995) developed the *Distance and Open Learning Environment Scale* (DOLES) to assess student perceptions of their learning environment when studying a distance education. The DOLES consists of the five core scales of Student Cohesiveness, Teacher Support, Personal Involvement and Flexibility, Task Orientation and Material Environment and Home Environment. These five scales were initially validated using 660 students (Jegede et al., 1995). In addition, the DOLES has the two optional scales of Student Centered Environment and Technology Resources, which can be used to customize the instrument to the teachers' or researchers' specific requirements.

2.3.2.11 Socio-Cultural Environment Scale (SCES)

The *Socio-Cultural Environment Scale* (SCES) was developed by Jegede and Okebukola (1988) to assess socio-cultural aspects of the learning environment in science classrooms in Africa. The SCES contains the five scales of Authoritarianism, Goal Structure, African World-view, Societal Expectations and Sacredness of Science. Responses to the 30 SCES items are scored on a three-point Likert-style scale, with some items being reverse scored.

In 1993, using a pretest-posttest format with 600 senior secondary science students in Nigeria, Jegede, Fraser, and Okebukola (1994) found the SCES to be valid and reliable. Results from this study suggest that socio-cultural factors can be addressed in classrooms, with a more positive attitude toward science being used as an outcome.

2.3.2.12 Distance Education Learning Environments Survey (DELES)

A recent addition to the available learning environments questionnaires is the *Distance Education Learning Environments Survey* (DELES) (Walker, 2003). The DELES was developed to afford researchers an instrument to measure the learning environment in post-secondary distance education settings. This unique web-based survey consists of 34 items over the six scales of Instructor Support, Student Interaction and Collaboration, Personal Relevance, Authentic Learning, Active Learning, and Student Autonomy. In a study involving 680 distance education students, the instrument was found to be valid and reliable (Walker & Fraser, 2005).

As the *What Is Happening In this Class?* (WIHIC) instrument and the *Questionnaire on Teacher Interaction* (QTI) were central in my research, they are discussed in detail in Section 2.3.3 and Section 2.3.4, respectively.

Table 2.2 is presented here as a summary of the learning environment instruments reviewed in this section. Included in the table is the classification of the various scales in the instruments according to Moos' (1974) scheme.

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

The WIHIC (Fraser et al., 1996) is a consolidation of scales extracted from existing learning environment questionnaires (see Table 2.1 above) into a compact and useful form. Of the original eight scales, seven have been retained in current versions. The current WIHIC is designed to measure seven dimensions of the psychosocial learning environment: Student Cohesiveness, Teacher Support, Involvement, Investigation, Task Orientation, Cooperation, and Equity. These seven dimensions span the three schemes of Moos (1974). Table 2.3 summarizes the focus of each scale and provides a sample item.

The WIHIC is composed of eight items in each of the seven scales. Participants respond to the 56 items using a five-point frequency response scale. The possible frequency responses are Almost Never, Seldom, Sometimes, Often, and Almost Always, with no reverse scoring utilized. The eight responses for a scale are

Table 2.2
Overview of Scales in 14 Classroom Environment Instruments

Instrument	Level	Items Per Scale	Scales Classified According to Moos' Scheme		
			Relationship Dimensions	Personal Development Dimensions	System Maintenance & Change Dimensions
Learning Environment Inventory (LEI)	Secondary	7	Cohesiveness, Apathy Friction, Favoritism, Cliqueness, Satisfaction	Speed, Difficulty, Competitiveness	Diversity, Formality, Goal Direction, Disorganization, Material Environment, Democracy
My Class Inventory (MCI)	Elementary	6–9	Cohesiveness, Friction, Satisfaction	Difficulty, Competitiveness	
College and University Classroom Environment Inventory (CUCEI)	Higher Education	7	Personalization, Involvement, Cohesiveness, Satisfaction	Task Orientation	Innovation, Individualization
Classroom Environment Scale (CES)	Secondary	10	Involvement, Affiliation, Teacher Support	Task Orientation, Competition	Order & Organization, Rule Clarity, Teacher Control, Innovation.
Individualized Classroom Environment Questionnaire (ICEQ)	Secondary	10	Personalization, Participation	Independence, Investigation	Differentiation
Science Laboratory Environment Inventory (SLEI)	Upper Secondary	7	Cohesiveness	Open Endedness, Integration	Rule Clarity, Material Environment
Constructivist Learning Environment Survey (CLES)	Secondary	6	Gender Equity	Investigation, Resource Adequacy	Innovation

Continued …

Table 2.2 *(Continued)*

Instrument	Level	Items Per Scale	Scales Classified According to Moos' Scheme: Relationship Dimensions	Personal Development Dimensions	System Maintenance & Change Dimensions
Computer Classroom Environment Inventory (CCEI)	Secondary	6	Satisfaction	Investigation, Open-Endedness	Material Environment, Organization
Questionnaire on Teacher Interaction (QTI)	Secondary Elementary	8	Leadership, Understanding, Helpful/Friendly, Student Responsibility & Freedom, Uncertain, Admonishing, Dissatisfied, Strict		
Cultural Learning Environment Questionnaire (CLEQ)	Secondary	5	Gender Equity, Collaboration, Deference	Competition	Teacher Authority, Modeling, Congruence
What Is Happening In this Class? (WIHIC)	Secondary	8	Student Cohesiveness Teacher Support Involvement	Investigation Task Orientation Cooperation	Equity
Distance and Open Learning Environment Scale (DOLES)	Tertiary	4–12	Student Cohesiveness, Teacher Support, Personal Involvement, Flexibility	Task Orientation, Material Environment, Technology Resources	Student Centre Environment, Home Environment
Socio-Cultural Environment Scale (SCES)	Secondary Elementary	6	African World View	Societal Expectation	Authoritarianism, Goal Structure Sacredness of Science
Distance Education Learning Environments Survey (DELES)	Post-Secondary	4-5	Instructor Support, Student Interaction & Collaboration	Personal Relevance, Authentic Learning, Active Learning	Student Autonomy

Based on (Fraser, 1998a)

Table 2.3
Description of Scales and a Sample Item for Each Scale of the WIHIC

Scale name	Description of scale	Sample item
Student Cohesiveness	Extent to which students know, help and are supportive of one another.	I know other students in this class.
Teacher Support	Extent to which the teacher helps, befriends, trusts and is interested in students.	The teacher takes a personal interest in me.
Involvement	Extent to which students have attentive interest, participate in discussions, perform additional work and enjoy the class.	I explain my ideas to other students.
Investigation	Emphasis on the skills and processes of inquiry and their use in problem solving and investigation.	I carry out investigations to test my ideas.
Task Orientation	Extent to which it is important to complete activities planned and to stay on the 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 by the teacher.	I am treated the same as other students in this class.

Source (Dorman, 2003, p. 243)

combined to produce a composite score for that scale that can be used for further data analysis. The WIHIC has a *'class'* form, which assesses students' perceptions of the class as a whole, as well as a *'personal'* form for their own perceptions. Taken together, these different forms offer a significant source of potential data that a teacher might use to make changes in the classroom learning environment (Fisher, Rickards, & Fraser, 1996).

Initial validation of the WIHIC took place by combining statistical analysis with interview data from 355 junior high school students in Australia (Fraser et al., 1996). Over the years, other studies (Dorman, Adams, & Ferguson, 2002; Fraser, 1998b) utilizing the WIHIC have found the instrument to be both reliable and valid. Further significant work confirming validation took place in a study of 1879 Taiwanese students and 1081 Australian science students (Aldridge, Fraser, & Huang, 1999). Cronbach alpha coefficients in the range of 0.90 to 0.96 indicate high internal consistency reliability. Even slightly stronger consistencies were reported by Dorman (2002) in a cross-national study of students from Australia, Canada and the United Kingdom. ANOVA results for the Taiwanese science students indicated that each WIHIC scale was able to significantly differentiate between classes. As Dorman indicated, "the scales do overlap but not to the extent that would violate the psychometric structure of the instrument" (2002, p. 505).

In what might be the single largest study involving the WIHIC, Dorman (2003) undertook to validate the scales from the WIHIC with 3980 students in Australia, Canada and the UK. The students were enrolled in Grade 8, 10 and 12 mathematics classes. Not only was the study cross-national but also encompassed several grade levels. The study found that the scales of the WIHIC had good internal consistency, good discriminant validity and were able to discriminate between within-school grade groups. This study is noteworthy because confirmatory factor analysis was used to support the WIHIC's sound structure and to show the invariance of this structure across countries, grade levels and genders. Results from the exhaustive data analysis clearly provided "substantive international validation of the WIHIC" (Dorman, 2003, p. 243).

The WIHIC has been used internationally, from the elementary-school level to adult learners, and has been shown to be valid across its many applications (Fraser, 1998b). Internationally the WIHIC has been shown to be useful and valid in Australia (Dorman et al., 2002), Brunei Durussalam (Khine & Fisher, 2002; Riah & Fraser, 1998), Canada (Dorman et al., 2002; Raaflaub & Fraser, 2002; Zandvliet & Fraser, 2005), India (Koul & Fisher, 2005), Indonesia (Margianti, Fraser, & Aldridge, 2004), Korea (Kim, Fisher, & Fraser, 2000), Singapore (Fraser & Chionh, 2000; Khoo & Fraser, in press), Taiwan and Australia (Aldridge & Fraser, 2000), and in the United States (Allen & Fraser, 2007; Martin-Dunlop & Fraser, in press; Moss, 2003; Ogbuehi & Fraser, 2007; Pickett & Fraser, 2004).

The WIHIC has been used successfully in a variety of settings. These successes include cross-national studies as well as applications involving multiple research methods (Aldridge et al., 1999). The feature that the WIHIC brings to a study of the impact of professional development on classroom learning environments is that it provides beta press data (i.e. students' observations about what is happening in the classroom – see Section 2.3.1), specifically about how learning opportunities are structured. For example, the Involvement, Investigation, and Task Orientation dimensions measured by the WIHIC give a strong indication of the emphasis placed on student learning versus teacher teaching (Aldridge et al., 1999). These would be valuable indicators for examining changes in a teacher's approach (e.g. as a result of a professional development experience). The information provided by the WIHIC is likely to contribute usefully to a more complete picture of changes occurring within the classroom.

Further evidence of the success of the WIHIC is found in two instruments which have their roots in the WIHIC. The WIHIC was used as the foundation for the Techonology-Rich Outcomes-Focused Learning Environments Inventory (TROFLEI). The TROFLEI consists of 80 items covering 10 scales. Three additional scales were added to the original seven (see Table 2.3) – Differentiation,
Computer Usage and Young Adult Ethos – in order to assess the technology-centered and individualized nature of emerging educational programs. Students respond to items using a frequency response format, with choices ranging from Almost Never to Almost Always. In a study of 772 Australian secondary students,

the TROFLEI was found to be a valid and reliable instrument (Aldridge, Dorman, & Fraser, 2004; Fraser et al., 1996).

The other relatively new instrument based from the WIHIC is the Outcomes-Based Learning Environment Questionnaire (OBLEQ). The OBLEQ was developed to address the new learning environments in South Africa brought about by the adoption of a new outcomes-based curriculum in 2005. The OBLEQ retains the WIHIC scales for Involvement, Investigation, Cooperation, and Equity. From the ICEQ (Fraser, 1990) and the CLES (Aldridge et al., 2000; P. C. Taylor et al., 1997), the scales for Differentiation and Personal Relevance were added. The eighth scale, Responsibility for Own Learning, was developed specifically for the new instrument. The response format consists of a five-point frequency scale, ranging from Never to Always. The pioneering work on this instrument produced the first published research of its kind in South Africa. The study, involving 2638 Grade 8 Science students, produced results indicating that the instrument, with the combining of two scales, is valid and reliable (Aldridge, Laugksch, Seopa, & Fraser, 2006).

The studies described in this section support the validity and reliability of WIHIC in portraying the nature of classroom environments. The studies consistently show that the WIHIC, and its progeny, are useful for gathering information from students for the purpose of improving teaching and learning. Given the wide use and effectiveness of the WIHIC, its applicability with middle-school mathematics classrooms, it was therefore chosen for use in my evaluation of a teacher professional development program.

2.3.4 Questionnaire on Teacher Interaction (QTI)

The QTI has its origins in empirical research in clinical psychology. Timothy Leary (1957) examined human communications as a means for identifying personality. He believed that the two strongest motivators in human behavior are the need to reduce fear and the corresponding need to maintain a certain level of self-esteem (Leary, 1957). He further believed that the choice of a certain interpersonal behavior is dictated by these needs and that success with a given communication pattern would cause that pattern to become characteristic. This ‘give-and-take’ model described by Leary is part of a larger outlook on communication which considers communications from a systems perspective. Family therapists, Watzlawick, Beavin and Jackson (1967), described the communications system as having a circular, or intertwined, nature. Such a system would be characterized by "communication processes [that] develop which not only *consist* of behavior but which *determine* behavior as well" (Creton, Wubbels, & Hooymayers, 1993, p. 2).

In his analysis of hundreds of patient-therapist dialogues, Leary found that the communications, representing interpersonal behaviors, could be arranged into 16 "mechanisms" (1957, p. 65), but were eventually distilled down to just eight. Deriving their conceptual framework from the Leary model, Wubbels, Creton, Levy and Hooymayers (1993) have represented their model for interpersonal teacher behavior in a two-dimensional plane. In the model for interpersonal teacher behavior, the two dimensions are labeled as Proximity (Cooperation – Opposition) and Influence (Dominance – Submission) as in Figure 2.1. Proximity describes the extent of cooperation or closeness between the communicating parties, where the Influence dimension is used to indicate the degree of control or influence over the communications process.

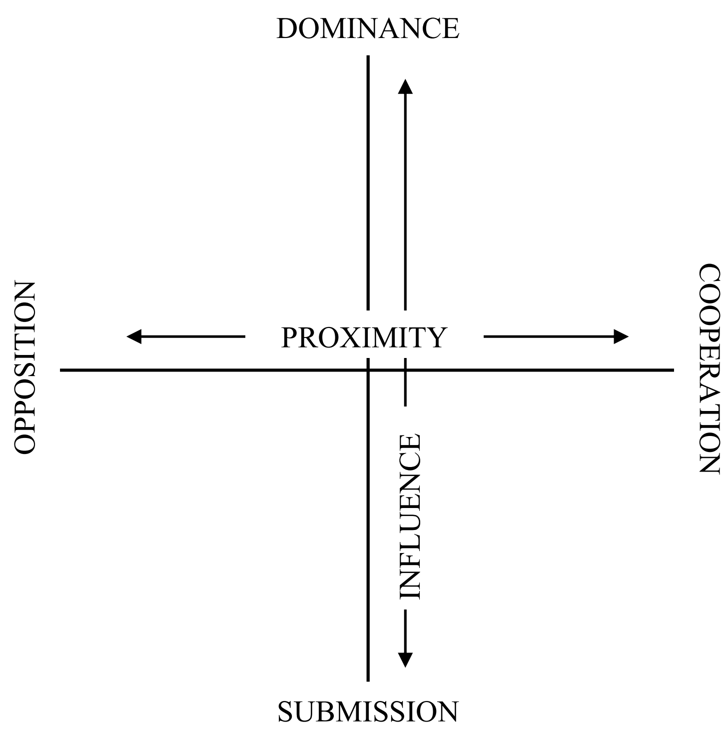


Figure 2.1 The Two-Dimensional Coordinate System of the Leary Model
(Source: Wubbels, Creton, Levy, Hooymayers, 1993, p. 15)

The model for interpersonal teacher behavior developed from the Leary model was found to be applicable in the classroom (Wubbels, Creton et al., 1993), but lacked an instrument capable of producing valid and reliable results. The instrument developed by Leary, the Interpersonal Adjective Check List (ICL), was found to be too difficult to use in a classroom context, but served as a basis for the QTI.

Developed from classroom research conducted in the 1980s in the Netherlands around the concept of student-teacher interactions (Wubbels & Levy, 1993), the Questionnaire on Teacher Interaction (QTI) was designed to efficiently assess student perceptions of the classroom learning environment from the perspective of interpersonal behaviors. The perceptions are grouped according to eight different behavioral constructs (Leadership, Helpful/Friendly, Understanding, Student Responsibility/Freedom, Uncertain, Dissatisfied, Admonishing, and Strict behavior). These constructs come out of the systems approach to human interaction, which examines teacher behavior from the perspective of the two polar domains of proximity and influence mentioned above. This was a unique approach and is one that has been found to be highly effective in a variety of settings (Fraser, 1998a).

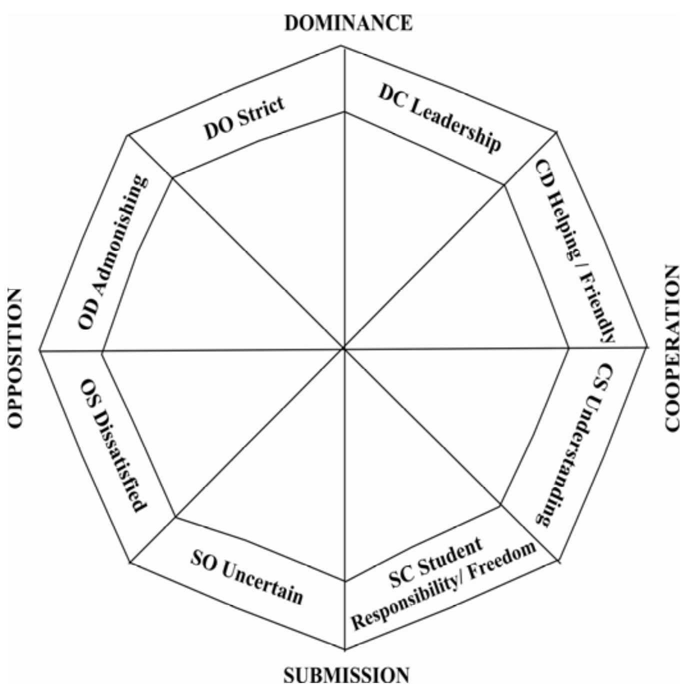


Figure 2.2 The Model for Interpersonal Behavior
(Source: Fisher, Fraser, & Wubbels, 1993)

The initial work with the QTI took place at the secondary level in The Netherlands (Wubbels et al., 1997) and originally consisted of 77 questions (Wubbels, Creton, & Hooymayers, 1992). Since that time, it has been shortened to 48 items, with six items for each of the eight scales, by Fisher, Fraser and Wubbels (1993). For each item, the student responds on a five-point frequency scale (Almost Never, Seldom, Sometimes, Often, Almost Always). Table 2.4 provides a scale description and sample item for each QTI scale.

Table 2.4
Description of Scales and Sample Item for each Scale of the QTI

Scale Name	*Description of Scale* (The extent to which the teacher …)	*Sample Item*
Leadership	… leads, organizes, gives orders, determines procedure and structures the classroom situation.	This teacher talks enthusiastically about his/her subject.
Helping/Friendly	… shows interest, behaves in a friendly or considerate manner and inspires confidence and trust.	This teacher helps us with our work.
Understanding	… listens with interest, empathizes, shows confidence and understanding and is open with students.	This teacher trusts us.
Student responsibility/ Freedom	… gives opportunity for independent work, gives freedom and responsibility to students.	We can decide some things in this teacher's class.
Uncertain	… behaves in an uncertain manner and keeps a low profile.	This teacher seems uncertain.
Dissatisfied	… expresses dissatisfaction, looks unhappy, criticizes and waits for silence.	This teacher thinks that we cheat.
Admonishing	… gets angry, expresses irritation and anger, forbids and punishes.	This teacher gets angry unexpectedly.
Strict	… checks, maintains silence and strictly enforces the rules.	This teacher is strict.

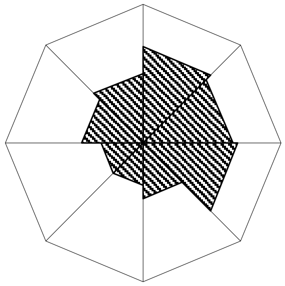
Source (Kim et al., 2000)

From the original Dutch form of the QTI, with 77 items, an American version was developed with 64 items (Wubbels & Levy, 1991). Following these two versions, an even more compact Australian version was created with only 48 items (Fisher et al., 1993). From these original forms, versions of the QTI were developed to assess not only how the students perceive the teacher's actual interpersonal behavior, but also what they perceive to be their 'ideal' or preferred, teacher. Running parallel to the students' forms are a similar pair for teachers for collecting data on how teachers perceive their own student-teacher interpersonal behaviors as well as what they believe to be an 'ideal' teacher.

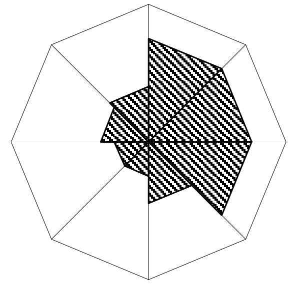
Validation of the QTI has taken place on a global scale, with data being collected most notably in the Netherlands, the United States, and Australia (Wubbels & Levy, 1993). The QTI has also been validated in both mathematics and science classroom studies in the United States (Wubbels & Levy, 1991), Asia (Goh & Fraser, 2000; Kim et al., 2000; Lee, Fraser, & Fisher, 2003; Lee & Kim, 2002; Quek, Wong, & Fraser, 2005b; Scott & Fisher, 2004), Australia (Fisher et al., 1996), and Greece (Kyriakides & Muijs, 2005). In the study described by Fisher, Rickards, and Fraser (1996), results confirmed the validity and reliability of the QTI when used with senior secondary biology students. In that study, it was found that the alpha reliability for different QTI scales ranged from 0.74 to 0.95 when the class mean was used as the unit of analysis. In addition, the QTI has been shown to be an effective instrument for assessing teacher perceptions of their own interactions. When teacher perceptions are considered in combination with student perceptions, the QTI constitutes an important tool for teacher self-reflection (Fisher et al., 1993).

Since its development in the 1980s the QTI has found application in a very diverse set of studies. The QTI has been utilized is such areas as science and mathematics classrooms, in computer classrooms, and even in teacher education classrooms (Wubbels & Levy, 1993). In addition to shedding light on the condition of the classroom, results from the QTI have been found to be very useful in promoting teacher self-reflection (Fisher et al., 1996) through application of such tools as the teacher typologies developed by Brekelmans, Levy and Rodriquez (1993).

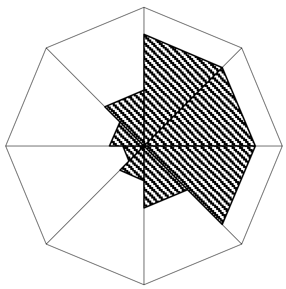
The teacher typologies are presented in Figure 2.3. The data to produce the profiles (i.e. radar plots) came from a number of Dutch, American, and Australian studies of



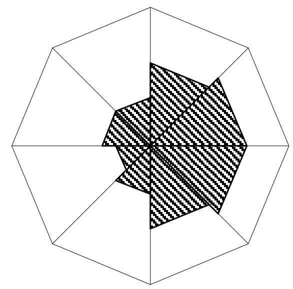
Type 1 Directive



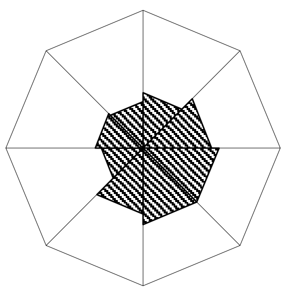
Type 2 Authoritative



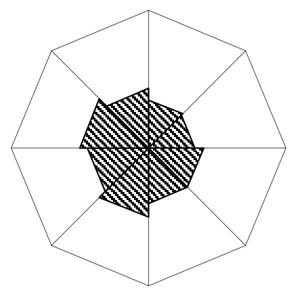
Type 3 Tolerant and Authoritative



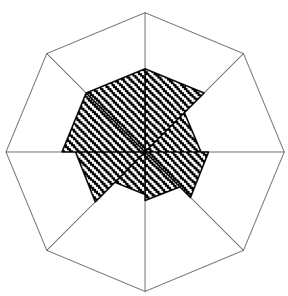
Type 4 Tolerant



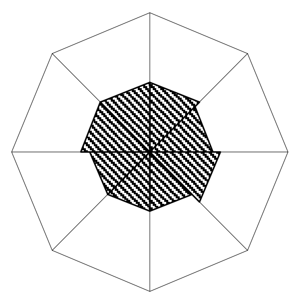
Type 5 Uncertain and Tolerant



Type 6 Uncertain and Aggressive



Type 7 Repressive



Type 8 Drudging

Figure 2.3 Typology of Teacher Interpersonal Behaviors

Source: (Brekelmans, Levy et al., 1993)

teacher interpersonal behaviors (Levy, Rodriguez, & Wubbels, 1992; Wubbels, Brekelmans, Creton, & Hooymayers, 1990; Wubbels, Brekelmans, & Hermans, 1987; Wubbels & Levy, 1991). Using cluster analysis, the class means for all of the teachers were grouped, resulting in the eight categories: Directive, Authoritative, Tolerant and Authoritative, Tolerant, Uncertain and Tolerant, Uncertain and Aggressive, Repressive, and Drudging. Further analysis of the data has validated the typologies and has shown associations between teacher communication style and classroom characteristics.

Additional studies have indicated that the learning environment is very closely related to student outcomes (Brekelmans, Wubbels et al., 1993). Brekelmans et al. found in their Dutch study of 66 Grade 9 physics classes that achievement on standardized physics test appeared to be strongly related to the interpersonal teacher behavior. And finally an obvious strength of the QTI is its ability to be used in conjunction with other research methods (combining quantitative and qualitative research) and in doing cross-national studies (Aldridge et al., 1999; den Brok, Fisher, Wubbels, Brekelmans, & Rickards, 2006; Fraser & Tobin, 1991).

As demonstrated by the successful, multiple applications of the QTI, it can be seen that it is a versatile tool for examining the interpersonal behavior component of a classroom. Much teacher professional development purports to bring about changes in the behavior of the teachers. With that in mind, a study might contribute significant knowledge by investigating the applicability of the QTI as a tool for evaluating the effect of teacher professional development on teacher interpersonal behaviors.

2.3.5 Past Studies of Learning Environments

In the four decades since learning environments research formally began, a large number and variety of survey instruments have been created. As can be seen from the previous sections, these instruments have been used to study a wide range of learning contexts and to make connections between many of the variables present in a learning environment. This section presents an overview of the areas of past research utilizing learning environments survey instruments.

As Fraser (1998a) points out, 12 distinct areas of research can be identified that have involved a wide variety of learning environments instruments. Table 2.5 provides a convenient summary of the 12 areas of research. The center column contains a brief description of the emphasis within each area of research. The far right column provides a short list of references for key studies within that a While not all 12 areas are relevant to my study, several of them are and therefore are highlighted in separate sections below. Because a focus of my study was the evaluation of the classroom learning environment in the context of a job-embedded, sustained teacher professional development project, it is most closely related to Fraser's (1998a) areas of Associations between Student Outcomes and Environment, Evaluations of Educational Innovations, Teacher Assessment, Student-Teacher Differences, and Combining Research Methods. Each of these is discussed in the subsequent sections.

2.3.5.1 *Associations between Student Outcomes and Learning Environment*

Chapter 1 identified the primary goal of the NO LIMIT (*New Outcomes, Learning Improvement in Mathematics Integrating Technology*) teacher professional development project to be that of enhancing student achievement in mathematics. The application of learning environment instruments to the study of the association between student perceptions of the classroom learning environment and affective and cognitive outcomes are numerous (Brekelmans et al., 1997; Brekelmans, Wubbels et al., 1993; Fraser, 1994; Fraser & Walberg, 1991; Goh & Khine, 2002; Goh et al., 1995; Pickett & Fraser, 2002). Results of these studies indicate that there is a positive association between the learning environment and student achievement.

Early studies of the associations between learning environment and student outcomes assumed independence of observations and so applied multiple regression analysis (Bock, 1989; Bryk & Raudenbush, 1992; Goldstein, 1987) to the data. More recently, educational and social-science researchers have employed multilevel analyses, such as hierarchical linear modeling, in recognition of the *nested* nature (Hox, 2002) of such contexts as classrooms. That is, students are *nested* in classrooms and classrooms are *nested* within the school. Two learning environments

Table 2.5
Areas of Past Research in the Field of Learning Environments and Their Emphases

Research Area	Main Emphasis of Research	Key References
Associations between Student Outcomes and Environment	Investigation of associations between perceptions of psychosocial characteristics of a classroom and students' cognitive and affective learning outcomes.	Fraser (1998a, 1998b, 2002), Margianti, Fraser, & Aldridge (2004), Webster & Fisher (2003)
Evaluation of Educational Innovations	Process criteria for the evaluation of educational programs are obtained via classroom learning environment instruments.	Maor & Fraser (1996), Spinner & Fraser (2005)
Student-Teacher Differences	Investigation of perceived differences between the students and teacher in a classroom situation. Differences could be between actual or preferred environments.	Fraser & Chionh (2000), Fraser et al. (1992)
Person-Environment Fit	Research into whether student outcomes depend on the similarity between preferred and actual classroom environment.	Fisher & Fraser (1983a, 1983b), Yarrow, Millwater & Fraser (1997)
Teacher Improvement	Instruments provide feedback information for use in reflecting upon, discussing, and attempting to improve classroom environment.	Nix et al. (2005), Pickett & Fraser (2004), Yarrow et al. (1997)
Combining Research Methods	Research involving the use of both quantitative and qualitative methods in the same study in order to identify salient features of the environment studied.	Aldridge et al. (1999, 2000) , Johnson & McClure (2004), Tobin & Fraser (1998b)

Source: Fraser (1998b)

Continued …

Table 2.5 Continued

Research Area	Main Emphasis of Research	Key References
School Psychology	Research instruments can be used to identify areas of classroom life and differences that impact the mental and emotional welfare of students.	Burden & Fraser (1993)
Links between Environments	Attempts to identify connections and influences of multiple environments involved in the education process, both in and out of the formal school. These also cover ergonomic aspects of the classroom environment, as well as the environments established through distance learning.	Walker & Fraser (2005), Zandvliet & Fraser (2005), Zandvliet & Straker (2001)
Cross-national Studies	Investigation of the similarities and differences between the educational environments of various countries, as well as to question the practices and beliefs of a given country.	Aldridge & Fraser (2000), Aldridge et al. (1999, 2000), den Brok et al. (2006), Dorman & Ferguson (2004)
Transitions between Grade Levels	Research on the effect of students moving from one level of education to another, such as from primary to junior high school. Also, research on interaction between the socio-cultural environment outside of the classroom and the classroom.	Ferguson & Fraser (1999), Jegede, Fraser & Okebukola (1994)
Teacher Education	Opportunities to include the topic of learning environments in programs for the preparation and training of future educators.	Johnson & McClure (2004), Nix et al. (2005), Yarrow et al. (1997)
Teacher Assessment	Dimensions of learning environments can yield insight into present teaching methods and focus, as well as possible effectiveness from the student perspective.	Baker et al. (2005), Nix et al. (2005), Yarrow et al. (1997)

Source: Fraser (1998b)

studies in particular have compared the results from multiple regression analysis and hierarchical linear modeling. Wong et al.'s (1997) study of 1592 Grade 10 students in 56 chemistry classes in Singapore examined the association between attitudes toward science and student responses to a modified Science Laboratory Environment Inventory (SLEI). Goh et al.'s (1995) study of the psychosocial climate and student outcomes with 1512 Grade 5 mathematics students in 39 classes in Singapore, using a modified My Class Inventory (MCI), related those responses to student achievement and attitude. In both studies, most of the statistically significant findings from the multiple regression approach were replicated using hierarchical linear modeling, both in terms of statistical significance and the direction of relationships.

Walberg (1981) identified nine factors (prior achievement, development, motivation, quantity and quality of instruction, psychosocial learning environment, home environment, peer group, and mass media) in his model of educational productivity. The model suggests that, to improve overall educational productivity, attention needs to be given to those factors that limit learning most. To test the educational productivity model, extensive meta-analyses of studies of the correlation of the factors to learning (Fraser, Walberg, Welch, & Hattie, 1987) and secondary analyses of large national data bases (Walberg, Fraser, & Welch, 1986) have been conducted. These studies found that classroom and school-level learning environments were strong predictors of student outcomes, even when other factors were controlled.

In a study examining associations between attitudes and psychosocial environment with 1021 Grade 9 and 10 science students in India, Koul and Fisher (2005) found that there was a positive association. Similar results were found for affective outcomes in a study in mathematics classes in Brunei Darussalam (Majeed et al., 2002). Similarly, associations between cognitive outcomes and learning environment have been investigated. In a recent study of technology-rich classrooms, the Technology-Rich Outcomes-Focused Learning Environment Inventory (TROFLEI), with 76 items covering 10 scales, was utilized (Aldridge, Fraser, Fisher, Trinidad, & Wood, 2003). For a subsample of 386 students, Aldridge et al. found that, for six of the ten scales, there were positive and statistically significant associations with student achievement.

If the NO LIMIT classrooms in my study were to show positive changes in the learning environment, there is sufficient evidence in the literature to suggest that student outcomes could be positively impacted as well. Using learning environment perceptions as an indicator would thus be a step toward attaining Guskey's (2000) fifth level of professional development evaluation.

2.3.5.2 *Evaluations of Educational Innovations*

Researchers have also employed learning environment instruments in the evaluation of teacher and curriculum enhancement efforts going as far back as Walberg's (1968) study of Project Physics. More recently, evaluations of a range of teacher enhancement programs have involved using learning environments instruments with students and teachers (Hurst, 1999; Maor & Fraser, 1996; Spinner & Fraser, 2005). In a recent study (Pickett & Fraser, 2004) involving evaluation of a science mentoring program for teachers, the WIHIC was used with students in a pretest and posttest modality. Pickett and Fraser indicated that the WIHIC did detect "relatively little, [but positive], difference between pretest and posttest results on classroom environment scales" (2004, p. 28).

Martin-Dunlop and Fraser (in press) assessed the effectiveness of an innovative science course for improving prospective elementary teachers' perceptions of laboratory classroom learning environments and attitudes towards science. The sample consisted of 27 classes with 525 female students in a large urban university in California. Students reported large and statistically significant improvements on all seven scales assessing the laboratory learning environment and attitudes towards science. The largest gains were observed for Open-Endedness and Material Environment.

Nix, Fraser, and Ledbetter (2005) used the Constructivist Learning Environment Survey (CLES) in evaluating the impact of an innovative teacher development program (based on the Integrated Science Learning Environment, ISLE, model) in school classrooms. Two separate response blocks of 30 items comprising five scales were presented in side-by-side columns to measure students' perceptions (using a five-point frequency response scale) of the extent to which certain psychosocial factors are prevalent in the science class taught by a teacher who had attended the ISLE program, as well as their perceptions of other science and non-science classes taught by other teachers in the same school. The five scales of the CLES are Personal Relevance, Uncertainty of Science, Shared Control, Critical Voice and Student Negotiation. The sample consisted of 1,079 students in 59 classes in north Texas. Students whose science teachers had attended the ISLE program perceived higher levels of Personal Relevance and Uncertainty of Science in their classrooms relative to the classrooms of other science and non-science teachers in the same school. Similar results were found when comparing the classroom environment perceptions of students whose science teachers had attended the ISLE program with the perceptions of students whose science teacher had attended alternative field trip programs.

Lightburn and Fraser (Lightburn & Fraser, 2007) evaluated the use of anthropometric activities among a sample of 761 high school biology students in Florida in terms of student outcomes (achievement and attitudes) and classroom environment (assessed with the Science Laboratory Environment Inventory, SLEI). The efficacy of using anthropometric activities was supported by pretest-posttest differences in achievement, as well as by a comparison with a control group's attitudes and perceptions of classroom learning environment. Overall, results provide a degree of support for the positive influence of using anthropometric activities in terms of students' attitudes and the classroom learning environment.

Mink and Fraser (2005) reported a one-year study of 120 fifth grade students whose teachers participated in a program entitled Project SMILE (Science and Mathematics Integrated with Literacy Experiences). The purpose of the study was to determine the extent to which the classroom implementation of Project SMILE positively influenced the classroom environment and student attitudes toward reading, writing, and mathematics. The implementation of SMILE was found to have a positive impact on the students of the teachers who participated in the inservice program in that students' attitudes towards mathematics and reading improved and there was congruence between students' actual and preferred classroom environment on the scales of Satisfaction and Difficulty.

While learning environments instruments have been utilized extensively in studying science classroom enhancement, a search of the literature has revealed sparse application in the area of middle-school mathematics. My study sought to add to that body of knowledge.

2.3.5.3 *Teacher Assessment*

Examination of classroom learning environments for the purpose of assessing teachers can be linked to the evaluation of educational innovations. In addition to the studies discussed in Section 2.3.5.2, other studies have utilized instruments in a multilevel approach. Nix, Fraser, and Ledbetter's (2005) investigation of a science teacher development program produced results at the classroom level that were used to assess the teachers' extent of implementation of the innovations. As these types of results are shared with teachers, then teachers have the opportunity to reflect more deeply on their own practice.

A study to examine the possibility of using student ratings of teacher behaviors as part of a teacher evaluation system was carried out with 1973 Year 6 Greek students (Kyriakides & Muijs, 2005). Kyriakides and Muijs found that "student responses to the Greek version of QTI helped explain variances on student achievement in both cognitive and affective outcomes of schooling" (p. 63). While additional work is needed to determine the relative importance of teacher interpersonal behavior factors, the potential does exist for integrating classroom learning environment measures into the assessment of teachers' performance.

As part of a study involving nearly 900 K–12 classroom observations in Washington State, Baker, Olzendam, and Arlington (2005) developed the *STAR Classroom Observation Protocol*. This *Protocol* includes indicators common to other learning environments instruments such as Student Involvement (from the WIHIC) and Personal Relevance (from the CLES). While this protocol relies heavily on acquisition of alpha press (i.e. external observer) data, it does provide both the observer and the classroom teacher with the opportunity to reflect on what goes on in the classroom. The *STAR Classroom Observation Protocol* has the potential to influence the way in which administrators observe and evaluate classrooms.

2.3.5.4 *Differences Between Students' and Teachers' Perceptions*

Also relevant to my study is past research of learning environments instruments, including the QTI (Fisher & Fraser, 1983a; Fraser, 2002; Wubbels, Brekelmans, Tartwijk, & Admiraal, 1999), in examining differences between student and teacher perceptions of the learning environment. The selected instruments typically have a *preferred version* to measure students' perception of their desired environment and an *actual version* to measure their perceptions of the actual environment. Parallel to these two versions are the teacher versions, which also could have *preferred* and *actual* forms. Results from these studies consistently point to the fact that students prefer a more positive learning environment than what actually exists (Fraser, 2002; Fraser et al., 1992; Margianti et al., 2004). Contrasted with this are the results from comparisons of student and teacher perceptions, which indicate that teachers tend to see their classes "through rose-coloured glasses" (Fraser et al., 1992, p. 6; Wubbels, Brekelmans, & Hooymayers, 1993). This area of research is significant for enhancing student achievement because, as Fraser points out the "actual-preferred congruence (or person-environment fit) could be as important as individualization *per se* in predicting student achievement of affective and cognitive aims" (1998b, p. 22).

2.3.5.5 *Combining Research Methods*

Upon examining the methods used in mathematics and science education research, Kelly and Lesh (2000) point out:

> We are now at a point where the growing maturity of mathematics and science education research has shifted attention from strict adherence to traditional experimental methods as the best path to scientific insight to renewed interest in the development of alternative methods for research. In the past few number of decades, educational researchers have moved into school systems, classrooms and workplaces and have found a complex and multifaceted world that they feel is not well described by traditional research techniques. In the past, educational phenomena derived their status by surviving a variety of statistical tests. Today, nascent

> educational phenomena are accorded primacy, and the onus is on research methods to describe them in rich and systematic ways. (p. 35)

In fact, in a recent survey taken by the National Science Foundation (US Government), Suter (2005) reports that 85% of mathematics and science education studies funded by the National Science Foundation incorporated at least two different research methods. This same trend, of using multiple methods, is true for research on classroom learning environment (Fraser & Tobin, 1991; Tobin & Fraser, 1998b).

Multi-method approaches to learning environments research have been applied in a variety of studies. In research reported by Fraser and Tobin (1991), higher-level cognitive learning in secondary science classrooms was investigated. The approach taken with the two science teachers was to conduct in-depth observations and interviews and to administer parts of the Individualized Classroom Environment Questionnaire (ICEQ) and the Classroom Environment Scale (CES). In the experimental design, the qualitative observations and interviews initially provided the information necessary to appropriately select the quantitative tools (ICEQ and CES). The combination of the two approaches allowed the researchers to conclude, among other findings, that students' perceptions of the classroom learning environment were related to teachers' knowledge and beliefs, and that teachers' attitudes and expectations were reflected in student perceptions of the learning environment. The quantitative data obtained from the students made sense in the context of the qualitative data obtained from the classroom, thus showing the value of the combined-method approach.

In another study reported by Fraser and Tobin (1991), researchers examined the practice of exemplary Australian mathematics and science teachers across a broad grade range in order to disseminate effective practices. The qualitative component of the research involved extensive observations, interviews, and the examination of instructional materials. Quantitative data were obtained by assessing students' perceptions of the learning environment with a variety of instruments. The study resulted in a number of findings, one of which connects strongly with my study. It was found that students' perceptions of the learning environment, coupled with the observations, allowed researchers to clearly distinguish between the classrooms of exemplary and non-exemplary teachers.

A third and extensive study involving the use of a multi-method approach was reported by Aldridge, Fraser, and Huang (1999), who produced a more complete and coherent picture of the learning environments in Taiwan and Australia. In this cross-national study, the quantitative component was comprised of the responses of 1081 students from Australia and 1879 students to the WIHIC (Fraser et al., 1996) and to the Test of Science-Related Attitudes (Fraser, 1981). The qualitative component consisted of classroom observations, interviews with teachers and students, and narrative stories written by the researchers. Unique differences between the two countries were observed: Australian students perceived their classroom environments more favorably than did their Taiwanese counterparts, while the converse was true when attitudes toward their science classes were considered. The observations and interviews, coupled with the quantitative results, allowed the researchers to come to a much deeper understanding of the classroom learning environments in the two countries and how each country might learn from the other. Taking a 'bricolage' approach produced a much richer picture of the learning environment (Denzin & Lincoln, 2005).

2.4 Summary

The goal of this review of the relevant literature was to present a succinct and coherent picture of teacher professional development, learning environments and such linkages as might exist between the two fields of study. Historically teacher professional development and learning environments research have shown little crossover but, as each has matured, mutually beneficial relationships have become apparent.

Purposeful teacher professional development has not occurred until relatively recently. As a consequence, many of the early practices produced significantly less than the desired outcomes. These 'hit and miss' episodes eroded not only the confidence of participating teachers but also that of funding agencies. From both the growth in understanding about the professional development process and from increased accountability have come powerful models for the design and evaluation of teacher professional development.

Many models for designing teacher professional development today suggest that planning should begin with the end in mind (Wiggin & McTighe, 1998) and establish what will count as evidence of success and plan from there. Where significant strides have been made in evaluating teacher professional development is in the recognition that the data must come from a multi-faceted approach. This approach includes not only information about the delivery of the professional development, but also about its impact on the cognitive, affective, and psychomotor aspects of the student.

In Section 2.3, it was shown that learning environments instruments have their roots in the early 20th century in studies directed at examining the social nature of learning. Since those early years, a wide range of instruments have been developed for a wide range of purposes, with nearly all of the instruments having been shared and used internationally. The specific learning environment questionnaires reviewed in this chapter were the: Learning Environment Inventory, My Class Inventory, College and University Classroom Environment Inventory, Classroom Environment Scale, Individualized Classroom Environment Questionnaire, Science Laboratory Environment Inventory, Constructivist Learning Environment Survey, Computer Classroom Environment Inventory, Cultural Learning Environment Questionnaire, Distance and Open Learning Environment Scale, Socio-Cultural Environment Scale, What Is Happening In this Class?, and the Questionnaire on Teacher Interaction. From this list of available questionnaires, the WIHIC and QTI were chosen as being most relevant and suitable for my study. Through international use of the WIHIC and QTI, these questionnaires have been shown to be valid and reliable across numerous countries, and in a variety of content areas and grade levels.

Additionally this chapter presented an overview of some of the major types of applications of learning environment instruments, including the QTI and WIHIC, in past research. The literature was shown to contain examples of extensive applications of these two learning environment instruments and their value in a wide range of settings. And finally the potential connections between the existing applications of these instruments and their application to a study of the impact of professional development were discussed.

It is clear from the literature that the QTI and WIHIC are valuable learning environments research tools, but it is also clear that they might have untapped potential as instruments for evaluating the effectiveness of teacher professional development. It is this untapped potential that is investigated in my study.

Chapter 3 describes the methodology used in my study, including details about how the study was conducted.

Chapter 3

Research Methods

3.1 Introduction and Overview

The preceding chapter provided a review of the literature associated with the design and evaluation of teacher professional development, as well as a discussion of the literature dealing with research on learning environments and interpersonal behavior. In Chapter 2, it was noted that there is a strong connection between the classroom learning environment and student outcomes, both affective and academic (Brekelmans, Levy et al., 1993; Brekelmans, Wubbels et al., 1993; Marzano & Marzano, 2003). Additionally, it was noted in Chapter 2 that, as educational reform is gaining momentum, there is a greater need for a more robust system for evaluation of teacher professional development activities that includes assessment of student affective outcomes (Guskey, 2000; Sparks & Hirsch, 1997).

From the discussion in Chapter 2, it is evident that the field of learning environments research has an extensive array of widely applicable instruments (Fraser, 1998b). These instruments have been utilized in a broad range of settings for assessing student and teacher perceptions of the classroom learning environment. Keeping in mind the connection between the impact of the learning environment on student outcomes and the extensive array of instruments available, my study undertook to evaluate the *New Outcomes: Learning Improvement in Mathematics Integrating Technology* (NO LIMIT) teacher professional development project using the What Is Happening In this Class? (WIHIC) and the Questionnaire on Teacher Interaction (QTI).

As identified in Chapter 1, the four research aims that guided my study of the NO LIMIT project were:

5. To investigate the validity and reliability of the WIHIC and QTI in the context of Washington State middle-school mathematics classrooms.

6. To investigate whether there are differences between how students and teachers perceive the learning environment.

7. To investigate whether changes occur in the learning environment over the course of sustained, job-embedded professional development.

8. To investigate whether boys and girls differ in their perceptions of the learning environment.

This chapter contains a discussion of the research methodology used in my study, including the population and sample size, data-collection procedures, the rationale behind the selection of survey instruments, and how the data was analyzed. The following section titles and numbers provide an overview of the chapter:

3.2 Research Design
3.3 Population and Sample
3.4 Data Collection
3.5 Instruments
    3.5.1 What Is Happening In this Class? (WIHIC)
    3.5.2 Questionnaire on Teacher Interaction (QTI)
3.6 Data analysis
    3.6.1 Reliability and Validity
    3.6.2 Perception Differences
3.7 Qualitative Investigation
3.8 Chapter Summary.

3.2 Research Design

Following in the footsteps of the many researchers who have adopted the approach of the *bricoleur* (i.e. creating things from existing materials), my study utilizes mixed methods (quantitative and qualitative) to evaluate the effectiveness of the NO LIMIT project (Aldridge et al., 1999; Denzin & Lincoln, 2005; R. B. Johnson & Onwuegbuzie, 2004). This approach to educational research has been widely recommended and the benefits continue to gain acceptance (G. Anderson, 1998; Punch, 1998; Tobin & Fraser, 1998b). In fact, as Kelly and Lesh point out:

> We are now at a point where the growing maturity of mathematics and science education research has shifted attention from strict adherence to traditional experimental methods as the best path to scientific insight to renewed interest in the development of alternative methods for research. In the past few number of decades, educational researchers have moved into school systems, classrooms, and workplaces and have found a complex and multifaceted world that they feel is not well described by traditional research techniques. In the past, educational phenomena derived their status by surviving a variety of statistical tests. Today, nascent educational phenomena are accorded primacy, and the onus is on research methods to describe them in rich and systematic ways. (2000, p. 35)

The quantitative component of my study was conceptualized in terms of 'focal' and 'reference' (Wiliam, 1998) variables, and the relationship between them, using a 'pretest-posttest' research design. The 'focal' variable consisted of the student and teacher perceptions of the classroom psychosocial environment, while the 'reference' variable was the ongoing teacher professional development provided as a part of the NO LIMIT project. Quantitative data were obtained from the pretest and posttest administrations of the WIHIC and QTI learning environments instruments.

Given that an increasing number of educational researchers are finding that "[m]any research questions and combinations of questions are best and most fully answered through mixed research solutions" (R. B. Johnson & Onwuegbuzie, 2004, p. 18), my study incorporated a qualitative component as part of the mixed-method approach. The qualitative aspect of my study employed a narrative case-study approach involving both classroom observations and teacher interviews. The interweaving of the researcher's classroom observations, teacher interviews, and student perceptions (i.e. triangulation) "opens up culturally rich ways that interviewers and interviewees generate plausible views of the world" (Silverman, 2000, p. 823).

In addition to mixing quantitative and qualitative methods, my study also utilized the construct of 'grain size' as a strategy for closer examination of learning environments (Aldridge & Fraser, 2000; Fraser, 1999; Fraser & Tobin, 1991). The quantitative instruments utilized a large sample size (grain size), while the narrative case-studies relied on the much smaller sample size found within individual classrooms. Actual sample sizes are discussed in Section 3.3. Quantitative results are presented in Chapter 5 for both grain sizes.

Both Fraser (1998a) and Guskey (2000) have pointed out the need for a more complete picture of what takes place in the classroom. As Fraser points out, more information is needed about the "subtle but important aspects of classroom life" (1998a, p. 556) in order to make decisions that impact students. In a parallel manner, Guskey suggests that the evaluation of teacher professional development needs to also incorporate measures of the impact on the affective outcomes of students, such as their "perceptions about learning or school in general" (2000, p. 213). In the same way that Fraser and Guskey suggest that multiple sources of information are needed about the classroom, my study incorporated a mixed method approach in the evaluation of the NO LIMIT project.

3.3 Population and Sample

When the *New Outcomes: Learning Improvement in Mathematics Integrating Technology* (NO LIMIT) project was initiated in 2001, more than 200 teachers were selected by competitive application to receive training and support over a two-year period. These approximately 200 teachers represented a typical distribution across the state of Washington. Schools from urban areas to remote and rural areas were proportionally represented, as were schools of different sizes.

In Years Three and Four of the project, approximately 250 new teachers were involved. These new teachers represented the same geographic, community size, and school size distributions as were present in Year One and Two. The "NO LIMIT! project has impacted approximately 24,000 students per year across Washington State, including those in urban and rural areas, from varied cultural, ethnic, and socioeconomic backgrounds, and also learning disabled children" (Popejoy et al., 2005).

From the entire sample of participating teachers across the state a nonprobability convenience sample of participating teachers was taken. This sample included 27 teachers and their 63 associated classrooms. Within the Educational Service District 113 (ESD 113) service area, there were 25 participating teachers, representing 58 classrooms. These teachers' schools also represented a large range of size. The smallest participating school had approximately 21 students per grade, while the largest had over 300 students per grade. All the other schools were evenly distributed between these limits. The two schools participating in the study from outside the ESD 113 service area were from geographically distant areas, but both fell within the given school size range.

The first research aim involves examination of the reliability and validity of the WIHIC and QTI questionnaires in the context of the participating middle-school mathematics classrooms in Washington State. For this portion of the study, the entire sample of 63 classrooms with 1212 students and 27 teachers was selected. The large size of the sample is advantageous in that it contributes to the generalizability of the results.

For the quantitative examination of the differences between student and teacher perceptions of teacher interpersonal behaviors, a subsample of 44 classes was utilized. This sample for Research Aim 2 was chosen from the larger pool in order to provide matched pairs of teachers' scores and class means of students' scores on the actual and preferred forms of the QTI (see Section 3.4).

The part of my study examining the changes in perception of learning environment and teacher interpersonal behavior over the course of the teacher professional development project involved a subsample of 57 classrooms with 879 students. In order to conduct the pretest-posttest analysis, only those classrooms with both pretest and posttest data could be used. The reason that all 63 of the original classrooms could not be used was due to incomplete data for a classroom either at the pretesting phase or posttesting phase. Despite the loss of some classrooms, the sample size was still quite large.

Research Aim 4 deals with gender differences in perceptions of the learning environment and teacher interpersonal behaviors. For this research aim, a subsample of 56 classes was chosen and analyzed. For the 56 classrooms, a within-class mean was utilized as the unit of analysis. See Section 5.2 for a description of the within-class mean.

For the qualitative portion of my study, observations were made in 21 teachers' classrooms, from the 25 teachers participating within the ESD 113 region. From the 21 teachers, a nonprobability convenience sample of 15 teachers was selected for sustained observations, coinciding with the duration of the NO LIMIT project of a school year. All teachers were given the opportunity to participate in the interview portion of the study. A subsample of ten teachers agreed to be interviewed and, of those, seven completed the interview process.

3.4 Quantitative Data Collection

3.4.1 Ethical Considerations

In preparation for conducting my research, approval of participating schools, teachers and students was requested and gained. Initially participating schools' principals were contacted and permission was requested to conduct the research in their buildings. All principals granted permission. In accordance with the ethics guidelines for research at Curtin University of Technology (Curtin University of Technology, 2007), a research consent form was then distributed to all participating teachers. This consent form outlined the goals of the research and provided the assurances of anonymity (see Appendix A). The consent form was printed on carbonless copy paper, allowing the teachers to retain a copy of their signed form.

To provide the students and their parents with the same level of information about the research and assurances, sufficient copies of a parent information letter, with the option of not participating in the research (see Appendix B), were provided to teachers to distribute to all participating students. Before the research began, all teachers were contacted regarding any students who were not going to participate. I was informed by the teachers that all students would be able to participate in the administration of the questionnaires. To further protect participants' confidentiality all data were entered in the databases with actual names being replaced with numerical identification codes.

3.4.2 Administration of Surveys

Collection of data about teacher perceptions of their interpersonal behaviors in class took place during the 'kick-off' session prior to the beginning of the school year. Along with the aims of the research, the teachers were presented with the research consent form. On the consent form was a unique identification number which was then used throughout the year with that teacher. Teachers then placed the identification number on the QTI questionnaire and were provided sufficient time to complete the questionnaire during the afternoon session. All completed questionnaires were collected.

Following administration of the questionnaire to the teachers, a set of instructions for how to administer the questionnaires was discussed with the teachers. They were also told that the same set of instructions would arrive with their class sets of questionnaires following the start of the school year, and again near the end of the school year. Having the opportunity to present the instructions to the entire group of participating teachers was not only efficient but also allowed them to ask any clarifying questions that they had in a timely manner.

For those teachers who were unable to attend the 'kick-off', the consent form, copies of the QTI and an explanatory cover letter were mailed to them. Upon return of all questionnaires, the data were entered into an Excel spreadsheet in preparation for analysis.

As my research design included a student 'pretest' phase, the student questionnaires were printed and prepared for delivery to schools near the beginning of the school year. As suggested by other researchers (Levy, den Brock, Wubbels, & Brekelmans, 2003; Wubbels, Creton et al., 1993), administration of the 'pretest' was delayed until students had been in class approximately six weeks and had had a chance to become familiar with their teachers. The administration took place in approximately the second week of October. Sufficient copies of the questionnaires were delivered to participating teachers using the existing ESD 113 courier service or the postal service. Teachers were asked to complete and return the questionnaires via the ESD 113 courier service, or postal service, within a two-week window. With the assistance of a reminder email, most questionnaires were returned within the prescribed time frame.

Upon receipt of the completed student questionnaires the data were entered into an Excel spreadsheet, by hand, in preparation for analysis. The process of hand entering was completed by myself and took approximately two months. Questionnaires which were substantially incomplete were marked for exclusion and not entered into the data file. Double marked items and blank items on questionnaires were entered as a '2' as a mid-range indicator. Data from randomly-selected questionnaires were re-entered into a data file and cross-checked with the original entries for accuracy.

For the posttest phase of the research teachers were polled regarding how many students were currently enrolled in their participating classes. Accordingly a second set of student questionnaires was printed and prepared for delivery to the teachers near the end of the school year. The questionnaires were scheduled to arrive by ESD 113 courier, or the postal service, in early May. Teachers were requested to administer the 'posttest' version to the students and return the finished questionnaires to ESD 113 by the end of May. With the assistance of a reminder email, most surveys were returned by approximately the first of June.

Following the same protocol that was established for the 'pretest' phase, the data from the 'posttest' questionnaires were entered into a new Excel spreadsheet in preparation for analysis. Again the data entry required approximately two months for to complete.

At the final NO LIMIT project gathering of the school year, the teachers were thanked for their participation. The offer was again extended that, when data analysis was complete, they could request to see a summary of the results for their classes.

3.5 Instruments

The instruments chosen for my study were a modified version of the WIHIC and the QTI. These instruments were chosen owing to their close alignment to the goals and desired outcomes of the NO LIMIT project. As the project directors describe it, the NO LIMIT "project combines high stakes professional development [professional development occurring in a high stakes assessment environment] and consequent teacher changes, successful implementation of national and state standards in mathematics, with appropriate infusion of technology to improve student learning" (Popejoy et al., 2005, p. 1). As defined in the NCTM standards, implementation of the standards involves applying the "Principles for School Mathematics" (2000, p. 10). Included in these principles is the 'Teaching Principle', a part of which indicates that "effective teaching requires a challenging and supportive classroom learning environment" (2000, p. 18). The two instruments, and their modifications, were chosen with these principles in mind and will be discussed in Section 3.5.1 and 3.5.2.

3.5.1 What Is Happening In this Class? (WIHIC)

As described in Chapter 1, the ESD 113 NO LIMIT project goals included a sharp focus on the original six 'Professional Standards for Teaching Mathematics' (National Council of Teachers of Mathematics, 2000):

- Worthwhile mathematical tasks
- The teacher's role in discourse
- The student's role in discourse
- Tools for enhancing discourse
- The learning environment
- The analysis of teaching and learning. (p. 17)

The What Is Happening In this Class? (WIHIC) learning environment questionnaire was chosen because it contains scales pertinent to the goals of ESD 113's version of the NO LIMIT project.

In addition, the WIHIC was chosen on the basis of its strength and extensive applicability. As stated in Chapter 2, the WIHIC was developed by Fraser, Fisher, and McRobbie in 1996 to bring together the best scales for measuring seven key psychosocial dimensions of the classroom learning environment: Student Cohesiveness, Teacher Support, Involvement, Investigation, Task Orientation, Cooperation, and Equity. The WIHIC is comprised of two versions which measure what the students perceive to be the *Actual* learning environment or their *Preferred* learning environment. Each version uses a five-point frequency response scale ranging from Almost Never to Almost Always.

The validity and applicability of the WIHIC have been demonstrated across many countries and in a wide variety of contexts. In one of the largest recent studies involving the WIHIC, Dorman (2003) used the WIHIC with nearly 4000 mathematics Grade 8, 10 and 12 students in Australia, Canada and the UK. Through reliability, exploratory factor analysis, and confirmatory factor analysis, Dorman found that the scales of the WIHIC had good internal consistency, good discriminant validity and were able to discriminate between within-school grade groups. In addition to Dorman's (2003) extensive study using the WIHIC, it has been used successfully in Australia (Dorman et al., 2002), Canada (Dorman et al., 2002; Raaflaub & Fraser, 2002; Zandvliet & Fraser, 2005), Indonesia (Margianti et al., 2004), Singapore (Fraser & Chionh, 2000), and the United States (Allen & Fraser, in press; Martin-Dunlop & Fraser, in press; Pickett & Fraser, 2004), to name a few. The studies cited above also represent the WIHIC in use at a range of grade levels and in both mathematics and science classrooms. These results, along with the others discussed in Section 2.3.3, clearly indicated that the WIHIC has validity and applicability in mathematics classrooms.

Because of its validity, applicability, and wide acceptance in learning environments research in mathematics classrooms, and because of the close alignment with the classroom indicators selected for the NO LIMIT project as illustrated in Table 3.1, I selected the WIHIC to assess the learning environment. To confirm that it was indeed valid and reliable in the context of Washington State middle-school classrooms, one of my research questions focused on validating the WIHIC. These results are discussed in Chapter 4.

For the purposes of my study, I chose to alter the original form of the WIHIC slightly by deleting two of the original seven scales. There were two reasons for reducing the number of scales from the original seven to five. The first reason was that the Equity scale and the Cooperation were not as closely aligned with the focus of the professional development with the teachers. We had chosen to place an emphasis on 'student engagement' (e.g. students engaging with the mathematics, students engaging with each other, and students and teachers engaging with each other). The second reason for reducing the number of scales was to minimize the disruption of instructional time caused by the administrations of the instruments. This five-scale version of the WIHIC will henceforth be referred to as the 'modified version'.

Table 3.1 presents a summary of the five scales retained in the modified version of the WIHIC used in my study. The center column contains a sample item from each of the retained scales and right-hand column contains sample indicators drawn from the ESD 113 Classroom Observation Tool (see Appendix D). It was the rich connection between the indicators that I was using in the NO LIMIT classrooms with teachers and the items assembled for the five WIHIC scales that led me to choose the WIHIC as one of the learning environment instruments in my study. The modified version of the WIHIC used in my study is found in Appendix C.

3.5.2 Questionnaire on Teacher Interaction (QTI)

Previous studies have shown that student achievement depends not only on teacher knowledge and skills, but also on teacher behavior (American Educational Research Association, 2005; Danielson, 2002; Schoen, Cebulla, Finn, & Fi, 2003). The learning environment research cited in Chapter 2 (Section 2.3.5.1) clearly demonstrates that both academic and affective student outcomes are influenced by the learning environment established in the classroom. A clear component of that environment is the teacher's interpersonal behavior. With this in mind, the Questionnaire on Teacher Interaction (QTI) was chosen as the second instrument in my study.

As described in Section 2.3.4, the QTI was developed by Wubbels and Levy in the 1980s to assess student perceptions of the classroom learning environment from the perspective of interpersonal behaviors. This instrument has its origins in the work by Leary (1957) who tried to systematize the description and analysis of human interactions. The perceptions measured by the QTI are grouped according to eight different behavioral constructs (Leadership, Helpful/Friendly, Understanding, Student Responsibility/Freedom, Uncertain, Dissatisfied, Admonishing, and Strict behavior).

The QTI has gained a reputation as an internationally validated learning environment instrument which has undergone extensive use in the Netherlands, Australia, and the United States (Wubbels & Levy, 1993). The QTI has been utilized in both mathematics (Goh & Fraser, 1998; Kyriakides & Muijs, 2005; Quek et al., 2005b) and science classrooms (Fisher et al., 1996; Lee et al., 2003) across a variety of grade levels. Internationally the QTI has been utilized in countries as diverse as Korea, Malayasia, and the United States (Lee & Kim, 2002; Scott & Fisher, 2004; Wubbels & Levy, 1991). In each of these settings, the QTI was found to be a valid and reliable instrument (see Section 2.3.4 for a more complete discussion of the QTI). While the reliability of the QTI in assessing key components of the learning environment led, in part, to its selection as an instrument in my study, I did conduct a validity and reliability analysis as reported in Chapter 4.

In addition to its reliability, the QTI was chosen because the teacher interpersonal behavior constructs measured closely match the intended outcomes of the NO LIMIT professional development project. As identified in the ESD 113 Classroom Observation tool (see Appendix D), a focus was placed on such classroom indicators as Classroom Culture, Worthwhile Mathematical Tasks, and Students' Role in Discourse. Changes in these classroom indicators should require a change in some of the teachers' interpersonal behaviors. It was on this additional basis that I chose the QTI as a second measure of the classroom learning environment.

Much like the WIHIC, the participants responded to the 48 items of the QTI (see Appendix C) using a five-point frequency scale (Almost Never, Seldom, Sometimes, Often, Almost Always). Also like the WIHIC, the QTI is available in both an *Actual* and *Preferred* form. I chose to use both the *Actual* and *Preferred* forms to enhance comparison of student and teacher perceptions of teacher interpersonal behaviors.

3.6 Data Analysis

All quantitative data were entered into Excel spreadsheets. A separate spreadsheet for identification information was maintained so that the confidentiality of the participants was protected. Prior to analysis, students' responses were matched so that the pretest–posttest analysis could be most efficiently conducted. Quantitative analysis associated with the four research aims was focused on two primary targets: reliability and validity (Research Aim 1) and differences between various subgroups (Research Aims 2 – 4). The following subsections described the analytic approaches used with these two targets.

3.6.1 Validity and Reliability of WIHIC and QTI

The first goal of my research was determine the reliability and validity of the WIHIC and QTI in the context of Washington State middle-school mathematics classrooms. The reliability of the instruments can be defined in terms of their internal consistency (e.g. the extent to which items within a given scale are measuring a common concept), while validity for the instruments can be considered to be the extent to which scales measures what they claim to measure. The following sections describe the methods used to accomplish this goal.

Internal consistency reliability for both the WIHIC and QTI were determined through the use of the Cronbach alpha coefficient (McKnight, Magid, Murphy, & McKnight, 2000). Cronbach's alpha has an upper limit of 1.0, with higher values indicating higher scale reliability (i.e. the more likely the scale is measuring a single underlying construct). The alpha coefficient was calculated for both the individual and class mean as the unit of analysis. Results for internal consistency reliability are presented in Section 4.2.2 for the WIHIC and Section 4.3.1 for the QTI.

One-way analysis of variance (ANOVA) was used to determine the discriminant validity for both the WIHIC and QTI. The eta^2 statistic from the ANOVA was utilized in determining whether each scale was able to differentiate between perceptions of students in different classrooms. Class membership was used as the independent variable. The eta^2 statistic is a measure the proportion of variance explained by class membership and is reported as the ratio of 'between' to 'total' sums of squares. Results on ability to differentiate between classrooms are reported in Sections 4.2.2 and 4.3.1.

A factor analysis was carried out to assess the internal structure for the WIHIC. Principal axis factoring with Oblimin rotation and Kaiser normalization was performed for both administrations of the instrument (i.e. pretest and posttest administration). Factor loadings for all 40 items were examined to determine that they were above the 0.40 threshold with their own scale, and only their own scale. Eigenvalues and total percentage of variance were also calculated as further measures of the structure of the instrument. Results for the WIHIC factor analysis are reported in Section 4.2.1.

Table 3.1
Description of Scales, a Sample Item for Each Scale of the Modified WIHIC, and Related Indicators from ESD 113 Classroom Observation Tool

Scale name	Description of scale	Sample item	Sample of Related Indicators
Student Cohesiveness	Extent to which students know, help and are supportive of one another.	I know other students in this class.	• Students question one another. • Students pay attention while another student is speaking. • Students respect each other students' thinking.
Teacher Support	Extent to which the teacher helps, befriends, trusts and is interested in students.	The teacher takes a personal interest in me.	• The teacher establishes an environment where students feel comfortable asking for help, seeking solutions, and learning from mistakes. • The teacher dignifies errors. • The teacher provides immediate, specific, and positive feedback.
Involvement	Extent to which students have attentive interest, participate in discussions, perform additional work and enjoy the class.	I explain my ideas to other students.	• Students are engaged with the tasks. • Students communicate about the math tasks at hand. • The teacher encourages participation of all students.
Investigation	Emphasis on the skills and processes of inquiry and their use in problem solving and investigation.	I carry out investigations to test my ideas.	• Students are engaged with the tasks. • Students look at problems and ideas in different ways. • Students are presenting and modeling their work.
Task Orientation	Extent to which it is important to complete activities planned and to stay on the subject matter.	I pay attention in this class.	• Students are engaged with the tasks. • Students pay attention while another student is speaking.

Based partly on Dorman (2003, p. 234)

Because the QTI is based on a two-dimensional circumplex model (see Section 2.3.4) for teacher interpersonal behavior, factor analysis is not appropriate. The eight scales within the QTI circumplex model are theoretically correlated with each other. To examine the strength of the structure of the QTI, an interscale correlation analysis was performed. High internal consistency would be indicated by adjacent scales having high positive correlations and opposing scales displaying high negative correlations. Results of the interscale correlation analysis are reported in Section 4.3.2.

3.6.2 Student-Teacher, Pretest-Posttest and Gender Differences

Once the reliability and validity of the instruments had been established for my context, the remaining research questions would be answered through analysis of the quantitative data for differences in perceptions of the learning environment. Research Aim 2 is concerned with differences between how students and teachers perceive the learning environment and teacher interpersonal behaviors, Research Aim 3 involves differences between the beginning and end of the professional development project, and Research Aim 4 addresses gender differences in perceptions of the learning environment.

For Research Aim 2 dealing with differences between student and teacher perceptions, Multivariate Analysis of Variance (MANOVA) was used. The five WIHIC scales and the eight QTI scales were used as the dependent variables in two separate analyses. The student/teacher served as the repeated-measures independent variable. When MANOVA results indicated statistically significant differences using Wilks' lambda criterion for the set of dependent variables as a whole, the corresponding ANOVA was interpreted for each scale. In addition, the effect size was calculated, to assist in determination of whether a difference might be educationally significant. The effect size was calculated by dividing the difference between two means by the pooled standard deviation, thus expressing the difference in the means of two samples in standard deviation units (J. Cohen, 1988). Results for Research Aim 2 are reported in Section 5.3.

Research Aim 3, dealing with pretest–posttest differences, was handled in much the same manner as Research Aim 2. For this research aim, the individual was selected as the unit of analysis. For both the WIHIC and QTI, MANOVA was applied and, when tests indicated significant overall pretest-posttest perception differences, the univariate ANOVA was interpreted for each scale. Again effect sizes were calculated in the same manner as for the analysis for Research Aim 2. Results for this research aim are reported in Section 5.4.

Differences between male and female students' perceptions of the learning environment (Research Aim 4) were analyzed using the same approach as the previous two research questions. The within-class gender mean was chosen as the unit of analysis because its use reduces the possible confounding when males and females are represented in different proportions in a classroom. The process of obtaining a within-class mean involves calculating for a class a mean for males and a separate mean for females, thus producing a matched pair of means that minimizes differences caused by unequal numbers of males and females within a class. MANOVA/ANOVA and effect size results for Research Aim 4 are reported in Section 5.2.

3.7 Qualitative Investigation

The role of a mixed-methods approach to research has been established by many researchers, as discussed in Section 3.2. Within the mixed-methods approach is the qualitative dimension. As Denzin and Lincoln (2005) point out, while both approaches to convey worthwhile information, there are five key differences between quantitative and qualitative approaches.

The first difference that Denzin and Lincoln (2005) point out is between the philosophical foundations underlying the positivist and post-positivist approaches to 'knowing'. Positivists would say that there is a truth out there that needs only be captured and described, while the post-positivist contends that we can only approximate the truth. Within the postpositivist stance is the need to examine reality from as many perspectives as possible.

The second difference is the acceptance of postmodern “sensibilities” (2005, p. 11) by qualitative researchers. The goal of research is to tell the story of what is taking place within a given context. The postmodern approach contends that it is telling a different story of the situation, not a better or worse story, just a different story. If research is to provide us with a more complete picture of classrooms, then more diverse perspectives provide a richer story.

More comprehensive inclusion of the viewpoint of the individual is the third difference identified. Often quantitative research relies on the collection of information from a large sample, and only then are inferential judgments made about individuals. While the information about the sample is useful at that level, a qualitative approach allows the research to capture the perspective of the individual and thus deepen the potential for understanding the story more completely.

The fourth difference as described by Denzin and Lincoln (2005) is that qualitative research is able to incorporate the constraints placed on a context, such as all of the realities of everyday life in a typical classroom. Oftentimes quantitative research examines only a small slice of the story, while the larger lens utilized by qualitative research allows for the inclusion of the details of specific cases within the context.

The final difference is that qualitative research involves the acquisition of as complete and rich a description of the sample under study as possible. It is an understanding of the richness of the context under study that is important to the qualitative researcher, as compared with the quantitative researcher’s desire to know about a small component in great detail.

Continuing with the theme of *bricoleur* begun earlier, it is evident that the inclusion of qualitative data will enrich the tapestry that is the classroom environments in the NO LIMIT project and that began with the quantitative results from the WIHIC and QTI questionnaires. To collect the qualitative data necessary to more fully tell the

NO LIMIT story, my research included both classroom observations and participant interviews.

Classroom observations were conducted as a part of my responsibilities with the NO LIMIT project. Each of the 15 classroom teachers was visited approximately 4 times a month during the 2003–2004 school year. Visits ranged from observing one mathematics class to spending a full day with a teacher depending on the teaching assignment for the particular teacher. Initially observations of classrooms were made with the goal of getting to know the classroom settings, the students, and the teachers' styles.

Additional classroom visits included general observations as well as observations directed toward the teachers' chosen areas for growth. General observations notes were kept in a journal along with notes on the ESD 113 Classroom Observation Tool (see Appendix D). Notations made on the observation tool were shared with teachers in post-observation discussions as both an accuracy check and as a source of information to further inform teacher practice.

Near the end of the school year, participating teachers were given the opportunity to participate in a follow-up interview. Interview by email was chosen as the interview method. Meho (2006) has identified many benefits and challenges to the use of email interviews. Included in the benefits are the cost effectiveness, efficiency associated with the asynchronous nature of email, access to remote participants, feeling of anonymity, and the time for participants to craft a careful response. Challenges included potential loss of rich face-to-face non-verbal cues, respondent's verbal versus written communication skills, and lack of ability to provide immediate feedback or clarification. Given the time available and geographic distribution of participants, the decision was made that the benefits of conducting the follow-up interviews by email outweighed the challenges.

From the sample of participating teachers who were invited to participate in an email interview, 10 agreed to proceed. A set of seven interview questions was emailed to the 10 participants as a word processing document (see Appendix E for interview questions). Following a reminder email, seven of the ten participants replied with their responses. Their responses were coded to assure anonymity and saved electronically for use in the case-study analysis. Additional emails were sent out to participants where clarification of a particular response was desired or needed. These responses were compiled with the original response and saved electronically.

The data collected from the classroom observations and email interviews were utilized in the construction of case studies. As detailed above, I chose to use case studies as a means of telling a richer story of the classrooms in which the students and teachers lived five days out of seven. While I wrote individual case studies as a process for capturing the viewpoint of the individual classrooms, the collection of case studies assists with the understanding of the mosaic that is the NO LIMIT project.

In his chapter on qualitative case studies, Stake (2005) identifies three basic types of case studies. The first is the *intrinsic case study*, which focuses on one particular case because of its singular importance. The second type is the *instrumental case study*, which uses a particular case to bring light to a particular circumstance or setting rather than the case itself. The final type is the *collective case study*, which utilizes several cases to illuminate a much larger context than any one of the individual cases. Within the framework provided by Stake (2005), I chose to employ a collective case study approach. And within the collective case study approach I chose extreme case sampling (Onwuegbuzie & Leech, 2007) because through a comparison of a range of classrooms the NO LIMIT story could best be understood.

3.8 Chapter Summary

The methodological foundation of my study is built on the premise that a mixed-method approach to the research questions will yield a much richer and colorful image of the complex mosaic that is the NO LIMIT project. Because the sample size is relatively large, the quantitative analysis has validity and reliability associated with it. Because the qualitative data draw from a diverse but representative subsample of the entire NO LIMIT project, it has the potential to provide the complementary perspectives necessary for a more complete understanding of the story.

This chapter presented the composition and characteristics of the sample of my study (i.e. the approximately 1200 Washington State middle-school mathematics students and their 27 associated teachers). In addition, this chapter has presented the unique mixed-methods approach to the evaluation of a job-embedded, sustained professional development project. Within those mixed methods, this chapter has identified that both qualitative and quantitative approaches were used, as well employing the construct of 'grain-size' in looking at the quantitative results both at the *entire sample size* level and then again at the *classroom size* sample level. While traditional quantitative analyzes such as MANOVA/ANOVA and effect sizes were detailed, the qualitative narrative case study approach was also described as the approach to more closely approximating the 'real' story of the NO LIMIT professional development project.

Development of the narrative case studies incorporated not only my observations of the classroom, but those of the teachers and students. The teachers' stories were constructed from classroom conversations that I had with the teachers and their responses to the follow-up interview questions. The student 'voice' came not only from my observations in the classroom, but also from the classroom-level perception data based on the WIHIC and QTI. The case studies were a blend of qualitative and quantitative observations from the classrooms. This method truly represents bricolage.

Chapter 4 reports in detail the analysis of the quantitative results for the purpose of establishing the validity and reliability of the WIHIC and QTI in the context of my study.

Chapter 4

Reliability and Validity

4.1 Introduction

The previous chapter presented the methodological approaches used in this study and the associated research aims. The research instruments selected, and their reliability and validity in prior studies, were also described in detail. This chapter presents results for my first research aim concerning the reliability and validity of the modified questionnaires (WIHIC and the QTI) when used in my study, by a sample of approximately 1200 Washington State middle-school mathematics students. Results for the QTI are reported for preferred perceptions, as well as for the actual form.

The first aim of this study was to establish the reliability and validity of the two instruments chosen to examine the learning environment in middle-school mathematics classrooms. The value associated with the analyses described below is two-fold. First, the availability of reliability and validity information for an instrument allows the researcher to have increased confidence in drawing conclusions based on data collected using this instrument. And, second, contributing a unique set of reliability and validity data to the already-extensive accumulation of information on these instruments further enhances confidence in these tools, allowing future researchers to apply them in new and different ways.

In the following two sections, the analysis of the validity and reliability for the two instruments is discussed. Section 4.2 discusses the analysis of the five-scale, 40-item version of the WIHIC and Section 4.3 discusses the analysis of the eight-scale, 48-item version of the QTI.

4.2 Validity and Reliability Analyses for WIHIC

The version of the What Is Happening In this Class? (WIHIC) questionnaire used in this study is a modification of the original form created in the late 1990s . The most widely-used form of the WIHIC contains 56 items covering seven scales. The WIHIC was chosen because of its parsimonious approach to assessment of classroom learning environments and because its scales closely align with most U.S. middle-school reform mathematics curricula and pedagogy. In addition, the WIHIC was chosen because it has been found to be valid and useful in a variety of contexts, including in the United States (Aldridge et al., 2006; Zandvliet & Fraser, 2005), in mathematics (Margianti et al., 2004; Raaflaub & Fraser, 2002), and at a variety of age levels (Dorman, 2003).

Of the original seven scales in the WIHIC, five of the scales (Student Cohesiveness, Teacher Support, Involvement, Investigation and Task Orientation) were selected as being most salient for inclusion in this study. These five scales were selected because they most closely align with both the desired learning environment and the intent of the professional development work being conducted in the classrooms in which my study took place. Refer to Chapters 2 and 3 (2.3.3 and 3.5.1) for additional information on the WIHIC and my reasons for selecting it. A copy of the wording of items in my version of the WIHIC can be found in Appendix C.

As described in Chapter 3, my methodology involved the use of a pretest and posttest administration of the WIHIC to the entire sample. Section 4.2.1 discusses the factor structure analysis for both administrations of the WIHIC, while Section 4.2.2 discusses the internal consistency reliability and ability to discriminate between classes for both administrations of the WIHIC.

4.2.1 Factor Structure of WIHIC

To study the internal structure of the modified version of the WIHIC with 40 items in five scales, a factor analysis was conducted using principal axis factoring with Oblimin rotation and Kaiser normalization. Results are presented separately for two administrations, namely, one before the sustained professional development began (pretest) and the one near the end of the school year (posttest). Results of this analysis for 1212 students in 63 classes on the pretest and 1054 students in 57 classes on the posttest are contained in Table 4.1.

Only factor loadings greater than the generally-accepted value of 0.40 are recorded in Table 4.1. The percentage of variance extracted and eigenvalue for each factor are also reported at the bottom of Table 4.1. In the factor analysis, an item was retained only if two conditions were satisfied. First, the factor loading of an item with its *a priori* scale had to be at least 0.40. Second, the factor loading of an item with each of the other four scales had to be less than 0.40. As Table 4.1 on the next page shows, each of the 40 items is my version of the WIHIC satisfied both of these conditions for both the pretest and posttest data. Therefore, all 40 items were retained in subsequent analyses.

The pretest percentage of variance ranged from 2.50% to 34.64% for different scales, with a total variance accounted for being 56.55%. Eigenvalues for those same scales ranged from 1.00 to 3.36. Posttest percentage of variance ranged from 2.64% to 34.19% for the different scales, with the total variance accounted for being 57.23%. Posttest eigenvalues ranged from 1.05 to 13.68 for the different scales.

The factor analysis results for the 40-item modified version of the WIHIC is consistent with results previously obtained (Aldridge & Fraser, 2000; Aldridge et al., 1999; Dorman, 2003; Zandvliet & Man, 2003, April). The *a priori* structure of the five-scale instrument was reproduced perfectly in that all of the items had factor loadings greater than 0.40 on their own scale and less than 0.40 on all other scales. From the results of this analysis, it can be concluded that the modified WIHIC used in this study can be used with confidence in testing the other research aims in my study.

Table 4.1
Factor Loadings for the WIHIC for Pretest and Posttest Data

Item No	Factor Loading									
	Student Cohesiveness		Teacher Support		Involvement		Investigation		Task Orientation	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
1	0.74	0.77								
2	0.66	0.61								
3	0.59	0.57								
4	0.86	0.86								
5	0.66	0.70								
6	0.43	0.41								
7	0.76	0.81								
8	0.54	0.60								
9			0.56	0.66						
10			0.73	0.80						
11			0.73	0.78						
12			0.62	0.66						
13			0.74	0.79						
14			0.77	0.82						
15			0.73	0.72						
16			0.62	0.66						
17					0.72	0.77				
18					0.82	0.89				
19					0.53	0.56				
20					0.72	0.71				
21					0.50	0.56				
22					0.57	0.64				
23					0.43	0.44				
24					0.43	0.43				
25							0.68	0.73		
26							0.56	0.67		
27							0.76	0.81		
28							0.62	0.67		
29							0.74	0.74		
30							0.79	0.82		
31							0.84	0.84		
32							0.81	0.76		
33									0.71	0.64
34									0.71	0.56
35									0.73	0.67
36									0.76	0.72
37									0.83	0.79
38									0.68	0.65
39									0.77	0.76
40									0.77	0.72
% Variance	5.99	8.68	5.02	6.68	2.50	2.64	34.64	34.19	8.40	5.04
Eigenvalue	2.40	3.47	2.01	2.67	1.00	1.05	13.86	13.68	3.36	2.02

Factor loadings less than 0.40 have been omitted.
N = 1212 students in 63 classes

4.2.2 Internal Consistency Reliability and Ability to Discriminate Between Classrooms for the WIHIC

The internal consistency reliability of the WIHIC was examined through the use of Cronbach's alpha coefficient. Internal consistency is an index of the extent to which items in the same scale assess a common construct. Table 4.2 presents the results of the analysis for the 40 items in the five scales of the WIHIC as used in this study. Reliability results are provided separately for pretest and posttest data and for two units of analysis, namely, the individual student and the class mean. Table 4.2 indicates that the alpha coefficients for the five scales are high for both the pretest and posttest administrations of the instrument. Using the class mean as the unit of analysis consistently yielded higher alpha reliability values than with the use of the individual as the unit of analysis.

The pretest alpha reliability values for different WIHIC scales ranged from 0.88 to 0.93 with the individual as the unit of analysis and from 0.92 to 0.96 for the class mean (Table 4.2). Posttest alpha reliability values ranged from 0.88 to 0.93 with the individual as the unit of analysis and from 0.92 to 0.97 for the class mean. These results are consistent with previous applications of the WIHIC (Aldridge et al., 1999; Fraser, 1998b). Taken together, these values support the internal consistency of the modified version of WIHIC as used in this study.

The ability of the actual form of each WIHIC scale to discriminate between perceptions of students in different classrooms is a desirable characteristic, and is often measured through the eta^2 statistic from a one-way analysis of variance (ANOVA) with class membership as the independent variable. ANOVA provides information about the extent to which students in the same class perceive the environment similarly, and the extent to which mean class perceptions vary from class to class. In Table 4.2, each scale's eta^2 statistic, which is the ratio of 'between' to 'total' sums of squares and represent the proportion of variance explained by class membership, show a significant difference ($p<0.01$) between students in different classes for each of the five scales in the modified WIHIC for both the pretest and posttest data. The values of the eta^2 statistic range from 0.10 to 0.18 on the pretest and from 0.11 to 0.22 on the posttest. Again these results are consistent with previous studies (Fraser, 1998b; Rickards, Bull, & Fisher, 2000) involving the WIHIC. These results indicate that the scales were able to differentiate between classes.

Table 4.2
Internal Consistency Reliability (Cronbach Alpha Coefficient) for Two Units of Analysis and Ability to Differentiate Between Classrooms (ANOVA Results) for the WIHIC for Pretest and Posttest

Scale	Unit of Analysis	Alpha Reliability		ANOVA eta^2	
		Pretest	Posttest	Pretest	Posttest
Student Cohesiveness	Individual	0.88	0.88	0.10**	0.11**
	Class Mean	0.92	0.92		
Teacher Support	Individual	0.90	0.92	0.18**	0.22**
	Class Mean	0.96	0.97		
Involvement	Individual	0.89	0.90	0.11**	0.12**
	Class Mean	0.94	0.95		
Investigation	Individual	0.93	0.93	0.13**	0.15**
	Class Mean	0.96	0.97		
Task Orientation	Individual	0.92	0.90	0.11**	0.18**
	Class Mean	0.94	0.96		

** $p<0.01$ $N = 1212$ students in 63 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.

Section 4.2 has reported results related to the validity and reliability of the WIHIC in the context of my study among middle-school mathematics students in Washington State. Every WIHIC item was retained because its factor loading was greater than 0.4 with its own scale and less than 0.4 with each of the other four WIHIC scales. Further support for the validity of the WIHIC came from the total proportion of variance accounted for (56.55% for the pretest and 57.23% for posttest), along with all eigenvalues being greater than 1.0. Further, analyses indicated that each WIHIC scale was internally consistent at two levels of analysis, with Cronbach alpha coefficients all exceeding 0.92, and that the WIHIC was able to discriminate between classrooms, as indicated by the eta^2 statistic ranging from 0.10 to 0.18 across the two administrations.

Section 4.2 has discussed the validity and reliability of the WIHIC in my study. As with the WIHIC, I used different forms of the QTI and employed those forms both in a pretest and posttest format (refer to Section 3.5.2). In addition I utilized a Student Preferred form during the pretest administration. Section 4.3 discusses the reliability and validity analysis for both forms of the QTI and both administrations.

4.3 Validity and Reliability Analyses for the QTI

As was discussed in Section 3.5, my study examined the learning environment using two instruments, the WIHIC and Questionnaire on Teacher Interaction (QTI). Whereas Section 4.2 discussed the reliability and validity for the WIHIC, this section discusses the reliability and validity for the QTI. The QTI was chosen to accompany the WIHIC because of its emphasis on teacher interpersonal behavior. As pointed out by several authors (Levy, Creton, & Wubbels, 1993), teacher interpersonal behavior plays a significant role in creating the classroom learning environment and, as such, could be used as an indicator of changes in the classroom brought about by professional development efforts. For the reasons discussed in Section 3.5, and given the applicability and proven validity of the QTI in a range of countries, content areas and ages levels (den Brock, Bull, Fisher, & Rickards, 2006; den Brock, Fisher, Wubbels, Brekelmans, & Rickards, 2006; Fisher, Waldrip, & den Brock, 2005; Goh & Fraser, 2000; Rawnsley, 1997; Wubbels & Levy, 1991), the choice of this instrument is well supported.

All eight scales in the QTI (Leadership, Understanding, Helping/Friendly, Student Responsibility, Uncertain, Admonishing, Dissatisfied and Strict) were used in this study. Refer to Chapters 2 and 3 (Sections 2.3.4 and 3.5.2) for background information on the QTI and my reasons for including it as an instrument in my study. Appendix C contains a copy of the wording of the items.

While the sample of approximately 1200 middle-school mathematics students was the same throughout the study, differences between the two instruments required slightly different validation approaches. As discussed in Section 2.3.4, the circumplex nature of the QTI assumes that adjacent scales should most highly correlate, while scales on opposites sides should be highly negatively correlated (Levy et al., 1993; Wubbels, Creton et al., 1993). Owing to this structure, factor analysis was not relevant and was replaced with inspection of the pattern of interscale correlations, looking for flanking sectors being highly positively correlated and opposing sectors being highly negatively correlated. Refer to Section 2.3.4 for greater detail on the structure of the QTI. In addition to the ‘actual’ form of the QTI for teachers and students, a ‘preferred’ form for students alone was utilized in my study both as a research tool and a teacher reflection tool, thus adding to the depth of analysis.

Section 4.3.1 discusses the internal consistency reliability and ability to discriminate between classrooms for the QTI in my study, while Section 4.3.2 presents support for the validity based on interscale correlations for the QTI.

4.3.1 Internal Consistency Reliability and Ability to Discriminate Between Classrooms for the QTI

As with the WIHIC data, the internal consistency reliability of each QTI scale was estimated through the use of Cronbach’s alpha coefficient. Table 4.3 presents the results of this analysis for the 48 QTI items in eight scales, when administered to the 1212 middle-school mathematics students in Washington State. The table includes results for two administrations, one before the sustained professional development began (both pretest preferred form and pretest actual form) and the one given near the end of the school year (posttest actual form only). Also the reliability of QTI scales was estimated for two units of analysis, namely the individual and the class mean.

Table 4.3 shows that the alpha reliability values for different QTI scales for the pretest preferred form ranged from 0.67 to 0.87 for individuals and from 0.78 to 0.94 for class means. For the pretest actual form, the alpha reliability values for different QTI scales ranged from 0.54 to 0.86 for the individual as the unit of analysis and from 0.53 to 0.94 for class means. Reliability for the QTI scales on the posttest actual form ranged from 0.64 to 0.88 for individuals and from 0.64 to 0.96 for class means. In almost every case, the class mean reliability values exceed the values of reliability with the individual as unit of analysis.

Table 4.3
Internal Consistency Reliability (Cronbach Alpha Coefficient) for Actual and Preferred Forms for Two Units of Analysis and Ability to Differentiate Between Classrooms for Actual Form (ANOVA Results) for the QTI

Scale	Unit of Analysis	Alpha Reliability			ANOVA eta^2	
		Pretest Preferred	Pretest Actual	Posttest Actual	Pretest Actual	Posttest Actual
Leadership	Individual	0.72	0.78	0.83	0.31**	0.28**
	Class Mean	0.78	0.92	0.94		
Understanding	Individual	0.85	0.81	0.86	0.28**	0.31**
	Class Mean	0.94	0.94	0.96		
Helping/Friendly	Individual	0.87	0.86	0.88	0.29**	0.31**
	Class Mean	0.92	0.94	0.96		
Student Resp/ Freedom	Individual	0.78	0.54	0.64	0.14**	0.12**
	Class Mean	0.81	0.53	0.64		
Uncertain	Individual	0.79	0.69	0.73	0.17**	0.17**
	Class Mean	0.90	0.85	0.87		
Admonishing	Individual	0.67	0.69	0.75	0.27**	0.28**
	Class Mean	0.82	0.85	0.88		
Dissatisfied	Individual	0.84	0.75	0.80	0.18**	0.25**
	Class Mean	0.94	0.89	0.94		
Strict	Individual	0.74	0.60	0.65	0.17**	0.18**
	Class Mean	0.85	0.76	0.78		

** $p<0.01$ N = 1212 students in 63 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.

Based on previous applications of the QTI, these values are consistent with expectations (Fisher, Fraser, & Rickards, 1997; Fraser, 1998b; Kim et al., 2000). These values indicate high internal consistency.

As was the case with the WIHIC, the eta^2 statistic was used as a measure of the ability of the actual form of each QTI scale to distinguish between perceptions of students in different classrooms within the study. That is, within a given classroom, the students' perceptions of the learning environment should be similar and, at the same time, distinguishable from other classrooms. ANOVAs, with classroom membership as the main effect, produced a significant difference ($p<0.01$) between classrooms for each of the eight scales of the QTI, for both pretest and posttest, as shown in Table 4.3. The eta^2 values, which are measures of the proportion of variance due to class membership, range from 0.14 to 0.31 for the different QTI scales for the pretest actual form and from 0.12 to 0.31 for the posttest actual form. Not only are these results comparable to prior studies (Fisher et al., 1993; Wubbels & Levy, 1991), but also they suggest that the QTI is consistently capable of distinguishing between perceived student-teacher interactions in the different classrooms in this study.

4.3.2 Pattern of Interscale Correlations for the QTI

As explained in Section 2.3.4, the QTI is atypical as a learning environment instrument in that it is based on a two-dimensional circumplex model for interpersonal teacher behavior. The model has a Proximity dimension and an Influence dimension (Figure 2.1) that, when plotted on a coordinate grid, yield eight sectors, each representing a scale of the QTI. Owing to the circumplex nature of the model, the scales theoretically are not independent of each other, but rather are correlated with each other. Therefore, factor analysis, as used with WIHIC, is not applicable.

Theoretically, the interscale correlations between adjacent sectors should be highly positive, and between sectors that are opposite should be highly negative. Many studies have provided empirical evidence in support of this model (Wubbels, Creton et al., 1993).

As support for the validity of the circumplex character of the QTI in this study, Table 4.4 provides the results for the interscale correlations arising from the analysis of the pretest actual form, with the student as the unit of analysis. Using data from Table 4.4, Figure 4.1 illustrates the circumplex character of the model for the Helping/Friendly scale by graphically displaying how inter-scale correlations are highest for adjacent scales and are negative for those scales farther away. Using the individual as the unit of analysis results as an example, the Helping/Friendly scale is highly and positively correlated with the scales for Leadership (0.73) and Understanding (0.75) behavior, is highly and negatively correlated with Dissatisfied

(-0.52) behavior, and has modest correlations with the other four QTI scales (Student Responsibility & Freedom, Uncertain, Admonishing, and Strict).

Table 4.4
QTI Interscale Correlations for Student as Unit of Analysis for Pretest Actual Form

Scale	***Interscale Correlations***							
	LD	***UD***	***HF***	***RF***	***UC***	***DS***	***AD***	***ST***
LD leadership	***1.00***	***0.77***	***0.73***	***0.15***	***-0.44***	***-0.46***	***-0.47***	***-0.21***
UD Understanding		***1.00***	***0.75***	***0.23***	***-0.41***	***-0.48***	***-0.55***	***-0.31***
HF Helping/Friendly			***1.00***	***0.32***	***-0.40***	***-0.52***	***-0.46***	***-0.33***
RF Student Resp/Freedom				***1.00***	***0.21***	***0.14***	***0.03***	***-0.13***
UC Uncertain					***1.00***	***0.61***	***0.57***	***0.30***
DS Dissatisfied						***1.00***	***0.58***	***0.50***
AD Admonishing							***1.00***	***0.44***
ST Strict								***1.00***

N=121

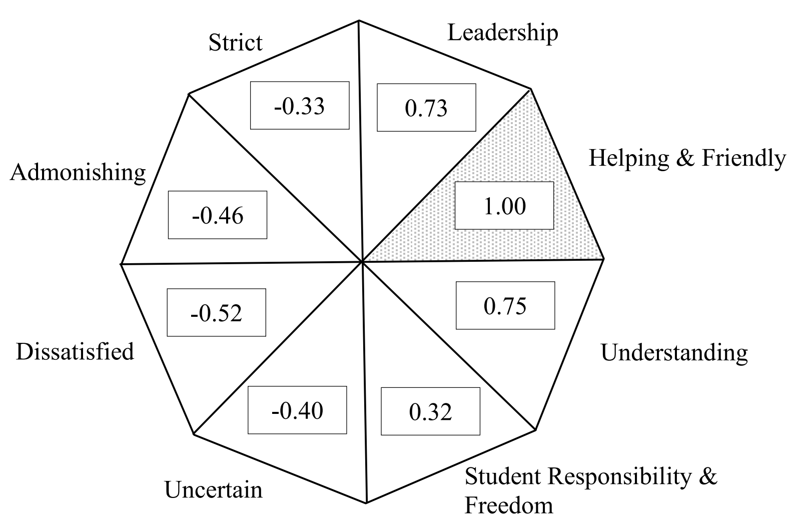


Figure 4.1 Correlations of the Helping/Friendly Scale with all Other QTI Scales

Once again, these results for the pattern of interscale correlations are consistent with previous studies (Fisher et al., 1996; Goh & Fraser, 2000) and support the validity of the QTI when used with middle-school mathematics students in Washington State.

4.4 Chapter Summary

This chapter has provided validity and reliability data for the two classroom learning environment instruments used in my study of over 1200 Washington State middle-school mathematics students. All scales of the 40-item modified What is Happening in This Classroom? instrument (WIHIC) and the Questionnaire on Teacher Interaction (QTI) were shown to have acceptable internal consistency reliability and the ability to discriminate between classrooms. These results are consistent with research previously published and support the use of these instruments in middle-school mathematics classrooms.

The *a priori* 5-scale 40-item structure for the modified WIHIC was confirmed as evidenced by the results of the factor analysis. This analysis indicated that all 40 of the items, in both the pretest and posttest form, yielded factor loadings of at least 0.40. Also the five WIHIC scales together, for either the pretest or posttest form, accounted for greater than 50% of the variance, and all scales had eigenvalues greater than 1.

The high values for the scale alpha coefficients for both the WIHIC and QTI (ranging from 0.88 to 0.93 for the WIHIC and from 0.64 to 0.88 for the QTI for posttest form each, with the individual as the unit of analysis) indicate that the WIHIC and QTI were reliable. Additionally the eta^2 statistic indicates that, for all five WIHIC scales and all eight QTI scales, on both forms, there was a significant difference between classrooms and thus both instruments were able to discriminate between classes. These results clearly indicate reliable and internally consistent instruments and are consistent with prior research.

Because the QTI is based on a two-dimensional model in which the scales are theoretically related, factor analysis was replaced with an examination of the interscale correlations. The pattern of interscale correlations of the QTI confirmed the circumplex nature of the instrument through the existence of strong positive correlations among adjacent sectors and strong negative correlations between opposing scales.

The results presented in this chapter for both the WIHIC and QTI indicate that they are valid and reliable instruments and have produced results which compare favorably with those of prior research. Therefore, the results in Chapter 4 justify the use of the two instruments in addressing my research aims with confidence in Chapter 5.

Chapter 5

Students' and Teachers' Perceptions of the Learning Environment

5.1 Introduction

The previous chapter reported the reliability and validity of the What Is Happening In this Class? (WIHIC) instrument and the Questionnaire on Teacher Interaction (QTI) as used in my study. This chapter presents the data generated by the application of the two instruments to research aims focusing on: gender differences in perceptions of interpersonal behaviors and of classroom learning environment, differences between student and teacher perceptions of interpersonal behaviors, and the effectiveness of ongoing teacher professional development in terms of learning environments and interpersonal behaviors in school classrooms.

This chapter also considers the qualitative data collected from a teacher follow-up questionnaire administered to a sub-sample of participating teachers. The qualitative data are interwoven with the characterizations derived from the WIHIC and QTI. The comments in the qualitative data serve to further illuminate and support the quantitative characterizations, and could provide markers for future studies.

The WIHIC and QTI were chosen because they provide data about different components of the classroom learning environment. The eight scales of the original WIHIC focus on classroom procedure features, as well as personal relationship factors. Of those eight scales, the five scales chosen for my study (Student Cohesiveness, Teacher Support, Involvement, Task Orientation, and Investigation) were hypothesized to most closely match the intended outcomes of the teacher's professional development experience and the learning environment embodied by reform mathematics in the United States (Kilpatrick, Swafford, & Findell, 2001; Popejoy et al., 2005). The QTI was chosen because it examines the learning environment from the perspective of teacher-student interaction. Because the QTI has both teacher and student forms, and each of those has an actual and preferred version, greater clarity about the learning environment is possible. While not utilized during the professional development project reflected in my study, results from these forms and versions can be used as professional growth reflection tools (Rickards, 1998).

This chapter reports the quantitative analyzes of gender differences in learning environment perceptions from responses to the WIHIC and in perceptions of interpersonal behavior as measured by the QTI (Section 5.2). Differences between how students and teachers perceived interpersonal behaviors are also reported here (Section 5.3). To round out the quantitative analyzes, the differences in learning environment perceptions and perceptions of interpersonal behaviors prior to, and following, the professional development experience are reported (Section 5.4). Finally, an examination of four classroom case studies is presented in an attempt to support and enrich the findings of the quantitative data analysis (Section 5.5).

5.2 Gender Differences in Learning Environment Perceptions

Differences between the perceptions of the learning environment of male and female students were tested for statistical significance, on each of the scales in the WIHIC and QTI, through the use of MANOVA for repeated measures. The eight QTI scales in one analysis and the five scales of the WIHIC in a second analysis served as the dependent variables, while gender was used as the repeated-measures independent variable. Also, effect sizes for male-female differences were calculated (see Section 5.2.1) in order to assess the magnitudes of any differences.

The unit of analysis for this test was the ‘within-class gender mean’. The within-class gender mean was chosen as a unit of analysis so as to avoid confounding caused by the fact that not all classrooms contain an equal number of males and females. The within-class gender mean provides a matched pair of means, one within-class mean for males and one within-class mean for females, thus minimizing differences caused by unequal group sizes.

It is worth noting here that the sample consisted of 1054 students in 56 classes. The overall composition of this sample was 522 males (49.5%) and 532 females (50.5%). These percentages closely match the distribution for all Washington State schools (Office of Superintendent of Public Instruction, 2004). Gender differences for the WIHIC and QTI are discussed in Sections 5.2.2 and 5.2.3, respectively.

5.2.1 Effect Sizes

Separate from the determination of statistical significance, the effect size is an index that describes the magnitude of the difference between two sets of data. Cohen (1988) defines the effect size as the difference between the two group means, divided by the pooled, or average within-population standard deviation ($d = (M_1 - M2)$ / SD_{pooled}). This definition of effect size essentially compares the difference between the two groups with an expected range of scores (i.e., the pooled standard deviation) (Long, 2000), and provides a measure of the educational importance of the difference, as distinct from its statistical significance.

According to Cohen (1992), effect sizes can be categorized as small, moderate and large. He states that a small effect is $d < 0.2$, a moderate effect is $0.2 < d < 0.8$ and a large effect is $d > 0.8$, but is reluctant to have researchers in the behavioral sciences and education adhere too rigidly to these parameters. Because effect size is independent of sample size, it affords us another indicator of whether the differences observed can be considered educationally significant or not. As Fan (2001) points out, statistical significance and effect size are two related sides of a coin, and function most effectively when used together. For these reasons, I report both effect size and statistical significance results below.

5.2.2 Gender Differences for WIHIC

Table 5.1 presents the average item mean and average item standard deviation for male and female students' scores on the five scales of the WIHIC used in my study. This administration of the actual form of WIHIC took place near the beginning of the school year. The average item mean is the scale mean divided by the number of items in the scale. For example, the average item mean of 3.12 for males on the Student Cohesiveness scale indicates that the average score of the eight items assessing this scale corresponds to just slightly higher than the 'Often' response. Because the WIHIC and QTI scales have different numbers of items and consequently different maximum total scores, the average item mean allows us to attach clearer meaning to any comparisons.

Table 5.1
Average Item Mean, Average Item Standard Deviation, and Difference between Males and Females (Effect Size and MANOVA Results) on the WIHIC Using the Within-Class Gender Mean as the Unit of Analysis

Scale	Average Item Mean		Average Item Standard Deviation		Difference	
	Male	Female	Male	Female	Effect Size	*F*
Student Cohesiveness	3.12	2.87	0.31	0.36	0.35	2.21**
Teacher Support	2.42	2.29	0.54	0.52	0.12	1.34**
Involvement	2.27	2.16	0.41	0.40	0.14	1.37**
Task Orientation	3.27	3.00	0.51	0.46	0.15	1.44**
Investigation	2.19	2.04	0.36	0.40	0.33	2.44**

*$p<0.05$ **$p<0.01$
N= 56 pairs of within-class means.
Average item mean=Scale mean divided by the number of items in that scale.

In order to test the statistical significance of gender differences on the WIHIC, MANOVA for repeated measures was undertaken. The set of five WIHIC scales constituted the dependent variable and student gender was the repeated-measures independent variable. As noted in Section 5.2, the unit of statistical analysis used was the within-class gender mean. Because the multivariate test revealed significant gender differences overall on the set of WIHIC scales using Wilks' lambda criterion, the univariate ANOVA for each individual WIHIC scale was interpreted and is reported in the last column of Table 5.1. These within-class gender differences were found to be statistically significant ($p < 0.01$) for all five of the WIHIC scales.

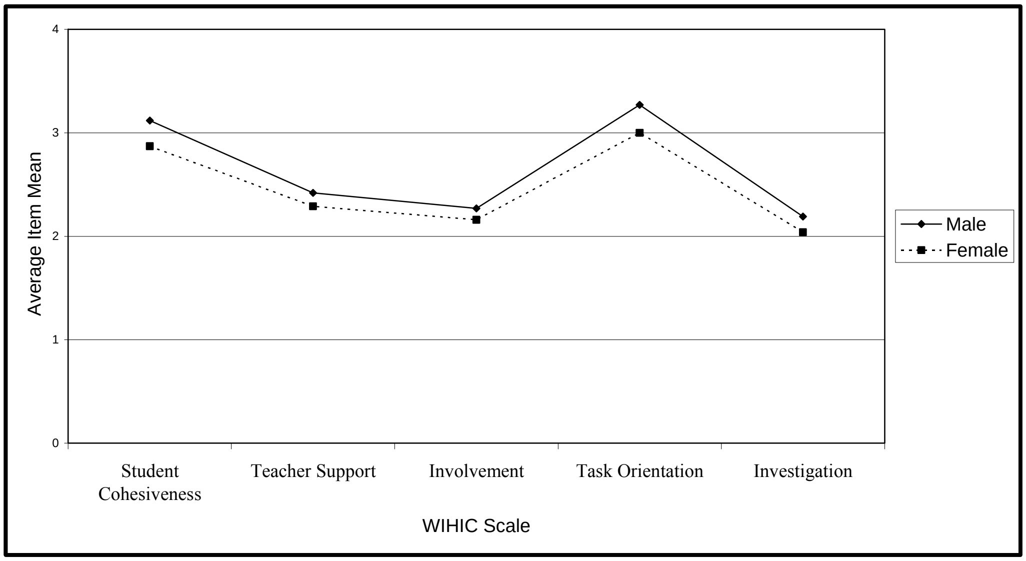


Figure 5.1 Within-class Gender Differences on the WIHIC Scales

As found in the second to last column of Table 5.1, the effect sizes for gender differences on all five WIHIC scales were in the low to moderate region (J. Cohen, 1992). Effect sizes for gender differences for Student Cohesiveness ($d = 0.35$) and Investigation ($d = 0.33$) fall within the moderate range, while for Teacher Support ($d = 0.12$), Involvement ($d = 0.14$) and Task Orientation ($d = 0.15$) fall in the low range of effect sizes.

The data presented in Figure 5.1 is based on the first two columns of Table 5.1 and represents the average item mean for the 5 scales of the WIHIC for males and females. As described in Chapter 2 (Section 2.3.3) students responded to a frequency scale (0 = Almost Never … 4 = Almost Always) for the WIHIC as a measure of the classroom learning environment. From Figure 5.1, it is clear that there is a modest difference between how males and females perceived the learning environment and that the males consistently perceived the learning environment more favorably.

The direction of differences and the results of statistical significance tests indicate that the gender differences observed in WIHIC scales are statistically significant and in a consistent direction (with males exhibiting more positive perceptions). Because the effect sizes are only modest in magnitude the differences are probably not of great educational importance.

Previous research (Fraser, Giddings et al., 1995) had found that females typically perceived the classroom learning environment more positively, with effect sizes in the moderate range, on the seven scales of WIHIC. Recent studies, however, have found there to be greater variability in gender perceptions. Taylor (2004), in a study of secondary mathematics students, found that males perceived the classroom environment more positively than females on the WIHIC Involvement and Investigation scales. In a study by Moss (2003) of secondary science students, males perceived more Involvement and Investigation. Other studies (Adolphe, 2002; Kim et al., 2000; Margianti, Fraser, & Aldridge, 2001) have found that, in addition to Involvement and Investigation, males also perceived the learning environment more positively than females on the Teacher Support and Task Orientation scales. The one scale for which the direction of gender differences appears unique in this study, with a moderate effect size (d = 0.35), is Student Cohesiveness. While other studies (Adolphe, 2002; Kim et al., 2000; Margianti et al., 2001) found that females perceived the learning environment more positively on the Student Cohesiveness scale, in my study, males perceived the learning environment more positively as measured by the Student Cohesiveness scale.

5.2.3 Gender Differences in Interpersonal Behaviors

The analyzes reported in Section 5.2.2 for the WIHIC were repeated for the QTI. The average item mean, average item standard deviation, effect size and the results of MANOVA for repeated measures for male and female students' scores on each of the eight scales of the QTI used in my study are presented in Table 5.2. This administration of the QTI took place near the beginning of the school year with both the actual and preferred forms of the instrument. Only the analysis for the actual form is included here, as it was the only form used in both the pretest and posttest administrations of the instrument. As described in Section 5.2.2, the within-class mean unit of analysis was chosen because it avoids potential confounding situations when males and females are represented in disproportional numbers within a class.

The results for the MANOVA analysis yielded significant results overall, using Wilks' lambda criterion, and therefore ANOVA results were interpreted for individual QTI scales and are reported in the last column of Table 5.2. Results indicate that there were significant gender differences for all eight scales of the QTI. As a further measure of educational significance, effect sizes were calculated and are reported in the second to last column of Table 5.2. As can be seen in Table 5.2, the effect sizes are predominately in Cohen's (1992) 'low' range.

Table 5.2
Average Item Mean, Average Item Standard Deviation, and Difference between Males and Females (Effect Size and MANOVA Results) on the QTI Using the Within-Class Gender Mean as the Unit of Analysis

	Average Item Mean		Average Item Standard Deviation		Difference	
Scale	Male	Female	Male	Female	Effect Size	*F*
Leadership	2.85	2.79	0.51	0.41	0.07	1.15**
Understanding	2.89	2.77	0.57	0.56	0.10	1.39**
Helping/Friendly	2.94	2.81	0.60	0.57	0.11	1.42**
Student Resp/ Freedom	1.53	1.66	0.32	0.29	-0.20	1.73**
Uncertain	0.68	0.85	0.35	0.35	-0.23	1.79*
Admonishing	1.18	1.34	0.50	0.52	-0.16	1.73**
Dissatisfied	0.80	1.05	0.49	0.52	-0.25	2.00**
Strict	1.60	1.72	0.39	0.41	-0.15	1.43*

*$p<0.05$ **$p<0.001$
N= 56 pairs of within-class means.
[a] Average item mean=Scale mean divided by the number of items in that scale.

The average item means represented in Figure 5.2 come from Table 5.2 and show that there are differences between the perceptions of male and female students, but that the direction and magnitude of the differences vary depending on the scale being referenced. Male students showed significantly more positive perceptions of the learning environment on the Leadership, Understanding and Helping/Friendly scales as indicated by their higher scores. However, female students perceived a significantly more positive learning environment based on higher Student Responsibility/Freedom scores. Male students also had significantly lower scores on the scales with negative connotations, namely, Uncertain, Admonishing, Dissatisfied and Strict, indicating again that males perceived the classroom environment more positively on these indicators.

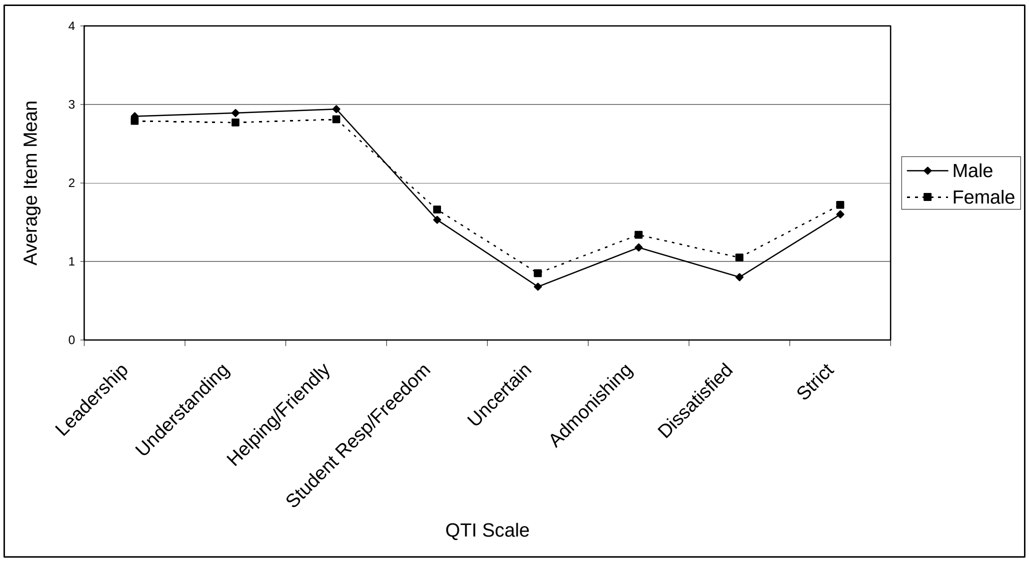


Figure 5.2 Within-class Gender Differences on the QTI Scales

While all eight scales of the QTI have statistically significant gender differences, an effect size analysis (J. Cohen, 1988) was again conducted in order to reflect the educational significance of gender differences in perceptions of the learning environment. Table 5.2 shows that all eight of the QTI scales produced effect sizes that are either very low or barely moving into Cohen’s (1992) moderate range. These effect size results, in concert with the statistically significant results, suggest that gender differences could exist in perceptions of teacher interpersonal behavior in the classrooms, but that these differences are small and are likely to be of limited educational significance.

In contrast to my results, previous studies (Fraser, Henderson, & Fisher, 1995; Levy et al., 2003; Quek et al., 2005b; Rawnsley, 1997; Rickards, 1998) in Australia, Singapore, and the United States found statistically significant differences between boys' and girls' perceptions of interpersonal behaviors in their classrooms on most of the QTI scales. Overall, they found that girls perceived their classrooms more positively than did boys. While these studies did find statistically significant differences, again the associated effects sizes were generally small.

The results for gender differences in QTI scores are not unlike those for the WIHIC discussed in Section 5.2.2. The results from both instruments indicate that males generally and consistently perceive teacher interpersonal behaviors and the classroom learning environment more positively than females on most scales, but that the effect sizes are generally small and probably of limited educational importance. These results from my study are consistent with the study of Korean students undertaken by Kim (2000) which found that, on the QTI scales of Leadership, Helping/Friendly, Understanding, Student Responsibility and Freedom, Dissatisfied and Strict, males perceived more positive interpersonal behavior than did females. In the same study, for the WIHIC scales of Teacher Support, Involvement, Investigation and Task Orientation, males perceived the learning environment more positively than females.

5.3 Differences Between Students' and Teachers' Perceptions of Interpersonal Behavior

Student data about interpersonal behaviors were collected through the use of the student actual form and student preferred form of the QTI. Teacher data about interpersonal behaviors were collected using only the teacher actual form of the QTI. A sample consisting of 44 matched pairs of teachers' scores and the class mean of students' scores was utilized in the analysis. Table 5.3 reports the average item mean and the average item standard deviation for each QTI scale's student actual form and teacher actual form. In addition, Table 5.3 presents the effect size and results of MANOVA for repeated measures (namely, the F ratio from each ANOVA), for differences between each QTI scale's student and teacher actual forms.

Results of the MANOVA for repeated measures (the eight scales of the QTI serving as the dependent variable and teacher/student serving as the repeated-measures independent variable) indicate that significant differences exist overall between teacher and student perceptions of interpersonal behavior, in terms of the Wilks' lambda criterion.

Table 5.3
Average Item Mean, Average Item Standard Deviation, and Difference between Teachers and Students (Effect Size and MANOVA Results) on the QTI Using the Class Mean as the Unit of Analysis

	Average Item Mean		Average Item Standard Deviation		Difference	
Scale	Teacher	Student	Teacher	Student	Effect Size	*F*
Leadership	3.07	2.79	0.47	0.47	0.50	1.83**
Understanding	3.16	2.79	0.48	0.53	0.77	2.27**
Helping/Friendly	3.44	2.85	0.41	0.57	1.12	2.73**
Student Resp/Freedom	1.46	1.57	0.45	0.29	-0.22	1.44
Uncertain	0.90	0.77	0.47	0.35	0.28	1.35
Admonishing	0.86	1.30	0.61	0.48	-0.87	2.41**
Dissatisfied	0.89	0.92	0.44	0.47	-0.06	0.43
Strict	2.09	1.68	0.41	0.31	1.05	-6.96**

*$p<0.05$ **$p<0.01$
N= 44 classes
[a] Average item mean=Scale mean divided by the number of items in that scale.

Because of these significant differences overall, the univariate ANOVA was interpreted and the results are reported in the last column of Table 5.3. ANOVA results indicate that statistically significant differences were present between student and teacher perceptions of interpersonal behaviors ($p < 0.01$) for five of the eight scales: Leadership, Understanding, Helping/Friendly, Admonishing, and Strict. Additionally, results of differences between students and teachers in their average item means are presented graphically in Figure 5.3.

Strong support for the existence of differences between students and teachers in their perceptions of the learning environment this study is found in the results of the effect size analysis. The results for effect size are reported in the next to last column in Table 5.3. For the five scales that had statistically significant differences between teachers' and students' perceptions, all had effect sizes that fell mainly into Cohen's (1992) large range: 0.50 standard deviations for Leadership, 0.77 for Understanding, 1.12 for Helping/Friendly, 0.87 for Admonishing, and 1.05 for Strict.

Based on the results presented in Figure 5.3, it can be seen that teachers generally rated themselves more favorably than did the students. For the three scales with a positive connotation for which teacher-student differences are statistically significant, teachers perceived more positive behavior than did students in terms of more Leadership, Understanding, and Helping/Friendly behaviors. For Admonishing behavior, teachers perceived the behavior less frequently (i.e. more favorably) than did the students. For the other behavior with a negative connotation, Strict, the students perceived the behavior less frequently than the teachers, indicating that they saw the teachers more favorably, as being less strict than the teachers themselves perceived.

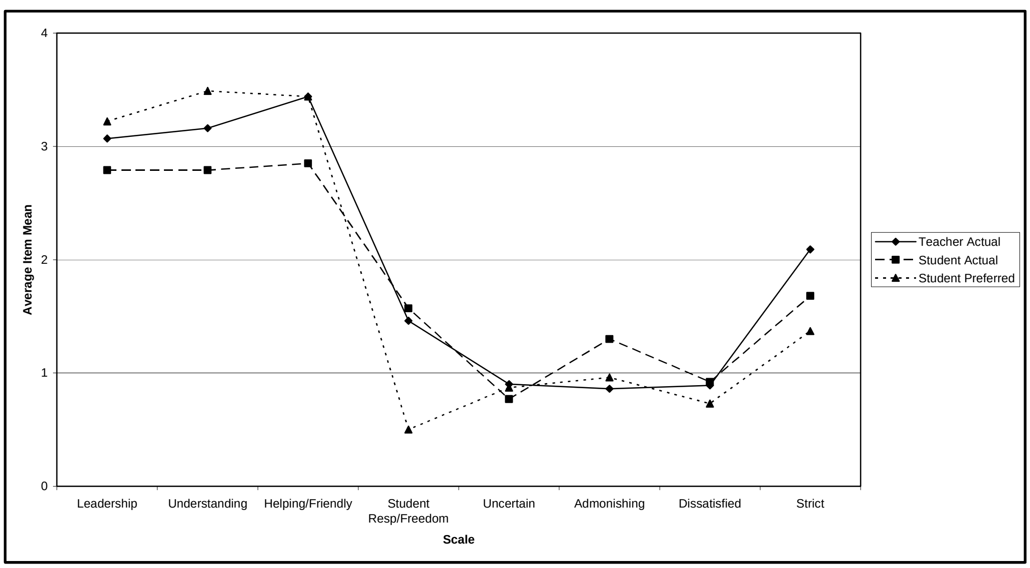


Figure 5.3 Student and Teacher Perceptions of Interpersonal Behavior

This section has reported differences between students' and teachers' perceptions of interpersonal behaviors in my study. The results indicate that there was a statistically significant difference, with moderate to large effect sizes, between how teachers and students perceive classroom interpersonal behaviors. The direction of these results would suggest that, on the whole, teachers perceive the classroom learning environment more positively than do the students. With regard to statistical significance, the direction of differences, and the effect sizes, these findings are consistent with previous studies (Kim et al., 2000; Rickards, 1998; Wubbels, Brekelmans et al., 1993). While the direction and magnitude of the teacher-student differences are consistent with prior studies, the magnitude of the differences found between teachers and students is much larger than that found for the gender differences in my study (as reported in Sections 5.1 and 5.2) using both the WIHIC and QTI. The combination of the statistical analysis results with the inter-study consistency would suggest that the differences might be of educational importance.

Section 5.4 discusses the analysis of differences between students' perception of the classroom learning environment and interpersonal behaviors near the beginning of the school year, and early in the teachers' professional development experience, and the students' perceptions at the close of the school year.

5.4 Pretest-Posttest Differences in Student Perceptions of the Learning Environment

As a potential indicator of the effectiveness of the teacher professional development in the *New Outcomes: Learning Improvement in Mathematics Integrating Technology* (NO LIMIT) project, differences in the students' perception of the learning environment between the start of the school year and the end of the year were studied. This time period corresponded to the time during which the teachers were involved with the NO LIMIT professional development. Changes in perceptions of interpersonal behaviors were investigated through the eight scales of the QTI, and changes in the classroom learning environment through the five scales of WIHIC. For both the WIHIC and QTI, the unit of analysis utilized was the individual student.

5.4.1 Pretest-Posttest Differences for WIHIC

The statistical significance of the differences found between the pretest and posttest administrations of the WIHIC were initially determined by employing a multivariate analysis (MANOVA) for repeated measures. The MANOVA simultaneously examined the five scales of the WIHIC as dependent variables, with the time of questionnaire administration (pretest and posttest) as the repeated-measures independent variable. Using Wilks' lambda criterion, differences for the set of WIHIC scales were found to be statistically significant and so univariate (ANOVA) results for each individual WIHIC scale were then interpreted.

Table 5.4 presents the average item mean and average item standard deviation for both the pretest and posttest administrations of the WIHIC questionnaire. In addition, the effect size and ANOVA results for the differences are presented.

Table 5.4
WIHIC – Average Item Mean, Average Item Standard Deviation, and Difference between Pretest and Posttest (Effect Size and MANOVA Results) Using the Individual as the Unit of Analysis

Scale	Average Item Mean[a]		Average Item Standard Deviation		Difference	
	Pretest	Posttest	Pretest	Posttest	Effect Size	*F*
Student Cohesiveness	3.04	3.05	0.77	0.77	0.01	0.42
Teacher Support	2.50	2.38	0.91	0.99	-0.12	1.87*
Involvement	2.30	2.24	0.92	0.97	-0.06	1.34
Investigation	2.19	2.12	0.97	1.01	-0.06	0.37
Task Orientation	3.18	3.17	0.81	0.79	-0.02	0.63

*$p<0.05$
N= 879 students in 57 classes.
[a] Average item mean=Scale mean divided by the number of items in that scale.

The average item means for pretest and posttest are presented in Figure 5.4. From this graphical representation, a slight pretest/posttest decrease in the average item means is evident. However, of the five WIHIC scales, only Teacher Support showed a statistically significant difference ($p < 0.05$) between pretest and posttest administrations, based on ANOVA results as reported in the last column of Table 5.4. The effect size for pretest/posttest differences in Teacher Support falls within Cohen's (1992) low range and is only 0.12 standard deviations. Despite there being a statistically significant difference for Teacher Support, such a small effect size suggests that the educational importance of this difference is quite low.

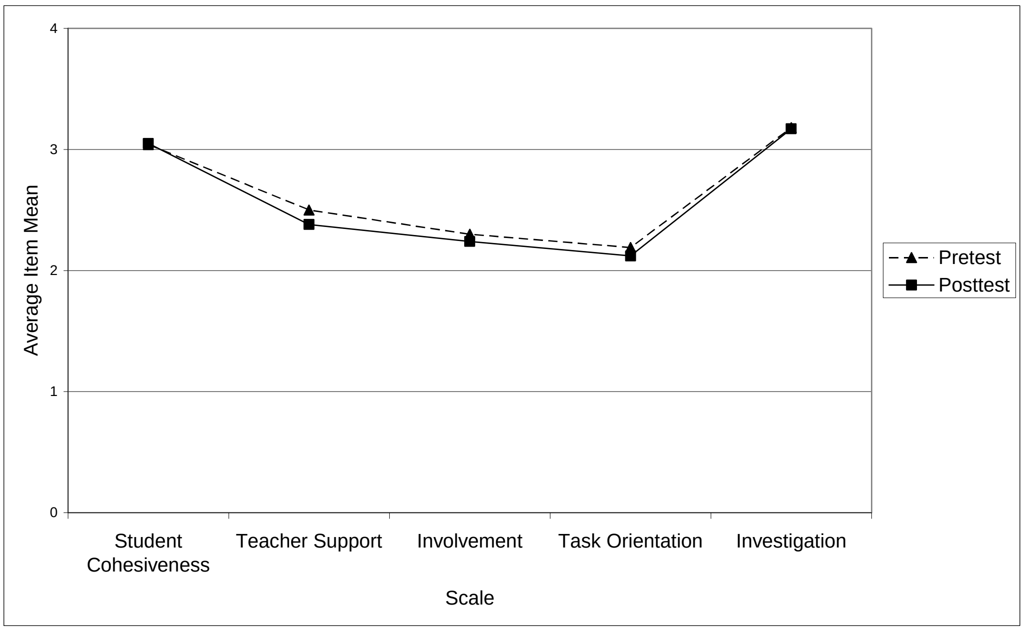


Figure 5.4 Pretest-Posttest Means on WIHIC Scales

5.4.2 Pretest-Posttest Differences as Measured by QTI

As was the approach with the data from the WIHIC, differences in the pretest and posttest administrations of the QTI were examined for statistical significance by the application of the multivariate analysis (MANOVA) and in terms of Wilks' lambda. The eight scales of the QTI comprised the dependent variable, while the timing of the administration (pretest and posttest administration) served as the repeated-measures independent variable. MANOVA results indicated significance and so the univariate analysis (ANOVA) was interpreted for each QTI scale. Table 5.5 presents the average item mean and average item standard deviation for both the pretest and posttest administrations of the QTI. In addition, the effect size and univariate analysis results (*F*) for the difference between pretest and posttest on each QTI scale are presented in the last two columns.

Table 5.5
QTI – Average Item Mean, Average Item Standard Deviation, and Difference between Pretest and Posttest (Effect Size and MANOVA Results) Using the Individual as the Unit of Analysis

Scale	Average Item Mean[a]		Average Item Standard Deviation		Difference	
	Pretest	Posttest	Pretest	Posttest	Effect Size	*F*
Leadership	3.01	2.85	0.70	0.81	-0.21	2.51**
Understanding	3.02	2.87	0.78	0.89	-0.18	2.31**
Helpful and Friendly	3.06	2.92	0.84	0.95	-0.16	2.17**
Student Resp/Freedom	1.58	1.59	0.62	0.66	0.02	0.56
Uncertainty	0.70	0.75	0.65	0.67	0.08	1.38
Admonishing	1.12	1.24	0.76	0.86	0.15	2.12**
Dissatisfied	0.81	0.88	0.75	0.82	0.09	1.55*
Strict	1.66	1.64	0.64	0.70	-0.03	0.72

*p<0.05 **p<0.01
N= 879 students in 57 classes.
[a] Average item mean=Scale mean divided by the number of items in that scale.

As presented in Table 5.5, ANOVA results indicate that, for five of the eight QTI scales, there is a significant difference between pretest and posttest administrations. The interpretation of the five significant pretest-posttest differences was that the teacher-student interactions deteriorated slightly in terms of less Leadership, Understanding and Helping/Friendly behavior and more Admonishing and Dissatisfied behaviors.

Table 5.5 also shows that the effect sizes are quite small for all of the QTI scales for which pretest-posttest differences are significant. These five effect sizes range from 0.09 to 0.21 standard deviations and are mainly classified as small according to

Cohen (1992). Effect sizes of this magnitude could suggest that the differences might not be of high educational significance.

The pretest and posttest administration results for the QTI are also presented graphically in Figure 5.5. From both Table 5.5 and Figure 5.5, it is clear that, for the five scales showing a significant difference, the magnitudes of changes were small and consistently in the direction of less favorable perceptions of interpersonal behaviors by the students.

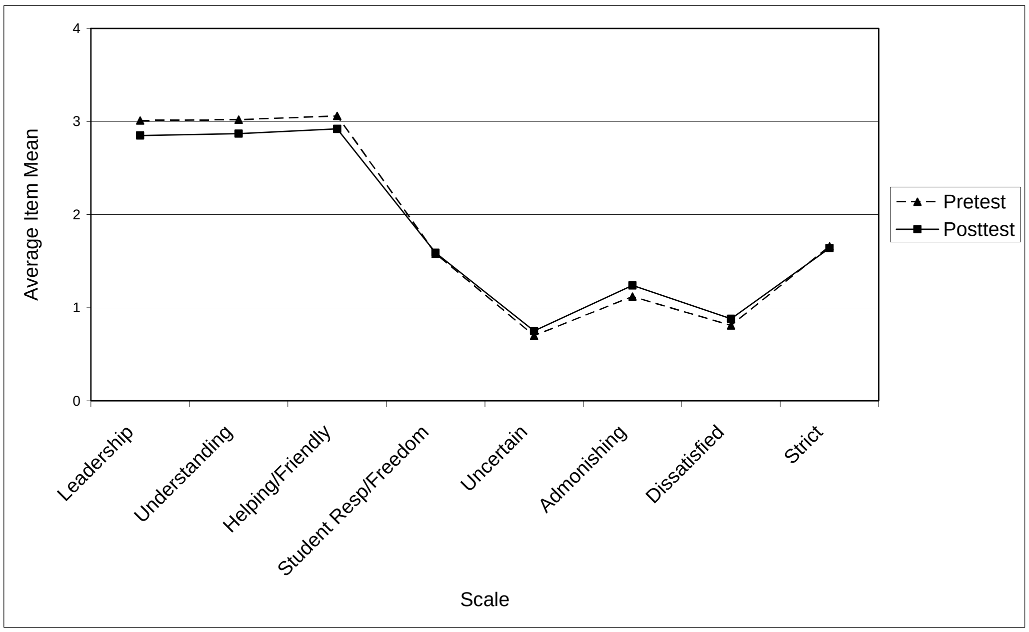


Figure 5.5 Pretest and Posttest Differences on QTI Scales

5.4.3 Discussion of Pretest-Posttest Results for WIHIC and QTI

The lack of statistically significant changes between pretest and posttest were not what I would have expected for either the WIHIC or QTI. Nor do they appear to support what the professional development project (NO LIMIT) had set out to accomplish (Popejoy et al., 2005):

> The project combines high stakes professional development and consequent teacher change, successful implementation of national and state standards in mathematics, with appropriate infusion of technology to improve student learning. (p. 1)

Speculations as to why the pretest-posttest results turned out as they did are reserved for a more extensive discussion in Chapter 6. As a means of further elucidating the context under study, qualitative research methods were employed as described in Chapter 3 and the resulting case studies are discussed in the next section.

5.4.4 Summary of Quantitative Results

Sections 5.2 – 5.4 reported the quantitative results from my study. Examination of both the WIHIC and QTI data for gender differences and pretest-posttest differences indicate that, while some scales showed a statistically significant difference, generally the effect sizes were small. The gender and pretest-posttest differences found in my study are likely of limited educational importance. In contrast, differences between teacher and student perceptions of interpersonal behaviors were significant for a number of scales, with the teacher generally perceiving the classroom learning environment more positively than the students. For the scales for which differences between teacher and student perceptions were statistically significant, effect sizes were generally large, thus further supporting that the teacher-student differences might be of educational importance.

The quantitative results only describe a portion of the tapestry that is made up of the students, teachers and classroom experiences in my study. As Tobin and Fraser (1998b) point out, "whereas the theoretical frameworks embedded in a selected learning environment instrument illuminate particular constructs to reveal trends and patterns, other constructs and associated patterns and trends are obscured" (p. 624). In an attempt to bring out other patterns, Section 5.5 discusses the same students, teachers and classrooms from a qualitative perspective. The narrative case studies presented include the quantitative data from the WIHIC and QTI, for individual classrooms, in order to provide a multilevel approach to my study.

5.5 Case Studies of Participating NO LIMIT Classrooms

As Anderson (1998) points out:

> Qualitative research is a form of inquiry… that tries to get at a deeper understanding, and bring meaning to an observed phenomenon. (p. 119)

While using qualitative methods in other fields of research is not new, and the concept of combining research approaches is clearly not new (Denzin & Lincoln, 2005), there is increasing support for combining quantitative and qualitative observations in mathematics and science education research (Aldridge & Fraser, 2000; Aldridge et al., 1999; Atweh, Forgasz, & Nebres, 2001; Fraser & Tobin, 1991; McLeod, 1992). Section 2.3.5.5 also provides additional background information about the use of qualitative methods in educational research.

In addition to other areas of inquiry, there is a strong indication for the need to interleave quantitative and qualitative data into research on teacher professional development programs (Guskey, 2000). The use of qualitative and quantitative approaches in this large study of middle-school mathematics classes in the United States is unique in that a combination of methods was used for mathematics learning environments and the evaluation of a teacher professional development program.

As mentioned in Chapter 3 (Section 3.7) ten teachers from the original nonprobability sample of 27 were selected for participation in follow-up interview. They were selected so as to represent a geographical spread across Washington State, the range of middle-school grades (i.e. Grades 6–9), and a range of project implementation successes as observed by me. Of the ten teachers selected, seven agreed to proceed with the follow-up interview. All seven of the selected teachers successfully completed the interview. The interview questions were designed to elicit from the teachers their perception of how their instructional practices changed and how they believed that the students responded to those changes. The follow-up interview questions can be found in Appendix E.

The basis for selection of a case-study approach was discussed in detail in the chapter on this study's methodology in Section 3.7. As Stake (1997) points out, the case-study approach is a respected method for studying education programs, in part by 'triangulating key observations' (Stake, 1994) with other data and thereby adding further validation to the findings. Through the presentation of alternative views of the classroom learning environment – including the teachers' perspective as drawn from their responses to the interview questions, my perspective as an observer and often as an active participant in the classroom, and finally the students' perspective as they responded to QTI and WIHIC items about their teacher – I hope to provide the necessary triangulation.

The four case studies presented in the Section 5.5.1 are drawn from the classrooms of four of the 56 participating teachers. These case studies were selected so as to represent the range of contexts experienced in my study. The classrooms represent a range of school sizes, demographics, geographic locations, levels of teacher engagement, and degrees of coherence between student and teacher perceptions. While the quantitative data tends to present a 'middle-ness' picture, these classroom 'extremities' serve to reveal greater detail about what was happening in the participating classrooms. As in Section 3.7, the conceptual framework under-girding these case studies was concerned with the issue of whether changes in the learning environment had taken place in middle-level mathematics classrooms in Washington State.

5.5.1 Case Study 1

Introduction: Fran Allen (not her real name) is a mother of two young children and wife of a teacher in a neighboring district, and has been teaching mathematics for 12 years. She has held her current position as a high school mathematics teacher for 10 years. Her first two years of teaching were in a middle school in the same district. Fran teaches in a relatively large high school of approximately 1000 students in Grades 9 – 12. The school year follows the statewide norm of extending from early September to the middle of June. This school is the only high school in a small city. Both the community and the school are quite homogeneous in ethnicity, with white being the dominant race. Recently the school and community have been experiencing an increase in Russian and Hispanic populations. Academically the school is underperforming as measured by the Washington Assessment of Student Learning (WASL) at the end of the 10^{th} grade. Only slightly more than 33% of 10^{th} graders achieved the score necessary to pass the WASL. A final note about Fran's school is that, just a year before this study commenced, the entire original portion of the school was lost to an arson fire. In conversations with staff, it was clear that they were still grieving the loss of their building.

School Context: Because she taught Freshman Algebra, Fran was eligible to apply to be a part of the *New Outcomes, Learning Improvement in Mathematics Integrating Technology* (NO LIMIT) statewide teacher professional development project. Fran was a participant of the NO LIMIT by her own choosing and was one of only two teachers in her building interested in participating. While teaching two sections of the Freshman Algebra course, she also had responsibility for two Geometry classes and a Second Year Algebra class. The Algebra course materials consisted of a 10-year-old book from a very traditional series of mathematics materials.

The mathematics department was fairly evenly staffed by men and women, but rarely did the two groups socialize or casually meet to discuss common concerns. Fran openly welcomed my presence in her classroom. When I visited for the first time in September, she introduced me to the class, explained the project with which she was involved, and clarified what my role might include. From that day on, I was a welcome part of her classroom. When asked in the interview process how she thought that her teaching had changed, she said: "NO LIMIT was a very positive experience. It has renewed my enthusiasm for teaching." This attitude was evident all year long as I took on predominantly the role of cognitive coach (Costa & Garmston, 2002) with her. As part of her participation in NO LIMIT, she chose to focus her efforts on shifting to a more student-centered and less teacher-centered approach to teaching. In addition, she wanted to be able to increase the frequency of student discourse in her classroom.

Classroom Observations: On my first visit to Fran's classroom, I encountered a very typical physical arrangement in the classroom with all of the student desks neatly aligned in straight rows. As students entered, they knew where their assigned seats were and that they were to take out the materials and assignments for that day. In a highly traditional format, the classes began with a review of the previous day's homework assignment. In very orderly fashion, the students then passed in their scored assignments and moved on to take notes as the teacher presented the new material. Individual students would shout out a question, while some raised their hands and waited to be called on. It was clear that the students trusted and relied on Fran to assist them with their understanding of the material.

During Fran's lunch break, we talked about what I had observed in class and, in light of the NO LIMIT observation sheet (see Appendix C), what her goals for the year might be. She recognized that the students were not very engaged with the mathematics and so she wanted to work on shifting the responsibility for understanding the mathematics to the students and to have them increase their reliance on one another (i.e. student discourse). The suggestion was made that she begin by teaching the skills of working in cooperative groups and then occasionally having the students work in small groups.

Over the course of the year, I spent an average of two days a month in her classroom. In that time, I saw remarkable structural changes and changes in her approach to teaching. By the close of the school year, the student desks were no longer in rows but were clustered in groups, with a group identification sign hanging from the ceiling above them. Group compositions were reassigned periodically including purposeful groupings and free-choice groupings. Fran's teaching slowly shifted toward a more student-centered approach so that, by the end of the year, over 50% of the class time was spent with students working in groups. During this group time, Fran would move among the groups to provide direction and support as needed.

The facilitation of student discourse was a more difficult task for Fran. She had to have the students create a set of norms that would allow respectful student-to-student talk. This was a skill that the students did not readily grasp and they had to be frequently reminded about the norms. Clearly the culture of learning for those students was that, because the teacher possessed the knowledge, questions should be directed to the teacher. In part because the students did not experience student discourse throughout their school day, in the end, the frequency of student discourse did not increase markedly. Fran indicated in conversation with me that she would continue to work on that and had hopes that, if she started off the next school year with those expectations, she would have more success.

Student Perceptions of the Learning Environment and Teacher Interpersonal Behavior: Student perceptions of the classrooms were gathered through the two learning environments instruments used in this study, the WIHIC and QTI. By using the WIHIC and QTI data at the classroom level, the 'grain size' (Fraser, 1999) under examination is reduced from the large scale results discussed in Sections 5.1 to 5.4. Examination with the same instruments, but at different 'magnifications', allows comparisons to be made between what happens in a given classroom and for the professional development project on the whole.

For the QTI data, a useful output format is the sector diagram (see Section 2.3.4 for complete discussion), as seen in Figures 5.6 and 5.7. Comparison of pretest and posttest profiles allows identification of any changes during the time between administrations. Sector diagrams represent profiles of student perceptions of their teacher. The extent of shading in each sector indicates student perceptions of that particular interpersonal behavior, ranging from zero to one, with a scale score of 'one' representing that all behaviors in a scale are always displayed. For example, in Figure 5.6 the shading in the Leadership sector corresponds to a scale score of 0.67.

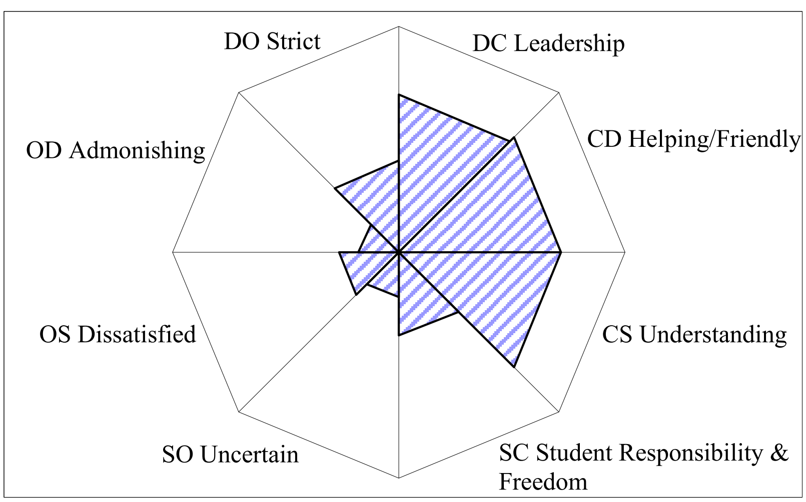


Figure 5.6 Students' Mean Pretest Perceptions of Fran Allen's Interpersonal Behavior

Student pretest and posttest perceptions of the interpersonal behaviors in Fran's class are shown in Figures 5.6 and 5.7, respectively. Comparison of the sector diagrams in Figures 5.6 and 5.7 indicates that the students perceived very little change in student-teacher interpersonal behaviors over the course of the school year. Student perceptions from the QTI match those of my observations. Interpersonally, Fran was and continued to be a very effective teacher.

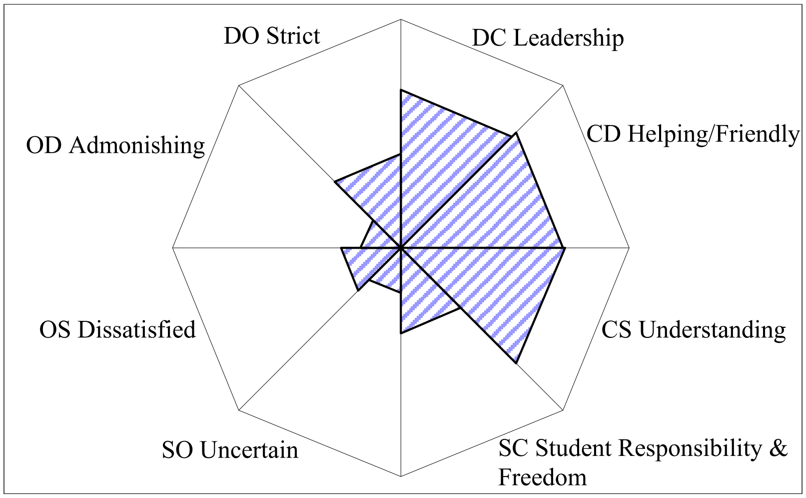


Figure 5.7 Students' Mean Posttest Perception of Fran Allen's Interpersonal Behavior

Student perceptions of the classroom learning environment were measured by five scales of WIHIC. Because the five scales are not theoretically related to each other as in the case of the QTI, the results are best represented on a line graph. Student pretest WIHIC and posttest WIHIC perceptions of the classroom learning environment for Fran's classroom are provided in Figure 5.8.

For the five scales of the WIHIC, the students did perceive some change in the learning environment over the school year. While Student Cohesiveness and Task Orientation showed virtually no change over time, the scales for Teacher Support, Involvement, and Investigation showed slight increases. These increases in scale scores suggest that the students felt that they experienced more teacher support and involvement in the class and that they had more opportunities for investigations.

The WIHIC results presented in Figure 5.8 are consistent with what I observed about changes in Fran's classroom over the course of the school year and are consistent with comments made in her follow-up interview. When asked in what ways she

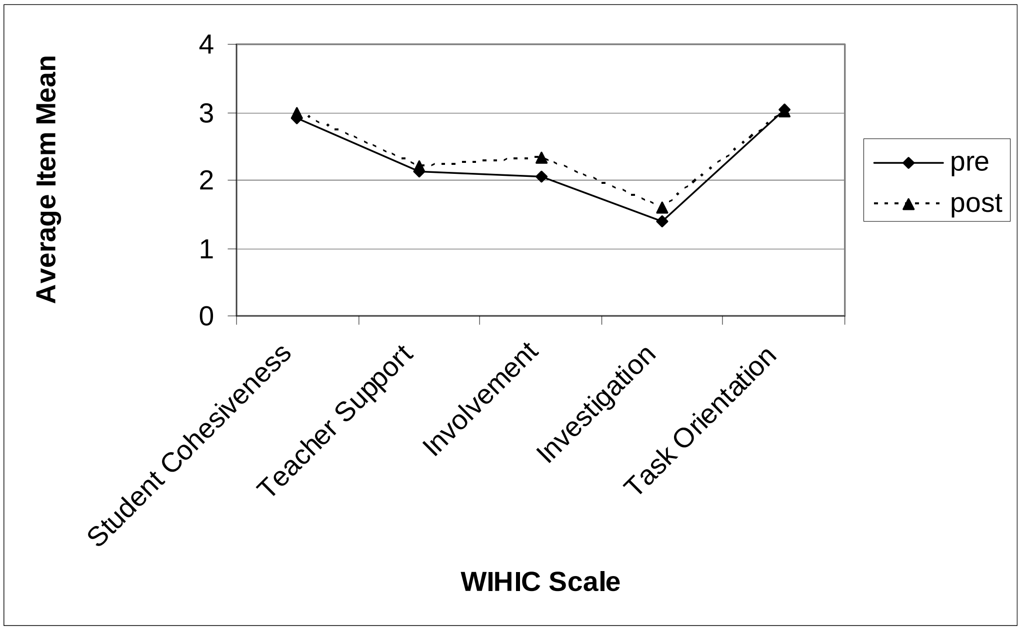


Figure 5.8 Students' Mean Perceptions of the Learning Environment in Fran Allen's Classroom

thought that the NO LIMIT project had impacted her classroom, she responded that "students in my classroom are now more involved in the learning process." It was her intention that her classroom move toward a more student-centered environment. Further evidence of congruence between observer-student-teacher observations is found in her response to the question about how she thought that students had responded to the changes that she had made in her teaching. She said: "Most of them reacted in a positive way. You can never please all of the students. I feel that the changes that I have made in the classroom have allowed more students to be successful and even made math fun."

In the case of Fran's classroom, there appears to be high congruence among the three 'observers'. The students perceived that there was some positive change in learning environment as evidenced by their responses on the WIHIC. Judging from Fran's comments, she believed that she had made some changes. Finally, through the observations that I made while in Fran's class, I was able to note changes in both the physical structure and instructional approach used in her classroom. While the changes might have been small, certainly smaller than either Fran or I had hoped for, the changes were observed by the students, Fran and me.

Taken in the context of the pretest/posttest results of the larger study discussed in Section 5.4, Fran's class appears as an anomaly. Collectively the classrooms in my study did not display changes in the learning environment which are large enough to be educationally important and, where small changes did occur, they suggested a less positive learning environment. Just as statistical significance testing was combined with effect sizes to establish quantitative educational importance, so too should these qualitative observations be combined with the quantitative results in an attempt to evaluate the impact of the professional development on Fran's classroom. Fran's classroom appears to have been positively impacted by her involvement in the NO LIMIT project.

5.5.2 Case Study 2

Introduction: Gregg Ikeda (not his real name) is married with young children and a working wife. Gregg has 15 years of teaching experience in mathematics and science. With the exception of about three years, Gregg has taught in the same school. Gregg teaches in a very small school consisting of a middle school and high school combined. The total student population in Grades 6 – 12 is approximately 250 students. As a consequence of the small school size, Gregg has a wide range of responsibilities. In addition to teaching mathematics, he also teaches some middle-school science, is head high school track coach and assistant basketball coach, and finally is the Knowledge Bowl (interschool academic competition team) coach.

Like most other school districts in the state, Gregg's school year begins in early September and runs until mid-June. The school is located in a very small town in rural Washington. To travel to the offsite NO LIMIT professional development trainings, Gregg had to drive for one and half hours. The staff at Gregg's school has an average of 12 years of teaching experience. The building that houses the middle-high school is older and in need of repair and modernization. For the last several years, the entire district has been in financial trouble and in danger of having to merge with an adjacent school district. The community surrounding the school is supported by the timber products industry. Other than the few migrant tree-planting families, the community's ethnicity is almost entirely white.

School Context: Because Gregg taught a 7th grade mathematics class, he was eligible to apply to be a part of the NO LIMIT teacher professional development project. Gregg chose to participate in NO LIMIT of his own volition, but with an additional outcome in mind. Because Gregg's school is in an economically-depressed district, and the NO LIMIT participants were to receive some classroom technology, Gregg's additional goal was the acquisition of technology for his classroom. In keeping with the overall conditions in the district, Gregg was using a very old, traditional middle-level mathematics textbook. In Gregg's school, just over 40% of the 7th grade students were achieving sufficiently well to meet the requirement to be considered passing on the WASL.

The entire mathematics department at Gregg's school consisted of himself and two other teachers, a male and a female. Gregg's male counterpart was also a NO LIMIT participant, but not one of the teachers with whom I worked. Gregg agreed to having me visit his classroom periodically, albeit somewhat reluctantly. When I would arrive for my visits, the primary focus was on how I might assist him in operating the newly-acquired technology. I saw that technology was going to be the angle from which I would have to approach the mathematics instructional improvement with Gregg. For his emphasis in the NO LIMIT project, he chose to try to make the mathematics activities more engaging for the students and to make certain that the activities were more closely aligned with the curriculum standards set out by the State of Washington.

Classroom Observations: A typical visit to Gregg's 7th grade mathematics class revealed students working more-or-less independently on their assignments. More often than not, the room was arranged in the typical fashion, with desks in nice straight rows. Because the mathematics materials were old and not well aligned to the state standards, Gregg chose to focus on providing more mathematically worthwhile tasks, and so would utilize lessons that he had found on the Internet which he thought would be more engaging and more of an investigative approach. While the lessons were more of a hands-on approach, the direction and support from Gregg was less than when he used a more teacher-centered approach. The students would often become frustrated and seek out his assistance. Unfortunately Gregg's other school responsibilities would often distract him from the current class and the students' questions would go unanswered. The impact of this neglect would reveal itself when the assignments were collected and, as Gregg shared in post-observation discussions with me, he was frustrated with the low level of student performance.

In conversations with Gregg about his teaching and teaching situation, he would often share his frustrations, including his desire to get out of teaching and out of the community. However, the dependable income which he was earning was keeping him from making drastic changes.

During the time that I spent in Gregg's classroom that year, I was able to observe that he had come to realize, that in order to improve student achievement in mathematics, his teaching needed to change. He clearly was attempting to move his teaching toward a more student-centered approach with the inclusion of more-engaging learning opportunities for the students. The students did appear to enjoy working on the more-engaging tasks. Gregg was able to observe this as well, as reflected in his brief response to the interview question about how students were responding to the changes: "The students were/are very excited."

The changes in Gregg's classroom can best be classified as primarily first-order changes (Cuban, 1988). Because Gregg recognized the need for a change, he brought in mathematics lessons that he thought would be more engaging and more aligned to the state standards. He also embraced the technology that was provided. However, what he was not able to do was to integrate the student-centered classroom into his own philosophy of teaching and use it to make the minute-by-minute decisions required of a teacher. I was able to observe this when he abandoned the use of the overhead projector for a more potentially-interactive document camera and digital projector. The tool had changed but the application had not.

Student Perceptions: As with all classrooms in my study, student perceptions of the learning environment were gathered through the used the WIHIC and QTI. Student pretest and posttest perceptions of the interpersonal behaviors in Gregg's class are shown in Figures 5.9 and 5.10. The sector diagrams in Figure 5.9 and 5.10 indicate that the students did perceive some change in student-teacher interpersonal behaviors over the course of the school year. They saw Gregg as displaying fewer Helping/Friendly behaviors and increasing in Dissatisfied behaviors. At the same time, they perceived more Understanding and fewer Admonishing behaviors.

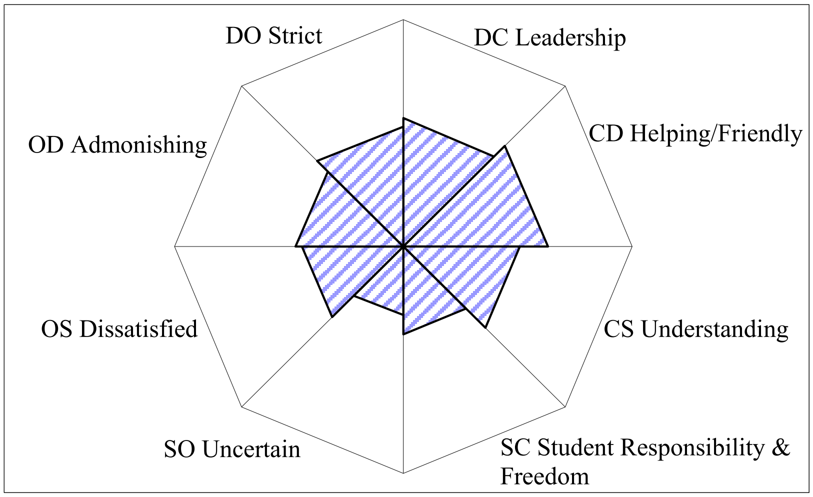


Figure 5.9 Students' Mean Pretest Perception of Gregg Ikeda's Interpersonal Behavior

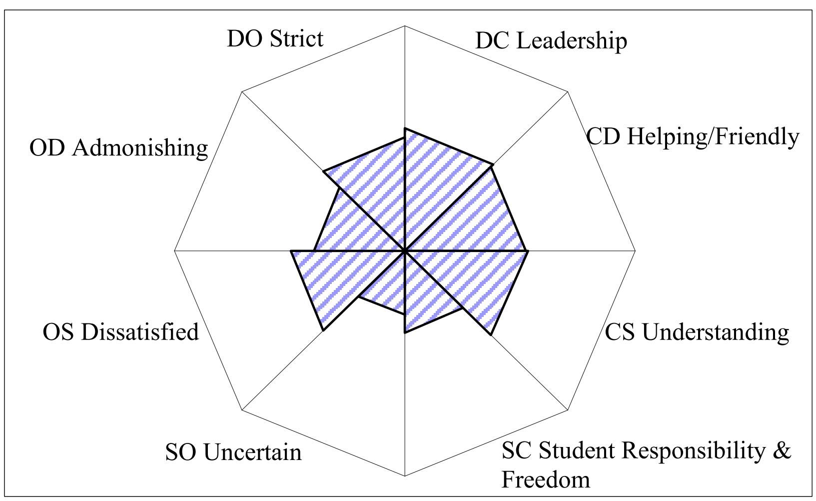


Figure 5.10 Students' Mean Posttest Perception of Gregg Ikeda's Interpersonal Behavior

From the WIHIC, the student pretest and posttest perceptions of the classroom learning environment for Gregg's classroom are provided in Figure 5.11. The 5 scales of the WIHIC indicate that the students did not perceive large changes in the learning environment over the school year. The students did perceive a slight increase in Student Cohesiveness and Task Orientation, while scale scores for Involvement and Investigation declined slightly.

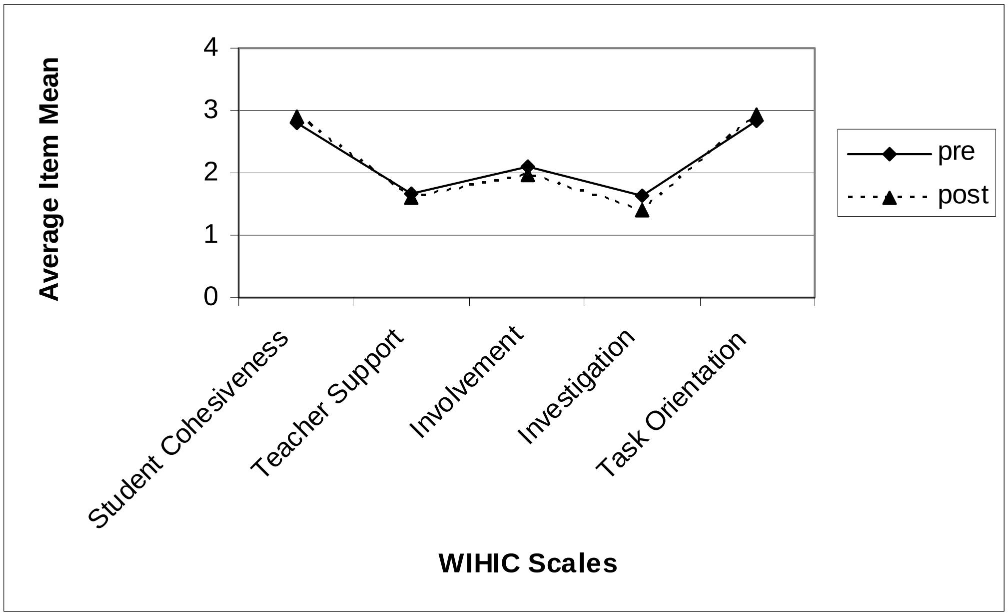


Figure 5.11 Students' Mean Perceptions of the Learning Environment in Gregg Ikeda's Classroom

It is interesting that, while Gregg was attempting to bring in more-engaging mathematical tasks, the students perceived a slight decline in their involvement in class and their level of engagement with investigations. A complicating situation here is the interplay between the use of the new technology and the application of new mathematical pedagogy. Gregg and his students adapted easily to the new technology, but the changes in mathematics instruction appeared to me to be more of a struggle. The separation between just using the technology and learning the mathematics was never clearly evident in Gregg's room.

The degree of correspondence between Gregg's perceptions of the learning environment in his classroom with those of the students, and with my observations, is weak. I saw a teacher deeply engaged with the new technology and the potential to engage students in mathematics more deeply, but struggling to make that fit his current paradigm. The students also did not see large positive changes in the learning environment. What does seem to fit well is that Gregg's remarks to me match with his QTI typology, that would be classified as 'drudging' (i.e. on the verge of burnout) by Brekelmans, Levy and Rodriguez (1993).

5.5.3 Case Study 3

Introduction: Dan Epp (not his real name) is a single young male. Dan has been teaching middle-level mathematics for three years, in the same the middle school that he attended as a student. Dan started his professional career as an electrical engineer but, because he did not find satisfaction in that, he returned to college to complete his teaching credentials. Dan's middle school is mid-sized, with an enrollment of approximately 500 students in Grades 7 – 8. As with the other case studies, this school follows the same school calendar, namely, from September to June. The school is located in a small town best characterized as being a cross between rural and suburban. The major north-south highway runs through the center of town. Economically, the community relies heavily on the timber industry and retail sales. Ethnically, the community is predominantly white but with an increasing Hispanic population.

Dan teaches in a building that is in excellent condition, in part due to modernization in the last 10 years. The rooms are well equipped for mathematics instruction, including all the latest instructional technology possible, as well a full compliment of mathematics manipulatives and supplies; thus allowing Dan to address a wide range of student learning styles. The school has a reputation for placing a high priority on academic excellence. Despite the focus on rigorous academic studies, the 7th grade students are still underperforming on the Washington Assessment of Student Learning (WASL). Less than 40% of the 7th grade students performed well enough on the WASL to attain a passing score.

School Context: Dan teaches both 8th grade mathematics and science. Two of his 8th grade mathematics classes have begun the high school mathematics curriculum, and were the two classes used in my study. The other 8th grade mathematics class consisted of students enrolled in the district-adopted middle-school mathematics course. This teaching assignment made Dan eligible to be a part of the NO LIMIT teacher professional development project. Dan was eager to obtain the professional development. Two of the three other mathematics department teachers had previously been involved with a different mathematics professional development opportunity and so had established a tone of professional growth in the department. Dan embraced this belief.

Both the middle-school and high-school curricula utilized course materials which had been developed with National Science Foundation funding. The school district had purposefully set out to adopt materials that were aligned from kindergarten through to 12th grade, correlated well with the state and National Council of Teachers of Mathematics (NCTM) standards (National Council of Teachers of Mathematics, 2000), and had a common philosophical underpinning. In addition, the mathematics department had spent time in carefully mapping the standards onto the curriculum materials.

As mentioned before, the mathematics teachers at Dan's school had a shared vision about the potential value of professional development. As a part of that, they were very open about having people observe in their classrooms. On my first visit to Dan's classroom, I was introduced and the project with which he was involved was explained to the students. Several students had questions for me about what it meant for them in their classroom. From that day on, I did not feel as an intruder in their classroom, but was treated as almost a co-teacher. Early on, Dan requested that I initially assume the role of cognitive coach, providing him with feedback on my observations and raising critical questions for him to consider. Like many other participants, for the NO LIMIT project, Dan's primary focus was improving student discourse. He also shared that he was having difficulty with engaging all students when they were heterogeneously grouped for tasks.

Classroom Observations: At first glance, Dan's classroom does not look unlike most other middle-school classrooms in the USA. Instead of individual student desks in the classroom, he has long tables that seat four students on one side. Typically at the start of one of Dan's classes the tables would be neatly arranged in rows. Clearly, there was a sense of order in the classroom, both structurally and behaviorally. Following the normal start-up routines, Dan would move right to discussing the work at hand. If a task had been completed as homework, or the day before, Dan would ask one of the groups to summarize what they had learned (as opposed to simply repeating the procedure). A brief class discussion normally followed. Dan would bring the discussion to a close with his own summarizing remarks. The classroom looked more chaotic than some others.

In conversations with Dan about his teaching, he shared that the curriculum materials that they had were very conducive to student engagement, but that he was having difficulty with making the flow feel natural. All through his own schooling, he had not experienced teaching in this manner and so it did not come easily. We decided that he should find ways to allow student discourse to occur more frequently and to work on strategies for engaging all students in a given work group. Because he believed in the philosophy of increasing students' responsibility for understanding the mathematics, it was the 'how' that he set as his goal.

While working with Dan that year, I was able to observe considerable effort on his part in changing the way in which he was approaching both student discourse and small-group dynamics. In true engineer fashion, he would develop a protocol for increasing student discourse, try it out while I was there to observe, and then ask for my feedback. On one such occasion, he had placed each student's name on a tongue depressor and placed them in a cup. When a student was called on to respond to a question, Dan would draw a student's name out of the cup. That student was then asked to respond to what the original student had said. This process could be used multiple times until he deemed it appropriate to move on. I observed that, at first, the students perceived it as part of his professional development and played along. Later, they did seem to be more attentive in class because of the possibility of being called on to respond to a fellow student. After a while, this protocol gave way to Dan simply calling on a student to respond. When it progressed to asking if anyone would care to comment on a student's statement, the system didn't work nearly as well. Discourse as a common practice was not yet a norm for the students.

It was evident from the classroom interactions that the students liked and respected Mr Epp a great deal. When I asked a student why the class cooperated with Mr Epp's efforts, the student responded that it was because students knew that he cared about them and that it must have been to their benefit. This level of trust between the students and the teacher allowed a climate of collaboration and growth.

Dan's work on small group dynamics was less visibly successful. Because he had not had any formal training in cooperative learning, I referred him to the Johnson and Johnson materials (D. Johnson & Johnson, 1991; D. Johnson, Johnson, Holubec, & Roy, 1984). He tried multiple strategies but with limited success. The school year was well underway and the students seemed to have already established their small group behaviors. The students did not readily adopt new group norms and behaved more independently than dependently. Dan vowed to start the next school off with a concerted focus on teaching students how to interact in a group.

Student Perceptions: Student responses on the two learning environments instruments (QTI and WIHIC) were gathered in Dan's two 8th grade high school mathematics classrooms in the course of my study. Pretest and posttest scores for student perceptions of interpersonal behaviors in Dan's class are shown in Figures 5.12 and 5.13. Initial examination of the sector diagrams in Figure 5.12 and 5.13 indicates that students perceived very little change in student-teacher interpersonal behaviors over the course of the school year. It is evident that the students continued to perceive high levels of Leadership, Helping/Friendly and Teacher Support behaviors throughout the year.

The QTI results presented in Figure 5.12 are generally consistent with the climate that I observed in Dan's classroom. Despite trying new approaches, sometimes more successfully than others, Dan maintained his overall demeanor throughout. I would characterize Dan as having a rather steady manner and not being easily flustered.

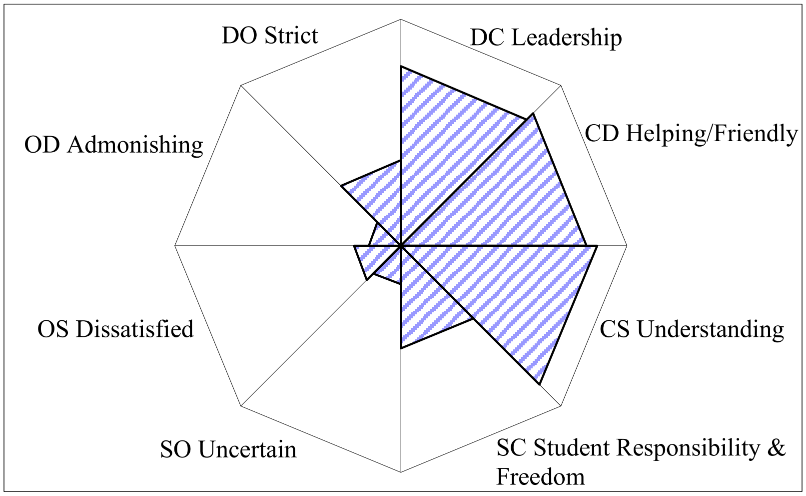


Figure 5.12 Students' Mean Pretest Perception of Dan Epp's Interpersonal Behavior

When he was asked how he thought that students had responded to changes that he'd made, he said:

> Student reactions were very positive to the changes in my teaching. My students over the two year span were more than willing to help me along with the new materials and try new ideas. I believe students will rise to the occasion if they sense the passion and willingness to improve from their teacher.

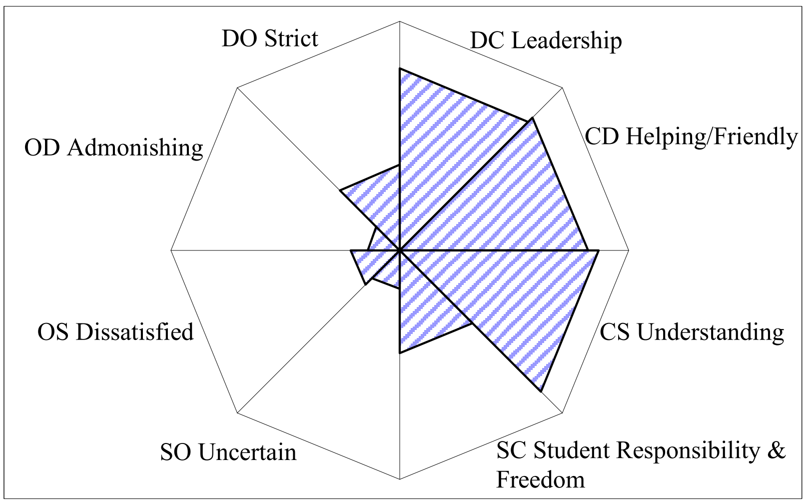


Figure 5.13 Students' Mean Posttest Perception of Dan Epp's Interpersonal Behavior

I observed that the students did see the passion and were willing to tolerate the changes that Dan was making. The quantitative and qualitative data sources appear to support each other.

Based on the five scales of WIHIC, student perceptions of the classroom learning environment in Dan's classroom were collected and presented in Figure 5.14 below. The students' perceptions of an overall decline in the WIHIC's classroom learning environment constructs in Dan's class were the most noticeable of the four case studies presented here. With the exception of Student Cohesiveness, the scales all showed a decrease.

The decline in the scale scores for the WIHIC presented in Figure 5.13 might seem inconsistent with the description of Dan's classroom. Dan's response to the interview question about how his teaching had changed was: "I believe that my teaching is much more structured as a result of … NO LIMIT." This is further evidence of the

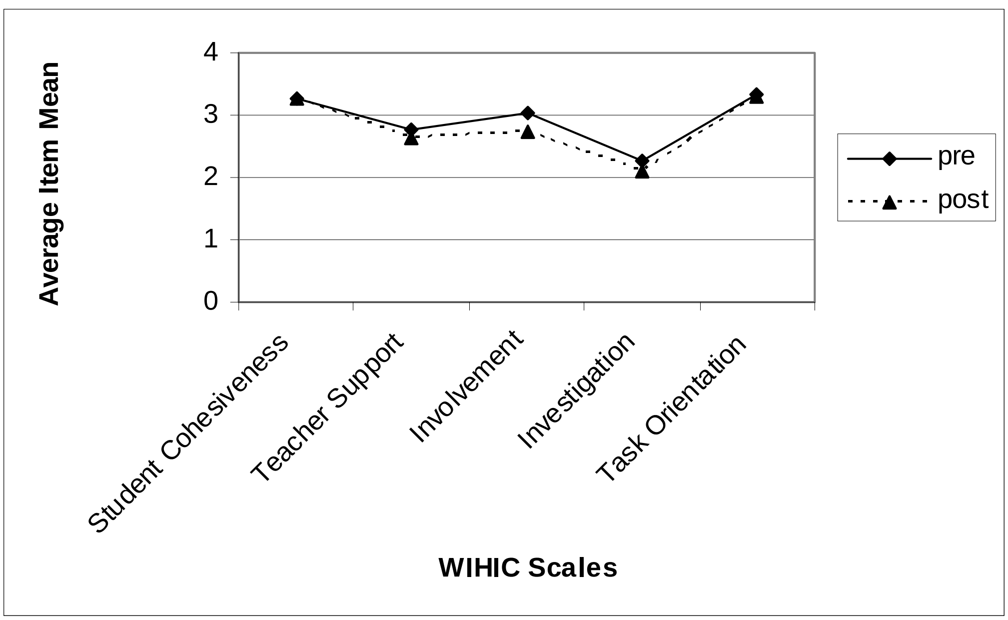


Figure 5.14 Students' Mean Perceptions of the Learning Environment in Dan Epp's Classroom

context present in his classroom. As an example, when he made the effort to increase the frequency of student discourse, he often resorted to contrived and mechanical means. The names on the tongue depressors is a prime example. While an excellent teacher, Dan clearly has a tendency to be very linear in his planning and delivery.

This pattern of declining WIHIC scale scores was born out in the overall study results as well and so this case study may be typical of what occurred in other classrooms.

5.5.4 Case Study 4

Introduction: Lynn Hull (not her real name) is a recently-married young teacher who is married to a banker in town. Lynn began her career in her current middle school 10 years ago. She grew up and went through high school in the same city in which she is now teaching. Her family members have been long-time residents of the area and are well known. In addition to her teaching, Lynn is pursuing a Masters degree in Curriculum. The enrollment at Lynn's middle school, Grades 6 – 8 is approximately 600 students. This is a mid-sized middle school. Lynn teaches mathematics only at the Grade 7 level. The 7th grade mathematics is taught using a very old, traditional textbook. Similarly the pre-algebra course for 7th graders uses an older textbook. The building in which she teaches is only about 10 years old (not the same one that she attended as a student in middle school) and the mathematics department is well supplied for teaching mathematics.

The community in which Lynn grew up and now teaches consists of a small city. Economically the community is heavily dependent on the timber industry and maritime shipping, as it has a large Pacific Ocean port. Timber exports are down and, as a result, the community is currently at an economic low point. The shift from what had been a thriving community to the present downturn has cast a pallor over the entire community. Additionally the demographics of the community are undergoing a change from predominantly White to a growing Hispanic and Russian population. The diversity of languages is now an issue in the schools.

Academically the students in Lynn's middle school are performing on par with students from around Washington State. They currently have just over 40% of the 7th graders scoring high enough on the WASL to be considered as passing. Because U.S. federal legislation requires that, by 2014, all 10th graders pass the 10th grade WASL, and because Washington State requires passing the 10th grade WASL in order to graduate, mathematics instruction is facing incredible pressure.

School Context: Teaching mathematics at the middle-school allowed Lynn to apply to be a participant in the NO LIMIT project. For both of her 7th grade mathematics classes and her pre-algebra classes, she chose to be a part of the project. As with some other case studies described here, she has a colleague in her school that is also participating in NO LIMIT. In part because all of Lynn's mathematics department colleagues are much older, she does not interact with them extensively. There are larger climate issues present besides the age difference. Given those contexts, it is no surprise that Lynn openly welcomed me to her classroom and actively sought support for her teaching through the professional development opportunity.

Partly because the age of the instructional materials with which Lynn had to work and because she was pretty much working in isolation, she viewed my presence as being a strongly supportive resource. Like many other participants she welcomed me to her classroom and did not hesitate to introduce me to her classes and explain the nature of the NO LIMIT project and what she anticipated that my role would be in her classroom. While she was pleased to have me there, she was not ready to allow me to model a lesson for her. We began our work with my providing her with replacement lessons that would help her to meet her goal of engaging the students more deeply in the mathematics. Additionally, she had a goal of making certain that what she was teaching was aligned with the state mathematics curriculum standards.

Classroom Observations: Lynn's classroom is laid out in very traditional fashion. The individual student desks are arranged in rows facing the front of the room. Based on the way in which students move around the room and treat each other, it is clear that she maintains an orderly environment. Early in my visits to Lynn's classroom it became clear that her primary style of teaching was a very teacher-centered approach. The assignments were corrected in class and handed in, the new material was presented, and the practice set assigned. Rarely did students interact with each other, aside from the normal social banter.

In conversations with Lynn, I shared my observations and asked what she wanted to do and what she felt comfortable starting with. She began by modifying a lesson in the textbook to require the use of manipulatives. The materials were passed out and chaos ensued. She quickly regained control and provided the clarification that the students needed in order to proceed. In watching the students in this circumstance, it is clear that they adored her and that, on that basis, would respond to requests for order to be restored. In the end, the lesson was successful and she was very pleased with the understanding the students had achieved. This led to more and more hands-on and engaging mathematics tasks for the students. After each lesson, we talked about what had gone well and the impact that it was having on student learning.

Over the course of the year that I spent time in Lynn's classroom, it was evident that she had changed her perspective on mathematics instruction. In response to the interview question about the impact of her involvement with NO LIMIT, Lynn said: "The project truly changed my philosophy on hands on learning…." When asked how her teaching had changed, Lynn responded:

> My teaching has changed immensely! The NO LIMIT truly helped me begin to shape the type of teacher that I want to be. I didn't want to be a teacher who has students strictly work out of a book and, without the NO LIMIT [project], I may have fallen into that pattern because it was the way I was taught. I now include more cooperative learning, project assessments, student discourse and hands-on learning than I believe I ever would have had I not had the opportunity of working with the No Limit grant.

Lynn's remarks clearly reflect her own sense of change. My observations do support her remarks in that, by the end of the year, there was far more engagement on the part of the students. However, visibly Lynn's classroom underwent very little change. At the end of the year, the desks were still in straight rows and she was still very clearly in control of the classroom. Further evidence of Lynn's change is found in a comment that she made to me about the future. She indicated that, in the future, she hoped to turn over more of the control to her students and have them directing the mathematical conversations.

Student Perceptions: All five of Lynn's 7th grade classes participated in responding to the two learning environments instruments (QTI and WIHIC). Students' responses were recorded and scale scores were calculated. Students' pretest and posttest perceptions of interpersonal behaviors in Lynn's class are shown in Figures 5.15 and 5.16. A first comparison of the two figures shows that students perceived very little change over the school year. The students continued to perceive high levels of Leadership, Helping/Friendly and Understanding behaviors. Slight decreases in Strict and Admonishing behaviors were noted.

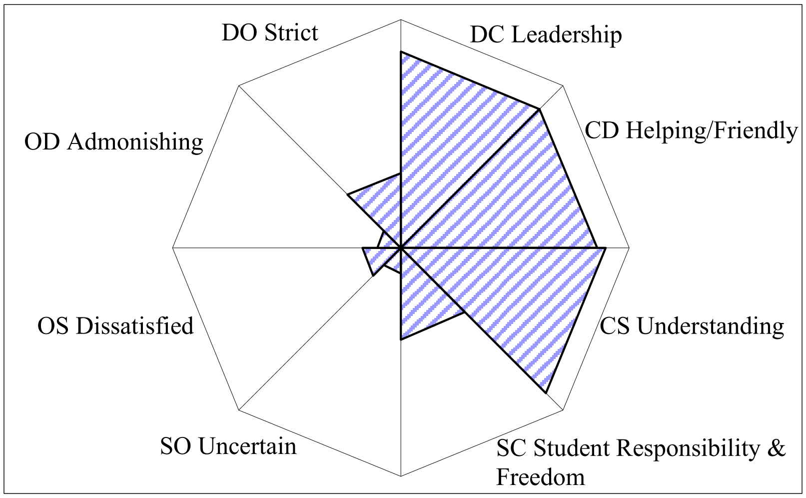


Figure 5.15 Students' Mean Pretest Perception of Lynn Hull's Interpersonal Behavior

The lack of a large change in Lynn's profile over time is consistent with my observations of the classroom. From the beginning of the year, the students adored her, in part because of her overall friendliness and willingness to help, while still holding them accountable.

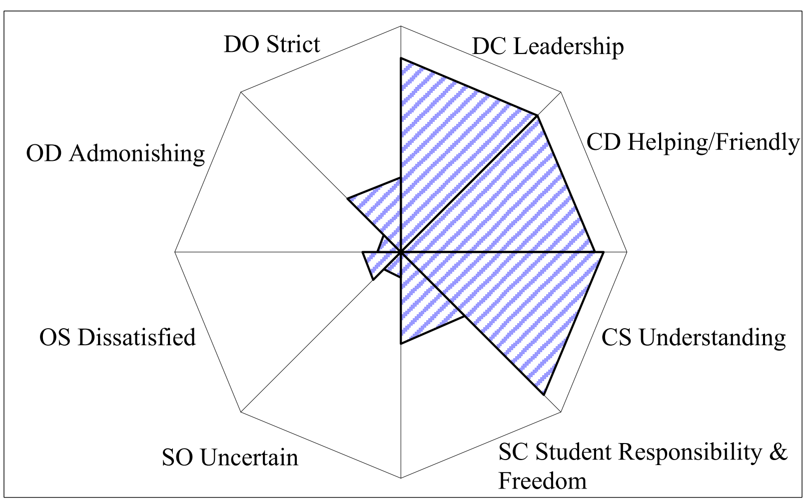


Figure 5.16 Students' Mean Posttest Perception of Lynn Hull's Interpersonal Behavior

Responses collected from Lynn's five classes to the items on the WIHIC are presented in Figure 5.17 below. In contrast, with the last case study, this case presents WIHIC results that consistently showed an increase in scale scores from pretest to posttest.

The pattern of increases in scores on each of the five WIHIC scales is very consistent with what I was able to observe in Lynn's classroom. She had changed the way in which she was teaching by including more tasks that required greater student involvement. The data indicate that the students also perceived the change. The fact that Teacher Support increased would seem to indicate that, while Lynn was able to transfer some of the learning responsibility over to the students, she took the opportunity to change the role that she played. She became more of a facilitator.

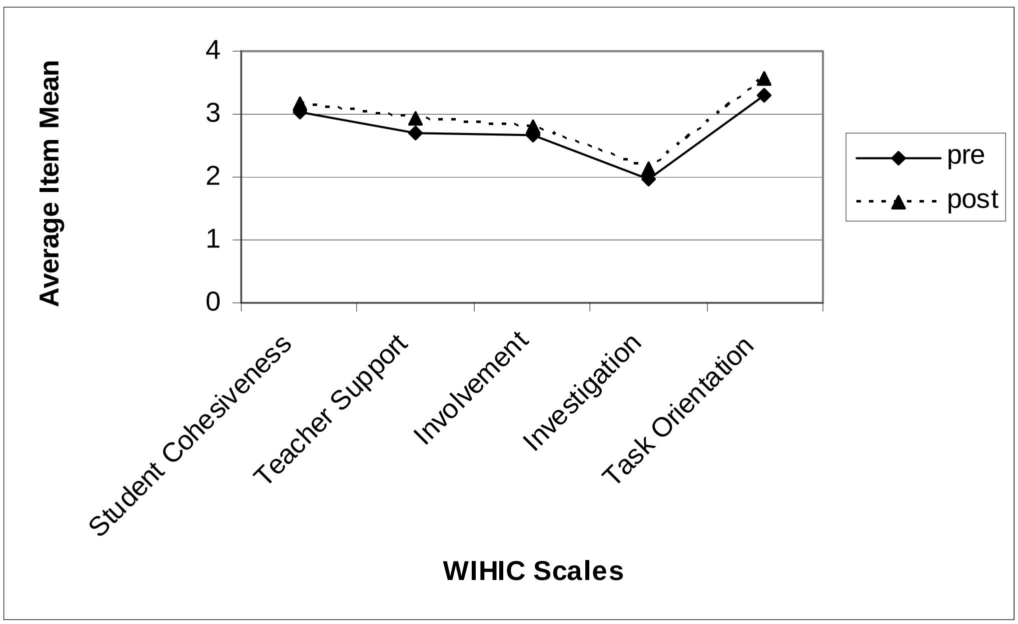


Figure 5.17 Students' Mean Perceptions of the Learning Environment in Lynn Hull's Classroom

5.5.5 Summary of Case Studies

Taken as a whole, the teacher interviews, my observations and the perceptions of the students in the case studies would seem to support some of the findings of the quantitative data. The potential causes for the observed cumulative decline in scores on several of the scales in the QTI and WIHIC are evident in the case studies when actual teacher behavior was observed. While the overall quantitative results showed a lack of statistically significant differences between genders and between pretest and posttest perceptions, and with relatively small effect sizes, the variability of these changes in classrooms was clear through my qualitative observations and the classroom-level responses on the scales.

As discussed in the introduction to the case studies and in Section 3.7, the conceptual focus for the qualitative methodology was to look for evidence of changes in the learning environment and teacher interpersonal behaviors in the classroom. The four case studies illustrate that there had in fact been changes taking place in the learning environment and teacher interpersonal behaviors. From the interviews with teachers, it is evident that they felt they had made changes in their classroom behaviors and routines. My classroom observations support that as well. When examined at a smaller grain size, the perceptions of the students also bear out that their teacher had made changes that impacted on the classroom learning environment. However, the magnitude and direction of the change were unique to each classroom.

What is evident from the case studies is that, in some classrooms, there were changes in the learning environment, both positive and negative. What is not evident from the case studies is the permanence of the change taking place in the students. My classroom observations suggested that the students were adapting to some of the changes made by the teacher. As examples, the acquisition of discourse skills was apparent in the students in two of the case studies, and the ability to engage more deeply with the mathematics in three of the case studies. The permanence of these changes remains an unknown as does the reasons why a given teacher changed more than another. Both of these questions are important areas for future research in evaluating professional development.

5.6 Summary

This study utilized both quantitative and qualitative approaches in examining the classroom learning environment of approximately 1200 middle-school mathematics students in Washington State. This combined approach was used to assess differences in perceptions between males and females, between students and teachers, and between the beginning and end of a teacher professional development project.

The quantitative results suggest that there are differences between the perceptions of male and female students in terms of interpersonal behaviors using the QTI and of the learning environment with the WIHIC. Generally, it was found that the males perceived the learning environment more positively than did female students. The results also suggest that there is a large difference between how the students and teachers perceived the interpersonal behavior in the classroom, with teachers generally viewing themselves more favorably than the students. Finally, the quantitative results indicated little overall difference between the beginning and end of the professional development project. Any differences tended to be small and probably of little educational importance.

Viewing the qualitative results in my study at the classroom level, through the use of case studies, suggests that there were learning environment changes occurring in the classrooms. These results showed both teacher and student changes. The case studies illuminated the uniqueness of the learning environment present in each classroom through the wide variability of responses to the professional development program present. While the impact of the professional development might not be evident solely through the quantitative results for the overall study, it was evident through the qualitative approach.

Chapter 6 addresses the four research aims first presented in Chapter 1. Based on the results presented Chapters 4 and 5, a number of conclusions are drawn concerning the learning environment in middle-school mathematics classrooms in Washington State in the context of a sustained, job-embedded teacher professional development program. A discussion of the implications follows. Chapter 6 concludes with a discussion of the limitations of the study and suggestions for future research efforts.

Chapter 6

Discussions and Conclusions

6.1 Introduction

The focus of my study has been the examination of the learning environment in middle-school mathematics classrooms whose teachers were experiencing a job-embedded, sustained approach to professional development. This professional development project was part of the State of Washington's effort to reform mathematics instruction. This effort is consistent with the national focus, described by Darling-Hammond and McLaughlin as far back as 1996.

> The vision of practice that underlies the nation's reform agenda requires most teachers to rethink their own practice, to construct new classroom roles and expectations about student outcomes, and to teach in ways they have never taught before – and probably never experienced as students. The success of this agenda ultimately turns on teachers' success in accomplishing the serious and difficult tasks of learning the skills and perspectives assumed by new visions of practice and unlearning the practices and beliefs about students and instruction that have dominated their professional lives to date. Yet few occasions and little support for such professional development exist in teachers' environments. (p. 597)

The NO LIMIT teacher professional development project was intended to provide the opportunity and support necessary to advance the practice of the participating teachers. As pointed out by many in the field of teacher professional development (Guskey, 2000; Loucks-Horsley et al., 2003; Shaha, Lewis, O'Donnell, & Brown, 2004; Sparks, 2002), a critical component of high quality professional development is its effective evaluation.

Inspired by Putnam and Borko's claim that "interactions with the people in one's environment are major determinants of both what is learned and how learning takes place" (2000, p. 5), together with the extensive research on classroom learning environments, my study undertook to evaluate the impact of the NO LIMIT project through a careful examination of the learning environments in the middle-school classrooms.

This chapter begins with an overview of my thesis in Section 6.2, with a chapter-by-chapter summary. Section 6.3 provides a thorough discussion of the major findings and conclusions of my study. Section 6.4 addresses the limitations present in my study, followed by its contributions in Section 6.5. Suggestions for directions that this research might take in the future are discussed in Section 6.6. Final, concluding remarks are offered in Section 6.7.

6.2 Overview of This Thesis

Chapter 1 of this thesis began with the mathematics education context within the State of Washington and within the Educational Service District in which I work as a professional development provider. The link between the statewide events and my work was described as the impetus for me to study the impact of sustained, job-embedded professional development from the perspective of the learning environment in the participating middle-school mathematics classrooms. The chapter also provided an overview the research questions and subsequent methodology. The limitations and implications of my study were also introduced.

Chapter 2 provided a review of the literature pertaining to the field of learning environments and the field of evaluation of professional development. The review was presented in three major divisions: 1) professional development, 2) learning environments, and 3) application of a mixed-methods approach to research. Section 2.2 discussed the history of teacher professional development, criteria for designing teacher professional development and the emerging elements of effective professional development evaluation. The evolution of the learning environment field, as well as the always-expanding body of instruments for probing learning environments, were presented in Section 2.3. This included a detailed discussion of the two instruments used in my study, the WIHIC and QTI. This section closed with a review of the use of mixed-methods in educational research. The theme present throughout this chapter was the potential linkages between learning environments research and the evaluation of professional development.

The research methods, sample sizes and selection, procedures for data collection, and survey instruments for my study were discussed in Chapter 3. In order to more thoroughly study the learning environments of the classrooms, a mixed-method approach was adopted (i.e. quantitative and qualitative methods), as recommended by Tobin and Fraser (1998b). Section 3.5 discussed the quantitative component which involved the administration of the widely-validated WIHIC and QTI to over 1200 students and 70 teachers. This section also included a discussion of how the learning environments instruments aligned with the observation protocols being used in the NO LIMIT classrooms. Section 3.6 introduced the data-analysis procedures used to address the research questions, and Section 3.7 presented how classroom observations and interviews were conducted to gather data for the qualitative component of my study.

Chapter 4 provided detailed discussion of the statistical analyses performed on the data collected in my study. Section 4.2 discussed the validity analysis for the WIHIC. Results of the factor analysis were reported for the 40 items in 5 scales. Internal consistency reliability and ability to discriminate between classrooms analyses were investigated through the application of Cronbach's alpha coefficient and the eta^2 statistic from ANOVA. The parallel set of analyses for the QTI was discussed in Section 4.3. Results for application of Cronbach's alpha coefficient and eta^2 statistic were presented for the QTI. As the QTI is based on a two-dimensional circumplex model and factor analysis is not applicable, this section provided a detailed discussion of the results of the interscale correlation analyses performed on the 48 items in 8 scales.

Results of the quantitative and qualitative examination of students' and teachers' perceptions of the learning environment from my study were presented in Chapter 5.

This chapter adopted a theme of looking for differences; between genders, between students and teachers, and between the beginning and the end of the year-long professional development experience.

Section 5.2 discussed the observed gender differences as measured by both the WIHIC and QTI. Results of the use of MANOVA for repeated measures provided the basis for discussion of statistical significance. Effect sizes were also reported to facilitate an understanding of the magnitude of differences between the genders, and therefore their educational importance. Differences between students' and teachers' perceptions of interpersonal behaviors as measured by the QTI were discussed in Section 5.3. Again results of MANOVA and effect sizes were provided. Section 5.4 described the statistical significance and effect sizes for the changes in student perceptions of the learning environment that occurred between from the beginning and near the end of the school year.

The qualitative component of my study was discussed in Section 5.5. This section presented the qualitative data gathered through classroom observations, discussions with teachers and selected teacher interviews. In keeping with the 'bricolage' approach these qualitative data were combined with the quantitative data at the classroom level. Using the classroom as the 'grain size', this section presented case studies of four representative classrooms. The combination of the qualitative and quantitative data, both at the entire project level and at the classroom level, provided a very thorough examination of the learning environments in these classrooms and a basis for the conclusions from my study.

6.3 Findings and Conclusions

The primary aim of my study was to examine the learning environment of middle-school mathematics classrooms whose teachers were participating in a sustained, job-embedded teacher professional development project. The examination of those classrooms provided a 'lens' for evaluating the NO LIMIT teacher professional development project. To accomplish this examination, four research aims were proposed and each is addressed in terms of their results below.

6.3.1 Factor Structure, Reliability and Validity of Instruments

The first research aim addressed in my study was:

> *To investigate whether the WIHIC and QTI are valid and reliable in the context of Washington State middle-school mathematics classrooms.*

6.3.1.1 Factor Structure, Reliability and Validity of WIHIC

The first results from my study, as presented in Chapter 4, were related to this research aim. Factor loadings for the individual questions on the WIHIC were determined through principal component factor analysis with Oblimin rotation and Kaiser normalization. The internal consistency of the WIHIC was determined using Cronbach's alpha coefficient. The ability of the WIHIC to discriminate between students in different classes was determined through the eta^2 statistic from the ANOVA, with class membership as the independent variable. The reliability and validity of the WIHIC were calculated separately using the individual and class mean as the units of analysis, and separately for pretest and posttest administrations.

The *a priori* five-scale structure of the modified WIHIC was found to be applicable to the data obtained in the middle-school mathematics classrooms. Each of the 40 items in my modified version of the WIHIC had factor loadings greater than 0.40 within its own scale and less than 0.40 on all other scales. These results allowed data for all 40 items to be retained for further analysis. This same factor structure was found for both the pretest and posttest administrations of the WIHIC. The total variance accounted for by the seven scales of the WIHIC was approximately 57% for both pretest and posttest administrations. These results are consistent with results obtained previously (Aldridge & Fraser, 2000; Aldridge et al., 1999; Dorman, 2003; Zandvliet & Man, 2003).

Cronbach alpha coefficients for the pretest administration of the five scales of the modified WIHIC ranged 0.88 to 0.93, with the individual as the unit of analysis.

Values of the alpha coefficient for pretest class means ranged from 0.92 to 0.96. Posttest alpha coefficient values ranged from 0.88 to 0.93 with the individual as the unit of analysis and from 0.92 to 0.97 for class means. The alpha coefficient values obtained in my study are quite high and as such indicate a high level of internal consistency. These results are consistent with previous results (Aldridge et al., 1999; Fraser, 1998b).

The ability of the WIHIC to discriminate between different classrooms is another desirable characteristic. A common measure of the ability of an instrument to discriminate is the eta^2 statistic obtained from a one-way analysis of variance (ANOVA), with class membership as the independent variable. Each of the five scales of the modified WIHIC, for both the pretest and posttest administrations, showed a significant difference ($p<0.01$) between classrooms. The values for the eta^2 statistic ranged from 0.10 to 0.18 on the pretest and from 0.11 to 0.22 on the posttest and indicated that the five scales of the WIHIC were able to differentiate between different classes in my study. The ability of the WIHIC to discriminate between classes in my study is consistent with other studies involving the WIHIC (Fraser, 1998b; Rickards et al., 2000).

6.3.1.2 Factor Structure, Reliability and Validity of QTI

As described in Chapters 2 and 4, the QTI, is slightly different from most other learning environment instruments. The QTI utilizes a circumplex structure which assumes that adjacent scales should be highly positively correlated and opposing scales should be highly negatively correlated. To examine the construct validity of the QTI an examination of the interscale correlations was undertaken. The interscale correlations, for the student as the unit of analysis, on the pretest actual form ranged from 0.73 for the adjacent scales for Helping & Friendly and Leadership to -0.52 for the opposing scales of Helping & Friendly and Dissatisfied. This pattern of interscale correlations indicates construct validity for the QTI as used in my study. These results are also similar to those reported by Fisher, Rickards, and Fraser (1996).

The internal consistency reliability of the QTI was examined using Cronbach's alpha coefficient. Alpha values for the eight scales of the QTI ranged from 0.54 to 0.86 for the pretest actual form, with the student as the unit of analysis, and 0.53 to 0.94 for class means. On the posttest actual form, alpha values ranged from 0.64 to 0.88 for students and 0.64 to 0.96 for class means as the unit of analysis. These values are consistent with results from previous studies (Fisher et al., 1997; Fraser, 1998b; Goh & Khine, 2002) and indicate, in the context of my study, an instrument with high internal consistency.

To assess how well the QTI was able to discriminate between classrooms, as was the case with the WIHIC, the eta^2 statistic was employed. Results of the ANOVA, with class membership as independent variable, indicated that for each of the eight QTI scales there were significant differences ($p<0.01$) between the classrooms. The eta^2 values ranged from 0.14 to 0.31 for different scales on the pretest actual form and from 0.12 to 0.31 for the comparable posttest form. Once again, these results are consistent with those from previous studies (Fisher et al., 1993; Wubbels & Levy, 1991) and strongly suggest that, when used in middle-school mathematics classrooms, the QTI is able to distinguish between classrooms.

6.3.2 Differences between Students' and Teachers' Perceptions of Interpersonal Behaviors

The second research aim addressed by my study was:

> *To investigate whether there are differences between how students and teachers perceive the learning environment.*

With the validity and reliability of the WIHIC and QTI established for middle-school mathematics classrooms in Washington State, attention was turned to examining differences between students and teachers. To measure the differences between students' and teachers' perceptions of interpersonal behaviors in the classrooms, the QTI was administered to both students and teachers at the beginning of the school year. Analysis was conducted using a sample of 44 matched pairs of teachers' scores and the class means of students' scores on the actual form of the QTI.

A MANOVA for repeated measures (with the eight scales of the QTI serving as the dependent variable and teacher/student serving as the repeated-measures independent variable) was conducted using the 44 matched pair sample. MANOVA results indicated that there were significant differences between student and teacher perceptions of interpersonal behavior, in terms of the Wilks' lambda criterion, for the whole set of QTI scales. Because significant overall differences were present, the univariate ANOVA was interpreted to show that significant differences ($p<0.01$) between student and teacher perceptions of interpersonal behaviors were present for five QTI scales: Leadership, Understanding, Helping/Friendly, Admonishing, and Strict. For the five scales showing significant differences, the effect sizes ranged from 0.50 standard deviations for Leadership to 1.12 standard deviations for Helping/Friendly.

For the five scales displaying significant differences and large effect sizes, I was able to conclude that the middle-school mathematics teachers and students perceived classroom interpersonal behaviors differently. The teachers perceived more positive behavior than did the students in terms of Leadership, Understanding, and Helping/Friendly behaviors. For the less positive behavior of Admonishing, teachers perceived that behavior less frequently than did the students. In terms of the other less positive behavior, Strict, the students perceived that behavior less frequently than did the teachers, indicating they thought the teacher less strict than did the teacher. The composite picture presented here is that teachers perceived themselves exhibiting more positive behaviors than did the students. Both in terms of the significance level and direction of differences these results are consistent with prior studies with the QTI (Kim et al., 2000; Rickards, 1998; Wubbels, Brekelmans et al., 1993).

Results based on the qualitative data collected, which were presented in the four case studies, further support the quantitative questionnaire results indicating that there were differences in how middle-school mathematics students and teachers perceived the same classrooms. It was particularly in the areas of Leadership and Strict that differences were noted between students' perceptions and what teachers thought was the situation. Through the change process, teachers perceived that they were exhibiting more leadership when in fact it was more likely just a different form of leadership, which could be a nuance possibly not discerned by the students. Also in the process of adopting the reform classroom practices, including having students direct more of their own learning, teachers felt they were giving up control (perceived by the students as strictness) while in fact the teachers felt that facilitating the student learning required more careful control; hence students perceived strict behavior less frequently than did the teacher.

The opportunity to share the results with participating teachers did not present itself during the course of my study and so the impact of that data on the teachers cannot be discussed here.

6.3.3 Pretest-Posttest Differences in Student Perceptions of the Learning Environment

The third research aim addressed by my study was:

> *To investigate whether changes occur in the learning environment over the course of sustained, job embedded professional development.*

Continuing with the theme of looking for differences, my study investigated changes in perceptions of the learning environment over time. With the individual as the unit of analysis, MANOVA was employed to detect statistically significant differences, using the scales of WIHIC and QTI as dependent variables and time of instrument administration as the repeated-measures independent variable. As further indication of educational importance, effect sizes were calculated.

6.3.3.1 Pretest-Posttest Differences for WIHIC

MANOVA results for the five scales of the WIHIC as dependent variable and time of administration (pretest or posttest) as the independent variable indicate that differences were statistically significant for the set of WIHIC scales. Interpretation of the results for individual WIHIC scales using ANOVA revealed that only the scale for Teacher Support showed a statistically significant pretest-posttest difference ($p<0.05$). The direction of the difference for this scale suggested decreasing teacher support. However the effect size for this difference was only 0.12 standard deviations, which is within Cohen’s low range and suggests low educational significance. Collectively the pretest-posttest results obtained using the WIHIC suggest that there was little or no educationally important change in the classroom learning environment during the time of my study.

6.3.3.2 Pretest-Posttest Differences for QTI

In a manner similar to the analysis of WIHIC data, MANOVA was applied to the data for the QTI pretest and posttest. MANOVA results indicated that there were statistically significant differences for the set of QTI scales overall and so ANOVA was interpreted for each individual WIHIC scale. ANOVA results revealed a significant difference between pretest and posttest administrations for five QTI scales: Leadership, Understanding, and Helping/Friendly behaviors showed declines, while the normally less positive scales of Admonishing and Dissatisfied behaviors showed increases. Taken as a whole, the results for these five scales might indicate a decline in positive learning environment over time. However, as the effect sizes for these five scales ranged from 0.09 to 0.21 standard deviations, it appears that any changes in the learning environment over time might not be educationally significant.

6.3.3.3 Differences as Discerned from Qualitative Case Studies

Results of the qualitative data analysis, through the case studies, collectively support the conclusions drawn from the quantitative data. While the teacher interviews seemed to suggest that teachers perceived changes in their classrooms, student responses to the two instruments, when examined at the classroom-level grain size, were widely ranging. My own observations of the classrooms are much more congruent with those of the students in those classrooms and further support the conclusion that, on the whole, the extent of change in learning environment was small. At the project level, there appear to have been limited changes in the learning environment of participating classrooms. Both quantitative and qualitative data suggest that, when examined on a classroom-by-classroom basis, changes could be observed. However, large-scale and consistent changes did not appear to take place.

The results indicating that changes in the learning environment, however small, were in the negative direction are not contrary to previous research. Fullan (1995), along with Hall and Hord (2001) have reported that during times of change teacher performance often declines. Research on learning environments conducted in the Netherlands (Wubbels et al., 1999) found that scores on the QTI went down in those classrooms where changes were being made. It is speculated that, when students perceive a change in their environment, they could find it unsettling and consequently perceive a decline in the quality of the learning environment. This might have been the case with certain classrooms within the NO LIMIT project, for which the teacher and I both observed positive changes in teacher behavior but the students reported declines in the learning environment scores. However, for the project as a whole, the effects sizes suggest that differences over time on both the WIHIC and QTI should be carefully interpreted.

6.3.4 Gender Differences in Learning Environment Perceptions

The fourth and final research aim was:

> *To investigate whether boys and girls differ in their perceptions of the learning environment.*

Gender differences were investigated using the within-class gender mean as the unit of analysis. The within-class gender mean was chosen as the unit of analysis because its use reduces potential analysis difficulties when males and females are represented in different proportions in a classroom. The within-class mean is obtained by calculating a class a mean for males and a separate mean for females, thus producing a matched pair of means that minimizes differences caused by unequal numbers of males and females within a class. MANOVA was applied using the scales of the WIHIC and QTI as dependent variables and student gender as the repeated-measures independent variable. Because the multivariate test showed statistically significant differences for the set of scales overall, results were interpreted using the ANOVA for each individual scale. Additionally, effect sizes were calculated for each of the scales.

6.3.4.1 Gender Differences for WIHIC

After a statistically significant difference was found for the set of WIHIC scales, interpretation of ANOVA results indicated statistically significant differences ($p<0.01$) for each of the five individual WIHIC scales (Student Cohesiveness, Teacher Support, Involvement, Task Orientation, and Investigation). It was found that males consistently perceived the learning environment more favorably than did the females. Effect sizes associated with these scales ranged from 0.12 to 0.35

standard deviations. While some of the effect sizes fall within Cohen's moderate range, they are still too modest to allow attachment of much educational importance to the gender differences. However, for Student Cohesiveness, the effect size for gender differences is 0.35 standard deviations, with males perceiving more Cohesiveness. This finding appears to be unique among previous studies.

The slight gender differences in perceptions of the learning environment, as revealed by the WIHIC, may be due to shifts in instructional approaches being used in many Washington State classrooms. The increased emphasis being placed on student discourse in classrooms may be encouraging more male students to engage with the teacher and other students, thus causing a narrowing of the gap between male and female perceptions of classroom learning environment factors such as student cohesiveness.

6.3.4.2 Gender Differences for QTI

Because MANOVA for the set of QTI scales yielded statistically significant differences between genders overall, ANOVA results were interpreted for each individual QTI scale. These results indicated that all eight scales of the QTI had statistically significant gender differences. Male students perceived significantly more positive classroom learning environment through higher scores on the positive behavior scales of Leadership, Understanding, and Helping/Friendly scales and through lower scores on the negatively-connoted scales for Uncertain, Admonishing, Dissatisfied, and Strict. Female students only perceived a more positive classroom learning environment on the scale for Student Responsibility/Freedom. Effect sizes for the eight scales yielded ranged from 0.07 to 0.25 standard deviations and would suggest that, while gender differences might exist, they are relatively small and could be of limited educational importance. These results are not unlike previous studies.

6.4 Limitations

Using perceptions of the learning environment in evaluating the impact of an ongoing teacher professional development project in middle-school mathematics classrooms has given rise to a number of limitations. The first limitation is the difficulty in generalizing my conclusions across the entire state of Washington. Because each of the Educational Service Districts utilized its own NO LIMIT project design, the results from my sample might not accurately represent the impact of the project on other classrooms. This same difficulty was identified early-on by the state-contracted evaluation group (Popejoy et al., 2005).

The second limitation of my study is that the professional development experienced by the participating teachers was not the only professional development going on at that time. Concurrently with the NO LIMIT project, the State of Washington, school districts and schools were conducting a wide range of teacher professional development in mathematics in an effort to improve student performance on the state mathematics assessment. As Thiessen points out, "[t]he classroom is more than a dependent variable patiently waiting to obstruct or welcome the passage of independent variables… into its midst" (1992, p. 88). Therefore, the results presented here must be interpreted with recognition of the complexity truly present in classrooms.

The third limitation arises from the duration of my study. While administering the learning environments instruments in the pretest and posttest format was very useful, the relatively short time between administrations could have contributed to the lack of statistically significant differences. Fullan and Stiegelbauer (1991) point out that two years of effort might be the minimum for measurable change. Using the QTI as a measure of the classroom learning environment, Wubbels, Creton, Levy and Hooymayers (1993) report that, in the absence of interventions, teacher interpersonal behavior is relatively stable over a school year, and so noticeable changes in the classroom learning environment could require several years of sustained effort.

A fourth limitation of my study is the absence of a comparison group. The overall design of my study would have been strengthened by also including an investigation of middle-school mathematics classrooms which were not participating in the NO LIMIT teacher development project. In fact, this design component was originally included in my study, but the low response rate from non-NO LIMIT classrooms prevented implementation of this part of the design.

A fifth potential limitation of my study is a lack of measures of student cognitive outcomes in mathematics. The original design of my study included the analysis of the relationship of classroom learning environments with student performance on the mathematics portion of the Washington Assessment of Student Learning (WASL). With WASL results not being made available to me by the State, I unfortunately had to eliminate that component of my study. My original intent had been to examine both cognitive and affective student outcomes, more in keeping with Level 5 of Guskey's (2000) professional development evaluation model. The inaccessibility of a common mathematics achievement measure forced me to narrow the scope of my work to the affective domain.

A sixth possible limitation within my study was restricting the interview process to the exchange by email. Allowing the respondent time to construct his or her response may have altered the content of the response. Responding to non-verbal cues and the opportunity to immediately follow up were not available, as they would be in a face-to-face interview.

The final potential limitation in my study is that qualitative data analysis always includes the bias of the observer (Doerr & Tinto, 2000). As I was the only one to conduct the classroom observations and discussions with teachers, any preconceived notions that I held might have been transmitted into the analysis. The mixed-method approach used in my study was an attempt to mitigate the impact of personal bias through 'triangulation' with other sources of data.

6.5 Implications

The results of my study have implications for researchers in both the fields of learning environments and professional development. My study represents one of a relatively small number of studies to utilize learning environments instruments to evaluate the impact of a professional development project (Mink & Fraser, 2005;

Nix et al., 2005; Pickett & Fraser, 2004). In addition, my study is one of a very few to examine middle-school mathematics classrooms in the United States.

From a learning environments research perspective, my study has implications for future research efforts focused on middle-school mathematics classrooms. From my study, it is clear that both the WIHIC and QTI are valid and reliable instruments for use in the middle-school mathematics classroom context. This will allow future researchers to conduct further studies of the impact of the learning environment on both cognitive and affective outcomes in middle-school mathematics classrooms in the United States. As educators continue to enhance their skills in helping *all* students achieve in mathematics, additional knowledge about the classroom, especially through the students' eyes, will be critical.

My study has also contributed to the larger body of knowledge about assessing classroom learning environments. Results of my study further support the international reliability and validity of the WIHIC and QTI. Based on these results, researchers will be able to use these instruments with confidence, knowing that they have been validated across grade levels, subject matters, and cultures, and with large sample size.

At a more local level, my research has implications within Washington State. The characterization of middle-school mathematics classrooms in Washington State produced by my results is believed to be the first in of its kind in Washington State. Because the educational standards set by the State of Washington have created a strong need for reforming mathematics instruction in the state, such learning environments instruments could play a key role in assisting with that reform effort.

Effective evaluation of professional development is still an emerging field. As we grow in our understanding of how to evaluate the impact of teacher professional development, the results of my study might have relevant implications. The 5-Level Guskey (2000) model calls for inclusion of measures of student outcomes, both cognitive and affective. My research focused on gathering student affective outcomes for this purpose.

Educational researchers and classroom teachers alike can benefit from my research as the results came from a tapestry woven of quantitative and qualitative data. Methodologically, the richness of the results confirms the importance of a mixed-methods approach (Tobin & Fraser, 1998b) to educational research. Pedagogically, the bricolage approach taken in my study models a process which, if adopted by a teacher, could enrich his/her sources of information about students, and perhaps ultimately lead to improved student achievement.

6.6 Future Directions

Future research involving the evaluation of professional development has considerable growth potential and the use of learning environments instruments could play an important role in that growth. Future evaluation studies could utilize the WIHIC and QTI in other teacher professional development projects which might incorporate investigation of other variables such as time between questionnaire administrations, or the classroom curricula being used, or examining relationships between changes in learning environment with changes in students' attitudes toward mathematics. With the continuing emphasis on enhancing mathematics instruction in the State of Washington, these instruments could readily be applied to the evaluation of any number of future professional development projects.

Future studies could be undertaken that would involve replication of my study in order to increase the validity and generalizability of findings in Washington State. Using the WIHIC and QTI again with another group of participants in a mathematics professional development project could add confidence in my results and in the questionnaires as tools for evaluating professional development.

Future studies could also include comparison of professional development participants with a comparison group of teachers not participating in professional development. The addition of a control group to the study could shed light on the sensitivity of these instruments to changes in the classroom and would serve to cross-validate the instruments.

In addition to inclusion of a non-participant comparison group, future studies might incorporate a longitudinal design. Increasingly teacher professional development is embracing a sustained approach, and so investigating the learning environment in classrooms over several years might provide useful information in evaluating the impact of professional development efforts. Coupled with this longitudinal design could be the inclusion of measures of student achievement in mathematics, thus allowing a more robust attempt at fulfilling Guskey's Level 5 for professional development evaluation.

Another potential area of study would involve incorporating the underlying constructs of the WIHIC and QTI, or other learning environment instruments, into the framework of the professional development project design (i.e. using scale descriptors from the instruments when establishing the goals of a professional development project). The study could then examine the effect of an overt effort to clarify the characteristics of a supportive learning environment on the changes actually produced in the classroom over time, including both students' and teachers' voices in the results.

This study could also be expanded to incorporate student interviews and face to face teacher interviews. My study incorporated only teacher-researcher conversations and interviews with NO LIMIT teachers. Inclusion of student interviews could contribute to a richer set of qualitative data, as well as enhancing the interpretations obtained through the quantitative components of the study.

Finally, the WIHIC and QTI are sufficiently easy to administer for teachers to utilize them in action research. If teachers were interested in studying the psychosocial environment in their classroom, these tools could produce results that would be easily interpreted. From the results, teachers could make changes in either classroom routines or their interpersonal behavior and monitor resulting changes in students' perceptions of the classroom learning environment.

The potential for future educational research utilizing the WIHIC and QTI is considerable. In addition to adding to academic knowledge about the very dynamic classroom system, learning environment instruments could be used by individual teachers to enhance their classroom performance and hence the cognitive and affective performance of their students.

6.7 Concluding Remarks

The results of my study suggest that, when viewed at the project level, the NO LIMIT teacher professional development project produced few, if any, detectable changes in the learning environment in participating teachers' classrooms. As one of the providers of the professional development, and as one who believes strongly in the importance of the classroom learning environment, I certainly had hoped to identify some changes in students' perceptions of the learning environment and teacher interpersonal behaviors. The evaluation team contracted by the state of Washington to evaluate the project also found "through testing in the fall with the ITBS [Iowa Test of Basic Skills], and limited access to classroom-level data from spring statewide administration, sixth graders in NO LIMIT! Classrooms have exhibited no higher scores than non-NO LIMIT! students" (Popejoy et al., 2005, p. 13). The evaluation team also found a "lack of real, systemic change across the program" (Popejoy et al., 2005, p. 15). Their findings appear to corroborate my results as obtained from a learning environment perspective.

It is my hope that this study will add to the international body of methodological and pedagogical knowledge about classroom learning environments. It is my most sincere hope that this study further illuminates the critical nature of reforming classrooms and that the evaluation of teacher professional development can be enhanced through examination of learning environments for the benefit of teachers and, most especially, students.

References

Adey, P. (2004). *The professional development of teachers: Practice and theory*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Adolphe, F. (2002). *A cross-national study of classroom environments and attitudes among junior secondary science students in Australia and Indonesia.* Unpublished doctoral thesis, Curtin University of Technology, Perth.

Aldridge, J. M., Dorman, J. P., & Fraser, B. J. (2004). Use of multitrait-multimethod modeling to validate actual and preferred forms of the technology-rich outcomes-focused learning environment inventory (TROFLEI). *Australian Journal of Educational & Developmental Psychology, 4*, 110-125.

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

Aldridge, J. M., Fraser, B. J., Fisher, D., Trinidad, S., & Wood, D. (2003, April). *Monitoring the success of an outcomes-based, technology-rich learning environment.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Aldridge, J. M., Fraser, B. J., & Huang, T. C. I. (1999). Investigating classroom environments in Taiwan and Australia with multiple research methods. *Journal of Educational Research, 39*, 48-57.

Aldridge, J. M., Fraser, B. J., & Sebela, M. P. (2004). Using teacher action research to promote constructivist classroom environments in elementary schools 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.

Aldridge, J. M., Laugksch, R. C., Seopa, M. A., & Fraser, B. J. (2006). Development and validation of an instrument to monitor the implementation of outcomes-based learning environments in science classrooms in South Africa. *International Journal of Science Education, 28*, 45-70.

Allen, D., & Fraser, B. J. (2007). Parent and student perceptions of classroom learning environment and its association with student outcomes. *Learning Environments Research, 10*, 67-82.

Allen, D., & Fraser, B. J. (in press). Parent and student perception of classroom learning environment and its association with student outcomes. *Learning Environments Research*.

American Educational Research Association. (2005). Teaching teachers: Professional development to improve student achievement. *Research Points, 3*(Summer), 1-4.

Anderson, G. (1998). *Fundamentals of educational research* (2nd ed.). London: Routledge Falmer.

Anderson, G. J., & Walberg, H. J. (1968). Classroom climate and group learning. *International Journal of Educational Sciences, 2*, 175-180.

Asghar, M., & Fraser, B. J. (1995). Classroom environment and attitudes to science in Brunei Darussalam. *Journal of Science and Mathematics Education in Southeast Asia, XVIII*, 41-47.

Atweh, B., Forgasz, H., & Nebres, B. (Eds.). (2001). *Sociocultural research on mathematics education: An international perspective*. Mahwah, NJ: Lawrence Erlbaum Associates.

Baker, D. B., Olzendam, A. M., Gratama, C. A., & Arlington, S. R. (2005). *The STAR classroom observation protocol: A measure of powerful teaching and learning in Washington State*. Seattle, WA: The BERC Group, LLC.

Bock, R. D. (Ed.). (1989). *Multilevel analysis of educational data*. San Diego, CA: Academic Press.

Bransford, J., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: brain, mind, experience, and school*. Washington, DC: National Academy Press.

Brekelmans, M., Levy, J., & Rodriguez, R. (1993). A typology of teacher communication style. In T. Wubbels & J. Levy (Eds.), *Do you know what you look like? Interpersonal relationships in education* (pp. 46-55). London: Falmer.

Brekelmans, M., van den Eeden, P., Terwel, J., & Wubbels, T. (1997). Student characteristics and learning environment interactions in mathematics and physics education: A resource perspective. *International Journal of Educational Research, 27*, 283-292.

Brekelmans, M., Wubbels, T., & Levy, J. (1993). Student performance, attitudes, instructional strategies and teacher-communication style. In T. Wubbels & J. Levy (Eds.), *Do you know what you look like? Interpersonal relationships in education* (pp. 56-63). London: The Falmer Press.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis method*. Newbury Park, CA: Sage.

Burden, R., & Fraser, B. J. (1993). Use of classroom environment assessments in school psychology: A British perspective. *Psychology in the Schools, 30*, 232-240.

Bybee, R. W. (1997). *The Sputnik era: Why is this educational reform different from all other reforms?* Paper presented at the symposium, Reflecting on Sputnik: Linking the Past, Present, and Future of Educational Reform, Washington, DC.

Clarke, D. (1994). Ten key principles from research for the professional development of mathematics teachers. In D. B. Aichele & A. F. Coxford (Eds.), *Professional development for teachers of mathematics* (pp. 37-48). Reston, VA: National Council of Teachers of Mathematics.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 12*, 155-159.

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). London: Routledge Falmer.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., et al. (1966). *Equality of educational opportunity*. Washington, DC: Government Printing Office.

Conant, J. B. (1963). *The education of American teachers*. New York: McGraw-Hill Book Company.

Costa, A. L., & Garmston, R. J. (2002). *Cognitive coaching: A foundation for renaissance schools* (2nd ed.). Norwood, MA: Christopher-Gordon Publishers.

Creton, H., Wubbels, T., & Hooymayers, H. (1993). A systems perspective on classroom communication. In T. Wubbels & J. Levy (Eds.), *Do you know* *what you look like? Interpersonal relationships in education* (pp. 1-12). London: Falmer Press.

Cuban, L. (1988). A fundamental puzzle of school reform. *Phi Delta Kappan, 69*, 341-344.

Curtin University of Technology. (2007). *Code of conduct for the responsible practice of research*. Retrieved. from http://www.policies.curtin.edu.au/documents/code_conduct_responsible_practice_research.doc.

Danielson, C. (2002). *Enhancing student achievement: a framework for school improvement*. Alexandria: ASCD.

Darling-Hammond, L. (1998). Teacher learning that supports student learning. *Educational Leadership, 55*, 6-11.

Darling-Hammond, L., & McLaughlin, M. W. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappan, 76*, 597-610.

den Brock, P., Bull, E., Fisher, D., & Rickards, T. (2006). Californian science students' perceptions of their classroom learning environments. *Educational Research and Evaluation, 12*, 3-25.

den Brock, P., Fisher, D., Wubbels, T., Brekelmans, M., & Rickards, T. (2006). Secondary teachers interpersonal behavior in Singapore, Brunei and Australia: A cross national comparison. *Asia Pacific Journal of Education, 26*, 79-95.

den Brok, P., Fisher, D., Wubbels, T., Brekelmans, M., & Rickards, T. (2006). Secondary teachers interpersonal behavior in Singapore, Brunei and Australia: A cross national comparison. *Asia Pacific Journal of Education, 26*, 79-95.

Denzin, N. K., & Lincoln, Y. S. (2005). Introduction: The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (3rd ed., pp. 1-32). Thousand Oaks, CA: Sage Publications.

Dhindsa, H. S., & Fraser, B. J. (2004). Culturally-sensitive factors in teacher trainees' learning environments. *Learning Environments Research: An International Journal, 7*, 165-181.

Doerr, H. M., & Tinto, P. P. (2000). Paradigms for teacher-centered, classroom-based research. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 403-428). Mahwah, NJ: Lawrence Erlbaum Associates.

Dorman, J. P. (2003). Cross-national validation of the What Is Happening In This Classroom? (WIHIC) questionnaire using confirmatory factor analysis. *Learning Environments Research: An International Journal, 6*, 231-245.

Dorman, J. P., Adams, J. E., & Ferguson, J. M. (2002). Psychosocial environment and student self-handicapping in secondary school mathematics classes: A cross-national study. *Educational Psychology, 22*, 499-511.

Dorman, J. P., & Ferguson, J. M. (2004). Associations between students' perceptions of mathematics classroom environment and self-handicapping in Australian and Canadian high schools. *McGill Journal of Education, 39*, 69-86.

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.

Elmore, R. F. (2002). *Bridging the gap between standards and achievement.* Washington, DC: Albert Shanker Institute.

Fan, X. (2001). Statistical significance and effect size in education research: Two sides of a coin. *The Journal of Educational Research, 94*, 275-282.

Ferguson, P. D., & Fraser, B. J. (1999). Changes in learning environment during the transition from primary to secondary school. *Learning Environments Research: An International Journal, 1*, 369-383.

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

Fisher, D., & Fraser, B. J. (1983a). 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., & Fraser, B. J. (1983b). Validity and use of the classroom environment scale. *Educational Evaluation and Policy Analysis, 5*, 261-271.

Fisher, D., 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., Fraser, B. J., & Wubbels, T. (1993). Interpersonal teacher behavior and school environment. In T. Wubbels & J. Levy (Eds.), *Do you know what you look like? Interpersonal relationships in education* (pp. 103-112). London: The Falmer Press.

Fisher, D., Harrison, A., Hofstein, A., & Henderson, D. (1998). Laboratory learning environments and practical tasks in senior secondary sciences classes. *Research in Science Education, 28*, 353-363.

Fisher, D., Rickards, T., & Fraser, B. J. (1996). Assessing teacher-student interpersonal relationships in science classes. *Australian Science Teachers Journal, 42*(3), 28-33.

Fisher, D., & Waldrip, B. (1997). Assessing culturally sensitive factors in the learning environment of science classrooms. *Research in Science Education, 27*, 41-49.

Fisher, D., & Waldrip, B. (1999). Cultural factors of science classroom learning environments, teacher-student interactions and student outcomes. *Research in Science & Technological Education, 17*, 83-96.

Fisher, D., & Waldrip, B. (2002). Measuring culturally sensitive factors of classroom learning environments with the CLEQ. In S. C. Goh & M. S. Khine (Eds.), *Studies in educational learning environments: An international perspective* (pp. 27-48). River Edge, NJ: World Scientific.

Fisher, D., Waldrip, B., & den Brock, P. (2005, August). *Primary teachers' interpersonal behaviour and cultural elements of the classroom environment.* Paper presented at the Fourth International Conference on Science, Mathematics and Technology Education, Victoria, BC.

Florida Department of Education. (2006). *Professional development system evaluation protocol*. Retrieved. from http://www.firn.edu/doe/profdev/pdf/pdsprotocol.pdf.

Fraser, B. J. (1981). *Test of science-related attitudes (TOSRA)*. Melbourne, Australia: Australian Council for Educational Research.

Fraser, B. J. (1990). *Individualized classroom environment questionnaire: Handbook and test master set*. Melbourne: Australian Council for Educational Research.

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

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. (1998a). Science learning environments: Assessments, 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. (1998b). Classroom environment instruments: Development, validity and applications. *Learning Environments Research: An International Journal, 1*, 7-33.

Fraser, B. J. (1999). "Grain Sizes" in learning environments 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 Publishing Company.

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

Fraser, B. J. (2002). Learning environments research: Yesterday, today and tomorrow. In S. C. Goh & M. S. Khine (Eds.), *Studies in educational learning environments* (pp. 1-25). Singapore: World Scientific.

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) (3rd vers.)*. Perth, Australia: Western Australian Institute of Technology.

Fraser, B. J., & Chionh, Y. H. (2000, April). *Classroom environments, 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., Fisher, D., & 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, NY.

Fraser, B. J., Giddings, G. J., & McRobbie, C. J. (1992). *Assessing the climate of science laboratory classes (What Research Says, No. 8)*. Perth: Curtin University of Technology.

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

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

Fraser, B. J., & Tobin, K. G. (1991). Combining qualitative and quantitative methods in classroom environment research. In B. J. Fraser & H. J. Walberg (Eds.), *Educational environments: Evaluation, antecedents and consequences* (pp. 271-291). London: Pergamon Press.

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., Treagust, D. F., Williamson, J. C., & Tobin, K. G. (1987). Validation and application of the college and university environment inventory (CUCEI). In B. J. Fraser (Ed.), *The study of learning envionrments: Volume 2* (pp. 17-30). Perth, Australia: Curtin University of Technology.

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

Fraser, B. J., Walberg, H. J., Welch, W. W., & Hattie, J. A. (1987). Synthesis of educational productivity research. *International Journal of Educational Research, 11*, 145-252.

Fraser, B. J., & Wubbels, T. (1995). Classroom learning environments. In B. J. Fraser & H. J. Walberg (Eds.), *Improving science education: International perspectives* (pp. 117-144). Chicago, IL: National Society for the Study of Education.

Fullan, M. (1995). *Change forces: Probing the depths of educational reform*. London: The Falmer Press.

Fullan, M., & Stiegelbauer, S. (1991). *The new meaning of educational change* (2nd ed.). New York: Teachers College Press.

Glenn, J. (2000). *Before it's too late: A report to the nation from the National Commission on Mathematics and Science Teaching for the 21st Century*. Washington, DC: U.S. Department of Education.

Goh, S. C., & Fraser, B. J. (1998). Teacher interpersonal behavior, classroom environment and student outcomes in primary mathematics in Singapore. *Learning Environments Research: An International Journal, 1*, 199-229.

Goh, S. C., & Fraser, B. J. (2000). Teacher interpersonal behavior and elementary students' outcomes. *Journal of Research in Childhood Education, 14*, 216-231.

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

Goh, S. C., Young, D. J., & Fraser, B. J. (1995). Pyschosocial climate and student outcomes in elementary mathematics classrooms. *Journal of Experimental Education, 64*, 29-40.

Goldstein, H. (1987). *Multilevel models in educational and social research*. London: Charles Griffin.

Guskey, T. R. (1994). Results-oriented professional development: In search of an optimal mix of effective practices. *Journal of Staff Development, 15*, 42-50.

Guskey, T. R. (1996). To transmit or to 'construct'? The lure of trend infatuation in teacher professional development (commentary). *Education Week, 16*(8), 34.

Guskey, T. R. (2000). *Evaluating professional development*. Thousand Oaks, CA: Corwin Press, Inc.

Guskey, T. R. (2003). What makes professional development effective? *Phi Delta Kappan, 84*, 748-750.

Guskey, T. R. (2005). Taking a second look at accountability. *Journal of Staff Development, 26*(1), 10-19.

Guskey, T. R., & Roy, P. (1995, December). *Evaluation of staff development.* Paper presented at the annual meeting of the National Staff Development Council, Chicago.

Hall, G., & Hord, S. (2001). *Implementing change: Patterns, principles, and potholes*. Boston, MA: Allyn & Bacon.

Henderson, D., & Fisher, D. (1998). Assessing learning environments in senior science laboratories. *Australian Science Teachers' Journal, 44*, 57-62.

Hofstede, G. (1984). *Culture's consequences*. Newbury Park, CA: Sage Publications.

Hopkins, D. (2001). *School improvement for real*. London: Routledge Falmer.

Hopkins, D., & Lagerweij, N. (1996). The school improvement knowledge base. In D. Reynolds, R. Bollen, B. Creemers, D. Hopkins, L. Stoll & N. Lagerweij (Eds.), *Making good schools: Linking school effectiveness and school improvement* (pp. 59-93). London: Routledge.

Hox, J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, New Jersey: Lawrence Erlbaum Associates

Hurst, R. W. (1999). Assessing the impact of a teacher enhancement program on classroom environment. Unpublished unpublished program evaluation report. University of Texas of the Permian Basin.

Jegede, O. J., Fraser, B. J., & Fisher, D. (1995). The development and validation of a Distance and opened That Learning Environment Scale. *Educational Technology Research and Development, 43*, 90-93.

Jegede, O. J., Fraser, B. J., & Okebukola, P. A. (1994). Altering socio-cultural beliefs hindering the learning of science. *Instructional Science, 22*, 137-152.

Jegede, O. J., & Okebukola, P. A. (1988). Educology of the socio-cultural factors in science learning. *International Journal of Educology, 2*, 70-85.

Johnson, B., & McClure, R. (2004). Validity and reliability of a shortened, revised version of the constructivist learning environment survey (CLES). *Learning Environments Research: An International Journal, 7*, 65-80.

Johnson, C. C., Kahle, J. B., & Fargo, J. D. (2007). A study of the effect of sustained, whole-school professional development on student achievement in science. *Journal of Research in Science Teaching, 44*, 775-786.

Johnson, D., & Johnson, R. (1991). *Learning together and alone* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Johnson, D., Johnson, R., Holubec, E., & Roy, P. (1984). *Circles of learning*. Washington, DC: Association for Supervision and Curriculum Development.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed research methods: A research paradigm whose time has come. *Educational Researcher, 33*, 14-26.

Kelly, A. E., & Lesh, R. (2000). Trends and shifts in research methods. In A. E. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 35-44). Mahwah, NJ: Erlbaum.

Khine, M. S., & Fisher, D. (2002, April). *Analysing interpersonal behavior 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. (in press). Using classroom pyschosocial environment in the evaluation of adult computer application courses in Singapore. *Technology, Pedagogy & Education*.

Killion, J. (2002). *Assessing impact: Evaluating staff development*. Oxford, OH: National Staff Development Council.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

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

Kim, H.-B., Fisher, D., & Fraser, B. J. (2000). Classroom environment and teacher interpersonal behavior in secondary science classes in Korea. *Evaluation and Research in Education, 14*(1), 3-22.

Kim, H.-B., & Kim, D. Y. (1995). Survey on the perceptions towards science laboratory classroom environment of university students majoring in education. *Journal of the Korean Association for Research in Science Education, 14*, 163-171.

Kim, H.-B., & Kim, D. Y. (1996). Middle and high school students' perceptions of science laboratory and their attitudes in science and science subjects. *Journal of the Korean Association for Research in Science Education, 16*, 210-216.

Koul, R. B., & Fisher, D. (2005). Cultural background and students' perceptions of science classroom learning environment and teacher interpersonal behavior in Jammu, India. *Learning Environments Research: An International Journal, 8*, 195-211.

Kyriakides, L., & Muijs, D. (2005). Drawing from teacher effectivess research and research into teacher interpersonal behaviour to establish a teacher evaluation system: A study on the use of student ratings to evaluate teacher behavior. *The Journal of Classroom Interaction, 20*, 44-66.

Leary, T. (1957). *Interpersonal diagnosis of personality*. New York: Ronald Press Company.

Lee, S. S. U., & Fraser, B. J. (2001, March). *High school science classroom learning environments in Korea.* Paper presented at the annual meeting of the National Association of Research in Science Teaching, St. Louis, MO.

Lee, S. S. U., & Fraser, B. J. (2002, April). *Laboratory classroom environments in Korean high schools.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

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

Lee, S. S. U., & Kim, H.-B. (2002). Curriculum development and learning environment research in Korea. In S. C. Goh & M. S. Khine (Eds.), *Studies in educational learning environments: An international perspective* (pp. 169-196). River Edge, NJ: World Scientific.

Levy, J., Creton, H., & Wubbels, T. (1993). Perceptions of interpersonal behavior. In T. Wubbels & J. Levy (Eds.), *Do you know what you look like? Interpersonal relationships in education* (pp. 29-45). London: The Falmer Press.

Levy, J., den Brock, P., Wubbels, T., & Brekelmans, M. (2003). Students' perceptions of interpersonal aspects of the learning environment. *Learning Environments Research: An International Journal, 6*, 5-36.

Levy, J., Rodriguez, R., & Wubbels, T. (1992, April). *Instructional effectiveness, communication style and teacher development.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

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

Lieberman, A. (1995). Practices that support teacher development. *Phi Delta Kappan, 76*, 591-596.

Lightburn, M. E., & Fraser, B. J. (2007). Classroom environment and student outcomes among students using anthropometry activities in high school science. *Research in Science and Technological Education, 25*, 153-166.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 140*, 1-55.

Logan, K. A., Crump, B. J., & Rennie, L. J. (2006). Measuring the computer classroom environment: Lessons learned from using a new instrument. *Learning Environments Research: An International Journal, 9*, 67-93.

Long, M. (2000). *The psychology of education*. London: Routledge Falmer.

Loucks-Horsley, S., Love, N., Stiles, K. E., Mundry, S., & Hewson, P. W. (2003). *Designing professional development for teachers of science and mathematics* (2nd ed.). Thousand Oaks, CA: Corwin Press, Inc.

Loucks-Horsley, S., & Matsumoto, C. (1999). Research on professional development for teachers of mathematics and science: The state of the scene. *School Science and Mathematics, 99*, 259 - 271.

Love, N. (2002). *Using data/getting results: A practical guide for school improvement in mathematics and science*. Norwood, MA: Christopher-Gordon Publishers.

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

Maor, D., & Fraser, B. J. (1993). Use of classroom environment perceptions in evaluating inquiry-based computer learning. In D. Fisher (Ed.), *A study of learning environments: Volume 7* (pp. 57-71). Perth, Australia: Science and Mathematics Education Centre, Curtin University of Technology.

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 at the university level in Indonesia.* Paper presented at the annual conference of the Australian Association for Research in Education, Fremantle, Western Australia.

Margianti, E. S., Fraser, B. J., & Aldridge, J. M. (2004). Learning environments perceptions, attitudes and achievement among private Indonesian university students [Electronic Version]. *International Journal of Private Higher Education*, *1* from www.xaiu.com/xaiujournal.

Martin-Dunlop, C., & Fraser, B. J. (in press). Learning environment and attitudes associated with an innovative science course designed for prospective elementary teachers. *International Journal of Science and Mathematics Education*.

Marzano, R., & Marzano, J. S. (2003). The key to classroom management. *Educational Leadership, 61*(1), 6-13.

McCoy, L. P. (2005). Effect of demographic and personal variables on achievement in eighth-grade algebra. *Journal of Educational Research, 98*, 131-135.

McKnight, C., Magid, A., Murphy, T. J., & McKnight, M. (2000). *Mathematics education research: A guide for the research mathematician*. Providence, RI: American Mathematical Society.

McLeod, D. B. (1992). Research on affect in mathematics education: A reconceptulization. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 575-596). New York: Macmillan Publishing.

Meho, L. I. (2006). E-mail interviews in qualitative research: A methodological discussion. *Journal of the American Society for Information Science and Technology, 57*, 1284-1295.

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.

Mizell, H. (2003). Facilitators: 10, refreshments: 8, evaluation: 0. *Journal of Staff Development, 24*(4), 10-13.

Moos, R. H. (1974). *The social climate scales: An overview*. Palo Alto, CA: Consulting Psychologists Press.

Moos, R. H. (1979). *Evaluating educational environments: Procedures, measures, findings, and policy implications*. San Francisco, CA: Jossey-Bass.

Moos, R. H., & Trickett, E. J. (1974). *Classroom environment scale manual*. Palo Alto, CA: Consulting Psychologists Press.

Moos, R. H., & Trickett, E. J. (1987). *Classroom environment scale manual* (2nd ed.). Palo Alto, CA: Consulting Psychologists Press.

Moss, C. H. (2003). *Improving learning environment and student outcomes in biology in North Carolina.* Unpublished doctoral thesis, Curtin University of Technology, Perth.

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

Nair, C. S., & Fisher, D. (2001). Learning environments and student attitudes toward science at the secondary and tertiary levels. *Issues in Educational Research, 11*, 12-13.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (2000). *Principals and standards for school mathematics*. Reson, VA: Author.

National Staff Development Council. (2001). *Standards for staff development* (Rev. ed.). Oxford, OH: National Staff Development Council.

Nix, R. K., Fraser, B. J., & Ledbetter, C. E. (2005). Evaluating an integrated science learning environment using the constructivist learning environment survey. *Learning Environments Research: An International Journal, 8*, 109-133.

Office of Superintendent of Public Instruction. (2004). *Washington State Report Card*. Retrieved October 19, 2005, from http://reportcard.ospi.k12.wa.us/

Office of Superintendent of Public Instruction. (2005). Washington State professional development planning guide part one: Teacher professional development. Retrieved September, 2005, from http://www.k12.wa.us/ProfDev/default.aspx

Ogbuehi, P. I., & Fraser, B. J. (2007). Learning environment, attitudes and conceptual development associated with innovative strategies in middle-school mathematics. *Learning Environments Research, 10*, 101-114.

Onwuegbuzie, A. J., & Leech, N. L. (2007). A call for qualitative power analyses. *Quality and Quantity, 41*, 105-121.

Physical Science Study Committee. (1960). *Physics*. Boston: D.C. Heath and Company.

Pickett, L. H., & Fraser, B. J. (2002, April). *The role of learning environments, 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 environments, student acheivement and attitudes, and teacher attitudes.*

Paper presented at the annual meeting of the National Association for Research in Science Teaching, Vancouver, British Columbia.

Pimentel, G. C. (Ed.). (1963). *Chemistry, an experimental science*. San Francisco: W.H. Freeman.

Popejoy, K., Myhre, O., & Carney, J. (2005). *New Outcomes: Learning improvement in mathematics integrating technology – NO LIMIT!* (Prepared for the Washington State Office of the Superintendent of Public Instruction).

Public Agenda. (2003). *Stand by me: What teachers really think about unions, merit pay, and other professional matters*. New York: Author.

Punch, K. (1998). *Introduction to social research: Quantitative and qualitative approaches*. London, England: Sage Publications.

Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher, 29*, 4-15.

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

Raaflaub, C. A., & Fraser, B. J. (2002, April). *Investigating the learning environment in Canadian mathematics and science classrooms in which laptop computers are used.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Rawnsley, D. G. (1997). *Associations between classroom learning environments, teacher interpersonal behaviour and student outcomes in secondary mathematics classrooms.* Unpublished doctoral thesis, Curtin University of Technology, Perth Australia.

Rentoul, A. J., & Fraser, B. J. (1979). Conceptualization of enquiry-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.

Rickards, T. (1998). *The relationship of teacher-student interpersonal behavior with student sex, cultural background and student outcomes*. Unpublished doctoral thesis, Curtin University of Technology, Perth Australia.

Rickards, T., Bull, E., & Fisher, D. (2000). *Application of the WIHIC in a study of school racial diversity and socio-economic status*. Paper presented at the annual conference of the Western Australian Institute for Educational Research, Perth, Australia.

Sashkin, M., & Egermeier, J. (1993). *Change models and processes: A review and synthesis of research and practice*. Washington, DC: US Department of Education.

Schoen, H. L., Cebulla, K. J., Finn, K. F., & Fi, C. (2003). Teacher variables that relate to student achievement when using a standards-based curriculum. *Journal for Research in Mathematics Education, 34*, 228 - 259.

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

Senge, P. (1990). *The fifth discipline: The art and discipline of the learning organization*. New York: Doubleday.

Shaha, S. H., Lewis, V. K., O'Donnell, T. J., & Brown, D. H. (2004). Evaluating professional development: An approach to verifying program impact on teachers and students. *Journal of Research in Professional Learning* Retrieved April, 2005, from http://nsdc.org/library/publications/research/shaha.pdf

Silverman, D. (2000). Analyzing talk and text. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 821-834). Thousand Oaks: Sage Publications.

Sparks, D. (2002). *Designing powerful professional development for teachers and principals*. Oxford, OH: National Staff Development Council.

Sparks, D., & Hirsch, S. (1997). *A new vision for staff development*. Oxford OH: Association for Supervision and Curriculum Development.

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.

Stake, R. E. (1994). Case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 236-247). Thousand Oaks, CA: Sage.

Stake, R. E. (1997). Case study methods in educational research: Seeking sweet water. In R. M. Jaeger (Ed.), *Complementary methods for research in education* (2nd ed., pp. 401-446). Washington, DC: American Educational Research Association.

Stake, R. E. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (3rd ed., pp. 443-466). Thousand Oaks, CA: Sage Publications.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York: Simon & Schuster.

Stigler, J. W., & Hiebert, J. (2004). Improving mathematics teaching. *Educational Leadership, 61*(5), 6-17.

Stockard, J., & Mayberry, M. (1992). *Effective educational environments*. Newbury Park, CA: Sage Publications.

Suter, L. E. (2005). Multiple methods: Research methods in education projects at NSF. *International Journal of Research & Method in Education, 28*, 171-181.

Taylor, B. A. (2004). *The influence of classroom environment on high school students' mathematics anxiety and attitudes.* Unpublished doctoral thesis, Curtin University of Technology, Perth.

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 a constructivist learning environments.* Paper presented at the Annual meeting of the American Educational Research Association, New Orleans, LA.

Taylor, P. C., Fraser, B. J., & Fisher, D. (1997). Monitoring constructivist classroom learning environments. *International Journal of Educational Research, 27*, 293-302.

Thiessen, D. (1992). Classroom-based teacher development. In A. Hargreaves & M. Fullan (Eds.), *Understanding Teacher Development*. New York: Teachers College Press.

Tobin, K. G., & Fraser, B. J. (1998a). 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.

Tobin, K. G., & Fraser, B. J. (1998b). Qualitative and quantitative methods in classroom environment research. 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 Pyschology, 65*, 93-102.

U.S. Department of Education. (2005). *The Secretary's fourth annual report on teacher quality: A highly qualified teacher in every classroom*. Washington, DC: Author.

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

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

Walberg, H. J., & Anderson, G. J. (1968). Classroom climate and individual learning. *Journal of Educational Pyschology, 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, 79*, 133-139.

Walker, S. L. (2003). *Development and validation of an instrument for assessing distance education learning environments in higher education: The Distance Learning Environments Survey (DELES).* Unpublished doctoral dissertation, Curtin University of Technology, Perth, Western Australia.

Walker, S. L., & Fraser, B. J. (2005). Development and validation of an instrument for assessing distance eduation learning environments in higher education: The distance education learning environments survey. *Learning Environments Research: An International Journal, 8*, 289-308.

Waller, W. (1932). *The sociology of teaching*. New York: Russell and Russell.

Watzlawick, P., Beavin, J. H., & Jackson, D. (1967). *The pragmatics of human communication*. New York: Norton.

Webster, B. J., & Fisher, D. (2003). School-level environment and student outcomes in mathematics. *Learning Environments Research, 6*, 309-326.

Weiss, I. R., & Pasley, J. D. (2004). What is high-quality instruction? *Educational Leadership, 61*, 24-28.

Wiggin, G., & McTighe, J. (1998). *Understanding by design*. Alexandria, VA: Association for Supervision and Curriculum Development.

Wiliam, D. (1998). A framework for thinking about research in mathematics and science education. In J. A. Malone, B. Atweh & J. R. Northfield (Eds.), *Research and supervision in mathematics and science education* (pp. 1-16). Mahwah, NJ: Lawrence Erlbaum Associates.

Wong, A. F. L., & Waldrip, B. (1996, November). *Science classroom learning environments and student attitudes in Singapore, Australia, and the South Pacific.* Paper presented at the annual conference of the Singapore Educational Research Association and Australian Association for Research in Education, Singapore.

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

Wubbels, T., Brekelmans, M., Creton, H., & Hooymayers, H. (1990). Teacher behavior style and learning environment. In H. C. Waxman & C. D. Ellett (Eds.), *The study of learning evironments* (Vol. 4, pp. 1-12). Houston, TX: College of Education, University of Texas.

Wubbels, T., Brekelmans, M., & Hermans, J. (1987). Teacher behavior: An important aspect of the learning environment. In B. J. Fraser (Ed.), *The study of learning environments* (Vol. 3, pp. 10-25). Perth, Australia: Curtin University.

Wubbels, T., Brekelmans, M., & Hooymayers, H. (1993). Comparison of teachers' and students' perceptions of interpersonal behavior. In T. Wubbels & J. Levy (Eds.), *Do you know what you look like? Interpersonal relationships in education* (pp. 64-80). London: The Falmer Press.

Wubbels, T., Brekelmans, M., Tartwijk, J. v., & Admiraal, W. (1999). Interpersonal relationships between teachers and students in the classroom. In H. C. Waxman & H. J. Walberg (Eds.), *New directions for teaching practice and research* (pp. 151-170). Berkeley, CA: McCutchan Publishing Corporation.

Wubbels, T., Creton, H., & Hooymayers, H. (1992). Review of research on teacher communicaton styles with use of the Leary model. *Journal of Classroom Interaction, 27*, 1-12.

Wubbels, T., Creton, H., Levy, J., & Hooymayers, H. (1993). The model for interpersonal teacher behavior. In T. Wubbels & J. Levy (Eds.), *Do you know what you look like? Interpersonal relationships in education* (pp. 13-28). London: The Falmer Press.

Wubbels, T., & Levy, J. (1991). A comparison of interpersonal behavior of Dutch and American teachers. *International Journal of Intercultural Relations, 15*, 1-18.

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

Wubbels, T., Levy, J., & Brekelmans, M. (1997). Paying attention to relationships. *Educational Leadership, 54*(7), 82-86.

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

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

Zandvliet, D. B., & Man, U. B. (2003). *Validation of the "What is happening in this classroom" questionnaire for the use in learning environments in Malaysion 'Smart School' classrooms.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Zandvliet, D. B., & Man, U. B. (2003, April). *Validation of the "What is happening in this classroom" questionnaire for the use in learning environments in Malaysion 'Smart School' classrooms.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Zandvliet, D. B., & Straker, L. (2001). Physical and psychosocial aspects of the learning environment in information technology rich classrooms. *Ergonomics, 44*, 838-857.

Appendix A

Participant Research Consent Form

ESD113/NO LIMIT Research Consent Form
A Study of Mathematics In-service Effectiveness

1. *What is the aim of the study?* The aim of this study is to investigate the impact of sustained, job-embedded staff development on the learning environment in middle-level mathematics classrooms.
2. *How were you chosen?* Teachers and classrooms in buildings where NO LIMIT is occurring were selected for participation in this study.
3. *What will be involved in participating?* Complete the Questionnaire on Teacher Interaction (QTI) yourself as a measure of the student - teacher interaction in the classroom, and administer the QTI and the What Is Happening In this Classroom survey (WIHIC) to your students. Both instruments will be administered in the fall and again in the spring.
4. *Who will know what you say?* Your name will not be connected with the data collected. You will be assigned an ID code and the correlation will be securely stored by SMEC at Curtin University, and in the data processing department at ESD113. At no time in the analysis and reporting will any completed survey form be connected with a name. Documents generated during the study will be retained in the superintendent's office of ESD113 and in the graduate studies filing cabinet in the offices of SMEC at Curtin University.
5. *What risks and benefits are associated with participation?* Your name will not appear in the transcripts. In any report or journal publication based on this study, all names, or other potentially identifying information pertaining to you, your school, or district will be omitted or changed. Where ever possible data will be presented in aggregate. Your contributions toward the study will be used to plan future professional development opportunities.
6. *What are your rights?* You may ask any questions regarding the research, and they will be answered fully. You may withdraw from the study at any time. Your participation is voluntary. Non-participation will not negatively effect your participation in the NO LIMIT project.
7. *What will be done with the results?* The results will be used to inform planning for future professional development programs. They may also be used in a doctoral dissertation, and may be published as a report in a professional journal.
8. *If you want more information, whom can you contact about the study?* Information concerning this study can be obtained from the Office of Graduate Research at Curtin University of Technology in Perth Australia. The lead researcher, Craig Gabler, can be reached at 360.586.1255 or email cgabler@esd113.k12.wa.us.

I have read the above and understand the terms of my participation.

__________________________ ___________________________________

Participant (Print) Participant (Signature)

Date__________________ Participant ID Code#_________________

Appendix B

Parent Consent Letter

October 7, 2003

Dear Parents,

As part of examining the effectiveness of a professional development opportunity for teachers, a survey will be conducted in your son or daughter's mathematics class. This survey will involve 2 short questionnaires. Both questionnaires will be given in October, and then repeated again in the spring of 2004. The questionnaires will ask students to respond to items directed at sampling their perceptions of their mathematics learning environment.

This research is being carried out in approximately 50 middle/high school mathematics classrooms throughout the state of Washington. The results will be used to further inform the professional development work done by mathematics specialists, and may be used as part of a PhD research project.

All of the questions relate to student perceptions of their mathematics classroom. Nothing of a personal nature will be asked of the students, and all information will be treated with the utmost confidentiality. No student, teacher, or school will be identified in any way in the report.

We would very much appreciate your support of this effort. Your son or daughter is free to withdraw at anytime without any classroom consequences. If you wish your son or daughter to not participate, please notify his or her teacher.

Thank you in advance, and if you have any questions concerning the research you can email me at, cgabler@esd113.k12.wa.us, or call the ESD at (360) 586-1255.

Sincerely,

Craig T. Gabler
Math Specialist & Lead Researcher
Educational Service District 113
Olympia, WA

Appendix C

Combined WIHIC and QTI Questionnaire

Questionnaire on Teacher Interaction

Your Ideal Teacher Questionnaire

The following questionnaire asks for your view of an ideal teacher's behavior. Think about your ideal teacher and keep this ideal teacher in mind as you respond to these sentences.

The questionnaire has 48 sentences about the ideal teacher. For each sentence, circle the number corresponding to your response. For example:

Never
Always
The teacher would express herself/himself clearly. 0 1 2 3
4

If you think that ideal teachers always express themselves clearly, circle the 4. If you think ideal teachers never express themselves clearly, circle the 0. You also can choose the numbers 1, 2 and 3 which are in-between. If you want to change your answer, cross it out and circle a new number. Thank you for your cooperation.

Part A: ACTUAL ID# (from front page)___________

		Almost Never	Seldom	Some-times	Often	Almost Always
1	This teacher talks enthusiastically about her/his subject.	⓪	①	②	③	④
2	This teacher trusts us.	⓪	①	②	③	④
3	This teacher seems uncertain.	⓪	①	②	③	④
4	This teacher gets angry unexpectedly.	⓪	①	②	③	④
5	This teacher explains things clearly.	⓪	①	②	③	④
6	If we don't agree with this teacher, we can talk about it.	⓪	①	②	③	④
7	This teacher is hesitant.	⓪	①	②	③	④
8	This teacher gets angry quickly.	⓪	①	②	③	④
9	This teacher holds our attention.	⓪	①	②	③	④
10	This teacher is willing to explain things again.	⓪	①	②	③	④
11	This teacher acts as if she/he does not know what to do.	⓪	①	②	③	④
12	This teacher is too quick to correct us when we break a rule.	⓪	①	②	③	④
13	This teacher knows everything that goes on in the classroom.	⓪	①	②	③	④
14	If we have something to say, this teacher will listen.	⓪	①	②	③	④
15	This teacher lets us boss her/him around.	⓪	①	②	③	④
16	This teacher is impatient.	⓪	①	②	③	④
17	This teacher is a good leader.	⓪	①	②	③	④
18	This teacher realizes when we don't understand.	⓪	①	②	③	④
19	This teacher is not sure what to do when we fool around.	⓪	①	②	③	④
20	It is easy to pick a fight with this teacher.	⓪	①	②	③	④
21	This teacher acts confidently.	⓪	①	②	③	④
22	This teacher is patient.	⓪	①	②	③	④
23	It is easy to make a fool out of this teacher	⓪	①	②	③	④
24	This teacher is sarcastic.	⓪	①	②	③	④
25	This teacher helps us with our work.	⓪	①	②	③	④
26	We can decide some things in this teacher's class.	⓪	①	②	③	④
27	This teacher thinks that we cheat.	⓪	①	②	③	④
28	This teacher is strict.	⓪	①	②	③	④
29	This teacher is friendly.	⓪	①	②	③	④
30	We can influence this teacher.	⓪	①	②	③	④
31	This teacher thinks that we don't know anything.	⓪	①	②	③	④
32	We have to be silent in this teacher's class.	⓪	①	②	③	④
33	This teacher is someone we can depend on.	⓪	①	②	③	④
34	This teacher lets us fool around in class.	⓪	①	②	③	④
35	This teacher puts us down.	⓪	①	②	③	④
36	This teacher's tests are hard.	⓪	①	②	③	④
37	This teacher has a sense of humor.	⓪	①	②	③	④
38	This teacher lets us get away with a lot in class.	⓪	①	②	③	④
39	This teacher thinks that we can't do things well.	⓪	①	②	③	④
40	This teacher's standards are very high.	⓪	①	②	③	④
41	This teacher can take a joke.	⓪	①	②	③	④
42	This teacher gives us a lot of free time in class.	⓪	①	②	③	④
43	This teacher seems dissatisfied.	⓪	①	②	③	④
44	This teacher is severe when marking papers.	⓪	①	②	③	④
45	This teacher's class is pleasant.	⓪	①	②	③	④
46	This teacher is lenient.	⓪	①	②	③	④
47	This teacher is suspicious.	⓪	①	②	③	④
48	We are afraid of this teacher	⓪	①	②	③	④

Part B: PREFERRED ID# (from front page)___________

		Almost Never	Seldom	Some-times	Often	Almost Always
1	The teacher would talk enthusiastically about her/his subject.	⓪	①	②	③	④
2	The teacher would trust students.	⓪	①	②	③	④
3	The teacher would seem uncertain.	⓪	①	②	③	④
4	The teacher would get angry unexpectedly.	⓪	①	②	③	④
5	The teacher would explain things clearly.	⓪	①	②	③	④
6	If students did not agree with the teacher, they could talk about it.	⓪	①	②	③	④
7	The teacher would be hesitant.	⓪	①	②	③	④
8	The teacher would get angry quickly.	⓪	①	②	③	④
9	The teacher would hold the students' attention.	⓪	①	②	③	④
10	The teacher would be willing to explain things again.	⓪	①	②	③	④
11	The teacher would act as if she/he did not know what to do.	⓪	①	②	③	④
12	The teacher would be too quick to correct students when they broke a rule.	⓪	①	②	③	④
13	The teacher would know everything that goes on in the classroom.	⓪	①	②	③	④
14	If students had something to say, the teacher would listen.	⓪	①	②	③	④
15	The teacher would let students boss her/him around.	⓪	①	②	③	④
16	The teacher would be impatient.	⓪	①	②	③	④
17	The teacher would be a good leader.	⓪	①	②	③	④
18	The teacher would realize when students did not understand.	⓪	①	②	③	④
19	The teacher would not be sure what to do when students fooled around.	⓪	①	②	③	④
20	It would be easy to pick a fight with the teacher.	⓪	①	②	③	④
21	The teacher would act confidently.	⓪	①	②	③	④
22	The teacher would be patient.	⓪	①	②	③	④
23	It would be easy to make a fool out of the teacher.	⓪	①	②	③	④
24	The teacher would be sarcastic.	⓪	①	②	③	④
25	The teacher would help students with their work.	⓪	①	②	③	④
26	Students could decide some things in the teacher's class.	⓪	①	②	③	④
27	The teacher would think that students cheat.	⓪	①	②	③	④
28	The teacher would be strict.	⓪	①	②	③	④
29	The teacher would be friendly.	⓪	①	②	③	④
30	Students could influence the teacher.	⓪	①	②	③	④
31	The teacher would think that students did not know anything.	⓪	①	②	③	④
32	Students would have to be silent in the teacher’s class.	⓪	①	②	③	④
33	The teacher would be someone students can depend on.	⓪	①	②	③	④
34	The teacher would let students fool around in class.	⓪	①	②	③	④
35	The teacher would put students down.	⓪	①	②	③	④
36	The teacher's tests would be hard.	⓪	①	②	③	④
37	The teacher would have a sense of humor.	⓪	①	②	③	④
38	The teacher would let students get away with a lot in class.	⓪	①	②	③	④
39	The teacher would think that students can't do things well.	⓪	①	②	③	④
40	The teacher's standards would be very high.	⓪	①	②	③	④
41	The teacher could take a joke.	⓪	①	②	③	④
42	The teacher would give students a lot of free time in class.	⓪	①	②	③	④
43	The teacher would seem dissatisfied.	⓪	①	②	③	④
44	The teacher would be severe when marking papers.	⓪	①	②	③	④
45	The teacher 's class would be pleasant.	⓪	①	②	③	④
46	The teacher would be lenient.	⓪	①	②	③	④
47	The teacher would be suspicious.	⓪	①	②	③	④
48	Students would be afraid of the teacher.	⓪	①	②	③	④

Part C: WHAT IS HAPPENING IN THIS CLASSROOM ID# (from front page)__________

		Almost Never	Seldom	Some-times	Often	Almost Always
	SC					
1	I make friends among students in this class.	⓪	①	②	③	④
2	I know other students in this class.	⓪	①	②	③	④
3	I am friendly to members of this class.	⓪	①	②	③	④
4	Members of the class are my friends.	⓪	①	②	③	④
5	I work well with other class members.	⓪	①	②	③	④
6	I help other class members who are having trouble with their work.	⓪	①	②	③	④
7	Students in this class like me.	⓪	①	②	③	④
8	In this class, I get help from other students.	⓪	①	②	③	④
	TS					
9	The teacher takes a personal interest in me.	⓪	①	②	③	④
10	The teacher goes out of his/her way to help me.	⓪	①	②	③	④
11	The teacher considers my feelings.	⓪	①	②	③	④
12	The teacher helps me when I have trouble with the work.	⓪	①	②	③	④
13	The teacher talks with me.	⓪	①	②	③	④
14	The teacher is interested in my problems.	⓪	①	②	③	④
15	The teacher moves about the class to talk with me.	⓪	①	②	③	④
16	The teacher's questions help me to understand.	⓪	①	②	③	④
	IN					
17	I discuss ideas in class.	⓪	①	②	③	④
18	I give my opinions during class discussions.	⓪	①	②	③	④
19	The teacher asks me questions.	⓪	①	②	③	④
20	My ideas and suggestions are used during classroom discussions.	⓪	①	②	③	④
21	I ask the teacher questions.	⓪	①	②	③	④
22	I explain my ideas to other students.	⓪	①	②	③	④
23	Students discuss with me how to go about solving problems.	⓪	①	②	③	④
24	I am asked to explain how I solve problems.	⓪	①	②	③	④
	IV					
25	I carry out investigations to test my ideas.	⓪	①	②	③	④
26	I am asked to think about the evidence for statements.	⓪	①	②	③	④
27	I carry out investigations to answer questions coming from discussions.	⓪	①	②	③	④
28	I explain the meaning of statements, diagrams and graphs.	⓪	①	②	③	④
29	I carry out investigations to answer questions which puzzle me.	⓪	①	②	③	④
30	I carry out investigations to answer the teacher's questions.	⓪	①	②	③	④
31	I find out answers to questions by doing investigations.	⓪	①	②	③	④
32	I solve problems by using information obtained from my own investigations	⓪	①	②	③	④
	TO					
33	Getting a certain amount of work done is important to me.	⓪	①	②	③	④
34	I do as much as I set out to do.	⓪	①	②	③	④
35	I know the goals for this class.	⓪	①	②	③	④
36	I am ready to start this class on time.	⓪	①	②	③	④
37	I know what I am trying to accomplish in this class.	⓪	①	②	③	④
38	I pay attention during this class.	⓪	①	②	③	④
39	I try to understand the work in this class.	⓪	①	②	③	④
40	I know how much work I have to do.	⓪	①	②	③	④

Appendix D

Mathematics Classroom Observation Checklist

Mathematics Classroom Observation Checklist		
Name: ____________________		Date: __________________
School: ___________________	Grade: 5 6 7 8 9	Class/time: _______________

	Worthwhile Mathematical Tasks	Comments
○	Students are engaged with the tasks	
○	Math tasks are meaningful and are directly related to the learning targets	
○	Conjectures, generalizations, and what if questions abound	
○	Misconceptions, limited understandings, and/or flawed reasoning surface.	
○	Students communicate about the math tasks at hand	
○	Requires teacher and students to use correct terminology	
○	Includes written communication as a part of classroom activities	
	Students' Role in Discourse	**Comments**
○	Students present solutions.	
○	Students question one another.	
○	Students pay attention while another student is speaking.	
○	Students use a variety of tools to reason, make connections, solve problems and communicate their thinking.	
○	Students make conjectures.	
	Tools for Discourse	**Comments**
○	Students are using "tools" to enhance discourse.	
○	Four kinds of tools are: written symbols, oral language, physical materials, previously acquired skills.	
○	Students are using the tools to: record, communicate, and think.	
○	Students are presenting and modeling their work.	
○	Students reflect on their learning.	
○	Students select tools that are appropriate.	
○	Students regularly use technology within their learning activities	
	Culture in the Classroom	**Comments**
○	Students look at problems and ideas in different ways.	
○	Students celebrate their AHA's.	
○	Wrong answers are viewed as worthwhile.	
○	Students are equitable in their spoken and unspoken messages about all students' mathematical potential.	
○	Students respect each other student's thinking.	

Mathematics Classroom Observation Checklist		
Name: ______________________		Date: ____________________
School: _____________________	Grade: 5 6 7 8 9	Class/time: ________________

	Classroom Management	Comments
○	Uses time efficiently and effectively	
○	Establishes an environment where students feel comfortable asking for help, seeking solutions, and learning from mistakes	
○	Encourages participation of all students	
○	Maintains appropriate standards of behavior and promotes fairness	
○	Students are aware of posted behavior expectations and consequences and take responsibility for their behavior	
	Questioning Techniques	**Comments**
○	Provides adequate wait time	
○	Solicits multiple approaches	
○	Asks students to explain and justify	
○	Includes all students	
○	Dignifies errors	
○	Provides immediate, specific, and positive feedback	
○	Asks higher level thinking questions, requiring higher level thinking and responses	
	Assessment	**Comments**
○	Uses a variety of assessments based on stated goals	
○	Uses assessment results to effect instruction	
○	Maintains an efficient record of assessment	
○	Assesses during instruction through listening, observing, and questioning	
○	Encourages students to analyze and correct errors	
○	Assessment addresses higher level thinking	

This checklist is adapted from the NCTM Teaching Standards, based on work by NO LIMIT Math Integration Specialists August 2002 and work done by Ruth, Chamberlin, ESD 112 MIS.
For an electronic version, more information and resources see the Administrator's Corner:
http://edtech.esd112.org/no_limit/administrators.html

Appendix E

Follow-up Questions

NO LIMIT! Follow-up Questions

Please respond to the following seven (7) questions about your involvement in the NO LIMIT! Project. Please provide at least a 3 or 4 sentence response so that your opinions can be clearly discerned.

Begin by clicking in the gray region and then 'typing' your response. The box will expand as you add text. When you are finished save this document and then attach it to me in an email. Upon receipt, your response will have a code number attached, but at no time will your name, building, district or ESD be associated with the responses. This information will be used assist in evaluation of the statewide NO LIMIT! Project.

Thank you for your participation.

1. **In what ways did the NO LIMIT project impact your classroom?**

2. **Following the NO LIMIT project, how is your classroom different?**

3. **How has your teaching changed as a result of NO LIMIT?**

4. **How did students react to the equipment that your participation in NO LIMIT brought with it?**

5. **How did students react to changes in your teaching that may have come as a result of NO LIMIT?**

6. **What are your professional goals now?**

7. **How much do you think your teaching changed as a result of your participation in NO LIMIT?**