Winter 1997

An Evaluation of the 2 + 2 for Teachers: Alternative Performance Appraisal Program in the Norfolk Public Schools PRIME Project

Alyce C. LeBlanc
Old Dominion University

AN EVALUATION OF THE
2 + 2 FOR TEACHERS: ALTERNATIVE PERFORMANCE APPRAISAL PROGRAM IN THE NORFOLK PUBLIC SCHOOLS PRIME PROJECT

by

Alyce C. LeBlanc

B.M. June 1972, Oberlin College
M.S. May 1993, University of Southern California

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy

Urban Services

Old Dominion University

December 1997

Approved by:

Dwight W. Allen, Ed.D.
Dissertation Chair

Wolfgang Pindur, Ph.D.
Member

Rebecca S. Bowers, Ed.D.
Concentration Area Director

Maurice R. Berube, Ph.D.
Member

Donna B. Evans, Ph.D., Dean
Darden College of Education

ABSTRACT

AN EVALUATION OF THE
2 + 2 FOR TEACHERS: ALTERNATIVE PERFORMANCE APPRAISAL PROGRAM
IN THE NORFOLK PUBLIC SCHOOLS PRIME PROJECT.

Alyce C. LeBlanc
Old Dominion University, 1997
Director: Dr. Dwight W. Allen

The current study was a utilization-focused implementation evaluation of the 2+2 for Teachers: Alternative Performance Appraisal Program (2+2). The 2+2 program is an experimental appraisal program designed to support substantial instructional reform in the Norfolk, Virginia based PRIME (Public school Restructuring through Innovative Mainstream Education) systemic reform project. Through frequent classroom observations by administrators, peers, and students, who offer two compliments and two suggestions for improvement at each observation, the 2+2 program provides a framework for teacher collaboration and professional growth. Ultimately, improved instruction is an expected outcome, but was not evaluated in this study. The formative evaluation of the initial 1996-97 implementation of 2+2 focused on adaptive program improvement, implementation processes, and how the 2+2 program made a difference to teachers. Both quantitative and qualitative methods were employed.

Research on school restructuring efforts of the last decade indicates that a collaborative culture among teachers can have a positive impact on teacher efficacy (certainty) and on systemic reform efforts. General agreement also exists in the education field that traditional teacher performance appraisal systems are largely ineffective in bringing about improved instruction, and are a source of anxiety for teachers.

The evaluation found overwhelmingly positive participant response to the 2+2 program. Most teachers reported they experimented with new strategies, experienced improved interaction with colleagues, were greatly encouraged by positive feedback, and preferred 2+2 as an appraisal system. Barriers to program implementation included uneven administrative support and time constraints. A need for a systemic perspective was indicated to sustain both the growth of 2+2 and the entire PRIME project.

To my parents,

George and Leona,

with love, admiration, and the deepest gratitude.

ACKNOWLEDGMENTS

Many people made major contributions to both the process and content of this dissertation. The evaluation of the 2 + 2 for Teachers: Alternative Performance Appraisal Program (2 + 2) was an outgrowth of a unique opportunity to work with the Norfolk Public Schools PRIME project over the past three years. I am deeply grateful to all the dedicated professionals with whom I had the honor and pleasure of working, especially the 2 + 2 participants. Their enthusiastic participation and candid feedback not only made this study possible, but was a source of inspiration. They are the vanguard of educational change and improvement.

I am most grateful to Dr. Dwight W. Allen, who provided the opportunity to become involved with the PRIME project. At the cutting edge of educational reform, Dr. Allen is the creator of the PRIME project, as well as the 2 + 2 for Teachers program. Most of the recent calls for breaking the mold of educational practice were advanced by Dr. Allen long ago. But beyond the challenge of his multiple alternative perspectives on teaching and learning, I am especially appreciative of his role as mentor and guide, this time in the dissertation process. As always, old mental models and hidden assumptions were exposed and new perspectives revealed. There is always a choice, once assumptions are made explicit.

Dr. Wolfgang Pindur provided invaluable guidance and insight regarding the evaluation of social and education programs. Best of all, in course work and conversation, he shared many of his evaluation experiences. His approach to evaluation, which combines theoretical rigor with the limitations imposed by real world circumstances, was especially enlightening. I am indebted for his many reminders of the need to focus, and his insistence on clarity and completeness. Many thanks are also due for the sending me in the right direction for current evaluation literature and thought.

I truly appreciate the support and feedback from Dr. Maurice Berube to whom I owe much of my perspective on current problems in education. His many insights in class, in print, and in conversation have altered and enriched my thinking processes tremendously.

I am also very grateful to Dr. Rebecca S. Bowers, director of the Ph.D. program in the Darden College of Education for her support and encouragement in bringing the process to completion. One could not wish for a more student oriented director.

To my colleague, Dr. Andrea E. Berndt, I wish to express deep appreciation for her many contributions to the data analysis of the questionnaire, of which she is co-author. Her many suggestions, support, and encouragement throughout the data analysis phase of the study were invaluable.

I wish to express appreciation to Chris Giersch, Patrick O'Shea, and Curtis Brown for their assistance with the 2 + 2 observation form analysis. Many thanks are also due Curtis Brown for his tireless assistance with library research and computer graphics.

I am also deeply indebted to Betty Levin for her generous assistance with last minute preparations of the final manuscript. Her willingness to give of her time, energy, and expertise will not be forgotten.

Beyond offering unwavering support and encouragement during the past three years, my husband, Bob, deserves much credit for the final product. A practitioner, teacher, principal, and assistant superintendent, he has helped balance my approach and shape my perspective. Always looking for ways to "break out of the box", he also keeps in touch with the current reality. He knows better than anyone that people are at the heart of systemic reform. I hope he knows he is always in my heart.

Cara and Andrew, your support was very special. Thank you so much for your patience, and for understanding that a dissertation takes time and energy away from our shared activities. I am glad you knew the "answer" to "2 + 2", and, no, there will not be a "3 + 3" sequel.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Overview

One important reason that educational reform efforts fail, according to Michael Fullan, is that "strategies that are used do not focus on things that will really make a difference. They fail to address fundamental instructional reform and associated development of new collaborative cultures among educators" (Fullan, 1993, p. 46). The PRIME (Public school Restructuring through Innovative Mainstream Education) systemic educational reform project in Norfolk, Virginia is engaged in an effort to address these issues. A major initiative of the PRIME project is the 2 + 2 for Teachers: Alternative Performance Appraisal Program (2 + 2). The 2 + 2 program is an experimental alternative to the district's teacher appraisal system. Based on frequent peer, administrator, and student observation and feedback, the 2 + 2 program was developed to address both teacher appraisal and professional development. By reducing teacher isolation, the program also seeks to foster a collaborative culture that will lead to an exchange and implementation of successful instructional strategies. The current study is a utilization-focused evaluation of the 2 + 2 program.

The PRIME Project

The following description of the PRIME systemic educational reform project will provide a context for the evaluation of the 2 + 2 for Teachers: Alternative Performance Appraisal Program. Sources include the PRIME Project Mission Statement and Guidelines, developed by PRIME participants in October 1995, and a funding proposal submitted on behalf of the PRIME Project to the National Science Foundation in August 1996. The funding proposal was written by the researcher under the direction of Dr. Dwight W. Allen, the PRIME Steering Committee, and PRIME school teachers.

The PRIME (Public School Restructuring through Innovative Mainstream Education) project is a major systemic initiative to improve education in six Norfolk Public Schools (NPS). PRIME, underway since 1994, is a unique top down/bottom up dynamic reform process being developed as a collaboration of faculty and staff from Old

Dominion University (ODU), Norfolk State University (NSU), Tidewater Community College (TCC), community members, teachers, building level and central office administrators, parents, and students of Norfolk Public Schools. The PRIME project aims to accomplish a research based, comprehensive redefinition of school roles and outcomes, challenging those closest to the classroom to re-examine every assumption about traditional schooling, and to dislodge the "program of the month" mentality that has afflicted educational reform in the past. Such a mindset refers to the steady stream of new programs which teachers are expected to implement, and which just as quickly evaporate when funding disappears, or quick results are not achieved.

The PRIME program empowers teachers and is responsive and adaptive to each school's environment. The purpose of the program is to create a new, innovative educational model to meet the needs of all students and provide college and career opportunities far beyond those currently available. PRIME was conceived by Dr. Dwight W. Allen, Eminent Professor of Educational Reform at Old Dominion University, in collaboration with Dr. Roy D. Nichols, Jr., Superintendent of the Norfolk Public Schools.

The PRIME Project is an experimental cluster of six schools, (three elementary, two middle and one high school). However, as a prototype restructuring project, PRIME aims to eventually be replicated throughout the entire Norfolk Public School system and serve as a model for other urban school districts. The project has a ten year commitment from Norfolk Public Schools, as well as the support of a 10 year waiver from state regulations. These local and state commitments guarantee continuation of the project independent of outside funding.

Both teachers and administrators at the school building level have been involved in planning PRIME from its inception. In 1994-95, some thirty study groups composed of teachers and administrators from all six PRIME schools, as well as university representatives and Norfolk Public Schools central office personnel, examined reform related topics and presented reports to large PRIME group meetings. The PRIME Steering Committee, established in 1995, includes administrators, teachers, parents, students, and university and business representatives. A nine member PRIME Advisory

Committee composed of math, science, technology, program evaluation, and organizational development faculty at Old Dominion University oversees and plans university level involvement with PRIME.

The goal of the PRIME Project, stated below, is supported by a wide range of PRIME initiatives which impact the educational process. In systemic reform, the concept of what constitutes an educational system expands beyond curriculum issues. Environmental forces such as poverty; parental involvement, the judicial system, and community attitudes affect students in a multitude of ways. Students' readiness to learn, motivation, discipline, and ability to concentrate may have constraints originating outside of the system. However well planned and no matter how great the merits of a particular math, science, or technology curriculum project, for example, a program focused exclusively on academic learning objectives will experience limited success sustaining growth over the long term. The systemic nature of educational problems mandates that staff development and teacher empowerment, pre-service teaching opportunities, mentoring, tutoring and volunteer activities, alternative scheduling, community service, and enhanced collaboration with social services, businesses and community be supported as integral to the long term success of large numbers of disadvantaged and at-risk youth.

Goal of the PRIME Project

The following goal of the PRIME project was formulated by the PRIME Steering Committee, and was officially adopted on June 20, 1997:

> To increase student achievement by implementing a K-12 public school, research based, systemic restructuring project that addresses all major interrelated educational processes by:
>
> > Transforming teacher instructional patterns so that they incorporate the best of educational practices.
> >
> > Creating the human and institutional support each individual requires for success.
> >
> > Fostering public understanding, to enable maximum commitment of all stakeholders.
>
> This statement of PRIME's goal was condensed from a previous list of 10 goals,

as part of the project's self-renewal process. The catalyst for a reexamination of the PRIME Project was a change in the district leadership on the PRIME Steering Committee in April 1997. The assistant superintendent now representing Norfolk Public Schools on the PRIME Steering Committee is seeking clarification of the project and a predictable level of commitment from PRIME schools.

The PRIME Project's Vision and Mission Statements

As a part of the same renewal process, the PRIME mission and vision statements were reviewed and slightly reformulated in the interest of clarity. Substantive changes were not made.

Vision Statement.

By the year 2000, PRIME schools will have an exemplary educational program that prepares students to meet the challenges of the twenty-first century. Students will complete a stringent curriculum that emphasizes academic rigor, problem solving, interdisciplinary instruction, and social skills. As a result, students will demonstrate continued improvement on all state and local measures of assessment. Through a compressed curriculum PRIME schools will accelerate student learning. As a result, PRIME high school students will have the opportunity to participate in college and career training. Collaborative partnerships among the schools, families, universities, businesses, and the military will result in an innovative educational model that incorporates the best of educational practices.

We believe:

- Parents should work in collaborative partnership with their children and teachers.

- Character development is an essential component of civic responsibility and real life skills.

- Students should learn how to question, think, and solve problems in order to adapt to a changing world.

- Integration of technology should take place in all subject areas.

- Students should look forward to coming to school each day.

- Change is a necessary, positive, and continuous process.

- Community satisfaction will increase parent requests for their children to be enrolled in PRIME schools.

Mission Statement. PRIME schools are committed to providing an innovative program that will allow students to acquire the academic, problem solving, and social skills necessary for continuous success throughout their school careers and adult lives. PRIME schools promote student success through high expectations and increased opportunity.

The 2 + 2 for Teachers: Alternative Performance Appraisal Program

Much of the following description of the 2 + 2 for Teachers: Alternative Performance Appraisal Program was taken from internal PRIME project documents and an article written by Dr. Allen, Dr. Nichols, and the researcher (Allen, Nichols, & LeBlanc, 1997).

The basis of the Alternative Teacher Performance Appraisal: 2 + 2 Program (2 + 2) is a series of regular classroom observations by teachers and administrators. The observer visits a classroom and makes two compliments and two suggestions for improvement or change. Observers stay as long as needed to make solid compliments and suggestions, usually 10 or 15 minutes. The 2 + 2 observation forms are completed before the observer leaves, or by the end of the day, allowing for almost immediate feedback to the teacher. The process is the same whether the observer is an administrator or teacher. The premise is simple and straightforward: maximizing professional interactions, decreasing teacher isolation, and increasing meaningful feedback will lead to improved instruction. As part of the district waiver for PRIME, teachers in PRIME schools may choose to participate in the 2 + 2 program in lieu of the traditional Norfolk Public Schools teachers appraisal system.

Objectives

Five main objectives of the program are to:

1. Increase peer feedback and encouragement.

2. Increase collaboration and build trust by reducing teacher isolation.

3. Improve instruction by extending teachers' knowledge of teaching behaviors through sharing multiple perspectives and modeling instructional strategies.

4. Improve the preparation of interns and new teachers by providing the opportunity to systematically receive feedback from their colleagues, as well as observe and offer

feedback to them.

5. Provide an effective alternative to traditional teacher performance appraisal systems that will better serve assessment and professional growth objectives.

History

The 2 + 2 program was first developed by Dr. Dwight W. Allen in Namibia in 1994, working with completely untrained teachers who had little access to trained supervisors. The program was then transported to China where it has been used in teacher education as a method for peer critiques of microteaching lessons. In the Chinese culture, a strong bias exists which tends to automatically devalue young or inexperienced people's opinions. In 2 + 2 training sessions, however, older professionals acknowledged the surprisingly high quality of peer critique exhibited by young teachers-in-training.

In the spring of 1995, the 2 + 2 protocol was amended and documented for introduction to PRIME schools. A few months later, an evaluability assessment, under the guidance of Dr. Wolfgang Pindur, was developed. The first 2 + 2 participants began observing their colleagues on an experimental basis during the 1995-1996 school year. In the fall of 1996, the 2 + 2 program was formally adopted by four of the six PRIME schools.

Program Description

Teachers may participate in the 2 + 2 program on a voluntary basis, subject to administrative approval. Teachers are then released from their classroom duties at least one lesson period every two weeks. During this time they observe at least two colleagues' classrooms. A visit lasts only as long as it takes to write two compliments and two suggestions. Typical observations take approximately 10 minutes. Before leaving the classroom, or at a convenient time later the same day, teachers complete the 2 + 2 observation form (Figure 1) in triplicate. One copy of the form is given to the teacher who was observed. The 2 + 2 observer also keeps a copy in his/her observation portfolio, and the third copy is filed with the administration. Administrators also participate as 2 + 2 observers and document their observations solely with the 2 + 2 form. Pre- and post-observation conferences with teachers or administrators are discretionary.

NOTE TO USERS

Page(s) missing in number only; text follows. Microfilmed as received.

7

UMI

2+2 Classroom Observation
Norfolk PRIME

Teacher: **School:**   **Supervisor:**   **Date:**

    The Class: **Grade:**   **Subject:**   **Size of Class and Arrangement:**

1. Compliments   **1.**

                          2.

2. Suggestions   **1.**

                          2.

3. Reflections

Figure 1. Sample 2 + 2 Classroom Observation Form

Prior to implementation of 2 + 2, participating teachers attend orientation sessions focusing on issues such as trust, fears, and logistics. Many teachers are reluctant to open their classrooms to peers for fear of exposing themselves to criticism. These initial orientation sessions are designed to reassure teachers getting started with 2 + 2, and emphasize the need for open dialogue and supportive peer interaction. Teachers are advised to expect some feedback to be inappropriate or not useful, and are encouraged to determine for themselves what is meaningful. Inappropriate or untimely feedback should be ignored or discarded, and never taken personally. Once teachers become comfortable with consistent feedback, it is expected that more meaningful, critical observations will be possible without leading to discontent or hurt feelings. The sessions also provide examples of how to phrase comments, give and receive meaningful feedback, and suggest specific focal areas for observations.

Schools are encouraged to use 2 + 2 as part of a systematic staff development program. For example, teachers can divide into self-study groups of four or five members each, systematically observing each other, with a focused staff development agenda. This approach can support school improvement efforts, interdisciplinary thematic approaches, and teaming initiatives, for example. Teachers are encouraged to observe classrooms at other grade levels and subject areas, and to visit other schools in order to gain feedback from multiple perspectives.

When an administrator has a serious concern about a teacher's performance that he or she feels cannot be addressed by the 2 + 2 process, the administrator may remove a teacher from the 2 + 2 program. The teacher is then placed in the district's performance appraisal system for further administrative action. Removal from 2 + 2 due to the administration's concern regarding teacher performance can occur at any time during the year. Documentation of a teacher's poor performance for use with the district's appraisal system, however, may only begin after a teacher is removed from 2 + 2. Therefore, all teachers in the 2 + 2 program can be assured they have a continuing vote of confidence from the school administration, unless they are otherwise informed.

Student involvement with 2 + 2 is entirely up to the discretion of the participating teachers. Students can be invited to give compliments and suggestions periodically, as a

group or individually, to give teachers access to student perceptions. The assumption is that even though such feedback may at times be misguided, it is important for the teacher to reflect on prevalent perceptions. A parallel objective is for students' confidence in the educational process to increase as they are asked to become actively involved in providing feedback. However, compliments and suggestions are provided for the benefit of the teacher, and the teacher alone retains the right to judge when comments are useful and when to discard them.

At the end of each semester, teachers prepare a self-reflection report on their 2 + 2 experience. Teachers review their 2 + 2 portfolio and determine which 10 compliments and suggestions were of most value to them. Teachers then reflect on which of the compliments and suggestions were most useful, and how reinforcement or improvement of their classroom teaching routines were affected. Any future plans for classroom experimentation or implementation which have emerged from the 2 + 2 process are also discussed. Teachers are thus empowered to abstract from the process what they learned, and give credit to their peers for excellent compliments and suggestions. The completed self-reflection reports are submitted to the administration for review. Using the office copies of the 2 + 2 observation forms for reference, an administrator endorses the summative appraisal when he or she is in agreement that it reflects the growth of the teacher. At this point, in rare cases, the administrator may choose to remove a teacher from 2 + 2 when he or she feels the 2 + 2 appraisal process has not been an accurate and/or reflective effort.

Providing release time for regular visitations is a matter requiring creativity, with varied approaches used in the different schools. Interns provide release time in some schools, while substitute teachers take over classes in others. Some PRIME elementary schools have created schedules with resource teachers in back-to-back configurations, allowing grade levels blocks of planning time several times per week. A portion of one planning block every two weeks could be used for 2 + 2 observations. Finally, some teachers have decided to use their own planning time for periodic observations, although the preferred approach is to find ways which do not invade the already limited planning time of teachers.

Program Rationale

The 2 + 2 program is at once an appraisal process, a staff development model, and a peer observation program, which, taken together, are expected to lead to improvement of instruction. The program moves away from the traditional appraisal approach, where, typically, lengthy checklists result in an attempt to focus on everything at once with limited impact. As an appraisal process, 2 + 2 shares the typical evaluation objectives listed in traditional teacher performance appraisal programs. These objectives include, for example, "to recognize professional strengths," "to help effect the transition of curriculum development and the acquisition of new instructional strategies into the classroom," or "to encourage creativity, innovation, and leadership."

But, in contrast to the traditional appraisal system, the 2 + 2 process is teacher centered and provides continuous feedback. Through 2 + 2 compliments, teachers will experience the giving and receiving of professional recognition. Through 2 + 2 suggestions, teachers will acquire and disseminate multiple instructional strategies. The value of communication on professional issues will be appreciated. 2 + 2 places teachers themselves at the center of the appraisal process, encouraging personal, intrinsic accountability and initiative. Teachers' portfolios, which will include compliments and suggestions given and received, as well as documentation of suggestions acted upon, will be a resource when making lesson plans. Through the reflective appraisals, teachers will be stimulated to reflect on and evaluate their own instructional strategies in a non-threatening context through the sharing of multiple perspectives. The expected result is that teachers' confidence in their own judgment will be strengthened, as well as their motivation to experiment with new strategies. The cumulative effect of the 2 + 2 appraisal process is expected to be improved instruction. No summative rating or comparison is required or desired.

In going beyond the traditional appraisal process, 2 + 2 performs not only an evaluative function, but a staff development role as well. One expected outcome is that 2 + 2 will help create an environment where teachers coach each other, and where there is intensified discussion on serious educational and instructional issues leading to actual change in classroom practice. Teachers are also encouraged to determine pertinent areas

of staff development focus as an outgrowth of their observations, or as a reflection of their school's needs. Schools may decide to focus their observations on a particular staff development issue such as instructional strategies, or classroom management. When a teacher or administrator observed another teacher, the compliments and suggestions would be relevant to the agreed upon topic(s). In addition, teachers would plan and conduct staff development sessions on the areas of focus and related topics.

As a peer observation process, 2 +2 is based on the philosophy that a person does not have to be sick in order to get better. In 2 + 2, teachers are expected to benefit from numerous, regular visits, and from the multiple perspectives gained from visitations to a wide variety of classrooms. An expected outcome is that 2 + 2 will help teachers become more reflective about their own teaching and to see a role for themselves in offering encouragement and feedback to their peers. 2 + 2 is designed to enhance positive communication between teachers and their colleagues, students, interns, and administrators. In the 2 + 2 process, the honing of observation skills should lead to a heightened awareness of the entire educational setting. Teacher isolation is expected to decrease, anxiety over the appraisal process diminish, and the enthusiasm for feedback and a variety of perspectives increase. These outcomes are expected to aid the emergence of a different, collaborative school climate. A new sense of professional identity is expected to give teachers confidence in their ability to positively affect the quality of curriculum and instruction in their schools, and, most importantly, to positively affect student outcomes across the entire PRIME project.

Statement of the Problem

A culture of teacher isolation coupled with largely ineffective mechanisms for the improvement of teaching, including teacher evaluation and in-service sessions, contribute to the lack of progress in systemic educational reform. The 2 + 2 for Teachers: Alternative Performance Appraisal Program, an initiative of the PRIME Project, seeks to address these issues. Related to problems in systemic reform efforts is the failure to systematically evaluate programs as they are being implemented. A short discussion of

the problem follows. A more detailed discussion of these issues can be found in Chapter II.

Reform in the United States educational system has been ongoing over the last decade. As discussed in Chapter II, restructuring efforts and systemic reform of schools have met with limited success. One of the reasons for negligible improvement has been the focus on changing organizational structures based on the assumption that improvement in teaching practice and student achievement would inevitably follow. While the introduction of such restructuring initiatives as site-based management or flexible scheduling can provide a first step toward creating an environment for innovative teaching practices, research indicates that, unless school norms, beliefs, and attitudes also change, teaching practice remains much the same.

Recently, researchers such as Eisner (1992), Fullan (1993, 1996), Sarason (1995, 1996), and Senge (1990) have identified the need for change in school culture to occur before lasting instructional change can take effect. At the root of changes in school culture is personal change. The premise is that the extent to which teachers are self-reflective, both individually and collaboratively, determines in large measure the capacity to build learning communities. Community building is seen to be crucial to systemic educational reform.

The culture of teacher isolation, then, is an important inhibitor of school improvement. Teachers working in isolation are not empowered either to exert control over educational issues beyond their own classrooms, or to accept responsibility for improvement. The status quo thus conspires to thwart the professionalization of teachers by limiting the extent to which teachers exercise control over their professional environment.

Existing mechanisms for improving instruction are flawed. Long assumed to be a vehicle for improving teaching performance, teacher evaluation as currently practiced, is viewed in many quarters as contributing to professional growth only in the few cases of truly marginal teachers. For the great majority of teachers, the evaluation process is a stressful, but unproductive process. Likewise, teacher in-service sessions, conceived to provide ongoing training for teachers, are typically short in duration and are often

unresponsive to the needs and concerns of teachers.

In contrast to the ubiquitous teacher appraisal programs, evaluation of educational programs is often missing altogether. Despite increasing demands by the public for accountability, an evaluation design is rarely developed in conjunction with an initial program design and implementation. The time and energy required to introduce and implement reform initiatives tax resources, often to the limit. Evaluation becomes an afterthought to be undertaken if or when time and funding permit. A consequence is that in the absence of an evaluation plan, programs cannot be modified or improved upon to successfully meet their objectives. Further, programs may be continued or discarded without knowing whether they have any effect on student learning.

The 2 + 2 for Teachers: Alternative Performance Appraisal Program was created in response to the problem of teacher isolation, and to provide a mechanism for teachers to learn and grow in their instructional practice. The PRIME Project expects the 2 + 2 program to facilitate the development of a collaborative culture supportive of systemic educational reform. At the same time, to address the need for program evaluation, a utilization focused evaluation was conceived and designed as 2 + 2 was being developed and implemented to provide a means for program growth and assessment.

Purpose of the Study

The purpose of the study was to investigate the implementation process and preliminary outcomes of the 2 + 2 for Teachers: Alternative Performance Appraisal Program from its inception through the first formal year of implementation. The main focus of the study was the PRIME high school because of its relatively large scale implementation of the 2 + 2 program. Two middle schools and one elementary school with smaller numbers of participants were also studied. The study was designed as a formative, utilization focused program evaluation. According to Patton (1997, p. 23): "Program evaluation is the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future programming.

Utilization-focused program evaluation (as opposed to program evaluation in general) is evaluation done for and with specific, intended primary users for specific, intended uses."

The purpose of the evaluation was threefold: (a) to examine what happens when a new initiative is implemented within the PRIME program; (b) to use gathered information to inform the 2 + 2 program in an "ongoing cycle of reflection and innovation" (Patton, 1997, p. 68); and (c) to help enlighten thinking about the issues of teacher evaluation and isolation in a general way. The evaluation served, therefore, to enable program modification and adjustment where necessary. Through participant feedback, program implementation was improved, emerging problems addressed and corrected where possible, and data gathered on outcomes where feasible after the initial implementation phase. Outcomes included participants' attitudes toward performance appraisal systems and toward the 2 + 2 program, the extent of actual participation in the program, attitudes of administrators toward the program, the participants' perception of the value of 2 + 2, types of suggestions and compliments made, whether teacher isolation and anxiety about observations were reduced, and whether participants made actual use of suggestions and ideas generated by their own observations. The study was a formative evaluation and was not intended as a summative evaluation of the program.

Research Questions

Three research questions provided the evaluation focus:

1. How was the 2 + 2 for Teachers: Alternative Teacher Performance Appraisal Program implemented?
2. What difference did the implementation of the 2 + 2 for Teachers: Alternative Teacher Performance Appraisal Program make?
3. What were teachers' perceptions of the benefits and drawbacks of the 2 + 2 for Teachers: Alternative Performance Appraisal Program?

The questions addressed both processes and outcomes. Although the questions provided a focus for the inquiry, they were intentionally open-ended. Similar to a case study, the current research "investigates a contemporary phenomenon within its real-life

context," an appropriate strategy "especially when the boundaries between phenomenon and context are not clearly evident" (Yin, 1994, p. 13).

Significance of the Study

Educational programs are often added in schools and then discarded due to the expiration of grant funding, due to poor support and follow-up, or due simply to a belief that the program did not work. Often the duration of the program is so short that to expect results in the allotted time frame is premature and unrealistic. The well-known and disparaging moniker "program of the month" refers to the introduction and subsequent demise of numerous programs, so short-lived they are referred to as fads. Rarely is there an attempt to systematically evaluate a new program, and even less often is a utilization focused evaluation attempted. The many reasons for this omission include lack of both time and human resources. Evaluation design and data collection take time, a desire to evaluate must exist, and school personnel must be committed to seeking information which can be used to modify the program as needs dictate.

The need is great for systematic feedback when a new program is implemented. This study allowed the examination of a new program from its inception through its first year of formal implementation. Issues of concern to teachers surfaced through monitoring processes and conducting interviews and focus groups. Systemic reform projects such as the PRIME project need to look carefully at their assumptions and be prepared to adjust program elements and implementation as a function of the dynamic and complex nature of educational systems. This evaluation served as a model for further PRIME initiatives.

The study was important because of the need to identify ways to reduce teacher isolation and to find alternatives to the current teacher evaluation systems, in pursuit of greater professional growth. Ultimately, improved instruction which leads to increased student performance is the desired outcome of systemic educational reform. It was, therefore, important to identify means of developing a culture of teacher collaboration in support of improved instructional practice. If all elements of systemic reform are not

only interrelated, but particularly dependent on a culture of teacher collaboration, the close examination of a program created to expose teachers to each others' professional strategies while encouraging communication and feedback was clearly important.

Research Design

A utilization-focused evaluation is an appropriate approach to evaluation of the 2 + 2 for Teachers: Alternative Performance Appraisal Program. A utilization-focused evaluation is highly personal and situational, and "does not advocate any particular evaluation content, model, theory, or even use" (Patton, 1997, p. 22). A utilization-focused evaluation has much in common with participatory action research (PAR). Both methods are based in the alternative paradigm of qualitative methodology, and both require evaluators, or researchers, to work actively together with implementers and decision makers in all matters related to implementation and evaluation. In PAR, the researcher responsible for the evaluation is also a participant in some way in the program and has a role as a change agent as well.

Historical Data and the Role of the Researcher

This researcher has served in a participant/consultant role with the PRIME Project since fall 1994, at the time the project formally came into being. The PRIME Project has been engaged in ongoing efforts to evaluate the various PRIME initiatives, including the 2 + 2 program. As a result, the development, implementation, and evaluation of 2 + 2 have evolved together since the program's inception.

The researcher has been involved from the beginning in developing an evaluation design. In 1995, as the 2 + 2 program description was being developed, an evaluability assessment was conducted under the guidance of Dr. Wolfgang Pindur, Professor of Public Administration at Old Dominion University. A proposal prepared by the researcher to fund the PRIME Project, including the 2 + 2 initiative, was submitted to the National Science Foundation in 1996. In preparation for the submission of a proposal, Dr. Pindur again guided the evaluation discussion. The development of the evaluability assessment, including evaluable objectives and implementation and outcome

measures, has itself helped to refine program definition.

The researcher also conducted orientation sessions, administered a questionnaire, and developed interview questions in fulfilling a role as data collector for the Norfolk Public School district on behalf of the PRIME project. This historical data was utilized in the evaluation. Other historical data included feedback from participants in the first pilots of the 2 + 2 program in the spring of 1995. Interviews conducted by PRIME interns with 2 + 2 participants during fall 1996 were also included in this evaluation. The current study was concerned with monitoring implementation, examining implementation barriers and constraints, and obtaining feedback from participants about attitudes and outcomes during the initial formal implementation year of the 2 + 2 program.

Data Collection and Analysis

The high school implementation, where there was large scale participation, was the main focus of the evaluation. Although data was collected at all four schools, limited participation at the two middle schools and one elementary school makes future studies desirable.

Data collection included interviews with 2 + 2 participants, focus groups, a teacher performance attitude toward evaluation questionnaire, the compliments and suggestions from the first semester of the school year, the written reflections on 2 + 2 at the end of the first semester of the school year (self-reflection reports), and an end of the year survey incorporating attitude and outcome measures. Conversations with administrators and participants throughout the implementation period added to the understanding of what actually occurred.

Interviews were conducted to assess participant expectations and understanding of 2 + 2 at the beginning of implementation. Questions related to teacher isolation, collegiality, appraisal, and self-reflection were discussed.

A focus group was conducted at each of the four participating schools during January 1997. A second series of focus groups meetings was held in April and May 1997. The reactions of 2 + 2 participants to their experiences with the program were solicited to monitor issues of implementation, but also to promote in-depth feedback on the perceived value of 2 + 2 to teachers. The focus groups were valuable in ascertaining

the depth of teacher feelings of isolation, attitudes toward appraisal, attitudes toward 2 + 2, and self-reflection practices. Discussion also focused on the value of teacher observation as a means of improving instruction and learning new teaching strategies.

The Teacher Performance Evaluation Attitude Questionnaire is a quantitative measure which was administered twice as a pre-and post-measure. The questionnaire was first administered in the fall of 1996, as part of the 2 + 2 evaluation design. The first administration measured teacher attitudes toward the Norfolk Public School evaluation system in place from 1983-1995/96. In May/June 1997, a follow-up survey was conducted to measure teacher attitudes toward the 2 + 2 performance appraisal model.

At the end of each semester, teachers who were in their summative evaluation year and opted to participate in the 2 + 2 program in lieu of the Norfolk Public School appraisal program submitted self-reflection reports based on their 2 + 2 experience. The written reflections addressed the usefulness of the 2 + 2 compliments and suggestions, and their utility in improving instructional strategies and introducing new perspectives. The written reflections were structured as four open-ended questions, and were submitted by participating 2 + 2 teachers at the end of the fall semester, in January 1997. The data was examined to evaluate the appropriateness of the reflection questions and the quality of the responses. The researcher also had access to each individual teacher's completed 2 + 2 observation forms.

A survey was administered at the end of the school year to receive feedback from all participants. The survey gathered information regarding the relative merits of the Norfolk Public School appraisal system and the 2 + 2 program, program outcomes, and recommendations for future program development.

Individual conversations which took place between the researcher, teachers and administrators constituted another part of the data collection process. These exchanges helped inform the process of determining actual practices and attitudes in the schools.

The evaluation design incorporated quantitative and qualitative methods. The methods employed were open-ended and flexible enough to allow exploration of issues that arose as the implementation proceeded. The qualitative nature of the evaluation allowed its design to be emergent and flexible even as the study progressed (Patton,

1990). The result is an evaluation shaped and informed by participants, culminating in a better understanding of how 2 + 2 works, and its effect on teacher isolation, collaboration, and attitudes toward appraisal that were the initial catalysts for the development of 2 + 2.

The data was analyzed utilizing both qualitative and quantitative methods. Interviews, focus groups, 2 + 2 observation forms, and self-reflection reports were analyzed through a process of coding and categorizing in a search for themes and patterns. Emerging themes, or observation categories in the case of the 2 + 2 observation forms, enabled the researcher to gain an understanding of how the 2 + 2 program was perceived by participants, and what difference the program made to teachers. The survey and questionnaire underwent quantitative analysis.

Definition of Terms

1. Utilization focused research: an approach to evaluative research with variable criteria for success, requiring that relevant decision makers and information users who are real and specific human beings (not agencies or organizations) be identified and organized, and that evaluators work together with them in all decisions about the evaluation; intended to make the act of evaluation useful and relevant to those implementing the program and to decision makers concerned with design modifications and the future existence of the program (Patton, 1978).

2. Participatory Action Research (PAR): Applied research where some members of the organization or program under study participate actively with the researcher throughout the research process; an active quest for information and ideas to guide future actions (Whyte, Greenwood, & Lazes, 1991).

3. Systemic reform: Fundamental school reform encompassing a wide range of interrelated issues, notably curriculum, teacher collaboration and empowerment, changes in organizational structures, parental involvement, and instructional practices; often used interchangeably with restructuring.

4. Restructuring: Fundamental rather than superficial change in schools (Eisner,

1992); "Restructuring activities change fundamental assumptions, practices, and relationships, both within the organization and between the organization and the outside world, in ways that lead to improved and varied student learning outcomes for essentially all students" (Conley, 1993, p. 8)

5. Peer observation: In this study, non-evaluative observations of teachers by teachers for the purpose of giving feedback, leaving teachers to make their own judgement about how to improve their teaching (adapted from Barber, 1990, p. 217).

6. Peer coaching: Peer coaching, an extension of peer observation, is a non-evaluative five step process involving 1) requesting a visit, 2) conducting a visit, 3) reviewing notes, 4) talking after visit, and 5) conducting a process review (Gottesman & Jennings, 1994, p. 99).

7. Peer review: "Peer review is a process in which teachers use their own direct knowledge and experience to examine and judge the merit and value of another teacher's practice" (Peterson, 1995, p. 100).

8. Teacher isolation: Teacher isolation is the lack of communication, collaboration, and contact among teachers due to the circumstance of being alone with students for nearly the entire professional workday. Giving and receiving feedback, and professional and reflective discussion is thus severely inhibited.

9. Collaborative culture: Characteristics of a collaborative culture are trust, openness, lack of defensiveness, support, and tolerance of a diversity of viewpoints, within a basic agreement on values developed over time within a teaching staff. Collaborative cultures are committed to continuous improvement, and look to external as well as internal sources for ways to improve (Fullan & Hargreaves, 1991).

10. Teacher collaboration: Sustained peer coaching, observation, action research, team teaching, interdisciplinary planning, and a willingness to engage in collective commitment and improvement; often used interchangeably with "collegiality"; for the purpose of this study, refers only to conditions where teachers can enter into stronger relationships of professional discourse, not to superficial social

relationships.

11. Teacher efficacy: "A person's perceived expectation of succeeding at a task or obtaining a valued outcome through personal effort; efficacy for teachers is based on their perceived ability to affect students' learning" (Lee, Dedrick, & Smith, 1991, p. 191).

12. Teacher empowerment: The engagement of teachers in the change process, with "authority to plan and monitor the quality of the educational process in their schools" (Eisner, 1992, p. 616). Areas of control can include curriculum, assessment, and teaching practice.

13. Teacher evaluation: The evaluation of the value, merit, or worth of a teacher's teaching, often through a limited number of classroom observations by a single administrator for the purpose of protecting children, reassuring teachers they are doing a good job, and making personnel decisions; often said to be used to improve teaching practice as well (Peterson, 1995, p. 30).

14. Teacher evaluation system: The particular standard evaluation procedure used uniformly in a school district.

15. Performance appraisal: Used interchangeably with teacher evaluation.

Limitations of the Study

This evaluation of the 2 + 2 for Teachers: Alternative Performance Appraisal Program was limited to an action research, utilization-focused look at the program from its inception through the school year 1996-97 through the eyes of the participants and the researcher. The purpose was to guide and inform further implementation and development of the program. A summative evaluation of the program's impact was not intended. Data analysis and participant feedback informed questions relating to teacher isolation, teacher evaluation, and dissemination of instructional strategies; the main purpose of the study, however, was to use the information in ways that would improve the program design and implementation.

The study focused on the first year of implementation. Further research will be

needed to explore the development of the process and outcomes of the program as it evolves. The high school implemented 2 + 2 in greater numbers than the two middle schools and one elementary school combined. Therefore, care needs to be exercised in drawing conclusions or forming extrapolations based upon the high school experience alone.

If this utilization-focused evaluation is to be judged on its usefulness to stakeholders and its collaborative evaluative processes, a caveat is in order. Although efforts were made to involve teachers and administrators in the planning of the evaluation, two factors precluded intensive collaboration. One was the lack of time available to teachers to complete 2 + 2 observations, much less to plan evaluation strategies. Secondly, as the implementation proceeded it became clear that for administrators, the 2 + 2 program was not a priority they were willing to spend extra time promoting. While most administrators believed the 2 + 2 program had merit, they were willing to allow the researcher to accept responsibility for nearly all monitoring and outcome evaluation activities. As implementation evolves during the school year 1997-98, it is hoped that the political climate will allow much closer coordination, including defining of evaluation goals, with teachers and administrators.

The issue of researcher involvement in the program was raised in an earlier section. An objection which might be voiced is that close relationships with stakeholders, or identification with and support of program goals may lead to overly favorable reporting and a lack of honest negative feedback (Patton, 1997). As Patton comments, "new participatory, collaborative, intervention-oriented, and developmental approaches are already being used...The issues are understanding when such approaches are appropriate and helping intended users make informed decisions about their appropriateness" (pp. 112-113). The challenge is to ensure that evaluation results are reported with the needs of the intended users uppermost in mind. The researcher has made every attempt to address all relevant information in conducting the study, with the knowledge that only full and fair disclosure can aid the development of a program dependent upon voluntary participation.

Urban Rationale

This study had significance for the urban teacher. PRIME is an urban systemic restructuring project. The six PRIME schools were chosen, in part, on the basis of characteristics typical of urban schools: high drop-out rate, low student achievement, high percentage of minority students, and discipline problems. Urban schools often have limited funds for teacher salaries, educational materials, and general maintenance of the educational environment. These characteristics contribute to difficulties retaining high quality teachers. According to a Clearinghouse on Urban Education (Ascher, 1991) report, high levels of collegiality and the breakdown of teacher isolation can lead to a greater sense of effectiveness, and greater retention of good teachers. In urban schools, help in generating new instructional ideas is especially needed in teaching a particularly disengaged student body. Peers who have experience in the same urban context are a prime source of ideas.

The 2 + 2 program is a cost-neutral initiative designed to reduce teacher isolation and disseminate instructional strategies. It was, therefore, especially important to examine what 2 + 2 participants experienced in the program. Only then can it be determined whether 2 + 2 exerts a salutary effect on the urban school environment.

Summary

In this chapter, the PRIME project and the 2 + 2 for Teachers: Alternative Performance Appraisal Program were described in detail. Educational reform issues associated with teacher evaluation and isolation were presented as further context for the current utilization-focused evaluation of the 2 + 2 program implementation in 1996-97. The significance of the current study, as well as its design and rationale, were discussed. The role of the researcher as participant was also reviewed. A statement of the relevance of the current study to the urban setting concluded the introductory chapter.

The literature review which follows in Chapter II further examines the relationship between educational change, and teacher collaboration, professional growth,

and teacher appraisal.

CHAPTER II

REVIEW OF RELATED LITERATURE

Introduction

The problems addressed by the PRIME project are well known to U.S. educators. A Nation at Risk, a 1983 report by the National Commission on Excellence in Education, chronicled the mediocrity of student achievement in the United States, including negative comparisons with foreign students and declining scores in science achievement. A landmark report, A Nation at Risk was the catalyst for a new wave of school reform known as the excellence reform movement (Berube, 1994). By 1990, math skills typically covered in the seventh grade were not mastered by over half of all U.S. twelfth graders (U.S. Department of Education, 1991). In 1991, studies showed that students in the United States had not closed the gap in international comparisons of achievement (Stevenson & Stigler, 1992). The good news is that in 1994, 73.1% of high school students graduated, not including those receiving the GED, the highest percentage since 1970 (U.S. Department of Education, 1994). On the other hand, more than 25% of U.S. children continue to drop out of school, and this percentage is much higher in many urban areas. Furthermore, 47% of Americans say they do not believe a high school diploma guarantees that a student has learned basic skills (Johnson J., 1995, p.19).

Reform Initiatives

Since 1983, most schools in America have attempted reform in some guise. Initial reforms were often top-down, mandated by states or districts, and focused on isolated aspects of the educational system. Restructuring and systemic reform are two terms sometimes used interchangeably to describe the need to achieve all encompassing reform. An ambiguous term with many meanings, restructuring connotes changing schooling in fundamental ways, as opposed to tinkering around the edges, or doing more of the same but with greater effort (Conley, 1993; Whitaker & Moses, 1994). Conley (1993)

distinguishes restructuring from renewal activities by saying that "restructuring activities change fundamental assumptions, practices, and relationships, both within the organization and between the organization and the outside world, in ways that lead to improved and varied student learning outcomes for essentially all students " (p. 8).

In operational terms, restructuring can be defined as a top-down, bottom-up reform process characterized by empowerment of teachers, and changes in roles, responsibilities, and rewards (Fullan, 1996; Whitaker & Moses, 1994). Restructuring has come to mean a move toward collaborative relationships among teachers, teacher empowerment, teacher involvement in decision making (site-based management), scheduling changes such as block scheduling, and curriculum changes with an interdisciplinary focus and an emphasis on "depth over breadth" (Goldberg, 1996). The concept of the educational system has been expanded to include educators at the state level, parents, social agencies, and businesses (Conley, 1993; Moore & Esselman, 1994; Whitaker & Moses, 1994; White, 1992).

The 1995 PRIME project mission statement also reflects the notion that restructuring is an examination of basic assumptions about the education paradigm:

> The PRIME project is a major school restructuring initiative to improve education currently underway in six Norfolk Public Schools. PRIME is a unique top down/bottom up reform process being developed collaboratively with faculty and staff from Old Dominion University, Norfolk State University, Tidewater Community College, community members, teachers, building level and central office administrators, parents and students of the Norfolk Public Schools. The PRIME project aims to accomplish a research based, comprehensive redefinition of school roles and outcomes, and will sustain reform at current funding levels supplemented by seed monies for technologies, and start up funds for specific initiatives. PRIME challenges those closest to the classroom to reexamine every assumption about traditional schooling, and to dislodge the "program of the month" mentality that has afflicted educational reform in the past. The ultimate goal of the program is to create a new, innovative educational model to meet the needs of all students and provide college and career opportunities far beyond those currently offered.

Restructuring efforts and systemic school reform over the last decade have been marked by uneven success. While some schools have successfully introduced structural changes, such as block scheduling or site-based management (SBM), many have

achieved minimal, if any, impact on instruction or student achievement (Allen, 1992; Elmore, 1995). Enabling teachers to share in decision making does not alone guarantee that greater attention is paid to improving curriculum and instruction, or that student achievement and success is the primary goal. To impact student learning, structural changes must be accompanied by changes in norms, skills, habits, and beliefs (Elmore, 1995; Fullan, 1993; Sashkin & Egermeier, 1993). Teachers must successfully alter instruction to give block scheduling, for example, a raison d'etre (Canady & Rettig, 1995). If the goal is to have students be effective problem solvers, creative thinkers, collaborative group members, and continuous learners, teachers must effectively model these roles. Teachers, however, teach in the same way that they were taught, and are typically isolated in their classrooms (Sarason, 1990).

One basic reason that educational reform efforts fail, according to Michael Fullan, is that "strategies that are used do not focus on things that will really make as difference. They fail to address fundamental instructional reform and associated development of new collaborative cultures among educators" (Fullan, 1993, p. 46). Schools may appear to be "restructured" without any significant change in teaching practice (Elmore, 1995). While structural changes are difficult to implement, they are often easier than changing school climate or culture, changing attitudes and beliefs of teachers, or building a sense of community. School systems, like all systems, work to preserve themselves (Eisner, 1992; Fullan, 1996; Sarason, 1995; Senge, 1990). Thus, the "system" in systemic reform is receiving new attention in the process of change with a greater focus on the interrelatedness of human and organizational elements. As a response to the intractability of education system problems, and the disappointing results of systemic reform efforts, a new focus on changes in attitudes, norms, and beliefs as a prerequisite to systemic reform has emerged (Anderson, 1995; Fullan, 1993; Senge, 1990). In the process, the importance of teachers' professional development in reaching the ultimate goal of improved student performance cannot be overlooked (Conley, 1993).

"Organizations learn only through individuals who learn. Individual learning does not guarantee organizational learning. But without it no organizational learning occurs"

(Senge, 1990, p.139). Successful school "restructuring" and improved student performance are not guaranteed by teachers who are reflective learners but without reflective learning, no improvement can take place.

<center>Teacher Staff Development</center>

<center><u>Teacher Inservice</u></center>

Traditional path to teacher learning has been through staff development, typically an inservice session or sessions designed as a one shot workshop. The purpose of inservice sessions is to improve instruction through the introduction of new practices or strategies. Teacher inservice has been criticized on several fronts as vehicles of teacher learning. Inservice seminars are "like a voice coach giving advice to a singer whom he or she has never heard sing. General recommendations go only so far" (Eisner, 1992, p. 614). The assumption that hearing about new practices naturally leads to implementation is a gross underestimation of what it takes for teachers to make use of new strategies and programs (Eisner, 1992).

Teachers are not often consulted on what type of assistance they would consider of value, adding to perceptions that staff development is a waste of time. An issue of <u>Horace</u>, a publication of Ted Sizer's Coalition of Essential Schools (CES) states:

> if the goal is fundamental school change from the bottom up... it must come about not through occasional advice from experts but in a more fluid, collaborative way, sustained by an active network of teachers sharing their own experiences. The Coalition's goal becomes not only to rethink classroom practices, then, but to redesign just as radically the entire process of professional development. For one thing, it is clear that school change doesn't necessarily happen schoolwide; it happens one person at a time. ("Practice Into Teaching", 1990, p. 2)

The path to personal change "could be as simple as teachers visiting each others' classes from time to time, and talking informally about the things they're trying to do" (p.3).

In an in-depth study of five of the CES, a reform initiative conceived and directed by Ted Sizer, one of the six major findings was that "gaining feedback from others provides staff with an important analytical tool, but it is not enough by itself. Equally

important is the ability to engage in tough, direct self-analysis" (Sarason, 1996, p.351). While self analysis often existed it didn't ask the tough questions or really focus on students. Rather, in schools where gains were not significant, logistical issues were addressed instead

Self reflective analysis is a "stepping back" from whatever action is undertaken to reflect on whether the effects of the program, reform, or initiative are focused on student gains (Sarason, 1996). Regular feedback from external sources is necessary to the process of self analysis. But teachers traditionally have few opportunities to see others teach, invite colleagues into their classrooms, and, except for summative evaluations, have little information about what or how they teach. Teachers, therefore, rely on their own self analysis, and researchers rely on teachers' self reports, which may reveal a picture based more on the school culture's promotion of successes rather than on critical, thoughtful analysis (Wasley, Hampel, & Clark, as cited in Sarason, 1996).

<u>Teacher Evaluation</u>

Conventional wisdom states that the traditional performance evaluation process is another road leading to teacher learning, or growth. Using the Norfolk Public Schools evaluation system as an example, the philosophy of a traditional evaluation system may be stated in the following terms: "Evaluations should include the process of defining goals and identifying, gathering, and using information for the purpose of improving instruction and professional performance. Evaluation should encourage continual professional growth with open communication and trust as the bases of this process" (Norfolk Public Schools, p. 1).

In the summative evaluation year of the Norfolk Public Schools teacher evaluation system, teachers are observed by evaluators three times using an evaluation instrument consisting "of a series of domains which delineate the essential qualities displayed by an effective teacher" (p. 6). Each observation is to be followed by a post-observation conference. At the end of the year, each teacher is ranked on each domain using the following ratings: ineffective, marginal, acceptable, proficient, and outstanding. The summative evaluation takes place every fourth year, with a requirement for professional development activities during the three years between evaluations.

Whether the traditional teacher evaluation process leads to individual learning is a subject of debate. School systems often emphasize professional growth and collaborative aspects of the process when discussing evaluation. But however much the rhetoric is slanted toward "professional growth", teacher appraisal systems tend to be punitive, and are often perceived as such by teachers.

> Imposing punitive appraisal schemes for all is like using a sledgehammer to crack a nut. It reduces 'appraisal' to the lowest common denominator. Appraisal schemes that implicate 100% of the staff in order to detect a small percentage of incompetents are a gross waste of time. Ironically, the anxiety they generate can also hold back the excellence of the many as they become reluctant to take risks for fear of punishment. (Fullan and Hargreaves, 1991, p. 10)

Teacher performance appraisal, with its observations, conferencing, and detailed documentation, is time consuming. Yet it appears to be a complex exercise of little value. Although traditional teacher performance appraisal often has negative connotations, rarely does either high or low quality teacher performance have any consequences. Despite this, administrators utilize valuable time in the standard appraisal process, and teachers feel threatened by it. "If the goal of teacher performance appraisal is to increase teacher effectiveness and lead to improved instruction, the typical performance appraisal system is hardly an ideal instrument" (Allen, Nichols, & LeBlanc, 1996, p.30).

Teacher performance has yet to be defined by a single set of criteria (Centra, 1977; Peterson, 1984; Rebell, 1990). Some teachers may not fit the framework for what constitutes good teaching on a given instrument, yet they may still impact student performance successfully (Peterson, 1995). Yet checklists for teacher evaluation are based on a wide variety of assumptions about effective teaching, through the use of experts, research bases, or committees (Good & Mulryan, 1990).

Teacher observations are controversial as a means of evaluation. "Using classroom visits...to evaluate teaching is not just incorrect, it is a disgrace" (Scriven, 1981, p. 251). Classroom interactions and dynamics are complex. Observation systems and instruments are, in most cases, too simple to be useful in reflecting the act of teaching and learning (Good, 1980). Many teacher behaviors which are appropriate for the context of their classroom are ignored or not visible during the observation period. The number of classroom observations is often insufficient to observe the wide range of

teacher behaviors and interactions, or generalize about the classroom or teacher (Peterson, 1995).

Other instruments have checklists or scales of tens of items, releasing the administrator from professional judgement and reducing the evaluation to meaningless quantification to give the impression of rigor. "Numbers, unfortunately, are no protection against nonsense" (Popham, 1987, p. 28). Flanders recommended that "the fewest number of ideas necessary to help a person develop and control his teaching behavior" be used in systematic observation systems. Limits are important because of the complexity of the teaching act and it is only possible to focus on a few key ideas at any one time. (Flanders, cited in Peterson, 1995, p. 159)

Current teacher evaluation systems are designed to assess minimal competency. In a study of 115 teachers, S. M. Johnson (1990) found that

> For...good teachers, schools offered no systematic way to productively review and improve their practice. The process of ...evaluation, supposedly meant for all teachers actually addressed the problems of only the weakest. Evaluators were seldom sufficiently skilled or experienced to offer constructive criticism in subject areas and frequently limited themselves to giving categorical praise. They concentrated on the procedural demands of the process that were subject to legal review in any dismissal case. These consumed enormous amounts of administrative time while diverting administrators' attention from the substance of most teachers' practice. (p. 274)

Peterson points out that administrators do have an important role in monitoring minimal teacher performance and severely compromised performance, typically involving very few teachers. Administrator observations, however, are traditionally the only source of teacher evaluation, even though administrators themselves may not have been exemplary classroom teachers (1995).

Teachers have long held a negative view of teacher evaluation. Peterson (1995) in researching the need for new evaluation practices states:

> Extensive interview studies of teachers show that they do not want to be evaluated, do not feel they need it to improve, or do not believe that it can be done. As it stands, evaluation is a threat to their livelihood and an intrusion on their time; they do not want or use the results of evaluation. No one wants to be made to look bad at doing something he or she cares about. Classrooms seem to go on well enough without it. There is little or no vision from teacher interview studies about how teacher evaluation could be changed so that it would be believable, credible, useful, and fair . (p. 25)

In summary, evidence that teachers or administrators grow professionally from the teacher evaluation process is nonexistent (Good and Mulryan, 1990). "Seventy years of empirical research on teacher evaluation shows that current practices do not improve teachers or accurately tell what happens in classrooms...Teacher evaluation as presently practiced does not identify innovative teaching so that it can be adopted by other teachers and used in teacher education programs. Finally current procedures do not reward exemplary teachers" (Peterson, 1995, p. 14).

Neither traditional staff development models, nor traditional teacher appraisal systems can adequately provide the support necessary to transform schools into learning organizations. A more collaborative culture is linked with schools that are changing, growing, and moving ahead. "It is assumed that improvement in teaching is a collective rather than individual enterprise, and that analysis, evaluation, and experimentation in concert with colleagues are conditions under which teachers improve" (Rosenholtz, 1989, p. 73).

Teacher Isolation

The prevailing norm in schools, however, is one of teacher isolation. Teachers work alone, without the benefit of adult referents. They are cut off from meaningful feedback about their teaching, leading to greater feelings of uncertainty about their performance (Fullan, 1991). Elliot Eisner cites teacher isolation as an obstacle to school reform because professional isolation leads to the necessity for teachers to learn on their own. Teachers are vulnerable to two types of ignorance in their reflective practice. One type is when they do not know something and are aware they do not know it. The second type is when teachers do not know something but do not know they don't know it. Teacher isolation fosters the second type of ignorance. "How can a teacher learn that he or she is talking too much, not providing sufficient time for student reflection, raising low-order questions, or is simply boring students?" (Eisner, 1992, p. 613).

Teachers are not likely to experiment and improve under conditions of isolation, thus perpetuating the status quo in education (Ashton & Webb, 1986). The tendency of

teachers to prefer to keep their ideas to themselves, to fear asking for help because they might reveal incompetence, or to fear offering help because they might be perceived as less than humble, works to "institutionalize conservatism" (Fullan & Hargreaves, 1991, p. 40). Help is also associated with early, unpleasant memories of evaluation. Thus, safe and non-risk taking teaching behavior is the norm, a strategy not likely to improve student performance. Fullan and Hargreaves (1991) point out that teacher individualism, defined as habitual patterns of working alone, is different from individuality. Individuality, where teachers are free to voice disagreement or be alone for the purpose of personal reflection, is necessary for growth. Individualism, however, leads to retrenchment.

Currently, individualism is prevalent due to high teacher expectations of themselves combined with insecurity about others' competence in relation to their own. Collaboration or peer observation becomes a potentially threatening situation fraught with risk from which teachers typically retreat in the name of lack of time. Isolation becomes a security, guarding against scrutiny. Collaboration should, therefore, be clearly distinct and separate from evaluation (Fullan& Hargreaves, 1991).

The fragmentation of the school day further exacerbates teacher isolation, making the giving and receiving of critical and supportive feedback a virtual impossibility. Especially critical is the lack of experience teachers have in the context of their colleagues' classrooms, where "the real business of education is played out" (Eisner, 1992, p.618). Schools, though, have a vested interest in arranging the school day so that teachers and schedules do not interact, simply because it is much easier to replace a teacher who functions as an individual unit (Ashton & Webb, 1986; Flinders, 1988).

Teacher Collaboration

Teacher collaboration is a term sometimes used interchangeably with collegiality. Little (1990) describes different kinds of collegial relationships among teachers. In some schools, teachers interact socially and relate stories, help each other, but only when asked, and confirm each other's ideas rather than debating or discussing new ideas.

These forms of collegiality can support stronger forms of collaboration, but are unlikely to serve as catalysts for growth on their own. A fourth type of collaboration involves sustained peer coaching, observation, action research, team teaching, or interdisciplinary planning, and implies a willingness to engage in collective commitment and improvement. Little warns that many examples of "collegiality" are superficial and do not create the conditions where teachers can enter into stronger relationships of professional discourse. Collaborative culture is less concerned with specific projects, or mandated meetings. Characteristics of a collaborative culture, rather, are trust, openness, lack of defensiveness, support, and tolerance of a diversity of viewpoints, within a basic agreement on values developed over time within a teaching staff. Collaborative cultures are committed to continuous improvement, and look to external as well as internal sources for ways to improve (Fullan & Hargreaves, 1991).

Fullan and Hargreaves (1991) identify three types of collaboration that, below the surface, are at best a waste of time, and at worst, detrimental to a collaborative culture: balkanization, comfortable collaboration, and contrived collegiality. Balkanization refers to groups of teachers whose loyalties lie with their particular cliques. A number of different groups can be in competition with each other, with limited communication between groups. Even groups of innovative teachers can isolate themselves in this way. High school subject area departments are particularly prone to balkanization. Interdisciplinary communication and curriculum planning, and cross 𝘰𝘳 𝘭e level cooperation is inhibited.

Measures such as exchanges of secondary school teachers and middle school teachers can help overcome balkanization (Hargreaves & Earl, 1990). Cross groupings of teachers from different grade levels can also help develop collaborative culture (Fullan & Hargreaves, 1991).

A second type of non-productive collaboration is referred to by Fullan and Hargreaves (1991) as comfortable collaboration.

> Comfortable collaboration is collaboration which does not embrace the principles of systematic reflective practice. In the prep time study, even within the most collaborative settings, there was much talk of sharing, exchanging, coordinating, celebrating and supporting. But there was virtually no talk at all about inquiring,

questioning, reflecting, criticizing and engaging in dialogue as positive or worthwhile activities. Research on site-based management also shows little evidence that this sort of collaboration results in instructional improvement in classrooms. (Levine & Eubanks, as cited in Fullan & Hargreaves, 1991, p. 56)

A third type of collaboration, contrived collegiality, can be useful in "getting started", in putting teachers in touch with each other. Curriculum planning, certain types of peer coaching, or mentorship can seem collegial, but if mandated with no support, can be counterproductive. On the positive side, contrived collegiality can serve to disrupt routines, and provide an opportunity and an environment for collaboration to develop.

Teacher Efficacy

Teacher collaboration has a positive effect on teacher uncertainty, and can raise levels of confidence (Rosenholtz , 1989). Ashton and Webb (1986) found that a teacher's sense of powerlessness is reduced and sense of efficacy is raised in a collaborative environment, a finding replicated by Raudenbush, Rowan, and Cheong (1992). The positive connection between student achievement and teacher efficacy has been well documented (Brophy & Good &, 1986; Moore & Esselman, 1994). A summary of efficacy research follows:

> Past research on the consequences of teacher efficacy reveals strong links with practice. Higher efficacy is associated with the use of teaching techniques which are more challenging and difficult, with teachers' willingness to implement innovative programs and with humanistic classroom management practices. The adoption of more effective teaching strategies is reflected in higher rankings by supervisors. There is also evidence to suggest that teacher efficacy contributes to the building of school consensus. Each of these findings suggests that higher teacher efficacy is associated with current conceptions of better teaching practice. Despite the consistency in the findings it is not clear that efficacy influences effectiveness or rather the reverse. (Ross, 1994, p. 23)

Teacher efficacy is defined as a "person's perceived expectation of succeeding at a task or obtaining a valued outcome through personal effort. Efficacy for teachers is based on their perceived ability to affect students' learning" (Lee, Dedrick, & Smith, 1991, p. 191). According to Rosenholtz (1987) four variables under school control had a direct effect on teacher efficacy: (a) collaboration with other teachers, (b) receiving positive feedback on teacher performance, (c) parent involvement in the school, and (d) school wide coordination of student behavior. The relationship between collaboration

and efficacy may be reciprocal, with efficacy contributing to increased teacher collaboration due to a sense on the part of the teacher that they have something worthwhile to give (Rosenholtz, 1989).

Teachers who interacted with peer coaches had higher general teaching efficacy (Ross, 1992). Hoover-Dempsey, Bassler, and Brissie (1987) found that teacher efficacy was higher among teachers who were more aware of the expectations of teachers in grades above and below them. One of the most frequently replicated findings is that teachers with higher teaching efficacy are more willing and likely to implement new instructional programs. This may be affected by collaboration - there was higher implementation associated with teacher efficacy only when there was substantial collaboration (Ross, 1994).

A sense of community can be established where collaboration is the norm, with similar implications for teacher efficacy. In fact, "The strongest predictor of teacher efficacy is community....Schools in which teachers feel more efficacious are likely to be environments in which human relationships are supportive" (Lee, Dedrick, & Smith, 1991, p. 205). Teachers need to understand the value and power of community, but a sense of community is not easy to achieve. It involves more than a commitment to professionalism where teachers collaborate when so inclined, and otherwise respect each other's autonomy (Kahne & Westheimer, 1992). As mentioned earlier, a collaborative community is one where willingness to give serious, ongoing thought to the real issues and to engage in critical self-reflection is the norm.

Peer Observation and Coaching

Peer observation is a form of collaboration directly related to the 2 + 2 program. Arguments for the use of peer observation include the fact that teacher colleagues are intimately familiar with real classrooms and with how children learn. Teachers also realize the limitations and demands faced by educators (Peterson, 1995). The term "peer observation" is most often used in the context of appraisal. Peer review is the term used for a more formal teacher evaluation of colleagues. Many professions outside of teaching use peer review with encouraging results. Peer coaching, on the other hand, is a non-evaluative staff development technique employing standardized techniques of visitations

and coaching sessions (Gottesman & Jennings, 1994).

The lines between evaluation and peer coaching, peer observation, and peer evaluation can be blurred in practice. Peer evaluation is, in fact, an underdeveloped and underresearched data source for teacher evaluation (McCarthey & Peterson, 1987). In a rare reference to peer evaluation in the literature, peer coaching and collegial support teams were the basis of a successful evaluation experiment in a Texas elementary school (Askins, 1994). In practice, the teachers in the program were responsible for compiling their own documentation of professional growth and measures of student achievement, so peer evaluation is a misnomer. Among the characteristics of the program were groups of two to four teachers with like goals (across grade levels) working together, engaging in monthly observations, and providing feedback to the observed teacher. Mandated traditional teacher evaluations were suspended for those volunteering for the program. The program was based on principles of peer coaching as a non-evaluative process based on classroom observations with feedback, and aimed at improving instructional technique (Ackland, 1991). Askins (1994) summarized teacher feedback from the Texas program:

> The professional bonding of teachers...was very rewarding; the communication/interaction of teachers from different grade levels was extremely helpful and...the need for curriculum alignment in certain areas became obvious; several teachers commented that they didn't realize how much they were working in isolation; the collegial, self-improvement process is not like the one-shot inservice training we usually receive; we now understand how we can collaborate to better solve some of our school problems. (p. 7)

The Texas experiment is similar in many ways to the 2 + 2 program, with results suggesting that the objectives of the 2 + 2 program can be supported.

Organizational Commitment

Teacher collaboration also has an impact on organizational commitment. It seems logical that organizational commitment is an important attribute of teachers engaged in school restructuring. As defined by Reyes (1992) teacher organizational commitment, "based on a review of recent research on school effects and the general theoretical literature on the workplace psychology" can be defined as follows:

> ...three core concepts comprise Teacher Organizational Commitment: 1) a belief

in the goals of the school, about what students should learn and be able to do as they become productive members of society; 2) and intention to remain an active member of the school; and 3) a willingness to exert extra effort that goes beyond personal interest on behalf of the school. (p. 1)

One of the most powerful predictors of organizational commitment, according to Reyes, are high levels of organizational collaboration. "Among the organizational conditions, collaborative efforts is the most powerful condition that must be present at school to enhance teacher organizational commitment" (p. 13).

Teacher Empowerment

Teacher empowerment is a characteristic of restructuring that directly affects commitment to change. In previous top-down change efforts, it was noted that those on the receiving end had limited commitment to a new program or policy when they had no role in its development. According to restructuring literature, school leadership should create conditions for teachers to have the authority to participate in decision making (Whitaker & Moses, 1994). Teachers may be given the latitude to define curriculum, assessment, instructional practice, and to initiate reform measures (Eisner, 1992). However, it is the teacher who must be willing to act out of a commitment to change.

Along with authority comes accountability. It is as yet unresolved how teachers can assume accountability for decisions related to curriculum or evaluation. In addition, most innovations are add-ons to what teachers already are doing. Teachers, as Shanker (1995) points out, are being asked to make changes in their values, beliefs, and cultures while "working 30 hours a day" (p. 82).

The 2 + 2 for Teachers alternative to traditional teacher appraisal in PRIME schools is a form of teacher empowerment necessitating the assumption of responsibility by teachers for their own performance. It attempts to minimize the "add-on" aspect of a new initiative by allowing release time for observations, and removal from the traditional appraisal process. The rationale is that empowerment will encourage commitment to changing the educational status quo, rather than a mere compliance with a reform agenda. Senge (1990) states that "the committed person brings an energy, passion, and

excitement that cannot be generated if you are only compliant, even genuinely compliant. The committed person doesn't play by the 'rules of the game'. He is responsible for the game" (p. 221). As teachers come to feel more responsible for reinventing schools, successful restructuring becomes more likely.

Experimentation

One of the most serious barriers to the dramatic transformation of K-12 education is the lack of systematic, coordinated, ongoing and sustained experimentation (Allen, 1992). Fullan and Hargreaves (1991) reveal a bias for action and, in supporting the need for experimentation, states "it is through informed experiments, pursuing promising directions, and testing out and refining new arrangements and practices that we will make the most headway. Therefore, action in trying out new approaches is imperative" (p. 63).

The 2 + 2 program is such an experiment, but, as the review of the literature demonstrates, not one ungrounded in theory. Athletes and artists know they cannot improve without constant feedback and critique. Athletes and artists are most often members of teams or orchestras and must collaborate to be successful. In 2 + 2, although the program takes the place of the school system's traditional teacher appraisal process, observations are not evaluative, but provide frequent feedback. The purpose of the 2 + 2 program is to build a truly collaborative culture, with frequent feedback from multiple observers. The evaluative element is provided by the teacher's own self-reflection. Cross grade level observations, especially the grade above and below the one being taught, may help enhance teacher efficacy, as mentioned in the discussion above on teacher efficacy.

In another departure from traditional peer coaching programs, observations are not necessarily followed by post-observation conferences. In a program where teachers receive 36 compliments and 36 suggestions each semester, conferences each week would be too time consuming, and are unnecessary as well. Some suggestions will not be useful to the teacher and require no follow up. Teachers take responsibility for following up on those that are useful. The collaboration element of 2 + 2 is reciprocal. Most sources

citing the advantages of peer observation for teachers never mention the benefit accruing to the <u>observing</u> teacher. If, as Eisner (1992) states, the context of their colleagues' classrooms is where "the real business of education is played out", teachers should be in them as much as possible (p.618).

Program Evaluation

"Inappropriate evaluation of an experiment is a common cause of its early demise" (Allen, 1992, p. 9). This citation is especially appropriate in the context of the evaluation of 2 + 2, since an approach other than an implementation and formative evaluation could leave the program without an opportunity to evolve. It would not be the findings (or lack of any findings) that would cause its demise, rather the program would be prone to deterioration without feedback from participants and close examination of its implementation.

As important as it is to learn whether a program has been effective, how it has made a difference, and its perceived strengths and weaknesses, it is just as valuable to know how and to what extent a program has been implemented as described. If outcomes are evaluated without knowledge of what happened during the actual implementation of the program, it is difficult to know what action to take since information is missing about what caused the outcome (Patton, 1990).

Deviations from program descriptions are common. In Rand's "Change Agent Study" of 293 federal programs supporting educational change, national programs were found to be implemented incrementally depending on local conditions and organizational dynamics. "Where implementation was successful, and where significant change in participant attitudes, skills, and behavior occurred, implementation was characterized by a process of mutual adaptation in which project goals and methods were modified to suit the needs and interests of the local staff and in which the staff changed to meet the requirements of the project (McLaughlin, as cited in Patton, 1990, p. 106). Patton (1990) concludes that where program implementation is a process of adaptation to local needs and interests, the means of studying implementation should be flexible, open-ended

and capable of describing program changes and development.

Utilization-focused evaluation does not preclude any methodological options. The evaluator is characterized as "active-reactive-adaptive" (Patton, 1997, p. 134), implying a dynamic, interactive assessment process. Intended users of the evaluation, in this case PRIME project schools, and the Norfolk Public School district, are the focus of the evaluation strategies. Action research is another way to characterize this evaluation of 2 + 2. In action research, "one is not only a researcher responsible for the research process, one is at the same time a participant and jointly responsible for the change process" (Karlsen, 1991, p. 145). The 2 + 2 teacher participants will also collaborate in the research. The idea of participatory action research (PAR) differentiates between participants who are especially knowledgeable or insightful about a new program, or the dynamics of the program and the organization. These participants become key informants, and active participants in the research. The flexible nature of action research allows the evaluator to take the opportunity to depart from the standard interview and have a different kind of conversation, eliciting different kinds of information (Whyte, 1991).

The evaluation of the 2 + 2 program will be based on qualitative methodologies especially suited for practical application. Patton (1990) states it very succinctly: "Pay attention, listen and watch, be open, think about what you hear and see, document systematically...and apply what you learn (p. 139). The following was voiced by John Goodlad: "Too often, proselytizing begins with a reform initiative's first accomplishments, well before a floor model is ready for display. While the advance team is on the road, visitors are dismayed to find in place so little of what is being marketed. Or, if time and energy are reserved until significant changes have been made, attempts at replication turn out to be far afield from the original (1996, p. 231). It is the intention of this study to avoid this lament by paying attention to what is happening in the 2 + 2 program from its earliest implementation, documenting it, and applying it.

Summary

This chapter provided a review of the literature tracing the development of systemic reform, the need for learning communities, and the findings that structural change does not necessarily lead directly to instructional change. In particular the need for the development of mechanisms for a reduction of teacher isolation, changes in the teacher evaluation system, and greater teacher collaboration were explored. The desirability of experimentation and the need to evaluate innovative initiatives, with specific reference to 2 + 2, were also discussed.

In Chapter III, the methodology to be used with the 2 + 2 for Teachers: Alternative Performance Appraisal Program evaluation will be further developed, including information about the research design, the data collection, and data analyses.

CHAPTER III

METHODOLOGY

Design of the Study

A utilization focused evaluation design was employed in conducting the evaluation of the 2 + 2 for Teachers: Alternative Performance Appraisal Program (2 + 2). Developed by Michael Quinn Patton (1978) over 20 years ago, "Utilization-focused evaluation is not a formal model or recipe for how to conduct evaluative research. Rather, it is an approach, an orientation, and a set of options. The active-reactive-adaptive evaluator chooses from among these options as he or she works with decision makers and information users throughout the evaluation process" (p. 284). Central to the utilization focused evaluation paradigm is that evaluation should be both useful and actually utilized. The intention of the researcher was for the 2 + 2 evaluation study to be perceived as useful by teachers and administrators participating in 2 + 2, and to be used in guiding informed decisions about program development and improvement.

Utilization-Focused Evaluation

In the mid-seventies, a recurrent concern of evaluators documented copiously in evaluation literature of the time, was that evaluations generally failed to affect decision-making in any significant way. Patton participated in a study conducted through the Minnesota Center for Social Research, University of Minnesota which found, in contrast, that "78% of responding decision makers, and 90% of responding evaluators felt that the evaluation had had an impact on the program" (Patton, 1978, p.28). The study revealed a distinct difference in perception between social scientists and program evaluators and decision-makers. The predominant view among social scientists was that the impact of evaluations should clearly be reflected in program and policy decisions, whereas the reality of most program environments is that "few major direction changing decisions" are ever made, and that "evaluation research is used as one piece of information that feeds into a slow, evolutionary process of program development" (Patton, 1978, p. 32). In light of these findings, Patton attempted to broaden the definition of evaluation utilization. "Utilization of research findings is not something that suddenly and

concretely occurs at some one distinct moment in time. Rather, utilization is a diffuse and gradual process of reducing decision maker uncertainty within an existing social context" (Patton, 1978, p.34).

Utilization, as a process of reducing uncertainty, can be ongoing during program implementation. Teachers' attitudes and expectations regarding the 2 + 2 program were particularly likely to be influenced by the availability of forums to give and receive feedback about the program. Teachers were interested in knowing how their colleagues reacted to being observed, how their colleagues found time for observations, how their colleagues experienced the 2 + 2 process, or what administrators thought about the program. They also needed to express concern about finding time for observations or difficulties making suggestions. A two-way information flow allowed reasonable suggestions to be made about how to overcome implementation problems, problems that might have remained hidden were feedback not available. Full disclosure of program policies and procedures was another means of reducing participant uncertainty and anxiety. The process of reducing uncertainty affects decision making, inasmuch as teachers and administrators are also decision makers. Their experiences with 2 + 2, whether negative or positive, will help shape future program development.

Another hallmark of the utilization focused evaluation methodology is the personal factor. The identification of the evaluation "audience" in utilization focused evaluation is different than in typical evaluation research, where written reports are aimed at organizations, or other impersonal targets. An evaluation is more likely to be utilized, and thus be useful, if "the information needs of a specific person or ...a group of identifiable and interacting persons" is targeted (Patton, 1978, p. 62). The audience in this evaluation was the superintendent of the Norfolk Public Schools, the PRIME school principals, the PRIME Steering Committee, the participating teachers, and Dr. Dwight W. Allen, developer of the PRIME project .

Utilization focused evaluation is not bound to any particular methodology (qualitative, quantitative, mixed), purpose (formative, summative, developmental) or focus (processes, outcomes, impacts, costs, and other possibilities). "Rather, it is a process for helping primary intended users select the most appropriate content, model,

methods, theory, and uses for their particular situation. Situational responsiveness guides the interactive process between evaluator and primary intended users" (Patton, 1997, p. 22).

<u>Formative Focus</u>

The approach utilized in the 2 + 2 evaluation research emphasized program development and improvement, that is, it had a formative rather than a summative focus. A formative evaluation is undertaken parallel to program implementation. The purpose is to create a feedback loop to inform participants and program administrators of difficulties, and unanticipated problems so that corrections and adjustments can be made where possible in "real time". A formative evaluation may be oriented toward improvement of a program. "Using evaluation results to improve a program turns out, in practice, to be fundamentally different from rendering judgment about overall effectiveness, merit, or worth...Improvement-oriented approaches tend to be more open ended, gathering varieties of data about strengths and weaknesses with the expectation that both will be found and each can be used to inform an ongoing cycle of reflection and innovation" (Patton, 1997, p. 68).

The researcher's objective in the current evaluation was to understand and document the process of program implementation, and to utilize participant feedback to inform program development and modification through an improvement orientation evaluation. Both monitoring and process evaluation approaches characterized the improvement oriented evaluation. Monitoring of implementation is crucial. Discovering what effects the program had on participants is important, but without knowing how and to what extent the program was implemented, it is difficult to know what action to take since information is missing about what caused the outcome (Patton, 1990). Monitoring 2 + 2 observations enabled a reasonable estimate of actual program activity to be made. Process questions focus on strengths and weaknesses of the program, and on the participant experience of 2 + 2 that will help define what 2 + 2 is. A process focus means looking at unanticipated interactions and informal patterns as well as anticipated, formal activities, and developing explanations of changes and events (Patton, 1990). "Process data permit judgments to be made about the extent to which the program or organization

is operating the way it is supposed to be operating, revealing areas in which relationships can be improved as well as highlighting strengths...[They are] also useful in permitting people not intimately involved in a program...to understand how a program operates (p. 95)." Perceptions of participants are usually included. The process data should help facilitate an informed expansion of the 2 + 2 program to other teachers and additional schools.

Research Plan

The first part of the evaluation addressed process questions through a case study approach. Case studies "are the preferred strategy when "how" or "why" questions are being posed, when the investigator has little control over events, and when the focus is on a contemporary phenomenon within some real-life context" (Yin, 1994, p. 1). A descriptive case study in response to the research question, "How was the 2 + 2 for Teachers: Alternative Performance Appraisal Program implemented?" was an appropriate methodology. Although the researcher had some control over events in the implementation of the 2 + 2 program, control did not extend into actual program policy making decisions. Control was limited to disseminating information, reacting to participant feedback, and proposing actions to decision makers.

In the second part of the evaluation, qualitative and quantitative methods will be used to examine the second and third research questions: "What difference did the implementation of the 2 + 2 for Teachers: Alternative Performance Appraisal Program make?", and "What were teachers' perceptions of the benefits and drawbacks of the 2 + 2 program?" A series of outcomes data was collected, all informing program implementation and improvement. The outcomes of the current study are teachers' perceptions of the implementation process and the value and usefulness of the 2 + 2 program. In addition, monitoring data and data obtained from the completed 2 + 2 observation forms are considered implementation outcomes. The evaluation does not address program outcomes such as improved instructional practices (see Evaluability Assessment below).

The following sections will discuss the background of the 2 + 2 evaluation, the evaluability assessment, data collection, data analysis, limitations and the researcher's role in the program, and issues of validity and reliability.

Background

The setting for this study is the PRIME (Public School Restructuring for Innovative Mainstream Education) Project in Norfolk, VA. All of the six PRIME schools (three elementary schools, two middle schools, and one high school) are urban schools. The 2 + 2 for Teachers: Alternative Performance Appraisal Program is one of twelve PRIME project reform initiatives in various stages of development. Initial orientation for the 2 + 2 program began during the summer field-based graduate level courses offered for PRIME teachers and administrators in summer 1995. Representatives from each PRIME school attended the courses and were introduced to the 2 + 2 concept at that time.

The development, implementation, and evaluation of 2 + 2 have evolved together since the program's inception. This approach is consistent with utilization focused evaluation, and with an action research typology. In action research, the program participants "solve problems by studying themselves" with the intent of making modifications, or solving problems as quickly as possible (Patton, 1990, p. 161). Evaluation and collection of feedback data must be synchronized with program development and implementation. The development of the program concurrent with the development of the evaluation plan allows for corrections to be made as the program is implemented.

The researcher has been involved in the development of the program and its evaluation since 1995. In spring of 1995, as the program description was being developed, an evaluability assessment was conducted by the researcher and another PRIME staff member under the guidance of Dr. Wolfgang Pindur (see Evaluability Assessment below). At the end of the 1995-96 school year, survey data was gathered from teachers participating informally in the initial 2 + 2 pilot implementation. In the summer of 1996, in preparation for the submission of a proposal to the National Science Foundation, an evaluation plan based on the original evaluation design was developed.

Four of the six PRIME schools (one elementary school, two middle schools, and one high school) implemented the 2 + 2 program during the 1996-97 school year. The current study focused on the 2 + 2 program implementation at Lake Taylor High School. The number of participants at the high school was more than twice that of the other three schools combined, and the 2 + 2 program more fully developed in its implementation. Data from the remaining schools was also collected and analyzed, but participation was limited. In the fall of 1996, the researcher made presentations on the 2 + 2 program to the PRIME schools and conducted orientation sessions for teachers opting for the program. The researcher has also served in the role of data collector on behalf of the Norfolk Public Schools PRIME Project, by administering the Teacher Performance Appraisal Attitude questionnaire (described below) and developing interview questions for 2 + 2 teachers.

Evaluability Assessment

In the spring of 1995, an evaluability assessment for the 2 + 2 program was conducted. An evaluability assessment is concerned with identifying the "program objectives, expectations, and causal assumptions of policy-makers and managers in charge of the program", the information needs of managers and to what extent measurement criteria have been established, and probable uses for information about the program's performance (Wholey, 1979, p. 18).

> Evaluability assessments, or "preevaluations," are designed to provide a climate favorable to future evaluation work and to create intimate acquaintance with an agency or program in order to facilitate the evaluation design. In addition, as systematic management consultation, such efforts may be utilized in and of themselves by program staff prior to further evaluation activities. Evaluability assessments can also reveal whether implementation corresponds to the program as defined by those who created its policy and operational procedures; if no, any evaluation that is undertaken will probably be useless. (Rossi and Freedman, 1993, p. 145)

In this phase of the evaluation, three PRIME school principals were interviewed, written documentation about the 2 + 2 process was reviewed, and an evaluable model in the form of a flow chart was developed. Figure 2 shows the evaluable model of the 2 + 2 program. Components of the 2 + 2 program not considered evaluable at this time are

2 + 2 for Teachers: Evaluable Model
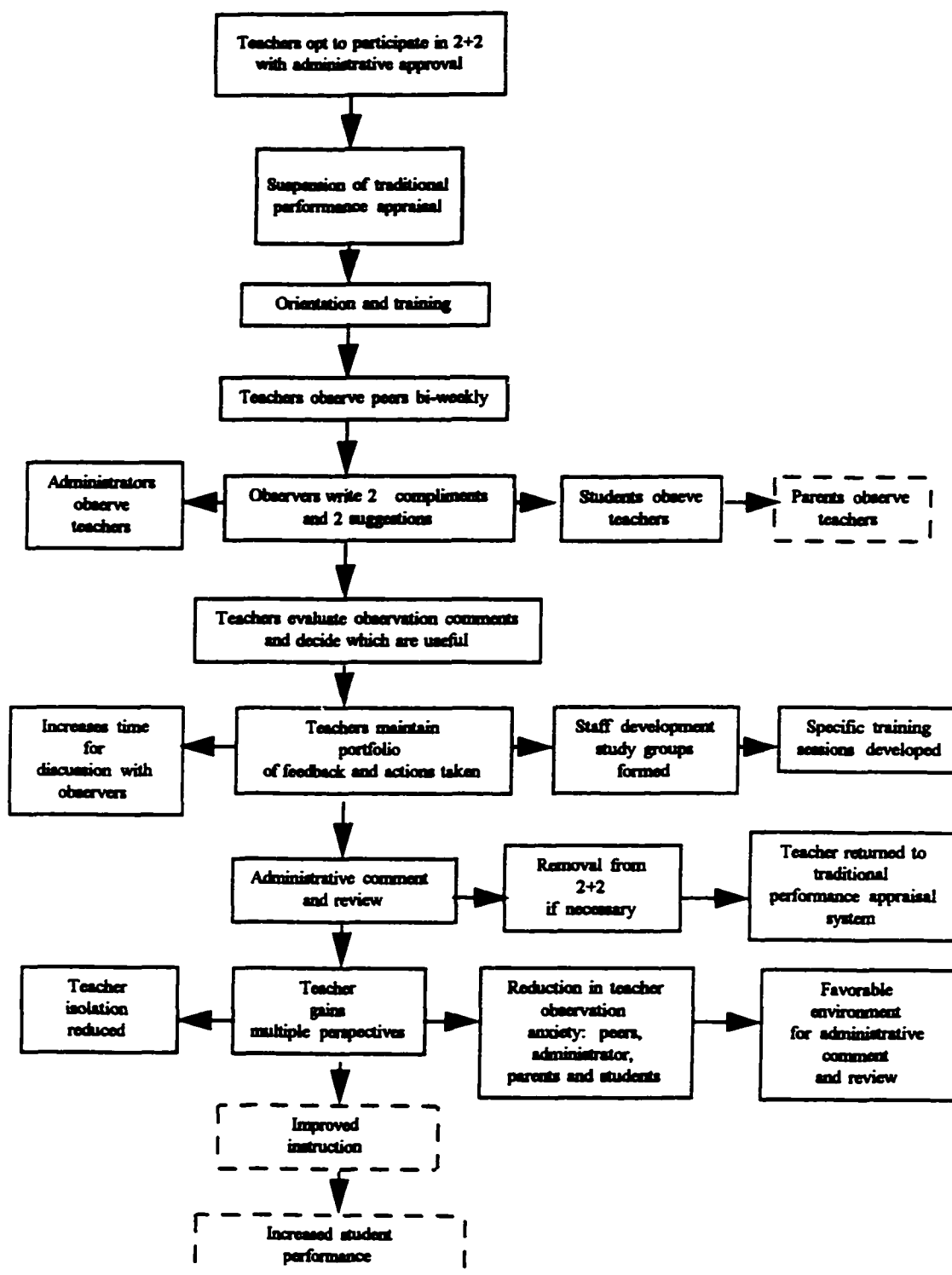
Figure 2. 2 + 2 for Teachers: Evaluable Model

represented by boxes with separated lines. For example, parent 2 + 2 observations were not implemented in the 1996-97 school year and could not be evaluated. In addition, improved instruction and increased student performance were not considered evaluable in the short term. The current evaluation is formative and focuses on program implementation. Although 2 + 2 is expected to contribute to improvement of instruction in the long term, a causal relationship between 2 + 2 and improvement of instruction cannot be determined in this evaluation. Where improvement of instruction is referred to as an outcome in this evaluation study, the term refers to teachers' perceptions, that is, self reports, of 2 + 2 as a vehicle for introducing different instructional strategies and establishing conditions under which improvement and professional growth could take place. Increased student performance was also not evaluated because too many factors influence whether student test scores increase or decline to establish a short term causal relationship with 2 + 2. The present evaluation design, however, lays the groundwork for future longitudinal studies of improvement of instruction and increased student achievement.

From the evaluability assessment, an evaluation design was developed. The evaluation design defined specific evaluation questions and created a measurement chart to provide data in answering the questions (Figure 3). The questions covered a variety of concerns, ranging from simple program monitoring to process and outcome issues. The evaluation questions essentially checked evaluable program components against the actual implementation process.

Research Questions

To recapitulate, three basic research questions focused the current study:

1. How was the 2 + 2 for Teachers: Alternative Performance Appraisal Program implemented?

2. What difference did the implementation of the 2 + 2 for Teachers: Alternative Performance Appraisal Program make?

3. What were teachers' perceptions of the benefits and drawbacks of the 2 + 2 program?

The basic research questions broadly encompass the evaluation design described

EVALUATION AND MEASUREMENT CHART

Evaluation Questions	Measurement Chart
1. Was adequate teacher release time available to all 2 + 2 participants?	Number of hours released. (One hour every two weeks available for observation of colleagues.)
2. Did teachers volunteer to participate in 2 + 2? Percentage of administrative approval?	Number of teachers who volunteered and number appoved by the administration.
3. Was the traditional evaluation suspended?	Consult administrators and teachers.
4. Were orientation and training sessions conducted?	Number of sessions conducted.
5. Were observations conducted by administrators, teachers, and students?	Review of 2 + 2 forms. Focus group reports.
6. Did observers follow the 2 + 2 protocol of two compliments and two suggestions?	Review 2 + 2 observation forms.
7. Did teachers increase time discussing instruction related topics with their colleagues?	Interview and survey teachers prior to start of 2 + 2; focus group interviews.
8. Did staff development study groups meet to develop observation foci?	Number of meetings held.
9. Were inservice sessions held as an outgrowth of number 8 above?	Number of sessions held.
10. Did the administration review and comment on 2 + 2 portfolios?	Interview administrators.
11. Were any teachers removed from 2 + 2?	Number of teachers removed.
12. Did teachers perceive a reduction in isolation from their peers?	Focus group meetings, survey, self-reflection reports.
13. Do the compliments and suggestions show a variety of instructional/educational perspectives?	2 + 2 observations forms, self-reflection reports.
14. Did teachers perceive a reduction in anxiety about being observed and critiqued?	Interviews/surveys before begin of 2 + 2, anxiety scale on attitude survey, self-reflection reports.

Figure 3. Evaluation and Measurement Chart for the 2 + 2 for Teachers: Alternative Performance Appraisal Program, developed as part of the evaluability assessment in June 1995.

above, while allowing for an open-ended approach to the evaluation. The current study, for example, also examines how teachers perceived the 2 + 2 program as an alternative to the traditional Norfolk Public Schools appraisal system, a topic not represented in the evaluability assessment or evaluation design above. The 2 + 2 program as a performance appraisal alternative was an inevitable issue of interest in the 1996-97 implementation, and could be addressed in reference to any of the three research questions.

<div align="center">Data Collection</div>

The design of the current study employed of a variety of data collection methods. Historical data, a questionnaire, interviews, focus group meetings, the 2 + 2 observation forms, the self-reflection reports, and a survey, in addition to conversations and PRIME Steering Committee meetings, were all data sources.

Historical data

As described in Chapter I, 2 + 2 is one component of the PRIME school restructuring project. The implementation of 2 + 2 nevertheless proceeded on an "ad hoc" basis. Orientations were held at schools which requested them, and a limited number of teachers engaged in peer observations at the high school and two middle schools with no prescribed number of observations or systematic implementation checks. A survey distributed to the mailboxes of PRIME high school and middle school teachers in June 1996 elicited 49 responses from both 2 + 2 participants and non-participants.

Prior to the survey, in December 1995, teachers and administrators were asked to comment on the 2 + 2 program for an article being prepared on 2 + 2. Their comments reflect perceptions of the 2 + 2 program during the pilot implementation year.

Questionnaire

The "Teacher Performance Evaluation Attitude Questionnaire", which measures attitudes toward the performance appraisal process, was administered twice in large group meetings to teachers in the four schools. The purpose was to compare attitudes toward the NPS traditional appraisal system and attitudes toward the 2 + 2 appraisal system. In the first administration of the questionnaire in October 1996, teachers were instructed to respond by considering only the traditional NPS district appraisal system in existence from 1983 through the 1995-96 school year. Only those teachers who had

experienced a summative year evaluation under that system were surveyed. Of these questionnaires, only those which indicated that the respondent was participating in 2 + 2 in lieu of the NPS summative year evaluation were included in this study (N=73).

In May and June 1996, the questionnaire was re-administered to teachers in the four PRIME schools who had chosen to participate in the 2 + 2 program in lieu of the district appraisal program. In the second administration of the questionnaire, some statements were slightly modified so that the statements referred to the 2 + 2 program. Teachers were instructed to respond with their perceptions of the 2 + 2 program. Fifty-three questionnaires meeting the above criteria were completed. Because of the optimal administration of the questionnaires in meetings, a response rate close to 90% was achieved.

The questionnaire measures attitudes on 16 subscales including anxiety, accuracy, achievement orientation, acceptance, fairness, feedback, supervisory knowledge, policies, procedures, performance obstacles, performance review meeting, purposes, management form, supervisory impact on the performance review meeting, supervisory trust, and performance rating. A valid and reliable instrument, the questionnaire was developed by Dr. Terry L. Dickinson and Andrea E. Berndt of Old Dominion University, and has been utilized in a study of Norfolk city employees' attitudes toward the appraisal process. A copy of the both versions of the questionnaire is included in Appendix G.

Interviews

Structured interviews were conducted in October/November 1996 with 48 participants of the 2 + 2 program. The interview protocol consisted of 12 questions. The purpose of the interview was to gather baseline information regarding teachers' expectations of the program, and their reasons for participation. They were also asked how they felt about the Norfolk Public Schools teacher appraisal system, how they felt about being observed, how they anticipated 2 + 2 would impact their instruction, and if they intended to ask for student 2 + 2 feedback.

Interns at the PRIME schools who were members of a university class on research methodology agreed to interview 2 + 2 teachers as part of a class assignment. The interns had an understanding of the 2 + 2 program and were expected to complete 2

+ 2 observations as part of their internship experience. It was felt they would benefit from the interview experience with more mature teachers, and better understand the 2 + 2 program as a secondary benefit. Teachers were expected to feel relaxed and cooperate with an interview request from an intern.

After a class session on survey and interview methodology, interns were given specific instructions on how to conduct the 2 + 2 interviews. Suggestions for maintaining the flow of the interview were discussed as well as the need to react in a supportive, but non-judgmental manner to interviewee comments. Interns were instructed to elicit further elaboration with responses such as "That's an interesting perspective. Can you tell me more about that?" Interns were given an interview protocol checklist to make sure they covered all questions in the course of the interview. Interviews were taped for accuracy and later transcribed. One teacher objected to being recorded and the interview was manually recorded.

Of the original 68 2 + 2 participants, 56 completed interviews. Eight of the 56 interviews were lost by a tape recording transcriber, while the remaining participants either did not respond to repeated requests for an interview, or were latecomers to the 2 + 2 program.

Focus Groups

Focus groups, more properly termed focus group interviews, are a very efficient qualitative data collection technique. Some measure of quality control is also achieved in that group members provide "checks and balances on each other that weed out false or extreme views" (Patton, 1990, p. 336). Focus groups shed light on what is actually happening in a program implementation.

> Specific implementation problems embedded in a new program or perceptions of the purpose (or potential threat) of a new program that may hinder implementation may be identified through focus groups. Second, focus groups can provide insights about the effects, advantages, and limitations of a program, curriculum, or set of materials after implementation...Useful information for revising or improving the program or materials also can be obtained. (Gredler, 1996, p. 86-87)

While focus group interviews are not simply problem solving sessions (Gredler, 1996), the flexibility to change the focus of the session to one of problem solving, should the group determine the need, is an advantage of action research methodology. Questions

about the program components, and problems involving the implementation process, common in new program implementation, were raised, discussed, and often resolved in the focus group meetings.

Focus group meetings were held at each of the four schools: (a) Lake Taylor High School on January 15, 1997 and April 23, 1997; (b) Lake Taylor Middle School on January 16, 1997 and April 23, 1997; (c) Azalea Gardens Middle School on January 29, 1997 and April 24, 1997; and (d) Little Creek Elementary School on January 30, 1997 and May 1, 1997. Except at the high school, all teachers participating in the 2 + 2 program were invited to the group meetings. Because the number of 2 + 2 teachers at the high school was so large, a different sample of fourteen teachers was invited to each meeting. The researcher consulted with the assistant principal to help assure a representative sample based on department and length of service. Although the researcher is well acquainted with several 2 + 2 teachers, most were not invited to the group meetings because other opportunities existed for the researcher to gain feedback from those teachers.

All meetings were tape recorded, with two exceptions. The April meeting at Azalea Gardens Middle School and the May meeting at the elementary school were not taped due to logistical considerations. Tape recording did not appear to have any inhibiting effect on the participants. The open-ended guiding questions were structured to address participants' experiences with the program to the date of the meeting. Logistics, implementation issues, student 2 + 2s, strengths and weaknesses, teacher isolation, and post-observation conferences were among the topics discussed.

The researcher conducted each meeting except for the May 1, 1997 meeting at the elementary school. A teacher participant led the meeting at the elementary school, and later briefed the researcher on the results.

Completed 2 + 2 Observation Forms

All completed 2 + 2 observation forms from the fall semester were collected from administrators at all four schools and photocopied. Only the 362 forms from the high school were analyzed. As explained earlier, the current study focuses on the high school 2 + 2 implementation.

Self-Reflection Reports

Thirty-one high school 2 + 2 teachers submitted a self-reflection report at the end of the fall semester. The report asks participants for a reflective response in four areas: (a) identification of the ten most useful/significant compliments; (b) identification of the ten most useful/significant suggestions; (c) reflection on ways which the compliments or suggestions were especially useful, and how they have led to reinforcement or improvement of classroom teaching routines; and (d) reflection on future agendas items which have emerged as a result of the 2 + 2 process, as well as any reflections about the entire 2 + 2 observation process, including an assessment of the value of observing others.

The self-reflection reports were collected from the high school assistant principal shortly after the beginning of the second semester, in February 1997.

Survey

A survey was distributed in May 1997 to gather participant responses on three aspects of 2 + 2: (a) direct comparisons of the old (1983-1995/96) district appraisal system with 2 + 2 program; (b) outcomes of the 2 + 2 program related to professional growth; and (c) programmatic issues such as whether 2 + 2 should be voluntary or mandatory, and what the optimum number of observations should be. The surveys were distributed in the four participating PRIME schools to all those who had participated in 2 + 2, either "officially", that is, in lieu of the district summative system, or "unofficially" by agreeing to making their classrooms open to 2 + 2 observers. Follow-up reminders were distributed to teachers' mailboxes, and a sign-off sheet for those completing the survey was posted in the high school office. The response rate was approximately 80%.

Meetings

PRIME Steering Committee meetings, individual meetings with principals, principals' meetings, and conversations with teachers were other sources of data.

Data Analysis

"The first step in qualitative analysis is description" (Patton, 1990, p. 374). "A...general analytic strategy is to develop a descriptive framework for organizing the case study" (Yin, 1994, p.104). In the first part of the analysis, the history of the 2 + 2

for Teachers: Alternative Performance Appraisal Program implementation within the context of the PRIME project will be described in detail. The case study approach to the question, "How was the 2 + 2 program implemented?", allows the researcher to chronicle the events over time. This, in turn, enables judgements to be made about why something did or did not occur, since the "basic sequence of a cause and its effect cannot be temporally inverted" (p. 116). The case study will describe events from the 2 + 2 program's inception through June 1997.

The second part of the data analysis will address outcomes of the 2 + 2 program implementation. Questionnaires, interviews, surveys, focus groups, 2 + 2 forms, and self-reflection reports were analyzed using quantitative and qualitative methods as described below.

Questionnaire

The questionnaire study was undertaken to better understand the following: (a) what teacher attitudes toward performance appraisal are in general; (b) how the 16 constructs are related to each other; (c) how the dependent variables, identified through psychology research as fairness, acceptance, and accuracy are predicted by the 13 other constructs measured in the survey; and (d) how teacher attitudes toward 2 + 2 performance appraisal differ from prior attitudes toward the Norfolk Public School traditional performance appraisal. The two data collection points are documented above. The following analysis is adapted from Berndt (1994).

The questionnaire underwent several phases of analysis. First, sample schools were tested for sample equivalence using a series of chi-square goodness-of-fit tests on relevant demographic characteristics. Because of the small sample size, composite variables were used rather than subscales. To include measurement error variance in the path analysis, error variance was set in the program for each composite variable.

An assessment was performed to examine how the hypothesized constructs were measured in terms of the observed variables, and to describe the reliabilities of the observed variables. This enabled the researcher to determine whether the relationships obtained on prior administrations of the survey to Norfolk city employees were consistent with those obtained with PRIME project teachers. In addition to comprising a

different sample, the PRIME project teachers had a different evaluation system than the city employees.

Third, it was considered how constructs were related, and then tested as to the strength of the relationship. A structural model assessment specifies the causal effects and relationships among the independent and dependent factors, or latent traits. The composite variables on the survey were considered observed variables. The squared multiple correlations were also examined for each hypothesized causal relationship in the model, as squared multiple correlation reflects the ability of the other latent traits to predict the dependent latent trait. A large squared multiple correlation indicates a strong relationship and suggests that the model is effective. Standard errors and measurement errors were also assessed. Standard errors are representations of the precision of each parameter estimate. Large standard errors indicate poor precision.

Following these tests, goodness-of-fit indices and chi-square statistics were used to investigate whether the sample covariance matrix provided a good fit for the solution. Parameter estimation of the measurement and structural models were undertaken with the goal of achieving goodness-of-fit estimates of .90 or higher. The root mean residual (RMR) should be less than .10, and the chi square value should be nonsignificant.

In addition, a General Linear Models (GLM) procedure was performed for each of the 16 composite variables: anxiety, accuracy, achievement orientation, acceptance, fairness, feedback, supervisory knowledge, policies, procedures, performance obstacles, post-observation and summative evaluation conferences, purposes, evaluation forms, supervisory impact on the summative conferences, supervisory trust, and performance rating. The GLM procedure is an analysis of variance (ANOVA) with cells of varying size. The Student-Newman-Keuls post-hoc test was performed where appropriate. School level was an independent variable with three levels, elementary, middle and high school. The second independent variable was teacher gender.

Interviews, Focus Group Meetings, 2 + 2 Observation Forms, Self-Reflection Reports

As Patton points out (1990), qualitative analysis does not have one single correct methodology. An appropriate approach to analyzing and interpreting qualitative analysis depends largely on the individual needs of the researcher with regard to the particular

problem or program under investigation. He comments, "...there are no absolute rules except to do the very best...to fairly represent the data and communicate what the data reveal given the purpose of the study" (p.372). The following data analysis descriptions were undertaken with the intention of providing relevant information to participants and implementers of the 2 + 2 program.

Entire texts and transcriptions were used in the analysis. As complete a set as possible of each data source was examined. Completeness was compromised only by non-existence of the data, for example, if a teacher did not submit a self-reflection report. Together, all data sources provided a rich description of the 2 + 2 implementation, and feedback for program modification.

Interviews and Focus Group Meetings

A process of inductive analysis was utilized to identify patterns and themes of interest in the oral and written data. Taped interview and focus group meetings were transcribed. "The first decision to be made in analyzing interviews is whether to begin with case analysis or cross-case analysis....Beginning with cross-case analysis means grouping together answers from different people to common questions or analyzing different perspectives on central issues" (Patton, 1990, p. 376). The interviews were analyzed on a cross case basis. Patton suggests that where the focus is on a particular program, a cross-case analysis is an appropriate methodology.

Individual responses to each interview question were examined, compared, and coded. The coding process itself was a "cut and paste" iterative process whereby conceptually similar responses were grouped into categories. Thus, responses from different teachers to each question were grouped together under categories which emerged from the distribution of the responses themselves after thorough reviews of the data.

Focus group transcriptions were analyzed to identify major themes and concerns emerging from open ended responses to the researcher's questions and ensuing discussion. The analysis of the focus group meetings, or, group interviews, was similar to that of the interviews. Each focus group transcription could be considered a "case", where a process of inductive analysis revealed patterns and themes. "Inductive analysis

means that the patterns, themes, and categories of analysis come from the data; they emerge out of the data rather than being imposed on them prior to data collection and analysis" (Patton, 1990, p.390). The focus group meetings were based on open-ended questions which allowed participants to determine issues to be discussed. The researcher categorized issues and responses arising from each group interview, employing an iterative process similar to that described for the interview analysis.

2 + 2 Observations Forms

Patton (1990) defines content analysis as "the process of identifying, coding, and categorizing the primary patterns in the data" (p.381). Classifying the data is a preliminary step in analyzing content, and facilitates "the search for patterns and themes within a particular setting or across cases" (p. 384). According to Patton, establishing a classification system can be a "simple filing system", a way to index the data by devising appropriate labels for different ideas represented in the data. More complex systems of coding, such as categorizing every paragraph in an interview with multiple coders, are "appropriate for very rigorous analysis of a large amount of data", not for "small scale formative or action research projects" (p. 384). The 2 + 2 program evaluation was a small scale formative evaluation project. In view of the intended purpose of the evaluation to inform and improve the program, a relatively simple classification system was employed.

For the analysis of the 2 + 2 observations forms, a process of categorizing, or labeling, 2 + 2 compliments and suggestions across cases was utilized. Suggestions and compliments were analyzed separately. To begin the process, potential categories of suggestions were constructed by the researcher, several graduate assistants, and Dr. Allen, all of whom are knowledgeable about the teaching profession. Teachers' 2 + 2 observations forms were examined and suggestions tentatively assigned to a category. As suggestions were found which did not fit in a category, a new category or sub-category was created. Categories were revised several times, as suggestions were reviewed. Suggestions were then once again examined, and assigned to categories in an iterative back and forth process. A similar process was utilized to classify compliments.

Self-Reflection Reports

Parts three and four of the self-reflection reports were analyzed in combination. In analyzing the self-reflection reports, content was classified and categorized to identify patterns and themes of teacher responses, using a similar coding of responses and development of thematic categories described above. This involved an iterative process of reviewing the data and identifying statements of significance, that is, those with direct bearing on the topics addressed by parts three and four of the self-reflection reports. Sometimes, for example, one sentence in a paragraph summarized a teacher's opinion. One hundred and three statements were gathered from the 31 self-reflection reports.

As the statements were compared and grouped together with similar statements in provisional categories, the writing prompts for parts three and four of the self-reflection reports were referenced. For example, comments were solicited in part four about the 2 + 2 observation process, and about the value of observing others. Therefore, it was natural that a number of statements in the data related to these themes. Themes, therefore, arose as a function of the self-reflection report prompts.

Survey

The survey administered at the end of the school year contained statements about 2 + 2 program evaluation elements, 2 + 2 program outcomes, and 2 + 2 programmatic elements which respondents were asked to rate on a Likert scale. The survey was analyzed using descriptive statistics including frequency, percent, mean, and standard deviation. Survey statements were also grouped in eleven categories for the purpose of correlation analysis.

Finally, the survey was analyzed using a General Linear Models Procedure (GLM). The GLM is an ANOVA with procedures to address cells of unequal size. Eleven GLM procedures tested the following null hypothesis: School level (high school, middle school and elementary school) made no difference in the number of observations made, number of observations received, number of observation considered ideal in a summative evaluation year, number of observations considered ideal in a non-summative evaluation year, positive responses to the 2 + 2 program, positive responses to the Norfolk Public School appraisal system, negative responses to the 2 + 2 program,

negative responses to the NPS appraisal system, positive outcomes of the 2 + 2 program, negative outcomes of the 2 + 2 program, and whether the 2 + 2 program should be mandatory or voluntary.

Triangulation

Multiple methods for gathering data and multiple data sources in this evaluation comprised a triangulation of data, a means of strengthening the study design, which provided checks on consistency of the data as well as a richness of perspective. Combinations of interviews, focus groups, and surveys make a study less vulnerable to the weaknesses of each particular method (Patton. 1990, p. 188). Yin (1994) writes:

> The most important advantage presented by using multiple sources of evidence is the development of converging lines of inquiry, a process of triangulation ...thus any finding or conclusion in a case study is likely to be much more convincing and accurate if it is based on several different sources of information, following a corroboratory mode. (p. 92)

A corroboratory mode was applied to the qualitative data collection methods of this utilization focused evaluation. Triangulation allows for multiple perspectives on the same reality, and provides a richness of data in that confirming as well as opposing views only add to the understanding of the program.

Issues of Validity and Reliability

Reliability is concerned with the consistency of the instruments used in measurement, and the minimizing "the errors and biases in a study" (Yin, 1994, p. 36). Patton contends that the dominant paradigm, that of quantitative research, emphasizes reliability, while the alternative paradigm, that of qualitative research, emphasizes validity. "Merton, one of the most prominent theorists in sociology has argued that the cumulative nature of science requires a high degree of consensus and therefore leads to an inevitable enchantment with problems of reliability" (Patton, 1978, p.229). However,

> The alternative evaluation paradigm makes the issue of validity central by getting close to the data, being sensitive to qualitative distinctions, developing empathy with program participants, and attempting to establish a holistic perspective on the program. (p. 230)

The alternative paradigm is the context for this study. The grouped together main concern should be getting as close as possible to what is actually happening, and listening carefully to participants. Then the researcher's data will have meaning in particular for the participants themselves.

Especially important, in view of the emphasis on empathy with those actually impacted by the program, is the face validity of the measurements used. "One of the best ways to facilitate decision maker understanding of and belief in evaluation data is to place a high value on the 'face validity' of research instruments" (Patton, 1978, p. 244). High utilization of evaluation information will depend on how credible the data is to decisionmakers.

Content validity is concerned with whether the instrument(s) accurately measure what it presumes to measure. The questionnaire used to measure teacher attitudes has been tested for reliability and validity.

Threats to internal validity are applicable to case studies with a causal logic, and not, as in this case, to a descriptive study. However, a more difficult aspect of the problem of internal validity is how inferences are made and rival explanations explored (Yin, 1994). Triangulation of data, and identification of patterns or themes are two ways in which this study attempts to justify its conclusions.

External validity in a case study refers to generalizability not to another population, but rather to a theory. "Case studies (as with experiments) rely on analytical generalization. In analytical generalization, the investigator is striving to generalize a particular set of results to some broader theory" (Yin, 1994, p. 36). More to the point, Patton discusses the philosophical issues which separate program evaluation from other types of scientific investigation. Evaluators are mainly concerned with information pertaining to a specific program, and providing feedback to inform users and decision makers affiliated with that program(s). But even when there are large samples on which statistical inferences are based, qualitative evaluators find generalizations which are void of specific context unimpressive.

Patton suggests the term "extrapolation" to replace "generalizability" to define a middle ground between extremes of tightly controlled experiments and too loosely

defined case studies. "Extrapolations are modest speculations on the likely applicability of findings to other situations under similar, but not identical, conditions. Extrapolations are logical, thoughtful, and problem oriented rather than statistical and probabilistic" (Patton, 1990, p. 489).

Limitations

The impact of the 2 + 2 program as documented in the evaluability assessment is the improvement of instruction and increased student performance. The present study is limited in that it cannot address the program impacts in the current time span. The current study aims to understand the program as completely as feasible, including attitudes toward the appraisal process and issues of teacher isolation, to enable decisions and alterations which will guide it in the direction of maximizing the impacts at a future date. No causality can be determined about the impact of the 2 + 2 program on teacher isolation, level of collaboration, or instructional practices.

It is assumed that the teachers participating in 2 + 2 are representative of PRIME faculties, and that, therefore, issues raised by participants, the wealth of information collected, and identified program modifications will serve to facilitate the expansion of 2 + 2 to greater numbers of teachers in the PRIME schools.

In this evaluation, the researcher is concerned with recognizing important and unique program elements, as well as with providing personal and useful feedback to participants. Some data may conflict, as interviews capture a person's response at one point in time. Another snapshot at another time might appear quite different.

Role of the Researcher

The involvement of the researcher in the program can be positive (Whyte, 1991). The researcher has been a part of the PRIME project since it began in 1994, and had good working relationships with many of the teacher participating in the 2 + 2 program. At various times the researcher facilitated seminars on transdisciplinary curriculum where

issues of collaboration, teacher isolation and overload were frequently raised. With regard to the 2 + 2 program, the researcher developed the evaluability assessment under the guidance of Dr. Pindur, briefed Norfolk Public School faculty, conducted orientation and training sessions, and monitored the implementation of 2 + 2. Because of this history, the researcher was very familiar with the school environments and with various positive and negative change forces operating within the context of the PRIME project and within the 2 + 2 program. Access to 2 + 2 participants and 2 + 2 observation forms by the researcher was freely given by participating schools.

While easy access to information facilitates data collection, the researcher is, as a proponent of systemic reform and the PRIME initiatives, also subject to bias in evaluating the 2 + 2 program. However, the potential for reporting exaggerated positive results is offset by a genuine desire to allow the 2 + 2 program to be driven by the intents and purposes of the program. That is, if the program in its present form shows no positive impact on teachers' perceptions of isolation, collegiality, or enhanced instructional repertoires, this evaluation may point the way to major modifications needed in the program. The program must be congruent with the needs and values of participants to support systemic reform. The researcher is also open to issues as yet unknown that may emerge from in-depth feedback from participants. Information and communication, without hidden agendas, are key to any talk of empowerment and organizational change, including programs such as 2 + 2. The researcher's unique knowledge of the PRIME project from the outset leads naturally to an informed perspective, but also a perspective aware of the pitfalls of bias. Because the values of an evaluator always influence his or her perspective, regardless of the evaluator's role as internal or external, results of any evaluation must be examined in the context of the needs of the stakeholders and the purpose of the evaluation. These considerations apply to this study as well.

Summary

In this chapter the methodology of the evaluation was discussed in detail. The type of evaluation being conducted, the origins of the evaluation design, and the audience were all identified. Data collection, including questionnaires, interviews, focus group meetings, 2 + 2 observation forms, self-reflection reports, and survey were discussed in detail, and data analysis was described. In Chapter IV, the analysis of the data will be reported in full.

CHAPTER IV

ANALYSIS OF THE DATA

Introduction

The 2 + 2 for Teachers: Alternative Performance Appraisal Program (2 + 2) was described in detail in Chapter I. The research strategy for this evaluation of 2 + 2 was explained in Chapter III, including sources and types of data collected. In the first section of this chapter, the 2 + 2 program implementation in 1996-97 will be examined utilizing a case study approach to address the research question, "How was the 2 + 2 for Teachers: Alternative Performance Appraisal Program implemented?" The description will begin with a review of the origins of the 2 + 2 program and historical data from the 1995-96 2 + 2 pilot program to establish the context of the current implementation. The case study will proceed to the 1996-97 implementation of the 2 + 2 program, described chronologically and organized around issues which emerged during the year.

The second section of this chapter addresses the outcomes of the 2 + 2 program implementation. Data collected through interviews, focus groups, the 2 + 2 observation forms, the self-reflection reports, a survey, and questionnaire were analyzed to illuminate two further questions, "What difference did the implementation of the 2 + 2 for Teachers: Alternative Performance Appraisal Program make?" and "What were teachers' perceptions of the benefits and drawbacks of the program?"

At times the issues raised in section one overlap portions of the data analysis in section two. The reader is referred to the relevant parts of section two where appropriate.

Case Study: Implementation of the
2 + 2 for Teachers: Alternative Performance Appraisal Program

This section begins with a review of the origins of the 2 + 2 for Teachers: Alternative Performance Appraisal Program (2 + 2), including historical data about the pilot year of the 2 + 2 program and its adoption as one of the PRIME (Public School

Restructuring for Innovative Mainstream Education) project components. A description of the implementation of 2 + 2 during the 1996-97 school year follows.

The History of the 2 + 2 Program

The 2 + 2 program had its start in two less developed countries on opposite sides of the earth, Namibia and China. 2 + 2, developed by Dr. Dwight W. Allen, filled a need to raise the level of teacher expertise in Namibia, where teacher supervisors often had no more than an eighth grade education. Based on a belief in the power of frequent feedback to stimulate professional growth, 2 + 2 was created under the motto: No teacher is so perfect that there is no room for improvement, and no teacher is so incompetent that there is nothing to praise. Fledgling teachers implemented 2 + 2 by regularly observing each other using the 2 + 2 protocol of two compliments and two suggestions. Results were so encouraging that the 2 + 2 program was transplanted to the United States in early 1995 by Dr. Allen. Having observed its success overseas, he recognized its potential as a staff development and school community building tool for the PRIME educational reform project. In consultation with the researcher and doctoral student/PRIME Project Liaison Robert C. Brinton, a 2 + 2 Program document was created describing the program's components and theoretical basis. As mentioned in Chapter I, the researcher then developed an evaluability assessment of the 2 + 2 program in the spring of 1995, under the guidance of Dr. Wolfgang Pindur.

During the summer of 1995, the PRIME project sponsored eighteen teachers and administrators in two graduate field courses offered through Old Dominion University (ODU). The 2 + 2 program was introduced to teachers and administrators at that time by Dr. Allen, along with R. Brinton and the researcher. Reception to the 2 + 2 program proposal ranged from enthusiastic to apprehensive. As part of the graduate course assignments, Action Group Reports were prepared by students on a wide range of PRIME initiatives in anticipation of possible implementation in the fall of 1995. The 2 + 2 Action Group Report, developed by two elementary school teachers, and one elementary and one high school administrator, outlined an implementation plan in great detail and with great optimism. The first point addressed in the report proved pivotal: "A crucial element for implementation is school and teacher willingness to use the 2 + 2

Program."

Pilot Year 1995-96

The school year 1995-96 was a pilot year for the 2 + 2 program. In September of 1995, the Superintendent of the Norfolk Public Schools called the 2 + 2 program a "very promising experiment", and authorized teachers to participate in the 2 + 2 program in lieu of the Norfolk Public Schools (NPS) performance appraisal system. Information about the 2 + 2 program was presented in fall, 1995, by Robert Brinton at faculty or small group meetings as requested by individual schools. Presentations were made in one elementary school, the two middle schools, and the high school. At the close of the 1995-96 school year, a small number of teachers at Lake Taylor High School (LTHS), Lake Taylor Middle School (LTMS), and Azalea Gardens Middle School (AGMS) were implementing 2 + 2. No teachers at elementary schools participated. A small group of teachers at the high school and the two middle schools participated, several as summative year teachers in lieu of the NPS appraisal system.

Early feedback from 2 + 2 participants was solicited in December 1995 for an article being prepared on the program. One LTHS administrator, commenting on 2 + 2 said,

> This gives them [teachers]... a vehicle to step forward and share in a meaningful way. Before, never, ever would you have another teacher come in and observe. As an administrator, I find that 2 + 2 provides additional help as we focus on improvement of teaching. It takes less time and we still have viable, effective observations - it is not burdensome. In the old system, you have lots of boxes to fill in, and you often find yourself making meaningless comments just to get the job done.

Among participating teachers, 2 + 2 was seen in a very positive light. A high school teacher commented: "After 25 years of teaching social studies, I finally made it off of my floor and out of my department. I have observed an outstanding TMR [Trainable Mentally Retarded] program...a counseling session, a technology class, and more. And I have taken away ideas that will make me a better teacher." One teacher in her summative evaluation year, who joined the 2 + 2 program in lieu of the NPS appraisal process, told the researcher that she had received more valuable feedback in

one year from 2 + 2 observations, than in her previous 20 years as a teacher, including summative year evaluations.

A gradual adoption of the 2 + 2 program was intended. Since 2 + 2 is a voluntary program, teachers needed information to become convinced of its value, and encouragement to participate. It was anticipated that "early adopters" of 2 + 2 would be a minority of the staff, and that as interest grew, a majority of teachers would participate.

Still, schools did not move as quickly to promote 2 + 2 as PRIME staff had hoped. One reason was that at the start of the 1995 school year, the new internship program was being piloted for the first time at the PRIME high school, one middle school and all three elementary schools. Innovation overload may have been one factor preventing early adoption of the 2 + 2 program on a large scale. A lack of time to actively promote a second new initiative was likely another factor.

A third factor was that, although administrators, especially at Lake Taylor High School, thought the 2 + 2 program would benefit teachers, 2 + 2 was not an administrative priority at any of the PRIME schools. Because PRIME staff was willing offer support in the 2 + 2 implementation by introducing the program to teachers, school administrators may have thought the program would remain an experiment implemented by external agents. The PRIME sponsorship of 2 + 2 was designed to support the initial phases of the innovation, with the schools gradually taking more responsibility for implementation and institutionalization. Support was also contingent upon requests by schools for briefings. By the end of the 1995-96 school year, though, some teachers at the PRIME schools still did not know about the 2 + 2 program, and promotion of the 2 + 2 program by administrators was perfunctory.

In June 1996, a one page 2 + 2 survey (see Appendix A) was distributed to the mailboxes of all teachers at the high school and both middle schools. The survey response rate was approximately 25%. Of the 47 teachers and three interns who responded, 37 indicated they had participated in 2 + 2, either in lieu of the NPS appraisal system, or informally by performing and/or receiving a number of observations. Thirteen respondents did not participate in 2 + 2.

Comments of respondents were overwhelmingly positive: "The program helped

me to see myself in another light. Enabled me to work more effectively!" One teacher wrote,

> "It is my opinion that the 2 + 2 program can do wonders to improve teaching styles and increase teacher's perceptions of what is really taking place in educating our students at Lake Taylor High...It is only by tearing down our walls of seclusion and offering our classrooms for others to observe that we can legitimately offer new ideas or receive new ideas to improve not only our personal performance, but the performance of our students....Tear down the walls! ...there is much to learn and much to share about quality teaching that we already have at Lake Taylor High. We have only to harvest that resource by sharing it amongst all our teachers."

Others commented that the 2 + 2 program "enabled me to improve my instruction and activities in class", "improves colleague interactions", and "is an innovative idea that is much better than the official NPS evaluation form". One teacher remarked, "Every time we observe someone else it helps us to look at ourselves". An intern liked "the system of cooperation it encourages between teachers", while another said 2 + 2 "comments were constructive, specific, and often were the only encouragement I received."

The only negative comments referred to lack of time for observations. Most teachers who indicated time was a barrier were positive about the program, but commented, for example, "Time - We need more time to do this." One teacher, however, was negative about the 2 + 2 program. The comment was "Not enough time for classes, now this!"

Ten of the non-participant respondents did not indicate a positive or negative stance toward the 2 + 2 program, answering only the preliminary questions. Reasons given by these respondents for non-participation in the 2 + 2 program varied: (a) lack of time, (b) teacher was not up for summative year evaluation, (c) teacher was on probation, that is, had less than three years experience in the NPS system, (d) teacher was never informed, and (e) teacher was not "assigned" to 2 + 2 . The feedback showed that some misunderstandings existed, and that some teachers were uninformed about 2 + 2.

No attempt was made to conduct a detailed analysis of the survey. Although the

low response rate precluded statistical analysis, much interest was generated by some of the very positive teacher responses. The primary aim of the survey was to obtain feedback from teachers about the program in the most feasible manner possible during the last week of school. No opportunity for follow-up existed, and the survey was the best option under the circumstances.

The results of the survey were shared at the PRIME Steering Committee meeting in June 1996. The Deputy Superintendent, at that time the school district's member and Chairperson of the Steering committee, agreed that 2 + 2 should be made an official PRIME component, to be implemented as a priority in the PRIME schools in 1996-97. However, although administrators at Lake Taylor High School and Little Creek Elementary School voiced support for 2 + 2 in principle, other administrative response was lukewarm. Another elementary school principal wondered how well 2 + 2 could serve an evaluative function if teachers were reluctant to give valid critical feedback to their colleagues. Doubts were expressed about the willingness of teachers to be observed, or to find the time to conduct observations. One principal expressed a reluctance to relinquish traditional control over the evaluation process. Nonetheless, a decision was made to print ten thousand 2 + 2 observation forms in triplicate for the upcoming school year.

An understanding was reached that the researcher would present 2 + 2 as a teacher performance appraisal option to faculties at the very beginning of the 1996-97 school year. The program would remain voluntary for teachers, but each school agreed to encourage participation. The field-based summer graduate courses in 1996 did not further address the 2 + 2 program because of the plan of action already in existence from the summer of 1995.

2 + 2 Program Implementation Year: 1996-97

The case study description of the 2 + 2 program implementation during 1996-97 is divided into two major sections. The Start-up Phase describes the faculty presentations, application procedures, and orientation sessions. The second section, Implementation Issues, addresses implementation concerns, adaptive reactions to problems, and levels of administrative support and leadership. Table 1 gives a time line

Table 1

Time Line: 2 + 2 Program Implementation 1996-97

Date	Activities/Data Collection	Implementation Issues	Adaptive Actions
August/ September/ October 1996	Faculty Presentations 2 + 2 Orientations Questionnaire Administration Interviews of 2 + 2 Teachers	Teacher eligibility question raised: Can probationary teachers (teachers with less than three years' experience) participate in 2 + 2 in lieu of the NPS appraisal system?	See November PRIME Steering Committee
November 1996	PRIME Steering Committee	2 + 2 Issue addressed: Teacher eligibility	Decision: probationary teachers may not participate in 2 + 2 in lieu of the NPS appraisal system.
	Conversations with teachers	Teacher feedback positive toward 2 + 2 program Issue: Suggestion Component- how to make suggestions 1) when class is already excellent, 2) that sound helpful, not critical.	
December 1996	PRIME Retreat/ 2 + 2 sessions	Teacher feedback positive toward 2 + 2 program Issues: 1) Suggestions Component-making meaningful suggestions; 2) Number of Observations- too many required; 3) Time-not enough, but also need to prioritize time for 2 + 2.	1) List of sample suggestions teachers have made distributed; strategies for making suggestions reviewed at retreat. 3) Elementary school principal hires monthly substitute teacher to allow 2 + 2 teachers time to observe colleagues.
January 1997	PRIME Steering Committee	2 + 2 Issues addressed: 1) Suggestion Component, 2) need for Staff Development focus	1) List of suggested observation foci distributed to principals for 2) adaptation as desired according to each school's staff development needs.
	Focus group sessions at four schools	Teacher feedback positive toward 2 + 2 program Issues: 1) Number of Observations-too many required; 2) Programmatic Questions; 3)Time-scheduling observations a problem of time and self- discipline; 4) high school 2 + 2 teachers request full group meeting for feedback and support; 5) Suggestions Component-making meaningful suggestions.	1) Proposal to require fewer observations added to March PRIME Steering Committee Meeting agenda; 2) Review of 2 + 2 program, protocol; 3) Teachers share strategies for completing observations; observations chart, distribution of teachers' schedules, central 2 + 2 form deposit box to observations implemented; 4) Administration asked to schedule 2 + 2 meeting; 5) Sample suggestions shared; suggestion making strategies reviewed.

74

Table 1 continued

Dates	Activities/Data Collection	Implementation Issues	Adaptive Actions
February 1997	Collection of 2 + 2 Observations Forms and Self-Reflection Reports.	Issue: Accountability - teachers have not completed required number of observations	High school 2 + 2 teachers receive individual reports noting the number of observations made and received
March 1997	PRIME Steering Committee	2 + 2 Issues addressed: 1) Number of Observations- too many required; 2) Accountability - teachers completing too few 2 + 2 observations.	1) Proposal to reduce number of observations to 15 to count five post-observation discussions toward the total requirement is passed. 2) Administrators briefed on situation, Steering Committee decides each school will address issue as needs dictate.
April 1997	Focus groups sessions at four schools	Teacher feedback positive toward 2 + 2 program Issues: 1) Time - remains obstacle; 2) Number of Observations - teachers receiving too few; 3) Administrator Observations -too few administrator observations of 2 + 2 teachers; 4) Need exists for: building 2 + 2 coordinator, scheduling support to balance observations received among all participants.	
	Meetings with individual principals	Principals give positive feedback about 2 + 2; Issues: More structure, goal setting for 2 + 2 teachers.	Principals briefed on issues needing resolution for 1997-98 implementation including: building coordinator, scheduling aid, administrator's role, teacher accountability.
May 1997	2 + 2 Survey administration Questionnaire administration PRIME Steering Committee	2 + 2 Issues: 1) Renewed controversy over 2 + 2 as alternative appraisal program; 2) Time line for completing observations	1) 2 + 2 Sub-Committee formed to establish written guidelines for the 1997-98 2 + 2 implementation; 2) 2 + 2 to mirror NPS appraisal system time line
June 1997	PRIME Steering Committee	2 + 2 Issue: Overt resistance by middle school principals to 2 + 2 as alternative appraisal system.	2 + 2 Sub-Committee report: Presentation of survey results and proposed 1997-98 2 + 2 guidelines by high school principal. Probationary teacher eligibility reinstated. Vote delayed until July.

75

and overview of the implementation of the 2 + 2 program in 1996-97.

Four of the six PRIME schools chose to adopt and support the 2 + 2 program as described in the program guidelines. Lake Taylor High School (LTHS), Lake Taylor Middle School (LTMS), Azalea Gardens Middle School (AGMS), and Little Creek Elementary School (LCES) were the participating PRIME schools. The two remaining elementary schools were not included in the evaluation. At one of the schools, the principal chose to evaluate all teachers using the new NPS appraisal system, while requiring each teacher to complete 15 2 + 2 observations for the year. (Teachers had been previously formally evaluated once every four years.) However, no follow-up was conducted on the 2 + 2 observation requirement by the principal. The principal had served on the core committee which designed the new NPS appraisal instrument, and possibly was, for this reason, more committed to that system of evaluation. He also expressed strong reservations about the potential of 2 + 2 to positively impact instruction. At the other PRIME elementary school, despite a presentation by the researcher, no teachers volunteered to participate in 2 + 2 due largely to a general mistrust of the process and lack of time. The principal explained that other priorities took precedence over implementation of the 2 + 2 program, and she felt her staff was disinterested. She, therefore, did not actively promote it.

Start-up Phase

Faculty presentations. The 2 + 2 for Teachers: Alternative Performance Appraisal Program implementation began prior to the beginning of the school year with informational presentations to the Lake Taylor High School (LTHS) and Lake Taylor Middle School (LTMS) faculties in August 1996. Little Creek Elementary School (LCES) and Azalea Gardens Middle School (AGMS) arranged for the researcher to conduct 2 + 2 presentations during teacher early-release days on September 19 and October 17, respectively. In all cases, the new NPS district teacher performance appraisal system was also introduced to teachers at the same meeting. Teachers who were in their summative evaluation year in 1996-97 had the opportunity to choose between the two programs. Those teachers not in their summative evaluation year were encouraged to participate in the 2 + 2 program as well.

The assistant principal at Lake Taylor High School, in particular, encouraged teachers in their summative evaluation year to opt for the 2 + 2 program for two reasons. First, he was convinced of the benefit of 2 + 2 observations for teachers. Second, approximately one-half of the teachers at the high school were in their summative evaluation year, and the assistant principal knew that each teacher opting for the 2 + 2 program would reduce the evaluation load for the administration. Such a reduction would be particularly welcome, since the new NPS appraisal system was still unfamiliar to administrators, and more complex than the previous system. His persuasiveness led more than half of the summative evaluation year teachers to opt for 2 + 2. Successful introduction of the 2 + 2 program at LTHS was thus aided by serendipity, that is, the concurrent implementation of a new NPS appraisal system.

2 + 2 application form/contract. Most teachers made application for the 2 + 2 program at the time of the 2 + 2 faculty presentations, or shortly thereafter. An example of the Application Form for 2 + 2 is included in Appendix B. Application forms constituted a contractual agreement and explained the conditions of participation as set out in the program guidelines. The school administration retained the right to terminate teachers' participation in the 2 + 2 program should an administrator have concerns about teacher performance.

The contract included a provision for teachers to be given one period of release time every two weeks to accomplish the observations. The release time for conducting observations was intended to be gained through the use of interns to cover teachers' classes. During the year, however, intern usage was viable only at the high school, due to the school's large number of interns. Schools with few interns could not easily schedule them to cover 2 + 2 teachers' classes. Teachers at all participating schools were responsible for organizing their own release time through intern usage, or by utilizing planning time.

The LCES principal handled the issue of release time by altering the contract to state that teachers would use one resource period once every two weeks, or use any other time teachers could arrange, to accomplish the 2 + 2 observations. Although the principal felt that she could not guarantee release time for teachers, she offered teachers

an incentive of up to $100 towards registration at a conference of the teacher's choice, and one daily substitute teacher if needed. No teacher claimed the incentive, perhaps because support for staff development was granted by the principal whenever possible anyway.

Orientation/training. Orientation sessions for teachers were held by the researcher at each of the four participating schools between September 19, 1996 and October 21, 1996. At Lake Taylor High School, the researcher conducted an orientation session each bell during one entire day to better accommodate the schedules of approximately 40 teachers who opted to participate in 2 + 2. Eight LCES teachers, 14 teachers from LTMS, and 6 teachers from AGMS also signed up for 2 + 2. By the end of October, approximately 68 teachers had opted to be 2 + 2 participants. Lists of 2 + 2 participants fluctuated due to administrative clerical error, late sign-ups, and occasional uncertainty regarding which teachers were in their summative evaluation year.

The interactive orientation at each school gave a more detailed description of the 2 + 2 program, and reviewed the theoretical basis for 2 + 2. Teachers practiced giving 2 + 2 feedback after viewing a 10 minute video segment of a class lesson appropriate to the school level. (Most high school teachers declined the practice video.) Teachers readily made compliments, but found suggestions more difficult to develop. By the end of the session, however, each group seemed to be more comfortable making suggestions, especially as the difference between offering suggestions for improvement and finding something "wrong" with the lesson was reinforced.

Nonetheless, teachers at all schools were concerned about how to make appropriate suggestions when they could find nothing "wrong" with a lesson, and how to formulate suggestions without creating hurt feelings. Framing suggestions as another, alternative strategy to "try out" was emphasized. As recipients of suggestions, teachers were reminded that they alone could determine whether to use the information to change behavior, or indeed, judge whether the feedback was relevant at all.

Finally, participants were advised to contact an administrator, or the researcher, if he or she was receiving no visits from teachers, was having difficulty finding time to conduct observations of other teachers, or encountered other problems.

Presentations about the 2 + 2 program to school faculties, and orientation sessions for participants were completed by the end of October 1996. Although two months' start-up time may seem excessive, schools were extremely busy during these first weeks. Meetings where teachers were introduced to the new NPS appraisal system were scheduled in compliance with the new NPS appraisal system guidelines. As mentioned earlier, 2 + 2 presentations to the faculties took place at the same meetings. Since the 2 + 2 program was an alternative to the NPS system, this was considered an acceptable time line.

Implementation Issues

This section of the 2 + 2 case study identifies issues related to the implementation of the 2 + 2 program in 1996-97. In focusing on the question, "How was the 2 + 2 for Teachers: Alternative Performance Appraisal Program Implemented?", this section does not address the strengths and weaknesses of the program itself, its outcomes, or how the 2 + 2 program made a difference to teachers. Those topics, including many exceptionally positive outcomes, are discussed below (see Outcomes of the 2 + 2 Program Implementation).

Instead, in this section, the actual process of program implementation is addressed. Implementation issues are defined as differences which arose between the 2 + 2 model and actual operations. As Patton (1997) points out, the implementation process "simply isn't that rational or logical. More common is some degree of incremental implementation in which a program takes shape slowly and adaptively in response to the emerging situation and early experiences...Design, implementation, and routinization are stages of development during which original ideas are changed in the face of what's actually possible" (p. 201).

Described in this section are implementation issues which emerged during the 2 + 2 implementation process. The question of teacher eligibility, the suggestion component of 2 + 2, the required number of observations, the level of administrative support, and the influence of the PRIME project on the 2 + 2 implementation are discussed, as are adaptive actions taken by Dr. Allen, the researcher, and the PRIME Steering Committee in response to the actual implementation process.

Examination of the 2 + 2 program implementation spills over into some of the difficulties inherent in managing systemic reform. One of the emerging issues was resistance to change, sometimes easy to overlook because it was usually not overt. A program such as 2 + 2 implies changes in degrees of control and power distribution. Teachers participating in the 2 + 2 program in lieu of the NPS appraisal process were exercising more control over their own growth, and were practicing a combination of peer and self-evaluation. Although administrators retained the authority to remove marginal or incompetent teachers from the 2 + 2 program, and could observe teachers as often as they wished, the power of evaluation was being shared to an extent not experienced before.

Another issue was governance of change, or change management. Implementation of an educational reform project, and innovative programs within the project, also require an effective mechanism to achieve consensus and ensure accountability. An example demonstrating both a control and change management issue arising in October was the question of teacher eligibility for the 2 + 2 program.

Teacher eligibility. The researcher experienced difficulty obtaining lists of participating 2 + 2 teachers from each school in October 1996. The problem was not unwillingness to share the lists, but that final lists had not been compiled. The administration of the 2 + 2 program seemed to proceed slowly. In addition, middle school administrators balked at allowing first year teachers, and teachers still on probation (that is, in their first, second, or third year of teaching) to participate, creating confusion regarding who should apply and who could be accepted.

Although consensus was reached the previous spring in a PRIME Steering Committee meeting that all teachers could opt to participate in 2 + 2 in lieu of the NPS appraisal system, administrators, in general, now seemed uncertain what the district's position was, and awaited word from the deputy superintendent, chair of the PRIME Steering Committee. The deputy superintendent was also unsure what the policy should be, and brought it back to the PRIME Steering Committee. The original decision reached in the spring seemed forgotten.

The confusion surrounding the issue reflected ambiguity on the part of PRIME

school administrators about what authority and autonomy the PRIME Steering Committee possessed to make and enforce decisions, and the nature of its relationship to the deputy superintendent and district office. Norms of accountability for implementing Steering Committee decisions had not been established, either.

In this environment, major decisions for school change were difficult to reach. All PRIME schools wished to appear "on board" to avoid repercussions from the central office. Therefore, administrators did not always share their concerns, or lack of commitment to a proposal or program at Steering Committee meetings, in spite of acknowledgment by the district that candor was a prerequisite for real change. Likewise, support voiced for a program, or report of school progress on a particular PRIME initiative was on occasion contrived, making in-depth consultation on issues affecting school change very difficult. Superficial acceptance of PRIME programs and decisions could and did mask resistance to those same programs.

Concerning the teacher eligibility issue, Dr. Allen, as originator and proponent of 2 + 2, argued that first year and probationary teachers had the most to gain from 2 + 2 because they could observe and receive feedback from more experienced teachers. The 2 + 2 contract clause allowing administrators to remove a teacher from 2 + 2 at any time performance warranted such action should have, in his opinion, been an adequate safeguard for administrators. At a PRIME Steering Committee meeting in November, however, even administrators who were not opposed to admitting new or probationary teachers to 2 + 2 did not voice support for this view. Sensing that the deputy superintendent was leaning their way, middle school administrators successfully argued for a policy of non-admittance to 2 + 2 for those teachers. Approximately five first year and probationary teachers, already accepted into the 2 + 2 program, were removed and placed on the district's performance appraisal system.

The decision to restrict teacher eligibility was indicative of a tendency for administrators to follow the lead of the Deputy Superintendent without candid discussion of all concerns and feelings. The proceedings demonstrate just how elusive establishing a norm of open consultation, leading to group ownership and subsequent accountability, can be. In retrospect it is clear that, although reservations were voiced,

such a norm had not been established, and administrators naturally agreed to do what they perceived was expected by the district office. A consequence for implementation was a continuum of commitment to the 2 + 2 program on the part of the administrators ranging from "interested" to "resistant".

Suggestion component of the 2 + 2 observations protocol. An implementation issue emerging early in the fall of 1996 was the suggestion component of 2 + 2. As can be seen in Table 1, teachers were often uncomfortable making meaningful suggestions for fear of appearing too critical, or were unable to think of a suggestion when a lesson was particularly effective. Teachers raised this concern at the orientation sessions, at the PRIME retreat, in focus group meetings, and in private conversations.

Making suggestions is a component of the 2 + 2 program, and it becomes an implementation issue if that component is perceived by teachers as difficult to accomplish. Not all teachers experienced the same degree of difficulty, however. For example, a review of the high school 2 + 2 Observations Forms for the first semester showed that of 50 teachers, 20% had made two suggestions for each observation. At the other end of the continuum, 22% had left half or more of the suggestion fields blank.

A few teachers were irritated by receiving suggestions, in particular those which referenced something outside of their control. An example might be an inadequate number of textbooks in the classroom. This issue is discussed in greater detail in the Outcomes section of this chapter, under Focus Group Meetings, and 2 + 2 Observations Forms.

Adaptive actions. In response to problems teachers experienced making meaningful suggestions, three 2 + 2 sessions were planned at the PRIME project retreat, held in Williamsburg, VA in December 1996. The LCES principal, assisted by the researcher, planned the 2 + 2 sessions primarily to address the issue of how to give meaningful feedback when making suggestions. Low level and high level suggestions were differentiated, along with strategies for developing more of the latter. Teachers particularly appreciated examples of suggestions which had been collected from actual 2 + 2 observation forms and discussed in the session.

As a result of positive feedback at the retreat sessions, sample suggestion sheets

were later distributed to 2 + 2 participants at focus group meetings, and in mailboxes of those not able to attend. Teachers found these lists of colleagues' 2 + 2 suggestions (names were removed to protect identity) to be particularly helpful, as well as conversations with other teachers, and focus group meetings which included discussion on how to formulate suggestions. Possibly this type of communication also served to alert the teachers to group norms about what type of suggestions were considered meaningful and acceptable to the rest of the staff.

Another strategy to assist teachers with observation feedback was to develop a guide teachers might use to focus an observation. At the January 8, 1997 PRIME Steering Committee meeting, Dr. Allen distributed a list of possible categories for 2 + 2 observation focus (Figure 4) which included skill areas commonly found in traditional evaluations. The generic list was intended to be a starting point for schools to develop a customized list of observation foci, although adopting the list "as is" was also an option. Administrators were asked to consider how the list might be altered to better reflect their individual school's agendas. For example, the list might be modified to include alternative assessment, a reading emphasis, or interdisciplinary instruction. 2 + 2 teachers could then consult the categories of focus as an aide in developing suggestions for improvement during their observations. Schools might include only the current areas of staff development focus, to reinforce those topics.

The idea received perfunctory attention at the meeting and, despite reminders at subsequent PRIME Steering Committee meetings, none of the PRIME schools modified the list, or adopted the list for use by their 2 + 2 teachers. Reasons for the inaction were unclear. By all appearances, harried administrators could not, or chose not to give the program implementation any additional attention.

Numbers of 2 + 2 Observations. Achieving the required number of observations was an implementation issue. In late January and early February 1997, the researcher collected and photocopied all 2 + 2 forms that had been submitted to the administration of the high school during the first semester. The purpose was twofold: (a) to analyze the types of compliments and suggestions teachers made, and (b) to chart the number of observations teachers had made and received. The analysis of the types of compliments

Possible Categories for 2 + 2 Observation Focus

1. *Use of Technical Skills*
    reinforcement
    questions
    examples
    teaching aids
    lesson structure
    group learning

2. *Pedagogical Skills*
    effective preparation
    learner attention
    learner interest

3. *Course Content*
    clearly identified concepts
    clear distinction between concepts and illustrations
    appropriate level of complexity

4. *Classroom Management*
    variety of control techniques
     positive and negative
     verbal and non-verbal
    efficiency of class administration
    use of students in administrative tasks

5. *Trial and Error Learning*
    appreciation of mistakes
    open to student correction
    sufficient repetition

6. *Classroom Environment*
    joy
    order
    best use of facility

7. *Language Skills*
    clear pronunciation
    good vocabulary level
    effective communication

8. *Evaluation*
    modification in lessons based on real time experience
    awareness of learners' success or failure
    assistance to weak students

9. *Administrative Issues*

Figure 4

and suggestions made by teachers can be found under 2 + 2 Observation Forms below. Determining the number of observations conducted by teachers was a monitoring task. Lower than expected numbers of observations were recorded.

Few teachers conducted the anticipated number of observations, but the trend over the course of the year indicated a rise in the number of observations. The required number of observations cited in the 2 + 2 application form was 20 per semester. That number was amended to 14 as the program did not officially get underway until October 1997. During the second semester, the number of observations was amended to 15, by the PRIME Steering Committee. Up to five post-observation conferences could count toward the total number of observations required. Among the issues discussed below which impacted the number of observations teachers conducted are time, accountability, administrative leadership and support, the researcher's role, and adaptive actions taken. Observation trends are also examined.

Table 2 summarizes the level of observation activity during the first semester at Lake Taylor High School. Of the 49 teachers whose observations were tallied,

Table 2

2 + 2 Observation Patterns at the High School, First Semester 1996-97

Number of Observations Made	Number of Teachers (N=49)	Number of Observations Received	Number of Teachers (N=49)
0 - 4	17	0 - 4	9
5 - 9	18	5 - 9	29
10 - 14	10	10 - 14	8
15 - 19	3	15 - 19	3
20 +	1		

approximately seven were unofficial 2 + 2 participants, and had no minimum number of observations to conduct. Numbers of observations teachers made ranged from 0 to 18,

with the exception of one department chair who made 67 2 + 2 observations. Nearly one-third (29%) of 2 + 2 teachers had made 10 or more observations at the close of the first semester. Forty teachers (82%) received at least five observations.

Time. Clearly, although 2 + 2 enjoyed widespread support among participants, the number of observations originally agreed upon was not being met. The reason cited most frequently was a lack of time (see focus group below). At the beginning of the second semester in February 1997, at least three high school teachers asked to be removed from the official 2 + 2 program because of lack of time to conduct observations. Different teachers at the high school had varying amounts of time for observations, a discrepancy which eventually undergirded the discussion on whether 2 + 2 should be mandatory for all teachers. Teachers felt that any mandatory program should impose similar burdens on each teacher, whereas finding time for mandatory 2 + 2 observations would be more difficult for some teachers than for others.

Teachers tended to both desire and resist release time. The simple paperwork required to schedule an intern to cover a class period, in addition to the lesson plan preparation which coverage entailed, seemed to many teachers "not worth the effort". Another reason teachers offered for resisting release time from classes was the feeling that classrooms were their primary responsibility, and spending time away from "their kids" constituted negligent behavior on their part. Teachers gave the impression that spontaneously taking advantage of free time during a planning bell was a more desirable alternative. Yet, complaints were also heard about the "necessity" of using valuable planning time for 2 + 2 observations. Thus, ambivalent feelings about how teachers' time should or could be used was probably an inhibitor of 2 + 2 observations for some teachers.

For several high school teachers who were successful in completing more 2 + 2 observations, planning a strategy in advance was key. One teacher, for example, identified potential times for 2 + 2 observations at the very beginning of a week. Another established a goal to visit teachers floor by floor. At the elementary school, 2 + 2 observations were successfully accomplished when a substitute teacher was organized by the principal and teachers signed up in advance for 2 + 2 observation time. For the

majority of teachers, though, it appeared that a strategy to prioritize time for 2 + 2 observations was absent, perhaps symptomatic, at least in part, of a general crisis driven approach to the school day.

Teachers at the elementary level had the most difficulty finding time to observe colleagues. Their day was fully scheduled, and they were required to assist resource teachers when not holding class. At the middle schools, grade levels had the same schedules, meaning that a sixth grade teacher could not observe other sixth grade teachers without a substitute. At one middle school, classes were staggered so that teachers could observe colleagues outside of their cluster only with great difficulty.

The reality of the situation was that time was also a concern of administrators. Time to spend promoting and monitoring the 2 + 2 program appeared virtually non-existent. Demands of time imposed by the educational system compromised the efforts of even the most supportive administrations both in the high school and elementary school. Were 2 + 2 not a sanctioned alternative appraisal system requiring some modicum of accountability, administrative support may have evaporated completely. Could administrators have made more time? What is possible is "more political and situational than rational and logical", according to Patton (1997, p. 201), and the situation of some and the politics of others, notably in the middle schools, precluded more time being devoted to the 2 + 2 program.

Adaptive action. As a result of feedback from the Lake Taylor High School focus group session at the end of January 1997, several actions were undertaken to improve ease of use at that school. First, a sign-up form was circulated to all teachers and interns to clarify who was participating, who still wished to participate, and whether they preferred prior/no prior notification of visits.

Second, bell schedules for 91 teachers and interns were then distributed to all 2 + 2 participants, which included prior notification preferences for each teacher and summative evaluation year status. Only 11 teachers indicated they felt a need for prior notification of a visit. Teachers were encouraged to observe those on summative evaluation, but were free to choose whom to observe at all times.

Third, a chart listing all participating 2 + 2 teachers, official and unofficial, was

posted in the teachers' mailroom, near a locked ballot box where all administration copies of completed 2 + 2 forms could be submitted. Teachers were encouraged to make a tally mark on the chart beside the name of the teacher observed, whenever they completed an observation. Consulting the chart, teachers could decide to visit colleagues with fewer observations so the system would be in better balance. The chart also served as a visual window on 2 + 2 activity in the school. Several teachers added their names at the bottom of the list, indicating they, too, wished to be considered for observations.

The implementation of the 2 + 2 program provides an interesting example of innovation diffusion. As the year progressed at the high school, more and more teachers became comfortable with the idea of opening their classrooms to peer observers. Participating on this level did not involve any paperwork or application form, making it easy to "try it out". The trade-off in terms of program implementation was that those teachers who were official 2 + 2 participants, i.e. those participating in lieu of the NPS appraisal system, received fewer observations than anticipated.

Accountability. Teachers in general were concerned about the numbers of observations being completed, in part because they had not fulfilled their own observation requirement as stated on the 2 + 2 application form, but also because many had not received the required number of visits. Consequences for failing to complete fourteen observations during the first semester were undefined, adding to teachers' uncertainty. Teachers' worries were tempered, however, by the fact that 2 + 2 teachers the preceding year were not held to any specific number of 2 + 2 observations, and because administrators showed minimal concern.

Administrative leadership and support. Principals at the four PRIME schools demonstrated varying degrees of support and leadership for the 2 + 2 program. Administrators at Lake Taylor High School and Little Creek Elementary School in particular were committed to successful implementation of 2 + 2. The LCES principal, who proactively supported 2 + 2, recognized the need during the first semester for a substitute teacher to be available once a month. Teachers signed up for time slots when the substitute could cover their class, allowing them to complete 2 + 2 observations. Although a supporter of the 2 + 2 program principles, the LCES principal found

elementary school teachers exhibited surprising resistance to peer observation, the enthusiasm of the small number of elementary level 2 + 2 participants notwithstanding.

At LTHS, teachers were encouraged to keep up with their observations with a briefing, memos, and an individual report at the semester change documenting the number of observations made and received. The assistant principal of the high school supported the 2 + 2 implementation fully, and voiced commitment to the principles of the 2 + 2 model. At the same time, the researcher initiated all feedback to participants, wrote all memos and individual reports, and arranged all meetings concerning 2 + 2. This was not unreasonable on one level, given that the researcher's role was to support the 2 + 2 implementation. Even at the high school, though, effective support of the 2 + 2 implementation was not always possible. Teachers at the January focus group meeting at the high school suggested that a large group meeting of all 2 + 2 teachers to share 2 + 2 experiences and ideas for suggestions would be helpful. Attempts to schedule such a meeting were not successful, due largely to logistical reasons. However, one meeting which had been scheduled was superseded by other agendas when the meeting date arrived, indicating the low priority 2 + 2 received among other school concerns.

At Azalea Gardens Middle School it was difficult to establish how 2 + 2 was supported by the principal. Teachers unanimously said they could not find time for 2 + 2 observations, in part because the interns were over scheduled. At the same time, assurances were given by the principal that the 2 + 2 program was functioning very well, and that a new schedule for the interns would allow them time to cover 2 + 2 teachers' classrooms. A follow-up visit to AGMS, however, revealed that interns were still not available for coverage to allow teachers to complete 2 + 2 observations. In addition, the assistant principal at AGMS continued to observe summative year teachers using the NPS appraisal instrument throughout the year, despite the principal's promises to brief her on the need to use the 2 + 2 protocol when observing 2 + 2 teachers.

Lake Taylor Middle School supported the 2 + 2 program by systematically tracking 2 + 2 observations using a database, including dates of observations and teachers visited. Administrative involvement did not appear to extend beyond monitoring observations. Due in part to difficulties obtaining class coverage with interns, teachers

faced logistical problems when conducting 2 + 2 observations. The administration suggested this was the teachers' problem, and that interns were, indeed, available. This impasse was typical of relations observed by the researcher between 2 + 2 teachers and the administration. However, full cooperation by the administration in organizing focus group meetings and distributing memos was granted the researcher.

Researcher's role. As the year progressed, the researcher's role as participant became more ambiguous. Faculty presentations and orientation sessions were clearly part of the researcher's role, as was data collection. As mentioned above, principals allowed the researcher to make presentations to the faculty and initiate the program in the fall of 1996. The cooperation of school administrations seemed sufficient at the outset, and an absence of administrative initiative in leading the way was not surprising given the large menu of programs needing attention, particularly at the year's begin.

During the second semester, as implementation issues arose, it became apparent that the authority to promote and lead the implementation of 2 + 2 rested with the schools' administrators. Therefore, barriers to the growth and success of 2 + 2 resulted when administrators did not take actions to assert their commitment to the 2 + 2 as a long and short term benefit to educational reform. At all of the participating schools with the exception of LCES, 2 + 2 was not mentioned at faculty meetings, 2 + 2 teachers were not complimented on the numbers of observations accomplished, and the program was not highlighted in ways to make it visible to the entire faculty. The agreement for the researcher to provide 2 + 2 orientation, evaluation and general support became an implementation issue in that it substituted, in practice, for active administrative promotion of the program.

Adaptive action. On March 5, 1997, the PRIME Steering Committee discussed two issues raised by the researcher in early February: (a) reducing the number of 2 + 2 observations required during the spring semester, and (b) the issue of teacher accountability for the number of observations completed.

The proposal to reduce the number of 2 + 2 observations from 20 per semester, as stated in the application form, was suggested by a teacher at a focus group meeting. Fifteen observations would be required, of which five could be post-observation

conferences, briefly documented by the teacher. The conferences, loosely defined as conversations where issues raised by the observation were discussed, could take place in the hall, at lunch, or anywhere convenient to the teachers. In approving the proposal, the Steering Committee noted that conferences served the objectives of the 2 + 2 program to reduce isolation and improve instruction through dialogue among teachers.

The issue of how teachers could be held accountable for meeting the required number of observations was not so easily resolved. On March 5, the Steering Committee voted to allow each school to decide upon a policy. As April went by with no formal decision, each school realized it was too late in the year to return 2 + 2 teachers to the NPS appraisal system, yet not fair to reprimand them in view of the lack of administrative attention. At the same time, many 2 + 2 teachers considerably increased the number of observations completed during the second semester. In the end, no negative repercussions were experienced by any 2 + 2 teacher.

The high school eventually decided to document the number of observations a teacher had made and received on the self-reflection report, the 2 + 2 summative evaluation document. The rationale was that, as part of a teacher's permanent file, the self-reflection report would speak for itself in terms of compliance with program guidelines, and that this amounted to a reasonable degree of accountability.

Trends in numbers of observations. Teachers' concerns resulted in a more concentrated effort during the second semester. Though still falling short of the intended target of 15 observations, the number of observations completed by summative year teachers at the high school rose dramatically from first semester totals. Of 28 summative year teachers whose observation totals were available in June 1997, 22 (79%) had completed 10 or more observations during the second semester. Eleven teachers doubled or tripled the number of observations completed during the first semester. Only three teachers completed the same or fewer observations than during the first semester. Table 3 shows the range of 2 + 2 observation activity for the 28 teachers.

Ten observations per semester eventually became the number proposed by the 2 + 2 Sub-Committee (see Sub-Committee Proposal for the 1997-98 2 + 2 Program below) as the requirement for the 1997-98 school year. As well as being the number

which emerged as a feasible minimum through teacher practice, 10 was the number of observations identified most often in a teacher survey in May as an optimal minimum number of observations per semester.

Table 3

2 + 2 Observation Patterns at the High School, Second Semester 1996-97
For Summative Evaluation Year Teachers

Number of Observations Made	Number of Teachers (N=28)	Number of Observations Received	Number of Teachers (N=28)
0 - 4	2	0 - 4	4
5 - 9	4	5 - 9	14
10 - 14	12	10 - 14	8
15 - 19	6	15 - 19	2
20 +	4	20 +	0

Impact of the PRIME Project on the 2 + 2 Implementation

Dynamics within the PRIME project also impacted the 2 + 2 program, providing a partial explanation of the apparent passiveness on the part of school administrators toward its implementation. Over the months, the ambiguity surrounding the role of the PRIME Steering Committee mentioned earlier had widened into a general leadership void. As a result, Dr. Allen became more active in the schools, and his efforts to monitor the project and offer ideas and encouragement began to create antagonism among some administrators. Well aware of the problem, Dr. Allen asked the deputy superintendent to call a principals' meeting to address concerns within the PRIME project, and to begin to renew and define the commitment to the project.

The purpose of the meeting on March 25, 1997 was, in the deputy superintendent's words, "to reassess the project and keep it on track" and to "ask what it's going to take to make the progress we want." The deputy superintendent called for

"frank and honest discussion" as principals responded to the question, "What is PRIME?" The Lake Taylor Middle School principal said that PRIME was no different than the district, "it's doing district stuff". The only difference was the interns. She questioned what the mission of PRIME was. Later she commented that 2 + 2 was a top down program. The Little Creek Elementary School principal perceived a lack of continuity among schools and not enough sharing of information or responsibility. As the LCES principal pointed out, no working mechanism existed to share information and ensure responsibility for project implementation.

The Azalea Gardens Middle School principal commented that PRIME was top down, lacked consistency, and didn't reach parents and teachers. He also commented that 2 + 2 was "outstanding". The LTHS principal said that PRIME was a plan for schools to restructure, especially the bottom 25% and then touched upon concerns relative to the district's quality schools initiative. "Regardless of what PRIME does", he said, "the state will issue a report within three years...There are too many conflicting expectations. We can't say we're doing PRIME when we don't know what it is."

The larger agenda of establishing accountability measures for PRIME initiatives, including the 2 + 2 program, and for PRIME schools was tabled until the next meeting. The ambivalence and, in some cases, resistance toward the PRIME project gives a picture of the political background against which the 2 + 2 program was implemented.

Meetings with Principals. To help give structure to the 2 + 2 discussion, a series of program issues needing resolution for the next school year's implementation of 2 + 2 had been compiled by the researcher (see Appendix C). The researcher met with the two middle school principals, and twice with the high school principal during April to solicit their input to these issues, and obtain general feedback about 2 + 2. Each principal was positive about the 2 + 2 program, especially as a means to reduce teacher isolation. Suggestions for program modification included incorporating the goal setting sessions employed in the NPS appraisal system into 2 + 2, and developing both more focus for observations, and staff development to help increase skill levels of teachers. The AGMS principal was concerned that a specific schedule, and a standard printed set of guidelines needed to be developed for the next year.

One area of concern to the LTMS principal was that some teacher behavior, such as relations with parents, or committee membership, was not addressed by 2 + 2. She felt 2 + 2 was incomplete as an evaluation alternative. The high school principal did not concur, noting that the administration should be aware of professional conduct regardless of the form of teacher evaluation used, and should handle any unprofessional behavior immediately on a separate, case by case basis. If there were no indications to the contrary, he felt, professional behavior should be taken for granted and need not be an element of teacher performance appraisal.

Interestingly, in making suggestions for future implementation, no mention was made of a greater role for administrators. Between the two middle school administrators, and despite a specified role for administrator observations, to this researcher's knowledge only one 2 + 2 observation had been performed during the first semester. One middle school principal commented that administrator observations were a part of 2 + 2 that had slipped her mind.

The high school principal was encouraged by the 2 + 2 implementation in his school, and was interested in developing guidelines based on the researcher's concerns list for the following school year. Evaluation data which later became available from the survey (see below), gave the high school principal convincing evidence that 2 + 2 made a positive difference. Among all the principals, he allowed his commitment to the 2 + 2 program to be contingent upon data, and was not resistant to the program for political reasons. Oddly, at a time when the principal was speaking up in support of 2 + 2, the assistant principal, who was leading the 2 + 2 implementation, inexplicably stated that he didn't think 2 + 2 was a reasonable alternative evaluation system. He acknowledged that he wasn't sure what his reasons were.

New District Leadership and the 2 + 2 Implementation. During April the district level leadership for the PRIME project changed hands, from the deputy superintendent to Dr. Lockamy, assistant superintendent for instruction. A series of four meetings were held in May to reassess the PRIME project. The 2 + 2 implementation became a secondary issue, overshadowed by intense discussion about the nature of the project itself.

Nevertheless, at the May 14 PRIME Steering Committee meeting, a number of issues regarding 2 + 2 were discussed. Dr. Lockamy, the Steering committee's chairperson, noted potential legal complications and union concerns, the first time these issues had been raised. The 2 + 2 observations needed, he said, to follow the NPS performance appraisal schedule, which meant that the time frame for making 2 + 2 observation should also end the next day, on May 15.

Also at the meeting, the two middle school principals and an elementary school principal questioned the suitability of the 2 + 2 program as an alternative evaluation program. The LTMS principal waved the 2 + 2 self-reflection report in the air and asked whether anything like it existed in the business world. Clearly implied was the idea that business evaluations are more objective and rigorous. The AGMS principal, who had assured the researcher that the revised intern schedule would allow release time for 2 + 2 teachers, reported being surprised to discover that his 2 + 2 teachers had not been accomplishing their observations. The same principal had also reported in the March 25 meeting that "2 + 2 is outstanding".

The LCES principal reported on the positive impact 2 + 2 had on the participating teachers, although logistics were a difficulty at the elementary level. A teacher at LCES, who had led a group meeting for 2 + 2 teachers earlier in May, 1997 was extremely positive about the program. An outcome of the discussion was the formation of a 2 + 2 sub-committee co-chaired by the LCES teacher and the high school principal. The researcher was also asked to be part of that committee.

The concerns voiced at the meeting on May 14 are surprising because 2 + 2 had been discussed with the principals many times, and documentation of 2 + 2 was disseminated widely among all PRIME school staff and administrators. Many PRIME administrators seemed unaware of which specific agreements were stated in the application form, which participating teachers and administrators had signed at the year's begin. Additionally, middle school principals had been interviewed by the researcher in April and had given positive feedback about 2 + 2.

Sub-Committee Proposal for the 1997-98 2 + 2 Program. The 2 + 2 sub-committee met in June, and the high school principal drafted a proposal (see Appendix

D) addressing all the issues raised by the teachers, principals, and the researcher. The proposal reinstated the eligibility of first year and probationary teachers, established guidelines for a 2 + 2 observation schedule, defined a role for a building coordinator of 2 + 2, and refined the summative evaluation report, formerly the self-reflection report. The number of observations to be required was based on the survey (see below) conducted by the researcher in May 1997. The proposal was presented to the PRIME Steering Committee on June 20, 1997, together with the survey results. One middle school principal was present and offered continued resistance to the 2 + 2 program.

The PRIME Steering Committee will meet on July 11 to consider the proposal. In response to the need for written guidelines, a group of teachers from the summer field-based graduate courses is reviewing the 2 + 2 program and plans to design a 2 + 2 handbook.

<div align="center">Summary</div>

The 2 + 2 program was officially implemented during the 1996-97 school year at four PRIME schools: Lake Taylor High School (LTHS), Lake Taylor Middle School (LTMS), Azalea Gardens Middle School (AGMS), and Little Creek Elementary School (LCES). After faculty presentations and orientation sessions early in the fall semester, 2 + 2 observations began during October 1996. Data collection in the form of questionnaires, interviews, focus group meetings, 2 + 2 forms, self-reflection reports, and a survey continued during the year (see Table 1 for time line). Evaluation of the 2 + 2 program focused mainly on the high school implementation, where about 40 summative evaluation year teachers participated in lieu of the NPS appraisal system and 50 more participated informally. Teachers participated informally by agreeing to be observed using the 2 + 2 protocol or completing a number of 2 + 2 observations on their own. Thirty-two summative evaluation year teachers eventually completed the year as 2 + 2 participants in lieu of the NPS appraisal system.

The 2 + 2 program implementation in 1996-97 greatly expanded the program's base of support among teachers, especially at the high school. As will be seen in the Outcomes section below, teachers were very enthusiastic about 2 + 2 observations. As implementation progressed, however, barriers to 2 + 2 implementation emerged.

Implementation difficulties included time to perform observations, formulating meaningful suggestions when conducting 2 + 2 observations, and passive, inadequate administrative support. Where possible adaptive action was taken to alleviate difficulties, including a PRIME Steering Committee decision to require fewer observations during the second semester.

Time for 2 + 2 observations was inherently difficult to find at the elementary school level, due to the nature of elementary school schedules. The principal moved quickly to hire a substitute teacher once a month, to provide release time for 2 + 2 observations. Although time was a constraint for most teachers, teachers generally agreed that making time for 2 + 2 was also a matter of prioritization. Only at Azalea Gardens Middle School did teachers insist that their overfilled day made 2 + 2 observations virtually impossible.

Another barrier to the success of the 2 + 2 program was the resistance of some school administrations. At LTMS, the principal found ways to highlight perceived advantages of the new NPS appraisal system. For example, at least one 2 + 2 teacher was observed by an administrator using the NPS form. At the summative conference, the teacher was told the glowing NPS appraisal could not be used because they were "on 2 + 2". Interns were told that they would be at a disadvantage when being interviewed for a job in the NPS district, if they were evaluated solely in the 2 + 2 format. These issues could have been handled in a way leading to modification of 2 + 2 if necessary, rather than by undermining the program.

At AGMS there was virtually no support for 2 + 2, it was rarely mentioned, and the assistant principal continued to use the NPS form "because the teachers thought it gave them better feedback". One teacher's response was that, since the assistant principal was writing positive things about her, she may as well stay with it. For her, as for others at AGMS, the 2 + 2 program had come to be an "add-on" to the NPS evaluation, and an especially difficult one due to exceptional time constraints.

The 2 + 2 program was most successful at the high school. More than half the teachers scheduled for summative evaluations participated in the 2 + 2 program in lieu of the NPS appraisal system. Nearly three-quarters of the school's staff participated in some

way. Teachers at the high school completed more 2 + 2 observations than at the other schools. The high number of participating teachers led to program recognition, that is, teachers knew what 2 + 2 was and that it was being implemented throughout the building.

The administration was supportive of the researcher, but did not proactively offer teachers encouragement or show commitment to the 2 + 2 principles except at the first of the year when summative evaluation year teachers were strongly encouraged to opt for the 2 + 2 program. Future implementations will need to have strong administrative support and commitment to move the program beyond a convenient appraisal alternative. Teachers will need encouragement to commit time for 2 + 2 observations in their non-summative years. The administration will also need to invest time and energy in monitoring and follow-up activities. At the present time, the principal of the high school has taken evaluation data into consideration and has drafted a strongly supportive proposal to better operationalize 2 + 2 for 1997-98.

Finally, the mission, mandate and accountability of the PRIME Steering Committee, and by extension of each individual PRIME school, needs to be clearly articulated. A vision of the interrelatedness of all the components of the reform effort is necessary for 2 + 2 to be more than an add-on initiative. But only the PRIME governing body can ensure that, at the minimum, compliance with decisions to implement programs is achieved at each school.

Outcomes of the 2 + 2 Program Implementation

The above description of how the 2 + 2 program was implemented addresses the processes at work in the PRIME project and its implementation of the 2 + 2 initiative. Two further questions address outcomes of the 2 + 2 program implementation: What difference did the implementation of the Alternative Teacher Performance Appraisal: 2 + 2 make? What were teachers' perceptions of the benefits and drawbacks of the 2 + 2 program? To address these outcome questions, the following data were collected and analyzed: teacher interviews, focus group meetings with teachers, completed 2 + 2

observation forms, self-reflection reports, a teacher survey, and a Teacher Attitudes Toward Performance Appraisal questionnaire.

Interview Analysis

Forty-eight middle and high school teachers were interviewed in October, and early November of 1996. The interviews were conducted and tape recorded by eight PRIME interns who were members of a research methodology course. The interview consisted of twelve structured interview questions (Figure 5).

Interviews were conducted to gather baseline data about teachers' expectations of the 2 + 2 program and their reasons for participating. The interviews were analyzed on a cross-case basis. Each question was analyzed separately, across all interviewees. Similarities of responses were noted and categories were developed around which similar responses could be clustered.

The interview data is summarized below. For each question, a table shows the range of teacher responses. Some questions received multiple responses from individual teachers, and occasionally a question was omitted during the interview session. Therefore, the total number of teachers does not always add to 48, and percentages included in each category do not always add to one hundred. To complement the tables and retain the richness of the data, excerpts of transcribed responses are noted under each category.

Question 1: Prior 2 + 2 Experience

Table 4

Categories of responses to the question, "Did you participate in 2 + 2 last year?"

Category of Response	Number of teachers	Percentage of teachers
No	31	65%
Yes	10	21%
Limited previous experience	7	14%
Total	48	100%

Interview Protocol for 2 + 2 Participants

Name of teacher:_____

Name of interviewer:_____

____1. Did you participate in 2 + 2 last year?

____2. Why did you opt for 2 + 2 this year instead of the NPS appraisal system?

____3. What are your expectations of the 2 + 2 program?

____4. What do you think are the objectives of 2 + 2?

____5. How do you feel about the NPS teacher appraisal system (the one in place from '83-'95)?

____6. How often do you discuss instructional strategies with colleagues?

____7. Do you wish for more opportunity to interact with colleagues professionally?

____8. How do you feel about having others observe you?

____9. Do you expect to develop new instructional strategies as a result of 2 + 2?

____10. Do you expect 2 + 2 to affect your teaching in other ways?

____11. Do you intend to ask your students for 2 + 2 observations? How do you feel about that?

____12. How important do you think your teaching strategies are in terms of student achievement?

Figure 5 Interview Protocol for 2 + 2 Participants was used by interns to interview 2 + 2 teachers in October 1996.

Table 4 shows the responses patterns to the first interview question, "Did you participate in 2 + 2 last year?" This table is self-explanatory. If the group with limited experience is included with those who participated last year, a total of 35% of those opting for the 2+2 program in lieu of the district appraisal process had prior experience with the program.

Question 2: Reasons for Participating in 2 + 2

Table 5

Categories of responses to the question, "Why did you opt for 2 + 2 this year instead of the Norfolk Public School appraisal system?"

Category	Number of teachers' responses	Percentage of teachers
1. 2 + 2 will improve instruction	23	48%
2. Respondent has a negative view of NPS appraisal system	11	23%
3. Respondent expects 2 + 2 to be easy, interesting, convenient	10	21%
4. Peers can help in ways administrators cannot	8	17%
5. Positive way to grow professionally	7	15%
6. 2 + 2 is brand new and different	4	8%
7. Teacher was encouraged to participate in 2 + 2	3	6%
Total - multiple responses were given	66	(over 100%)

One teacher's response, classified in category one of Table 5, was that she opted for 2 + 2 because she thought it would "improve instruction in the long run...it's kind of a hot line for teachers to move strategies from one classroom to the other". Another teacher felt that the "only way to improve instruction is to see other teachers work". Or, as summed up by another teacher, "A surgeon or dentist can watch surgery or dental procedures, but teachers don't know how teachers teach."

Regarding comparisons with the district appraisal system (category two), teachers said: "I get more feedback [from 2 + 2] and it is a more useful tool than the

monstrosity that they have built [new district appraisal system]"; "...[traditional appraisal] doesn't do much to improve performance...2 + 2 gives the means to improve your teaching." Teachers also commented on peer help: "...peers can give you a different look...working with you in the classroom...that even administrators might not see".

Question 3: Expectations of 2 +2

Table 6

Categories of responses to the question. "What are your expectations of the 2 + 2 program?"

Category	Number of teachers' responses	Percentage of teachers
1. New ideas to improve teaching	30	63%
2. Quick, honest, mutual feedback	13	27%
3. Alternative evaluation system	6	12%
4. Start a dialogue among teachers	3	6%
Total - multiple responses were given	52	(over 100%).

Under category one of Table 6, one teacher expected 2 + 2 to "help with renewal - get back feelings of creativity and enthusiasm". Others said: "I hope to gain more methods of teaching which are better than what I'm doing now"; "I hope to get accurate information about how to improve".

The second category is related closely to the expectation that 2 + 2 will improve teaching, but was created because of the specific emphasis on the value of feedback. In category two (Quick, honest, mutual feedback), teachers made comments such as "I want some real honest feedback", or "...seeing the feedback possibilities and the professional excitement. I thought it was neat". Regarding 2 + 2 as an alternative evaluation system, one teacher expected to have a "valid assessment of what I'm doing". All comments in this category reflected the expectation that the 2 + 2 program would be

"friendlier" and "more helpful" than the traditional evaluation system. Finally, one teacher commented, "To me, it's going to cause teachers to start talking amongst themselves."

Question 4: Objectives of 2 + 2

Responses to this question were, not unexpectedly, consistent with those to question three. Each teacher gave only one response to this question.

Table 7

Categories of responses to the question, "What do you think are the objectives of the 2 + 2 program?"

Category	Number of teachers	Percentage of teachers
1. Increasing professionalism	22	46%
2. Improvement of instruction	20	42%
3. Evaluation	6	12%
Total	48	100%

Table 7 shows how teachers' perceived the objectives of 2 + 2. Decreasing isolation among teachers was seen as a way to increase professionalism, as well as the development of teamwork through observation and communication. For example, a teacher commented, "End the isolation of teachers; it will increase the professionalism of teachers because they know at any moment someone could be walking in..." Teamwork was mentioned several times. "Encourages teamwork, especially among the group who're in this system [2 + 2]. I think it's to make us all more professional."

Comments in category two, improvement of instruction, included the following: "Enhance academic rigor; our main job is that we are here for the students, and anything that can help the teachers to help the students to learn should be implemented." An objective for another teacher was "to give you a broad base of instructional strategies which you may not be using". Category three, alternative evaluation, included comments

such as "easier, more effective way to get evaluations done" and "find an alternative way to evaluate a person that's more effective, less threatening."

Question 5: Attitudes toward the NPS appraisal system

This question produced a wide range of responses (shown in Table 8), but it was possible to categorize them broadly in terms of positive, negative, or neutral attitudes toward the Norfolk Public Schools (NPS) appraisal system. Responses were considered negative when a critical comment was made and nothing positive was mentioned. Responses were classified neutral if pros and cons were weighed, and positive if comments were generally favorable. For example, a negative comment was "I don't think it was helpful...it was basically a checklist and there was not feedback". Negative comments

Table 8

Categories of responses to the question, "How do you feel about the Norfolk Public Schools appraisal system that was in effect from 1983-1995/96?"

Category	Number of teachers	Percentage of responding teachers
1. Negative comments	18	40%
2. Neutral comments	10	22%
3. Positive comments	11	25%
4. Had no prior experience with the NPS system	5	11%
5. "No comment"	1	2%
Total	45	100%

about the NPS appraisal system raised both issues of fairness and the evaluation's purpose. A neutral comment was, "...the NPS system was sufficient in some aspects, but it was deficient in others." An example of a positive comment was, "It has worked fine for me. I feel they are looking to do the best that they can to improve instruction for our

students." Other positive comments had qualifications. "I think everything was fine. It depended on how it was administered."

Three teachers were not asked the question, resulting in 45 responses. Although the question did not ask for a comparison with the 2 + 2 program, seven teachers compared 2 + 2 favorably with the NPS appraisal system. Of these seven responses, three were from teachers offering negative comments about the NPS appraisal system, three were from teachers with positive comments about the NPS appraisal system, and one simply said, "Well, I'll just say this, I like 2 + 2 better". The teacher of the year, 1996/97 at the high school, responded, "I didn't particularly care for it [the former Norfolk Public Schools evaluation system]...with 2 + 2, there's more observation, more interaction, and I think it will be much more valid".

Question 6: Frequency of Discussions about Instructional Strategies

Table 9

Categories of responses to the question, "How often do you discuss instructional strategies with colleagues?"

Category	Number of teachers	Percentage of teachers
1. Daily, or almost daily	21	44%
2. Frequently, quite often, a lot, constantly	8	17%
3. 2-3 days per week	7	15%
4. Several times per month	6	12%
5. Other	3	6%
6. No time	3	6%
Total	48	100%

A majority of teachers (59%), those in categories one and three (see Table 9), said they discussed instructional strategies at least two days a week. Another 17% of teachers responded that they "frequently" or "quite often" discussed strategies with their

colleagues. While it is impossible to know quantitatively how many interactions that may be, one teacher defined "frequently" as "not daily, but several times a week". Thus, a large majority of teachers considered themselves already quite actively involved in professional dialogue with colleagues.

Four teachers differentiated between interactions between a limited group of teachers, and other colleagues. For example, "Now within the same area of computer concepts, we're constantly talking about the methods of doing things or teaching, but in general very seldom." One high school teacher said, "I think there's a pretty good exchange of ideas, at least within the department. Maybe not across the academic areas." Another ten teachers implied that their interactions occurred within departments or clusters: "Us in Spanish are always getting new ideas from each other".

At the opposite end of the spectrum, one teacher said: "Not very often. To be honest with you, there's not even enough time at lunch to talk."

Question 7: Desire for More Professional Interaction with Colleagues

Table 10

Categories of responses to the question, "Do you wish for more opportunity to interact with your colleagues professionally?"

Category	Number of teachers	Percentage of teachers
1. Yes	39	81%
2. No, present amount of interaction is enough	5	11%
3. Time is a barrier to further interaction	3	6%
4. Objects to use of release time to interact with colleagues	1	2%
Total	48	100%

Table 10 indicates a strong desire for more professional interaction. Teachers

responding positively commented: "Absolutely"; "I would like to not only go to senior high schools and talks to some of those teachers...but also elementary people to find out what they're doing" (middle school teacher); "Within my own department, no, but outside of the department, yes"; "Definitely. I see that there are opportunities for even more interaction because of the 2 + 2 program, and I think it's an important part of our day."

Five teachers felt they interacted with colleagues enough. One teacher indicated no need for more interaction because "we meet in cluster meetings, or in the hallways or after school, or at lunch, so I feel that there is enough time spent with colleagues during the day." Three teachers indicated time was a barrier to further interactions, and one teacher objected to having substitutes or interns provide release time to interact with colleagues.

Question 8: Attitudes Toward Being Observed

Table 11

Categories of responses to the question. "How do you feel about others observing you?"

Category	Number of teachers	Percentage of teachers
1. Very comfortable	33	69%
2. Some degree of anxiety or nervousness	14	29%
3. Other	1	2%
Total	48	100%

A majority of teachers reported they were very comfortable about peer observation (see Table 11). Comments commonly noted included:"It doesn't bother me in the least", "I welcome it". "I feel like we've been doing it here for two years now

[including the high school pilot year] and most people are used to it. Students don't turn around anymore or wonder why we're here." "I'll take whatever feedback I can get."

Fourteen teachers (29%) felt varying degrees of anxiety. Of the fourteen, three didn't mind too much. "I don't mind as long as people don't expect to see something spectacular every single time." Eleven of the fourteen teachers were apprehensive or nervous. "Nervous", said one teacher, "but keep in mind they're not there to tell you things aren't going well, but to give opportunities to grow." "Some anxiety, but now, these days, there's more trust. You see young interns and teachers so used to people coming in and out of their rooms." The response in the "Other" category was "no one has come to see me yet."

Question 9: Developing New Instructional Strategies

Table 12

Categories of responses to the question, "Do you expect to develop new instructional strategies as a result of 2 + 2?"

Category	Number of teachers	Percentage of teachers
1. Yes, definitely	27	56%
2. Hopefully, probably, maybe	17	36%
3. No, but maybe add-on	4	8%
Total	48	100%

As indicated in Table 12, most teachers did expect to develop new instructional strategies as a result of 2 + 2. Twenty-seven teachers were quite definite in their expectations. Many comments included: "Yes, I certainly do" or "Of course. Already have". One teacher said, "The best teachers don't come up with their strategies on their own - they pool from four or five teachers and put it into one lesson."

In the second category, fourteen teachers indicated they hoped or thought they might develop new instructional strategies as a result of the 2 + 2 program, while three

of the teachers responded with a "maybe". One teacher commented that "I hope so. What I want is to look at the suggestions at the end of the evaluation period and build from there." Four teachers had comments such as "No, not new ones, but maybe add-on to the ones I already use."

Question 10: Other Outcomes

Table 13

Categories of responses to the question, "Do you expect 2 + 2 to affect your teaching in other ways?"

Category	Number of teachers	Percentage of responding teachers
1. Yes	29	64%
2. Possibly	7	16%
3. Unsure of question or misinterpreted question	5	11%
4. Not really	4	9%
Total	45	100%

Table 13 shows teachers' expectations regarding the potential of 2 + 2 to affect their teaching in additional ways. One teacher saw 2 + 2 as a way to model continuous learning: "I want it to help them [students] too. That they will see things differently...that if we're willing to change, we're willing to do more creative things, I hope they would also." Another teacher thought 2 + 2 would open "an atmosphere of innovation...if you're an innovative teacher, there's some professional jealousy, and they're [other teachers] afraid that's going to force them to have to change. I think with 2 + 2 everyone's growing, so it's a win-win situation." Another comment was, "I think it will help me as a teacher. Not only instruction wise, but in general. Yes, just feedback for me personally."

Other teachers were hopeful that 2 + 2 would have effects in areas other than instructional strategies, or adopted a "wait and see" attitude. "It's too early to tell," was one teacher's comment. Another said, "I'm hoping that it will help me ensure a closer relationship with my co-workers, not just in my department, but throughout the building." Five teachers seemed unsure of the question. One teacher responded, "If it's teaching and it's not affecting other aspects of instruction, what other types of affects could it have? I'm at a loss". Three teachers were not asked this question.

Question 11: Student 2 + 2 Observations

Table 14

Categories of responses to the question, "Do you intend to ask your students for 2 + 2 observations?"

Category	Number of teachers	Percentage of teachers
1. Yes	22	46%
2. Possibly	14	29%
3. Not a good idea	6	13%
4. Already solicited feedback but not through 2 + 2 format	4	8%
5. Don't know how to do it	2	4%
Total	48	100%

Table 14 shows a range of attitudes toward student 2 + 2 observations. Twenty-two teachers (46%) responded that they did intend to ask students for 2 + 2 feedback. Comments included: "They are the most honest and will give you the most accurate feedback"; "Sometimes there's a wall between the teacher and the students. This will break down that wall. Let them know they have some input"; "It puts some responsibility on the students". Fourteen teachers (29%) were considering asking students for 2 + 2 feedback. Several were concerned about asking only the more mature

students for feedback.

Six teachers (13%) thought it was not a good idea. One teacher said, "No. Their goals and mine are usually most definitely not in line." Another teacher was concerned about the maturity and/or functional level of the students. "I don't think they'd be fair in their observations, considering the kids I'm dealing with." Of the remaining two teachers, one was a dean of students and one was a counselor. They both felt they did not know how student 2 + 2 observations could be accomplished.

Question 12: Importance of Teaching Strategies

Table 15

Categories of responses to the question, "How important do you think your teaching strategies are in terms of student achievement?"

Category	Number of teachers	Percentage of teachers
1. Very important	43	90%
2. Important, but students have a role too	5	10%
Total	48	100%

Every teacher reported that teaching strategies were very important (see Table 15). Comments from teachers included the following: "Teachers are beginning to change strategies and most of that's because of PRIME." "I'll probably be shot for this, but I think that strategies are more important than knowledge of the content area, because you can learn it as you go if you can transmit it to the students." "I think students should be in control of the class, how they learn...you know, I've already learned this stuff, they haven't."

Some teachers (five) noted that teaching strategies were important, but added comments about the students' role in successful achievement. "But it's 50% students'

responsibility. There's two sides to student achievement"; "But first they have to attend class. Once they're here my teaching strategies play a huge role"; "The strategies are working. The problem is students who don't want to be here. Kids have a lot of problems unrelated to academics. Once we can concentrate on the academic problems we can be successful with the strategies."

Summary

The interviews yielded useful baseline data about the expectations and motivations of 2 + 2 participants. The interview data showed that teachers had high expectations that 2 + 2 would be of value to them, and that they looked forward to the opportunity to observe other teachers. Sixty-three percent of teachers interviewed expected to receive new ideas to improve teaching, and 91% thought it at least probable that they would develop new instructional strategies as a result of 2 + 2. These responses dovetail with the answers to question 10, where 90% of teachers reported that teaching strategies were a very important, or the most important factor in student achievement. Individual teachers mentioned the value of observing others, receiving honest, frequent feedback, and growing professionally, and 48% cited improving instruction as a reason for choosing to participate in 2 + 2.

Most teachers did not feel they were isolated, although the interactions of many were confined to department or cluster members. Seventy-six percent of teachers said they discussed instructional strategies with their colleagues at least twice weekly. At the same time, 81% said they would welcome more opportunity for professional interaction, including many who verbalized a belief that 2 + 2 would provide such an opportunity.

Most teachers (69%) felt comfortable about other teachers observing them, and several mentioned trust building as a salutary consequence of the 2 + 2 process. Trust in the process was at lower ebb concerning student 2 + 2 observations. Less than half the teachers (46%) reported that they intended to solicit student 2 + 2 feedback.

Only 23% of teachers mentioned a negative view of the NPS appraisal system as a reason for opting to participate in 2 + 2, and less than half (40%) voiced a negative view of the NPS appraisal system when asked how they felt about that system. Positive responses toward the NPS appraisal system came from 25% of teachers, indicating that

the 2 + 2 program is not necessarily a refuge for those who harbor dissatisfaction toward the traditional appraisal system. In fact, responses to this question corroborate the view captured by responses to other questions that 2 + 2 attracts teachers who are primarily interesting in growing as professionals.

Focus Group Analysis

Introduction

Two focus group meetings were conducted by the researcher at Lake Taylor High School, on January 15, 1997, and April 23, 1997. The purpose of the meetings was to give teachers an opportunity to ask questions, offer feedback, and raise issues of concern Thirteen teachers attended each meeting, for a total of 26 2 + 2 teacher participants. The teachers, representing a cross-section of participants, were on summative year evaluation and had elected 2 + 2 in lieu of the Norfolk Public School appraisal system. Meetings took place during the school day, with interns covering teachers' classes, and were tape recorded.

The first focus group meeting began with introductions of the researcher and teachers. Perhaps indicative of interdepartmental isolation, many teachers needed introductions to their colleagues. The use of the term "evaluation" was discussed briefly, to reassure teachers that they were not being evaluated or judged on the basis of their feedback. The groups seemed comfortable with the taping of the session, and were eager to share their opinions.

Focus group sessions were loosely structured around an initial open-ended question, "Describe your experiences with the 2 + 2 program so far". The researcher occasionally asked directed questions to gather feedback about feelings of teacher isolation, the required number of observations, barriers to 2 + 2, post observations conferences or conversations, and staff development. Most of these topics were raised by the participants themselves, with teachers offering personal experiences and advising each other, in a type of cooperative learning session. As one teacher commented, "I think we learn [about 2 + 2] from each other. I just discovered that I can possibly do two [2 + 2 observations] in one bell and that's okay". The second meeting in April covered many of the same concerns, with additional discussion about student 2 +2 s,

accountability for completing required numbers of observations, and scheduling.

Analysis

"Inductive analysis means that the patterns, themes and categories of analysis come from the data; they emerge out of the data rather than being imposed on them prior to data collection and analysis" (Patton, 1990, p. 390). The researcher categorized issues arising from each group interview, issues determined by the group's responses to open-ended questions. From the transcribed tapes, themes emerged which were of particular interest or concern to participants. Comments relating to particular themes were grouped together in a cut and paste procedure. The process was largely one of referring to text, assigning selected text to a particular category, and repeating the process for the entire meeting. The process is iterative, as text may need to be recategorized as new possible categories are suggested by discussion further into the sessions. The analysis will report on five main themes of discussion: (a) responses to the 2 + 2 program, (b) program clarification, (c) program implementation, (d) the suggestion component of 2 + 2, and (e) other issues.

Results

Responses to the 2 + 2 program. When asked to describe their experiences with the 2 + 2 program, teachers were unanimously positive, to a degree that the researcher felt it necessary to actively solicit divergent viewpoints several times. Teachers had nothing critical to report about the 2 + 2 program and its rationale, and the value of 2 + 2 to teachers was never challenged. Typical of comments on 2 + 2 were comparisons with the traditional appraisal approach: "I think the old method...was ineffective and not very helpful, and I think 2 + 2 is, because your peers are making the observations, and when they come in and suggest working on my questioning techniques, they're not saying they do it better, they're just saying it's something to work on and I agree with that, and it's appreciated...I wish I'd had this at the beginning as a new teacher." "The difference in me as a teacher 20 years ago, and today, the things I do well, I did not learn from an administrator. Or a department head...If 2 + 2 had been in place then, there's no telling what I could have gotten from being the observer...in my cluster which is predominantly math and science, it's been really helpful seeing how they do things

differently from social studies." One teacher said, "I think this has the greatest potential to help teachers grow, whether they're really competent of even if they're marginal..."

Teachers also remarked on the contrast between peer and administrator visits. One teacher commented, "If another teacher comes in, they're [students] much less likely to pay attention to the person...If a principal comes in, the students are likely to act differently and throw you off because they're interacting with you differently." Another put it this way: "To have peers come in is not as intimidating, and I feel I can continue working, not making a lot of changes in what I'm doing for the benefit of the observer, and continue as the 'real me'..., then I can work with what's going on in my real teaching...".

Programmatic clarification. Misunderstandings regarding 2 + 2 program elements surfaced during discussion, which prompted other teachers to offer clarification. For example, some teachers were uncertain whether a teacher needed to notify a colleague in advance of a visit (they did not unless it was specifically requested), or if it was acceptable to visit a class midway through a bell (yes). Clarification was also sought on whether administrators and department heads, some of whom continued to use the traditional evaluation checklists, should use the 2 + 2 forms during their observations of 2 + 2 teachers (yes they should).

Teachers were reminded of the provision allowing them to decline a 2 + 2 visit, and several commented that this option was important to them, even if never utilized.

Program implementation. Program implementation issues raised by teachers included completing the required number of observations, receiving the anticipated number of observations, and giving and receiving suggestions.

Completing observations. Teachers shared the following strategies for performing regular observations: (a) set a goal to visit a "floor " at a time, (b) set a goal to visit each teacher on the list, (c) plan ahead at the beginning of the week for particular visits, (d) observe during the planning bell. One teacher lamented, "There's no reason why I shouldn't be getting out more. I'm stuck in my room!" He suggested peer pressure might help. Incentives and encouragement for those completing their required number of observations were suggested, as was a meeting of all 2 + 2 teachers in the fall of the

1997-98 school year's implementation to create support and "peer pressure".

Teachers considered the planning bell the most convenient time to accomplish the observations, although interns were available for class coverage. A teacher commented that at the end of the semester, when 2 + 2 observations have fallen behind, "you don't want to go through that extra step" and do the paperwork involved in scheduling an intern. Another teacher commented, "It's not that I don't want to use the interns, but the bottom line is that it's easier for me to teach my class than it is for me to set it up for somebody else to teach the class who doesn't know anything about what I'm teaching."

Receiving observations. Teachers were concerned about not receiving the anticipated fourteen observations per semester. To help teachers to self-regulate observation patterns, a chart to show how many observations each teacher had received was proposed. Teachers would make a tally mark beside each teacher's name after that teacher was observed. A box, located in proximity to the chart in the mailroom, was also proposed to reduce confusion about where to submit administration copies of the 2 + 2 forms.

At the second focus group meeting later in the year, teachers agreed that some type of scheduling would be necessary to more evenly distribute observations. The issue of receiving observations grew in importance because many non-2 + 2 teachers were encouraged to try 2 + 2 during the year on an informal basis. By February 1997, 91 of the school's 120 member staff had signed up to be available for 2 + 2 observations, and were listed on the 2 + 2 schedule indicating their willingness to be observed. The 2 + 2 summative year teachers became a minority among participants, and were not receiving the numbers of visits that were part of the 2 + 2 program design.

Various scheduling suggestions mentioned at both meetings included assigning teachers for observations, sign-up schedules for observations, and interim deadlines for observations during each semester. Because of varying teaching loads, subject area requirements, and personal preferences, no agreement was reached on a specific proposal.

Suggestions component of 2 + 2. Teachers typically had some difficulty both

receiving and giving suggestions. Although suggestions were generally received in a spirit of trust and appreciation, two types of suggestions were considered irksome: (a) suggestions that referenced situations beyond the teacher's control, and (b) suggestions that addressed something that was implemented by the teacher before or after the visit. In addressing a positive approach to these concerns, one teacher summarized, "When I read the 2 + 2 forms and read what was positive, I can make up my own mind whether I agree with that or not. And I take a look at the other comments and then see how they [the observer] could have taken it that way, and then if it appears they misunderstood I just ignore it. Or sometimes it's good to make contact with the teacher and elaborate some, just for the sake of sharing information". As an indication that some suggestions were a continued, if minor, source of irritation, teachers at the second focus group meeting in April also mentioned the same complaints.

It appeared, in fact, that perceived misunderstandings were one impetus for post-observation conversations and discussion, though not the only one. Commented one teacher, "I think that's part of the value of [2 + 2], the informal feedback or comments that you might make to each other, it could be several days later, or week later." While no one was in favor of requiring post-observation conferencing, teachers agreed that more discussion was taking place. Said one teachers, "Everyone who's observed me, we've just talked, you know, in passing. I think it's just natural...after one person it seemed like we just talked about it and brought out so many different ideas."

Giving suggestions also proved difficult for teachers. Some argued that there was not always a need to write suggestions when the lesson was excellent and none came to mind, and objected to writing something down just because it was required. Some teachers left the suggestions section blank, or wrote to "continue with..." and inserted the observed teaching strategy.

At the same time, teachers wished to improve their skills at making suggestions. The researcher circulated a list of teaching skills teachers could focus on when observing colleagues, and a list of sample suggestions compiled from actual 2 + 2 forms. Teachers found both handouts helpful, and suggested the list of teaching skills be incorporated on the 2 + 2 form itself. None of the participants at either meeting thought a formal training

session was necessary, but interest was expressed in arranging meetings for 2 + 2 participants to review techniques for making suggestions, and to share program experiences. The researcher mentioned that PRIME Steering Committee members had received the list of teaching skills to modify as each school found appropriate, and that feedback was pending.

Other issues. Teachers at the second focus group meeting suggested the need for a building coordinator for 2 + 2, to develop structure, give more frequent reminders, and handle scheduling issues with summative evaluation year teachers. Teachers also felt that detailed, written 2 + 2 program guidelines would be helpful. The need for better communication was highlighted when several teachers indicated they were unaware of a change in observation requirements for the second semester, despite a memo to that effect. (The number of observations was amended to 15 in March by the PRIME Steering Committee. Post-observation discussions could substitute for up to five observations.)

A new issue raised at this meeting was a perceived need to set a minimum length of time for a visit. Teachers reported colleagues were appearing for just a few minutes, presumably to catch up quickly with their 2 + 2 observation requirement.

Teachers were also not entirely opposed to making 2 + 2 a mandatory program at the high school. One department chairperson commented, "Until people try it, they will never know how much benefit there is". The consensus was, however, that 2 + 2 would remain a voluntary program for the coming year.

The topic of student 2 + 2 feedback was never brought up by teachers during the year, but it was discussed at both meetings at the instigation of the researcher. Teachers were generally lukewarm in their response, although some teachers indicated they asked students for "feedback" without using the 2 + 2 protocol.

Other schools. Focus group meetings were also held at the two middle schools, and at the elementary school. Teacher responses did not differ significantly from those at the high school. Along with praise for 2 + 2, finding time for observations, scheduling problems, and making suggestions were major issues raised.

A lack of time for 2 + 2 observations was a particular source of frustration at

AGMS. Interns were not available to cover classes, since they were scheduled the entire day.

A divergence of opinion was noted at LTMS, where teachers considered the use of a teaching skills list to help focus observations too reminiscent of the traditional evaluation system. Teachers agreed that the lists might be useful, but were acceptable only if utilized on a voluntary basis. Lake Taylor Middle School teachers were the only group openly receptive to the idea of 2 + 2 student feedback, and shared strategies for eliciting useful student comments. In contrast, the idea of student 2 + 2 feedback met with near hostility at the second focus group meeting at AGMS. Teachers felt their type of at-risk students were incapable of giving meaningful feedback.

Summary

Focus group meetings provided an opportunity for teachers to give feedback on the 2 + 2 program. Teachers appeared to enjoy the opportunity to discuss their experiences, and share strategies and programmatic knowledge. Meetings seemed to reinforce commitment to the program because teachers realized they had common concerns and positive experiences alike. Feedback from the focus group meetings indicated that teachers found 2 +2 observations interesting and valuable. Despite repeated efforts to solicit other, less positive viewpoints, no teacher offered such an opinion. Even the most frustrated group of 2 + 2 teachers, those at AGMS, who could not find time in their day to observe, were unanimously in favor of retaining the 2 + 2 program and searching for solutions to the implementation barriers.

Program clarification issues, program implementation difficulties, and concerns about making and receiving suggestions were major topics of interest to participants at the first series of meeting. By the second set of meetings in April, programmatic issues and questions were minimal. Problems of time for observations, and concerns about making suggestions were continued topics of discussion. Teachers, especially high school teachers, pinpointed a need for a building coordinator, some type of scheduling or sign-up plan, and more encouragement and feedback from administrators.

As a result of focus group feedback, changes in the required number of observations were made by the PRIME Steering Committee , and, at the high school,

logistics support was increased through the use of a chart and 2 + 2 submission box. The second series of focus group meetings were arranged largely in response to participants' requests to have more opportunity to share 2 + 2 experiences. Efforts to hold a full group 2 + 2 meeting at the high school were not successful, due to scheduling difficulties.

2 + 2 Observation Form Analysis

2 + 2 observation forms were collected from the high school administrative office and photocopied after the close of the first semester, in February, 1997. A total of 362 forms were collected, representing the observations of 50 teachers. The content of the 2 + 2 observations forms was then analyzed to gain information about what kind of 2 + 2 feedback teachers were giving and receiving.

Analysis

Patton (1990) defines content analysis as "the process of identifying, coding, and categorizing the primary patterns in the data" (p.381). Classifying the data is a preliminary step in analyzing content, and facilitates "the search for patterns and themes within a particular setting or across cases" (p. 384). According to Patton, establishing a classification system can be a "simple filing system", a way to index the data by devising appropriate labels for different ideas represented in the data. More complex systems of coding, such as categorizing every paragraph in an interview with multiple coders, are "appropriate for very rigorous analysis of a large amount of data", not for "small scale formative or action research projects" (p. 384). For the analysis of the 2 + 2 observations forms, a process of categorizing, or labeling, 2 + 2 compliments and suggestions across cases was utilized.

The purpose of the analysis was to examine the content of compliments and suggestions to obtain information about the quality and pattern of comments being made as a result of 2 + 2 observations. To begin the process, potential categories of suggestions were constructed by the researcher, several graduate assistants, and Dr. Allen, all of whom are knowledgeable about the teaching profession. These categories, included "Objectives", "Curriculum", "Presentation", and "Materials". Teachers' 2 + 2 observations forms were examined and suggestions tentatively assigned to a category. As

suggestions were found which did not fit in a category, a new category or sub-category was created. Categories were revised several times, as suggestions were reviewed. Suggestions were then once again examined, and assigned to categories. The procedure was an iterative back and forth process. The major categories eventually established were "Objectives", "Curriculum", "Instructional Strategies", "Presentation Variables", "Materials", "Assessment", and "Non-productive". Suggestions and compliments were analyzed separately. A similar process was utilized to classify compliments.

<u>Results</u>

All the suggestions made by an individual teacher were assigned to a category and recorded on a single coding form (Figure 6). All the category assignments of compliments made by a particular teacher were recorded on a second coding form. A total of 362 2 + 2 observation forms were analyzed. A 2 + 2 observation form contains fields for two compliments and two suggestions (see Figure 1), for a total of 724 compliments and 724 suggestions. In the analysis, however, 764 compliments and 746 suggestions were recorded, due to instances where, for example, a compliment actually contained two different comments. Aggregate results were calculated and are represented in Table 16. Figures 7 and 8 also show graphically how compliments and suggestions were distributed among categories. Individual results and aggregate results for each category and sub-category may be viewed in Appendix E.

<u>Compliments</u>. Over half of all compliments (63%) were categorized under "Instructional Strategies", while the remaining compliments were distributed relatively evenly among all the categories. Within the "Instructional Strategies" category, "Student engagement/On task" was the largest sub-category, accounting for 22% of compliments in that category. "Classroom management/Strategies" (13%), "General Instructional Strategies" (11%), "Use students' names/Rapport" (10%), and "Questioning" (7%) were the next largest sub-categories. "General Instructional Strategies" was a category for specific compliments not able to fit well into other sub-categories. For example, "Excellent use of guided practice to help with completion of an independent practice assignment".

The smallest category of compliments was the "Non-productive" category.

OBSERVER'S NAME:
REVIEWER'S NAME: Reviewer's notes:

OBJECTIVES	
Student awareness of objectives	
Learning objectives general	
CURRICULUM	
Time management	
Warm up	
Quality of examples	
Interdisciplinary study	
Group study/Cooperative learning	
Classroom discipline/Strategies	
Class participation	
Planning and Preparation	
Use students' names	
Student encouragement	
Move around the room	
Change seating	
Student engagement/On Task	
PRESENTATION VARIABLES	
Give feedback	
Slow down	
Speak up	
MATERIALS	
Overheads	
Chalkboard	
Handouts	
Technology	
Material preparation	
ASSESSMENT	
Assessment general	
Monitor student progress	
NON PRODUCTIVE	
Blank	
Continue teaching style	
Good job (blowing sunshine)	
TOTAL Suggestions	
NUMBER OF FORMS	

Suggestions not found under headings:

Figure 6. Sample of original coding form for 2 + 2 suggestions

Table 16

2 + 2 OBSERVATION FORMS: CUMULATIVE TOTALS IN EACH CATEGORY

	Compliments	Suggestions	Total
OBJECTIVES	40	12	52
Learning objectives general	40	12	52
CURRICULUM	46	29	75
Curriculum General	4	1	5
Builds on previous lessons	8	4	12
Warm up	21	15	36
Quality of examples	11	8	19
Interdisciplinary study	2	1	3
INSTRUCTIONAL STRAT.	483	331	814
Instructional Strategies General	52	78	130
Relate to real world/Applications	19	6	25
Clarity of instructions	29	8	37
Classroom displays	8	7	15
Group study/Cooperative learning	27	26	53
Classroom discipline/Strategies	61	78	139
Class participation	24	38	62
Planning and Preparation	32	3	35
Use students names/Rapport	46	9	55
Positive feedback/Reinforcement	26	7	33
Move around the room	12	6	18
Use games	4	4	8
Questioning	35	40	75
Peer helping	4	13	17
Student engagement/On Task	104	8	112
PRESENTATION VARIABLES	50	50	100
Vivid/Captivating	11	3	14
Presentation Variables General	2	6	8
Time management/Flow	30	27	57
Slow down	0	9	9
Speak up	7	5	12
MATERIALS	44	71	115
Materials General	5	30	35
Visual aids	23	31	54
Handouts	7	6	13
Technology	5	3	8
Material preparation	4	1	5
ASSESSMENT	62	31	93
Assessment general	8	9	17
Student involvement	8	4	12
Monitor student progress	46	18	64
NON PRODUCTIVE	39	222	261
Blank	6	158	164
Continue teaching style	1	42	43
Good job (blowing sunshine)	32	22	54
TOTAL Compliments/Suggestions	764	746	1510
NUMBER OF FORMS	362	362	362

Total Compliments (N=764)
Percentages by Major Category
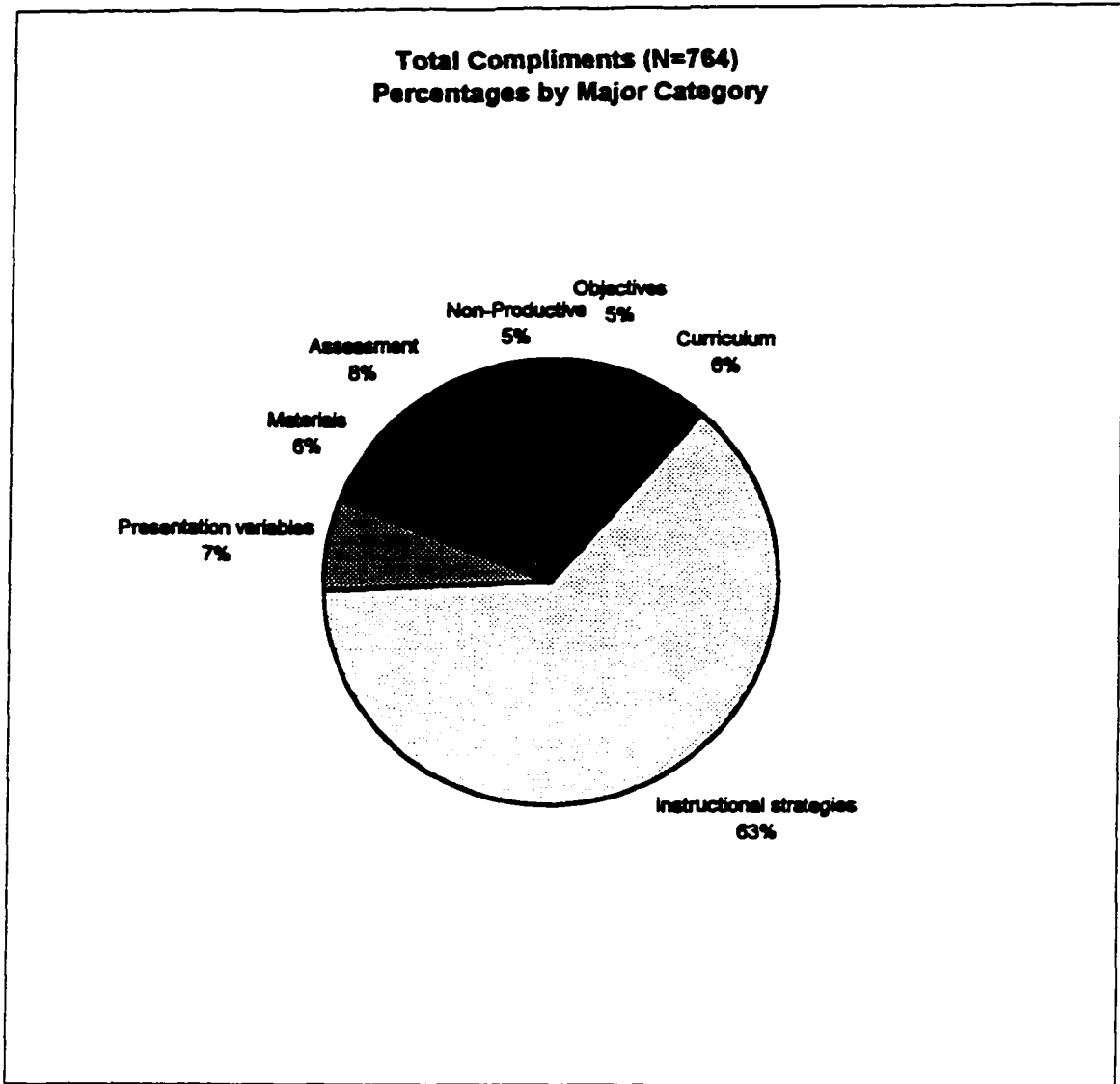
Category Breakdown	Number
Objectives	40
Curriculum	46
Instructional strategies	483
Presentation variables	50
Materials	44
Assessment	62
Non-productive	39
Total compliments	764

Figure 7. Total Compliments (N=764): Percentages by Major Category

Total Suggestions (N=746)
Percentages by Major Category
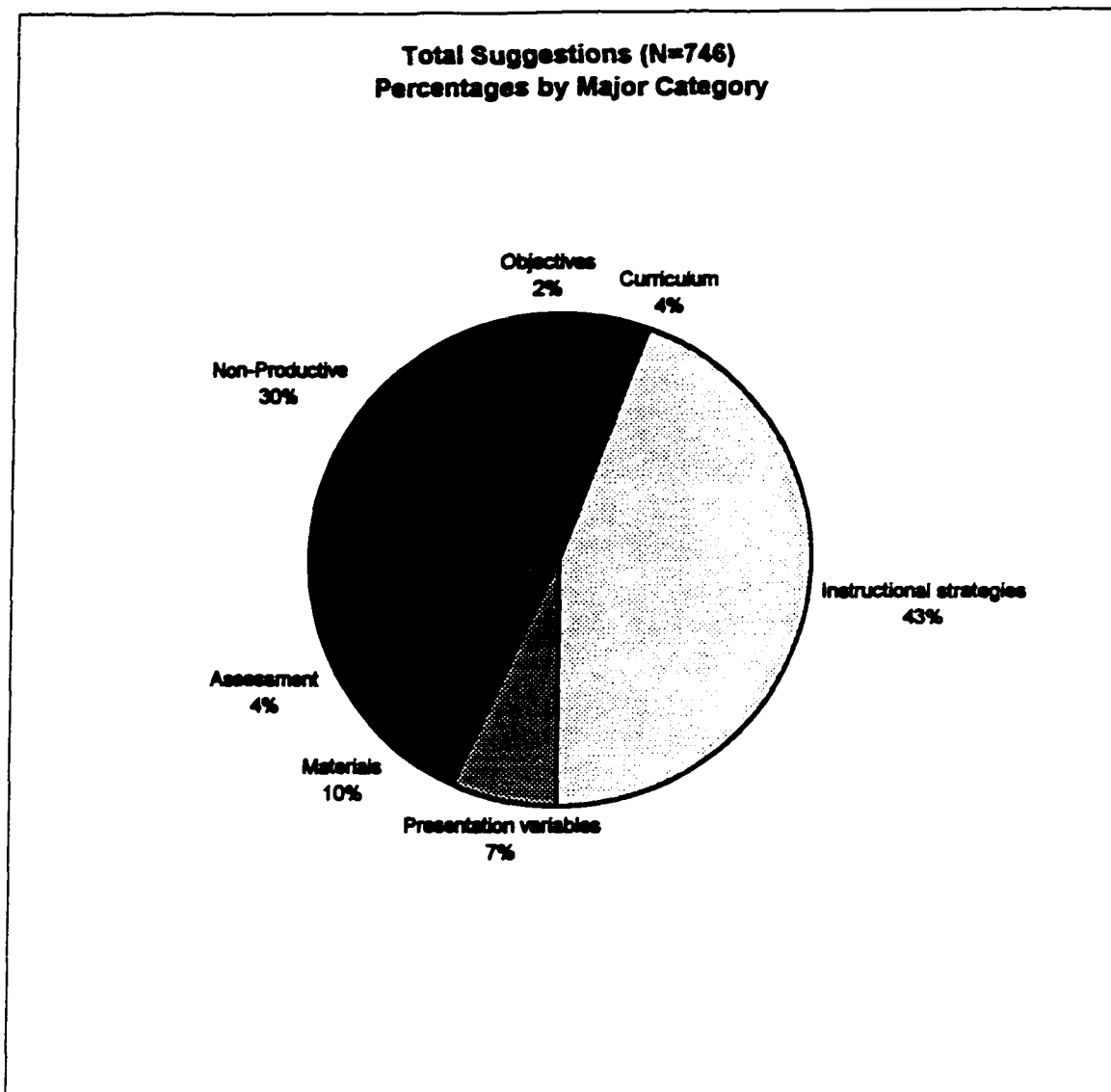
Category Breakdown	Number
Objectives	12
Curriculum	29
Instructional strategies	331
Presentation variables	50
Materials	71
Assessment	31
Non-productive	222
Total suggestions	746

Figure 8. Total Suggestions (N=746): Percentages by Major Category

Only 39 compliments out of 764 were considered "Non-productive". Compliment fields which teachers had left blank were tabulated and categorized as "Non-productive". Global compliments such as "Excellent job with quality of students" were also categorized as "Non-productive". Of the total "Non-productive" compliments, 6 were blank compliment fields, and 33 were non-specific global praise.

Each of the remaining categories, "Objectives", "Curriculum", "Presentation Variables", "Materials", and "Assessment" contained between five and eight percent of the compliments.

Suggestions. Most suggestions for improvement were categorized under "Instructional Strategies" (331 of 746 total suggestions or 43%), and under "Non-productive", the next most frequent category (222 of 746 total suggestions or 30%). The two categories together accounted for 73% of all suggestions.

The largest sub-categories under Instructional Strategies were "General Instructional Strategies" and "Classroom discipline/Strategies". Together they accounted for nearly half of the suggestions classified as "Instructional Strategies". "General Instructional Strategies" was a category for specific suggestions not able to fit well into other sub-categories. For example, "Incorporate activities that allow student movement (board work, writing on an overhead)" or "Give advanced work for faster learners" were instructional strategies not fitting any established sub-category. The large number of suggestions in the "General Instructional Strategies" category is indicative of the wide range of instructionally related comments.

On 362 2 + 2 observation forms, with 724 expected suggestions (746 suggestions were actually recorded), 158 (22%) suggestion fields were left blank. These blank fields were categorized as "Non-productive", and accounted for 71% of that category. The many missing suggestions corroborate teachers' feedback that suggestions were often difficult to make.

Two other "Non-productive" sub-categories were "Continue teaching style" and "Good job". These sub-categories included suggestions phrased as "Continue..." as in "Continue with energetic style of instruction", and global suggestions such as "Keep up the good work", which was assigned to the sub-category "Good job". Comments in

the "Continue teaching style" and "Good job" sub-categories, although classified as non-productive, were sometimes cited in the list of the most useful or meaningful suggestions in the self-reflection reports. Some teachers apparently found these comments affirmatory, suggesting that they are not entirely unproductive, especially if the teacher thereby refocused on a useful instructional strategy.

Table 17 shows the percentage of each teacher's total suggestions which were classified as "Non-productive", and the number of teachers with that percentage of "Non-productive" suggestions. Of the 50 teachers whose 2 + 2 observation forms were analyzed, 10 teachers (20%) had no suggestions classified as "Non-productive" while an additional 17 teachers (34%) had fewer than 30% of their suggestions classified as "Non-productive. The remaining 46% of teachers had 30% or more of their suggestions classified as "Non-productive".

Table 17

Percentages of Non-productive Suggestions and Number of Teachers per Category

Non-productive Suggestions as Percentage of Total Suggestions Made	Number of Teachers
0	10
1-9%	6
10-19%	7
20-29%	4
30-39%	6
40-49%	0
50-59%	9
60-69%	3
70-79%	3
80-99%	0
100%	2

A higher number of observations did not guarantee more useful suggestions. However, the data is useful because it suggests that some teachers make more productive suggestions than others, and implies a potential for improvement of observation skills by those less adept or experienced. For example, 16 teachers with fewer than 10% "Non-productive" suggestions may be able to coach their colleagues in suggestion making techniques. That this type of "training" would be useful and desirable is corroborated by focus group feedback

Summary

In general, the results of the 2 + 2 observation forms analysis are promising. Ordinary teachers with no special training were able to offer useful suggestions to their colleagues 70% of the time, and compliments 95% of the time. Triangulation of the data with results of the self-reflection reports analysis which follows in the next section, and focus group meetings (see above) suggests that 2 + 2 feedback was professionally relevant and useful to teachers almost all of the time.

Teachers found it easier to give compliments than to make suggestions. Teachers focused on instructional strategies when making compliments, with 63% of total compliments classified as instructional strategies. There were over five times fewer "Non-productive" compliments as suggestions.

Teachers also focused on instructional strategies when making suggestions, with 43% of suggestions classified as "Instructional Strategies". The data indicates that teachers felt comfortable making suggestions related to classroom discipline/strategies. For example, "Insist that all students remain seated while information is being given". Yet a broad range of higher level suggestions was also represented in the analysis. One teacher suggested using more "higher level questions - synthesis and evaluation". Another wrote, "If a student does not agree on a certain approach, have them explain how they approached the problem".

However, 30% of the suggestions on the 2 + 2 observation forms were assigned to the "Non-productive" category, with blank suggestion fields accounting for over 70% of the category total. Of 724 suggestion fields on the 2 + 2 observations forms analyzed, 22% were left blank. These findings corroborate focus group feedback that

suggestions were difficult to make.

Implications include the potential for those teachers with less than 10% "non-productive" suggestions to hold training sessions with other 2 + 2 teachers. These teachers represented 32% of teachers whose forms were analyzed, a sizable minority. The idea that teachers could help teachers with the 2 + 2 process, including formulating suggestions, was proposed by teachers in focus group meetings, making such an idea viable and acceptable to teachers.

This analysis was conducted to gain information on the types of compliments and suggestions teachers made in general. In the 2 + 2 program, every observer is important. However, more study is needed to determine patterns of suggestions and compliments made by teachers. Differences in the teacher population should be taken into account. Teachers had varying degrees of teaching experience. Some teachers were participating in lieu of the NPS appraisal system, while others were not. Finally, no attempt was made to differentiate between those with supervisory experience and those with none.

Self-Reflection Reports Analysis

Thirty-one first semester self-reflection reports were collected at Lake Taylor High School in February 1997. Self-reflection reports were only submitted by teachers participating in the 2 + 2 program in lieu of the NPS appraisal system and constitute part of the teacher's permanent evaluation record. The self-reflection report form was divided into four parts: (a) a list of the ten most useful/significant compliments, (b) a list of the ten most useful/significant suggestions, (c) an explanation of how the compliments and suggestions led to reinforcement or improvement of classroom teaching routines, and (d) a list of any future agenda items which have emerged as a result of the 2 + 2 process, as well as any reflections about the entire 2 + 2 observation process, including the value of observing others (see Appendix F for the self-reflection report form [titled 2 + 2 Summative Teacher Report Form]). The current study focuses on the implementation of the 2 + 2 program at the high school, and the following analysis was completed utilizing high school data only. Ten first semester self-reflection reports were also completed at the two middle schools. An analysis of these self-reflection reports may be found in

Appendix F.

Parts one and two of the self-reflection reports contained lists of compliments and suggestions each teacher received. An analysis of teachers' compliments and suggestions was undertaken previously in the section on 2 + 2 observation forms. Therefore, the analysis of the self-reflection reports focuses on parts three and four, which were considered together for this purposes.

Analysis

The purpose of analyzing parts three and four of the self-reflection reports was to gain insight into how compliments and suggestions had reinforced or improved the teaching process from the teachers' perspective, and how teachers viewed the entire 2 + 2 process. Most teachers addressed both sections briefly. In analyzing the self-reflection reports, content was classified and categorized to identify patterns and themes of teacher responses. This involved an iterative process of reviewing the data and identifying statements of significance, with direct bearing on the topic. Sometimes, for example, one sentence in a paragraph summarized a teacher's opinion. One hundred and three statements were gathered from the 31 self-reflective reports.

The statements were compared and grouped together with similar statements in provisional categories, keeping in mind the writing prompts for parts three and four of the self-reflection reports. For example, comments were solicited in part four about the 2 + 2 observation process, and about the value of observing others. Therefore, it was natural that a number of statements in the data related to these themes. Statements were reviewed several times, compared, and grouped together under the appropriate theme.

As Patton points out (1990), qualitative analysis does not have one single correct methodology. An appropriate approach to analyzing and interpreting qualitative analysis depends largely on the individual needs of the researcher with regard to the particular problem or program under investigation. He comments, "...there are no absolute rules except to do the very best...to fairly represent the data and communicate what the data reveal given the purpose of the study" (p.372).

Results

Table 18 provides an overview of the five identified themes, with sample

Table 18

Patterns of Response from Parts Three and Four of 31 High School Self-Reflection Reports,
February 1997

Themes from Self-Reflection Reports (N=31)	Number of Statements (N=103)

1. Teachers encouraged by praise and compliments from other teachers   26

"It is great to have colleagues let you know what they feel is good about your
    teaching/classes, and some of their suggestions I have already put into practice...".
"I feel that the compliments listed gave me a lot of positive reinforcement. They
    also gave me encouragement to try additional innovative and creative activities..."
"All compliments, if the teachers truly knows he/she is doing the cited behavior/technique,
    reinforce instruction by causing you to want to keep on doing what you are doing well."

2. Teachers provided specific suggestions to change instructional practice   24
"The suggestion has made me aware...that positioning students when working in
    cooperative learning teams is certainly significant"
"I will make specific point of getting students in front of class to present material."
"I have incorporated more warm-up questions at the appropriate time for students"
"I'm trying to become a better questioner during instruction."
"I have used small group activity before, but I will increase that. Also, I want the
    students to question each other and ask why instead of turning to me."

3. Observing other teachers provided insight, perspective, and ideas   22

"This process has given me an even greater appreciation of the work my colleagues do."
"The 2 + 2 program has enabled me as a seasoned teacher to gain meaningful
    insights and new ideas as I visit my peers."
It was helpful observing a "'troubled student' in another classroom situation...
    to see what works or does not work for that particular individual."
"By observing others, I get reinforcement that what happens in my classes is normal and
    that I'm on target in my planning and implementing strategies that result in success."

4. 2 + 2 encouraged collegiality and collaboration among teachers   17

"The 2 + 2 concept decreases my isolation and makes me feel more a part of a team effort."
"[2 + 2] has led me to a better understanding of the entire working of our faculty as a whole."
"Observing other teachers has led me to see how my department [media center]
    might better interact and enhance instruction in other departments."

5. Suggestions and comments about the 2 + 2 program   14

"The more 2 +2 s you do, the more relaxed you are in doing them."
"The 2 + 2 process is an excellent evaluative tool for assessing teacher performance."
"[2 + 2 should] stress the importance of teachers observing colleagues across all disciplines."

responses. A more complete compilation of teachers' statements may be found in Appendix F.

Theme one: Teachers were encouraged by praise and compliments from other teachers. Statements pertaining to the first theme were mentioned most often in the self-reflection reports. This theme is captured by one teacher's almost poignant comment, "It was good to see that other teachers appreciated my work and saw it as worthy". In other words, teachers very much valued comments from colleagues that led to reinforcement (and a reaffirmation) of classroom routines. Many teachers cited particular compliments which they viewed as especially pertinent. For example, "I believe I do have a good rapport with my students and will work to maintain a high level of respect", or "The compliment of [well-paced] instruction is significant. I never want to run through information just to say it has been taught". Other comments included, "Comments have reinforced my belief that I am able to effectively communicate with students...as a guidance counselor" and "I use my sense of humor to 'humanize' myself to students".

Theme two: Teachers provided peers with specific suggestions to change instructional practice. Teachers cited many specific suggestions and some more general in nature, which teachers reported led to reinforcement or improvement of classroom teaching routines. Table 18 lists several examples of this theme. A more general response was, for example, "All suggestions were considered and weighed in reference to how I could use them to improve my planning, delivery, and assessment". Some teachers found 2 + 2 feedback useful in establishing future agendas: "I do need to utilize my students more as monitors/technicians. They know a lot more than I do on occasion."

Theme Three: Observing other teachers provided insight, perspective, and ideas. This theme included several sub-themes. 2 + 2 compliments and suggestions were viewed as a means to (a) access fresh ideas, (b) discover new appreciation for the work other teachers accomplish and (c) raise the confidence level of the observer.

Most responses centered on the idea that seeing new practices in colleagues' classroom was of value, and provided motivation to try new ideas. "In all of the visits I made, I found I learned something and came away with ideas which I could adapt for implementation in my own classes." A department chair wrote: "The compliments and

suggestions I noted when observing other teachers will be the most help to me as I change classroom teaching routines."

At the same time, teachers also commented on how observing others validated their own teaching. "Observing others gives a person confidence in his or her teaching duties", wrote one teacher.

Theme four: 2 + 2 encouraged collegiality and collaboration among teachers. Theme four covered a range of comments, including the ideas of teachers helping teachers, teachers sharing what they know with each other, interdisciplinary projects, decreased isolation, and empowerment of teachers. Said one teacher, "2 + 2 has got to be the very best way of teachers making teachers better." The following insight was offered by another teacher: "This program gives teachers a choice which was not possible before. Choice promotes empowerment, creativity, and good morale...the 2 + 2 program opens lines of communication among teachers which could possibly lead to cooperative teaching endeavors."

Theme five: Suggestions and comments about the 2 + 2 program. The last theme included suggestions on logistics and implementation. Of the nine comments about logistics, three suggested requiring fewer observations per semester, and three indicated that lack of time or scheduling of observations was an issue. Two teachers expressed a desire for small group meetings for 2 + 2 participants. One teacher perceived a need for more practice and additional in service in making comments, especially suggestions.

Other general comments related to the 2 + 2 process, not specific enough to be categorized elsewhere, were included under this theme. "I think the 2 + 2 process is helpful to all teachers in many ways", and "The section on suggestions gives the teacher the opportunity to see himself as others see him"are examples of comments categorized here.

Summary

The analysis of the self-reflection reports shows thoughtful and, for the most part, specific comments about how 2 + 2 has made a difference in instruction. The responses of teachers regarding compliments were as enlightening as their comment on useful suggestions. Teachers obviously valued positive feedback. The emphasis on

positive feedback was somewhat surprising because the 2 + 2 program focused strongly on providing suggestions for improvement. It is easy to forget the power and value of positive descriptive feedback as a means to motivate and improve performance. Often the tendency is to focus exclusively on "corrective" measures or alternative methods of instruction. Teachers, however, appeared very much encouraged, not only by the praise they received from colleagues' visits, but by visiting others' classrooms and finding they were not inferior to their fellow professionals.

The comments provided a valuable window on the 2 + 2 implementation. The wide range of teachers' responses indicated that the 2 + 2 program was progressing toward its objectives of decreasing isolation, increasing collegiality, and improving instruction. The data was also a source of triangulation corroborating results of the survey and focus group data.

One limitation is that teachers may have felt an obligation to respond positively in writing the self-reflection reports. No comments which were negative toward the 2 + 2 process appeared in the self-reflection reports. At the same time, though, comments not included in the themes above because they were unrelated to 2 + 2 could be construed as critical of the school district. For example, a teacher mentioned there was "a real need for new books, software, and hardware for this [advanced Office Systems] course". The possibility nevertheless exists that, since the self-reflection reports were to become part of their permanent file in the central office, some teachers inflated the value of the process to underscore the benefits they received during their summative evaluation year.

2 + 2 Survey Analysis

Participant responses were solicited in a survey addressing three aspects of 2 + 2: a) direct comparisons of the old (1983-1995/96) NPS district appraisal system with the 2 + 2 alternative appraisal program, b) outcomes of the 2 + 2 program related to professional growth, and c) programmatic issues such as whether 2 + 2 should be voluntary or mandatory, and what the optimum number of observations should be. The surveys were distributed in May 1997 to all those who had participated in 2 + 2, either "officially", that is, in lieu of the district summative system, or "unofficially" by agreeing

to open their classrooms to 2 + 2 observers. A total of 76 surveys from the Lake Taylor High School, Lake Taylor Middle School, Azalea Gardens Middle School, and Little Creek Elementary School were collected. The survey was analyzed using descriptive statistics, correlation analysis, and General Linear Models analysis of variance.

<u>Descriptive Statistics</u>

Table 19 shows response percentages for each survey statement. Part one of the survey shows clear acceptance of the 2 + 2 program as an evaluation alternative leading to improved instruction. Seventy-nine percent of teachers preferred the 2 + 2 appraisal program over the traditional district appraisal instrument, and 67% felt 2 + 2 was a better appraisal program (statements 1 and 3). Improvement of instruction was thought to be more likely as a result of the 2 + 2 program than the traditional appraisal system by 72% of respondents (statement 7).

In part two, results indicate progress toward more experimentation with different teaching strategies, and enhanced collegial relations. Over 84% of teachers indicated they had experimented with new instructional strategies as a result of 2 + 2 (statement 1). Other statements confirm this result. For example, 96% of teachers reported implementing at least one 2 + 2 suggestion (statement 4). Statement 8 shows progress toward increased collegiality. Over 72% of teachers indicated the quality of interaction among their colleagues had improved as a result of 2 + 2.

In part three, little support was indicated for a mandatory 2 + 2 program. Only 16% of respondents agreed that 2 + 2 should be mandatory for all PRIME teachers every year, although 30% agreed it should be mandatory during summative evaluation year (statements 6 and 3). The issue of 2 + 2 feedback from students also appears to be a topic for future consultation, with 29% of teachers agreeing that it is an important component of the 2 + 2 program, and 29% disagreeing with that statement (statement 4). Only 43% of teachers agreed that regular administrator participation in 2 + 2 was important, indicating that a review of the 2 + 2 program principles may be needed.

Table 19

2 + 2 Program Survey Results (N=76)

Statements	Percentage of Responses in Likert Categories				

Part One: Evaluation Alternatives	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. I prefer the 2+2 system to the traditional teacher evaluation system.	0.0	2.6	18.4	38.2	40.8
2. The traditional evaluation system gave me the best feedback for my growth as a teacher.	14.5	42.1	30.3	9.2	3.9
3. The 2+2 system is a better appraisal alternative than the traditional teacher evaluation system.	0.0	5.3	27.6	42.1	25.0
4. I prefer the self evaluation summative form to the traditional teacher evaluation form.	0.0	11.8	27.6	36.8	23.7
5. I like to have my teaching performance ranked by an administrator.	2.6	9.2	60.5	23.7	3.9
6. The traditional evaluation system is the best system for evaluating teacher performance.	16.0	46.7	32.0	2.7	2.7
7. Improvement of instruction is more likely to occur as a result of 2+2 than the traditional system.	0.0	3.9	23.7	51.3	21.1
8. The 2+2 system does not lead to professional growth.	26.3	63.2	9.2	0.0	1.3
9. The traditional Norfolk Public School evaluation system does not lead to professional growth.	3.9	42.1	39.5	11.8	2.6

Part Two: Outcomes of the 2 + 2 Program					
1. As a result of 2+2 participation, I have experimented with new instructional strategies.	1.3	3.9	10.5	65.8	18.4
2. During 2+2 observations, I have seen strategies I wanted to try with my own students.	0.0	2.7	10.8	59.5	27.0
3. More training would make 2+2 feedback more useful.	2.6	11.8	28.9	36.8	19.7
4. I have implemented at least one 2+2 suggestion.	1.3	1.3	1.3	71.1	25.0
5. I have implemented strategies I observed in other classrooms.	1.4	2.7	11.0	61.6	23.3
6. 2+2 observations have helped me gain perspective on my own teaching abilities.	2.7	4.0	10.7	60.0	22.7
7. The 2+2 program has helped me gain confidence as a teacher.	5.3	7.9	22.4	43.4	21.1
8. The quality of interaction among my colleagues has improved as a result of 2+2.	1.3	3.9	22.4	43.4	28.9
9. 2+2 has had no impact on my professional growth.	36.8	44.7	11.8	3.9	2.6

Table 19

Continued

	Percentage of Responses in Likert Categories				
Statements					

Part Three: Programmatic Issues	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. Participation in 2+2 should remain voluntary.	0.0	5.3	17.1	51.3	26.3
2. Only a few observations should be required during a teacher's non-summative years.	2.6	15.8	15.8	50.0	15.8
3. The 2+2 system should be mandatory for all PRIME teachers during summative evaluation years.	10.5	32.9	26.3	22.4	7.9
4. Student feedback is an important component of 2+2.	13.2	15.8	42.1	26.3	2.6
5. Regular administrator observations are an important component of 2+2.	1.3	22.4	32.9	36.8	6.6
6. The 2+2 system should be mandatory for all PRIME teachers, every year.	21.3	32.0	32.0	9.3	5.3

Which specific changes would you make in the 2 + 2 program? (Multiple suggestions were given.)

No comment	51.3%	Fewer required observations	10.5%
More training	9.2%	Clearer guidelines	9.2%
More release time needed	7.8%	Scheduling suggestions	6.6%
More 2 + 2 meetings	2.6%	Other suggestions	17.1%

Summary of responses (cited are the three most frequent responses)

I believe the optimum number of observations in the summative evaluation year per semester should be:

Number of observations	Percentage of respondents
10	37%
5	14%
15	13%

I believe the optimum number of observations in the non-summative evaluation year per semester should be:

Number of observations	Percentage of respondents
5	31%
10	17%
2	16%

Correlation analysis

To better understand how different aspects of 2 + 2 interrelate, a correlation analysis was performed. For example, the relationship between the number of observations teachers made, and attitudes toward the 2 + 2 program as an alternative appraisal system was certainly of interest. If no correlation, or a negative correlation existed, it might have indicated a need to revise the number of observations, or the type of information teachers were receiving about 2 + 2.

A total of eleven categories of interest were created from the survey data. The following four categories consisted of raw data where an appropriate response was a single numeric answer: (a) number of observations received, (b) number of observations made, (c) optimal number of summative year observations, and (d) optimal number of non-summative year observations.

Four additional categories were created by grouping several similar Likert survey statements together. For example, the category, "Positive responses to the 2 + 2 alternative appraisal program" was a combination of four Likert survey statements from part one of the survey. The statements, items one, three, four and seven, were all positive toward 2 + 2 as an alternative appraisal system. The other three categories formed in this manner were: "Positive responses to the NPS appraisal program", "Positive 2 + 2 outcomes", and "Positive 2 + 2 as mandatory program". The statements which are combined to form each category can be seen in Table 20.

Three categories were also formed based on the responses to one Likert survey statement. These were: "Negative responses to the 2 + 2 alternative appraisal program", "Negative responses to the NPS appraisal program", and "Negative 2 + 2 outcomes". Refer to Table 19 to see the individual survey statements. Table 20 shows descriptive statistics, including means and standard deviation, for each category.

In Table 20, the "Max. Score" (maximum score) refers to the highest score possible for each category. Since the Likert scale had five possible responses, the number five was the highest score for a single survey statement. When two or more statements were combined to form a category, five was added to the maximum possible score for each additional statement. If a category was composed of four statements, for example,

the "Max. Score" possible was 20. The "Range" refers to the actual range of individual scores calculated in the analysis. The first four categories were not Likert items, and had no maximum score.

Table 20

Descriptive Statistics for 11 Survey Categories

Category	N	Max. Score	M	SD	Range Min.	Max.
1. Observations received	46	N/A	12.0	7.7	2	35
2. Observations made	46	N/A	13.7	9.0	0	32
3. Number summative year observations	71	N/A	9.1	4.8	0	30
4. Number non-summative year observations	71	N/A	5.5	3.6	0	18
5. Positive 2 + 2 appraisal system Scale 1: Statements 1, 3, 4, 7	76	20	15.6	2.7	10	20
6. Positive NPS appraisal system Part 1: Statements 2, 5, 6	75	15	7.9	1.9	3	13
7. Negative 2 + 2 appraisal system Part 1: Statement 8	76	5	1.8	.7	1	5
8. Negative NPS appraisal system Part 1: Statement 9	76	5	2.6	.8	1	5
9. Positive 2 + 2 outcomes Part 2: Statements 1, 2, 4, 5, 6, 7, 8	73	35	27.9	4.5	10	35
10. Negative 2 + 2 outcomes Part 2: Statement 9	76	5	1.9	.9	1	5
11. Positive 2 + 2 as mandatory program Part 3: Statements 3, 6	75	10	5.2	2.1	2	10

A Pearson correlation analysis of the 11 categories showed significant

relationships among several categories. Of interest is the strong, positive correlation of .65 between the number of observations made by a teacher and positive ratings of 2 + 2 as an alternative appraisal program. Similarly, a strong, positive correlation (.56) was observed between the number of times a teacher was observed and positive ratings of 2 + 2 as an alternative appraisal program. The correlation between the number of times a teacher was observed, and positive ratings of 2 + 2 outcomes was .40. Positive ratings of 2 + 2 outcomes also correlated, not surprisingly, with positive ratings of 2 + 2 as an alternative appraisal program (.56). Also interesting is that positive ratings of 2 + 2 outcomes, positive ratings of 2 + 2 as an alternative appraisal program, and the higher optimal number of 2 + 2 observations in a summative evaluation year, all correlated with a higher interest in making 2 + 2 a mandatory program (.29, .29, and .27, respectively). Finally, the more positive the ratings of 2 + 2 as an alternative appraisal program, the less likely were positive ratings of the Norfolk Public Schools appraisal system (.39).

The correlations above provide additional information about how teachers perceive the 2 + 2 program. While cause and effect relationships cannot be determined, it is still interesting to note, for example, that the more observations a teacher made, the more likely that teacher had a positive attitude toward 2 + 2 as an alternative appraisal system. On the other hand, a teacher may have had positive attitudes toward the 2 + 2 program as an alternative appraisal system apriori, which provided motivation to accomplish a higher number of 2 + 2 observations. In any event, it would appear that conducting a higher number of observations does not lead to a more negative view of the viability of 2 + 2 as an alternative appraisal system

The correlation analysis results are not surprising, but useful, in that important assumptions of the 2 + 2 program are corroborated. It was hoped, for example, that positive responses 2 + 2 outcomes, and positive views of 2 + 2 as an alternative appraisal system would encourage consideration of 2 + 2 as a mandatory program at some time in the future. Also, the dichotomy between the 2 + 2 program and the NPS appraisal system was highlighted, a source of data triangulation with the questionnaire below.

General linear model procedures

The General Linear Model Procedure (GLM) can be utilized to perform analysis of variance (ANOVA) procedures when there are cells of unequal size. Eleven GLM procedures were calculated (see Table 21). The independent variable was school level with three conditions, elementary, middle and high. The dependent variables were the eleven variables categorized for the correlation analysis. The post-hoc procedure examined, when needed, was a Student-Newman-Keuls test.

Results confirm that significantly more observations were made and received at the high school than at the middle and elementary school levels. High school teachers had a mean score of observations made at 16.1, and a 13.8 mean score of observations received. Both mean scores were over twice as high as those of the middle or elementary teachers. For middle school teachers, the mean scores of observations made and received were 8.0 and 6.7, respectively. Elementary teachers' mean scores of observations made and received were both 5.3.

The results are not unexpected. Despite the availability of a substitute teacher once each month for elementary 2 + 2 teachers, significant time constraints remained. Elementary school teachers also indicated less willingness to leave their classrooms in the hands of interns or substitutes. Middle school teachers' scores were affected by severe time constraints at one school, and scheduling problems at the other.

Consistent with reports from focus groups and other feedback related to greater time constraints, elementary school teachers suggested significantly fewer observations be required for the 2 + 2 program with a mean score of 5.3. Middle and high school teachers suggested 9.8 and 9.6 observations be required, respectively. The middle school results seem to contradict the findings above, that middle school teachers conducted significantly fewer observations than high school teachers. One possible explanation is that enthusiasm for the 2 + 2 program was very high at Lake Taylor Middle School. Even at Azalea Gardens Middle School, teachers did not want to end the 2 + 2 program implementation. Therefore, the relatively high number of suggested observations may indicate a desire to find ways to successfully implement the program because it is considered worthwhile, rather than reduce the number of observations to a smaller, more

convenient, but less effective number.

Table 21

General Linear Model Analysis of Variance for 2 + 2 Outcomes

Source	df	F
Observations received	2	4.56*
Observations made	2	6.54*
Optimal number of summative year observations	2	3.44*
Optimal number of non-summative year observations	2	1.35
Positive 2 + 2 appraisal system	2	3.34*
Positive NPS appraisal system	2	3.91*
Negative 2 + 2 appraisal system	2	1.74
Negative NPS appraisal system	2	1.99
Positive 2 + 2 outcomes	2	.35
Negative 2 + 2 outcomes	2	.08
Positive 2 + 2 as mandatory program	2	2.73*

Note. *$p < .05$.

A significant difference was found in school level for the variables "positive 2 + 2 appraisal system" and "positive NPS appraisal system". These variables refer to how favorably each system is viewed as an instrument of appraisal. High school teachers were significantly more likely than elementary school teachers to consider 2 + 2 a better appraisal system. Consistent with this finding, elementary school teachers were also significantly more likely to rate the NPS appraisal system favorably.

More study is needed to explore further the differences between high school and elementary school teachers and their attitudes toward appraisal systems. Of interest are the results of the questionnaire regarding teachers' attitudes toward the NPS performance appraisal system in effect from 1983-1995, which is analyzed in the following section. Elementary school teachers were found to be significantly more likely to consider their supervisors knowledgeable about teacher performance than were high school teachers. The findings suggest that high school teachers are more likely to value peer feedback than are elementary school teachers. Implications might also include differences in attitude toward authority, with elementary school teachers more likely to accept traditional hierarchies of supervision.

Finally, teachers at the two middle schools were more likely to consider a mandatory implementation of 2 + 2 in a positive manner than were elementary school teachers. These differences among school levels cited above may be limited by two mediating factors: elementary school teachers were under represented in the survey sample, and time to conduct 2 + 2 observations was more readily available at the high school level.

Summary

The survey gave significant feedback to 2 + 2 implementers. A clear majority of teachers felt the 2 + 2 program was a better appraisal program than a traditional evaluation system. Confirmative findings about positive outcomes of the program were also indicated. Ninety-six percent of teachers implemented at least one 2 + 2 suggestion, and 85% implemented strategies they observed in colleagues' classrooms. Only 6% of teachers agreed that 2 + 2 had no impact on their professional growth.

Teachers were less than unanimous in their support of 2 + 2 as a mandatory program. A small minority agreed 2 + 2 should be mandatory for every teacher each year, but 30% agreed on a mandatory program during summative evaluation years. Collectively, teachers were ambivalent about student 2 + 2 observations. Over 40% of teachers were neutral on this question, with the remainder evenly divided for and against student 2 + 2 feedback as an important component of the program. A review of original 2 + 2 program guidelines may also be indicated by the response of teachers to the

importance of 2 + 2 observations by administrators. Only 43% of teachers agreed that administrator observations were an important component of 2 + 2. Yet, the 2 + 2 program was not conceived exclusively as a peer observation program. Rather, the program emphasizes feedback from a variety of sources to better inform and improve instruction.

The correlation analysis provided additional information about relationships between 2 + 2 outcomes, the numbers of observations teachers made and received, positive views toward 2 + 2 as an alternative appraisal system. Results showed that the more observations a teacher made and received, the more positive their responses to statements about 2 + 2 as an alternative appraisal system, and 2 + 2 outcomes. Expected relationships were achieved, including one indicating that the traditional evaluation system and the 2 + 2 appraisal system are philosophically opposed. The more a teacher valued the 2 + 2 program, it appeared, the less likely the teacher was to rate the NPS appraisal system positively.

The analysis of the survey raised an interesting question about the difference in high school and elementary school teachers' attitudes toward the appraisal process. High school teachers were significantly more likely than elementary school teachers to consider 2 + 2 a better appraisal system than the NPS appraisal system. The small number of elementary school teachers responding to the survey may have biased the results. However, it may, in part, explain the low level of 2 + 2 participation at the elementary level. This finding may indicate a need to spend more time at the elementary level encouraging teachers to experiment with 2 + 2.

The survey, which was administered late in the spring, provided significant confirmation that teachers valued the 2 + 2 process. The survey was also instrumental in gaining credibility for 2 + 2 among non-participants and administrators, primarily because it delivered quantitative results which were easy to assess.

Teacher Attitude Toward Performance Appraisal Questionnaire

The Teacher Attitudes Toward Performance Appraisal questionnaire was administered twice, first in October 1996, and again in May 1997 (see Appendix G for questionnaires). The first questionnaire (pre-2+2) asked teachers about their attitudes

toward the teacher performance appraisal system in place in the Norfolk Public Schools from 1983 through the school year 1995-96. Teachers were asked to indicate on the pre-2+2 questionnaire whether it was their summative evaluation year, and, if so, whether they were opting for the 2 + 2 program alternative to the NPS performance appraisal system. A total of 73 respondents met both criteria. In May, after a year participating in the 2 + 2 program, teachers were asked to return a second questionnaire (post-2 + 2) which solicited information about teachers' attitudes toward the 2 + 2 performance appraisal alternative. Only teachers from the four PRIME schools, who indicated they had participated in the 2 + 2 program in lieu of the NPS appraisal system, completed the post-2+2 questionnaire (N=53).

The pre- and post-2 + 2 questionnaires were analyzed separately. Although matching through the use of a four digit identifier was attempted, teachers had difficulty recalling their identifiers in May. Because only 27 matches were achieved, analysis was not based on matching. Chi-square goodness-of-fit tests on relevant demographic characteristics were performed to test sample equivalence between pre- and post-2+2 questionnaire respondents. A review of the results indicated no significant differences between the two groups.

Model

The questionnaire was selected to determine whether there was a difference in teacher attitudes toward the NPS appraisal system and the 2 + 2 appraisal program. The questionnaire was designed to uncover causal relationships among variables, with an overall aim of determining which variables contributed to the acceptance of each appraisal system. The instrument has 16 scales. The authors of the instrument designated the following scales as dependent latent traits (dependent variables): acceptance; accuracy; fairness; feedback; performance review meeting. The 11 independent variable scales were: achievement orientation; knowledge (how well the rater or observer understood teaching performance); teacher evaluation forms; performance obstacles (whether teachers believed they were held accountable for circumstances beyond their control); post-observation and summative conferences; evaluation policies; evaluation procedures; evaluation purposes; supervisory impact (action and reactions by the

to question the fairness of the process.

Males were also more likely to accept the 2 + 2 appraisal process than females. Although female teachers perceive 2 + 2 to be as fair as males perceive it to be, their acceptance of the program as an appraisal system appear may not be as great. Feedback and recognition in the 2 + 2 program is of a different kind than in a traditional evaluation system. Female teachers may need more positive strokes from administrative or other supervisory sources than currently provided by the 2 + 2 program. Were the findings to be corroborated in a more rigorous setting, an explanation for these differences in attitudes might be theorized which would allow a more differentiated approach to teacher performance appraisal by gender.

Table 22

Factorial Analysis of Variance (GLM Procedures) for Pre/Post 2 + 2 Questionnaires

Pre-2+2 Questionnaire	F		
Source	School Level[a]	Gender[b]	Interaction[b]
Fairness	1.70	4.72*	.87
Knowledge	3.61*	1.04	.55

Post-2+2 Questionnaire	F		
Source	School Level[a]	Gender[b]	Interaction[b]
Acceptance	.13	5.73*	.81

Note. [a]df = 2. [b]df = 1.

*p<.05.

General linear models

For the pre-2+2 questionnaire, sixteen separate General Linear Models (GLM) procedures were conducted in a 3 x 2 factorial analysis. School level with three conditions (elementary, middle, and high) and gender with two levels were the independent variables, and each scale variable was a dependent variable. Similarly, fifteen GLM procedures were performed for the post-2+2 questionnaire. The sixteenth variable, teacher ratings, was irrelevant to the post-2+2 questionnaire. The Student-Newman-Keuls post-hoc test was included in the procedures.

Most findings were insignificant at the $p < .05$ level. The few exceptions follow (Table 22). On the pre-2+2 questionnaire, a difference in the perception of "fairness" by gender was significant. Male teachers were more likely to perceive the former NPS appraisal system as fair than were female teachers. A significant difference was also present between school level for the "knowledge" variable. Also on the pre-2 +2 questionnaire, elementary school teachers were significantly more likely to believe that supervisors were knowledgeable about teacher performance than high school teachers. On the post-2+2 questionnaire, a significant difference was observed for the "acceptance" variable by gender. Acceptance of the 2 + 2 appraisal alternative was significantly greater for male teachers than for female teachers.

The findings that male teachers were more likely than female teachers to perceive the former NPS appraisal system to be fair, and that male teachers were more likely than female teachers to accept 2 + 2 as an alternative appraisal system need further study. A theoretical basis for these unanticipated results was not discovered in the literature, and these outcomes are most likely due to the statistical procedures used, the small sample size, and varying cell sizes. For example, only two elementary male teachers participated in the survey. However, to speculate on other reasons for the findings, the possibility exists that female teachers, raised in a culture where females are taught to please others, are more dependent on extrinsic rewards such as a positive performance appraisal, than are male teachers. Female teachers would therefore be more dependent on the results of a performance appraisal for validation. In such a scenario, female teachers would be more likely to be disappointed with less than expected appraisal results, and more likely

supervisor during the conferences); supervisory trust; summative evaluation ratings (for the pre-2+2 questionnaire only). In general, the questionnaire was designed to examine whether higher positive levels of acceptance of performance appraisal, and perceptions of fairness, accuracy, feedback, and performance review meetings are influenced by the 11 factors (independent variables) stated above.

The questionnaire was originally designed for and administered to city employees. To reflect differences in the business and education settings, some questionnaire terminology was slightly modified for this study. Modifications were all approved by one of the instrument's authors. In the pre-2+2 questionnaires, for example, the dependent variable "performance review meetings" was changed to "post-observation and summative evaluation conferences". Further modifications were necessary in the post-2+2 questionnaire. Because there were no formal post-observation or summative evaluation conferences in the 2 + 2 program, the variable was changed to "post-observation discussion".

Some questionnaire statements were edited as well. For example, in the pre-2 + 2 questionnaire, a statement in the "knowledge" scale read, "The person who conducts my teaching evaluation has a complete understanding of my teaching performance." In the post-2 + 2 questionnaire, the statement read, " The individuals who observe my teaching have a complete understanding of my teaching performance".

Correlation analysis

A correlation analysis was performed for each scale. Descriptive statistics were calculated for each scale item, and a Cronbach's Coefficient Alpha reliability rating was calculated for each scale. An alpha coefficient of .6 to .7 or greater is desirable. Only the "purpose" scale in the post-2+2 questionnaire, with an alpha coefficient of .34, was not considered reliable. A review of the questionnaire showed that teachers knew 2 + 2 was being implemented to enhance teacher growth and development. At the same time, they disagreed that the purposes of peer observation had been explained to them. A possible explanation is that the wording of the two items caused confusion, with respondents uncertain whether to equate 2 + 2 with peer observation.

The finding that elementary teachers were more likely than high school teachers to believe that supervisors were knowledgeable was consistent with a survey result. In the survey, elementary school teachers were significantly more likely to rate the NPS appraisal system favorably than high school teachers. Since the NPS appraisal system depends on the judgement of supervisors, these findings seem related. Further questions, including whether differences exist in teacher autonomy at different school levels, are beyond the scope of this study, but are significant if potential differences result in different performance appraisal system needs.

Because only three of 31 GLM procedures demonstrated significant differences, it was reasonable to treat the questionnaire samples as homogeneous with regard to gender and school level. Thus, the complete sample for each questionnaire administration was utilized in a path analysis model.

Path analysis

The path analysis model included the following steps: model specification, parameter estimation, and model evaluation. Model specification requires that causal relationships be determined by creating a theoretical model. The model depicts causal relationships by connecting variables with lines, using arrows to show directionality of influence. Also to be noted was that any scales that were negative in the questionnaires were reverse scored

A correlation analysis produced a squared multiple correlation for each causal relationship in the model (Table 23). The dependent variables were acceptance, fairness and accuracy, feedback, and post-observation conferences. In the path model, the variables "fairness" and "accuracy" were combined because they were highly correlated (> .7). The R^2 value in the table represents the total variance accounted for by contributing factors to each of the four dependent variables.

Table 23

<u>Path Equations</u>

Pre-2 + 2 Questionnaire

	R^2
Fair/Accuracy = .73 Perf. Obstacles + .39 Rating	.68
Acceptance = .85 Fair/Accuracy + .14 Anxiety	.66
Post-Conf. = .46 Knowledge + .52 Supervisor Impact	.73
Feedback = .40 Post-Conf. + .37 Supervisor Impact	.83

Post 2 + 2 Questionnaire

	R^2
Fair/Accuracy = .89 Post-Disc. + .55 Perf. Obstacles + .63 Supervisor Impact	.75
Acceptance = .91 Fair/Accuracy + .22 Knowledge + .49 Perf. Obstacles	.80
Post-Disc. = .64 Feedback + .20 Knowledge + .34 Supervisor Impact	.86
Feedback = 1.37 Achievement Orientation + .37 Supervisor Impact	.75

Figures 9 and 10 show the path models for the NPS appraisal system (pre-2+2 questionnaire), and the 2 + 2 program (post-2+2 questionnaire) respectively. The two models, despite several similarities, are clearly very different. The NPS appraisal system model (Figure 9) consists of two separate sections which do not interact. The achievement orientation variable does not exist in the model. In contrast, the 2 + 2 appraisal system model is an integrated model with the achievement orientation variable directly contributing to the need for feedback.

Pre-2+2 model. In this model, the independent variables included performance obstacles, rating, anxiety, supervisory impact, and knowledge. The first section of the model shows that acceptance by teachers of the NPS appraisal system is influenced in large measure by teachers' perceptions of its fairness/accuracy and, to a lesser extent, by anxiety. Contributing directly to the perception of fairness/accuracy of the appraisal system are the summative ratings and performance obstacles.

The second section of the model is unconnected to the first and appears to have

Pre-2+2 Questionnaire Path Model
Norfolk Public Schools Appraisal : 1983-1995/96
(N=73)

Figure 9. Questionnaire path model with unstandardized parameter estimates.

Post-2+2 Questionnaire Path Model
2+2 For Teachers: Alternative Performance Appraisal Program
(N=53)

Achievement

Supervisor Impact

Performance Obstacles

.37

1.37

.63

.55

.49

Feedback

.64

Fairness/Accuracy

.34

.91

Acceptance

Post-observation discussion

.89

.33

.20

Knowledge
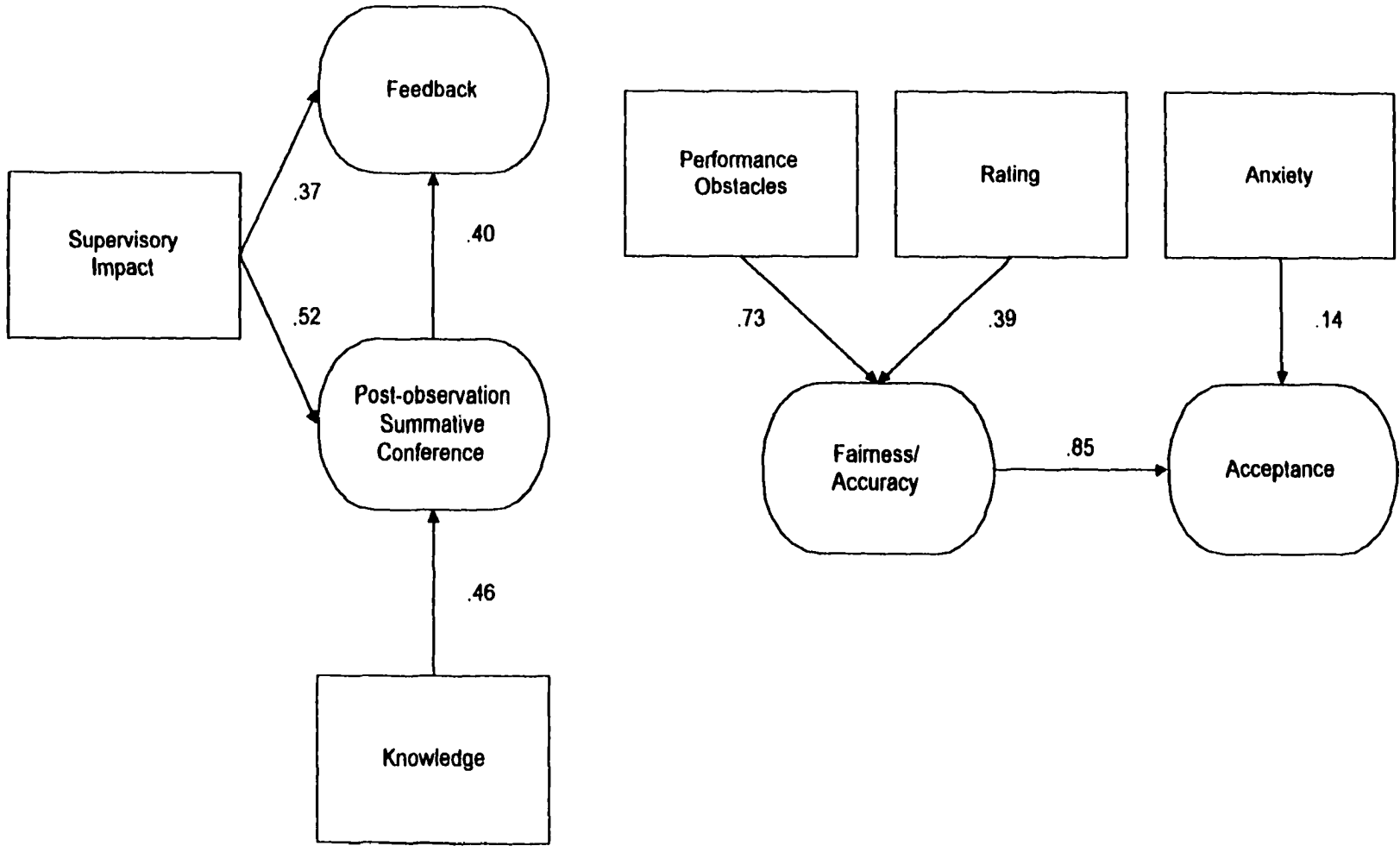
Figure 10. Questionnaire path model with unstandardized parameter estimates.
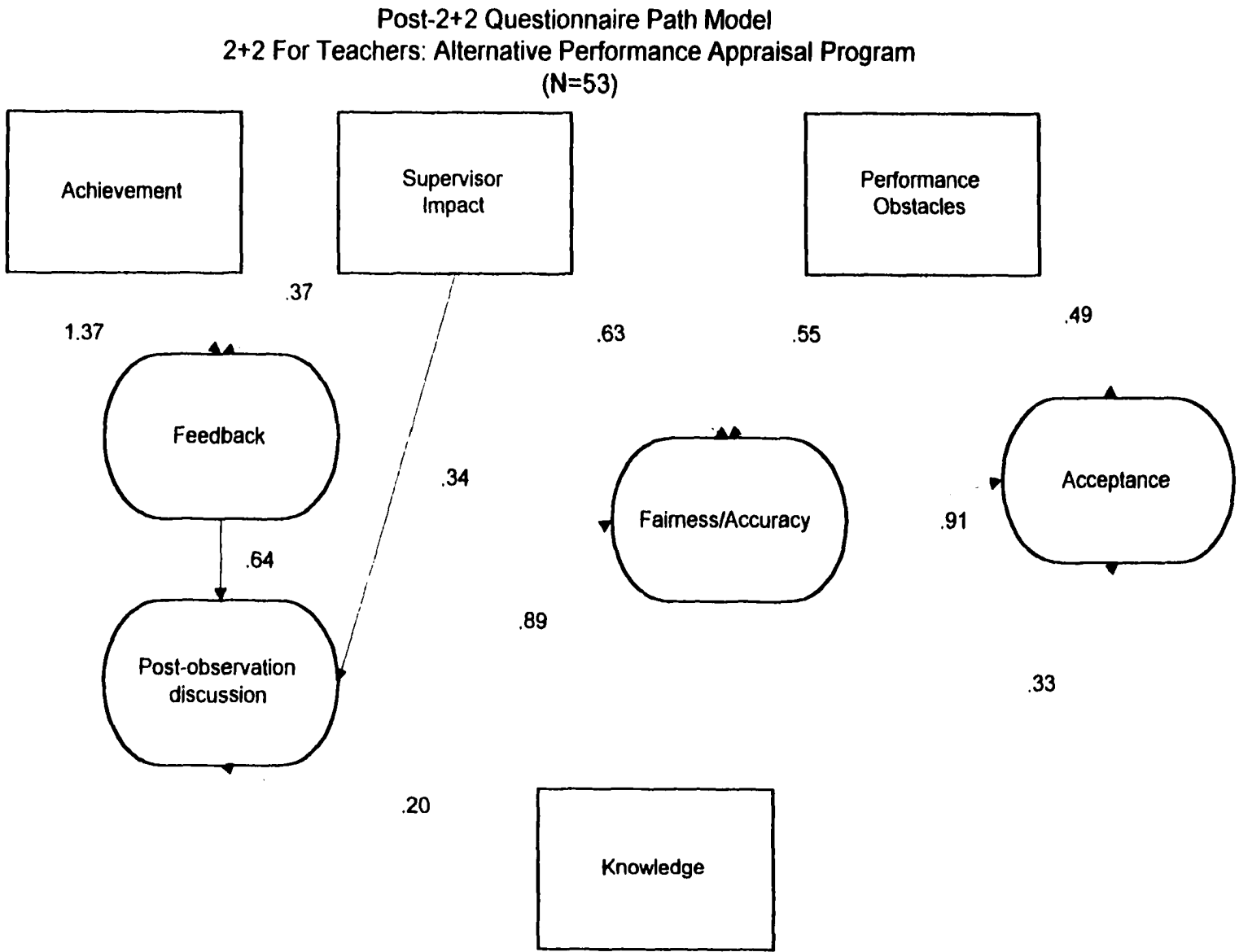
no impact on acceptance of the appraisal system. This section shows the interaction of the following constructs: feedback, supervisory impact, post-observation conference, and knowledge. Attitudes toward feedback were directly influenced by supervisory impact and the post-observation conference. Supervisory impact refers to the reactions and actions of the observer as they took part in post-observation discussion. Supervisory impact, along with the knowledge construct, also contributed to a teacher's positive perception of the post-observation conference. How teachers felt toward the feedback they received, however, did not influence any other aspect of the appraisal process. Thus, contrary to expectations, feedback and post-observation conferences appeared to be unrelated to teachers' acceptance of the NPS appraisal system and to their perceptions of its fairness/accuracy either directly or indirectly.

The complete separation of feedback, in the NPS model, from perceptions of fairness/accuracy and acceptance of the appraisal system led to questions about the objectives of the appraisal system. Ratings contributed to perceptions of fairness/accuracy, but feedback did not. If feedback were felt to be an important and valuable element of the appraisal system, it should logically have an impact on perceptions of fairness/accuracy. The question arose whether the rating was the operationalized objective of the appraisal process, instead of useful performance feedback for the improvement of instruction.

Goodness-of-fit indices supported the NPS pre-2+2 model. The chi-square statistic was statistically non-significant ($X^2 = 62.62$, $p = .28$). The root mean square residual was .043 ($< .10$ is desirable), and the comparative fit index was .99 ($>.90$ is desirable).

Post-2+2 model. In the post-2+2 questionnaire path model (Figure 10), independent variables included achievement orientation, knowledge, supervisory impact, and performance obstacles. The dependent variables were feedback, post-observation discussion, fairness/accuracy, and acceptance. In this model, achievement orientation, which did not appear in the NPS model, directly impacts feedback. The perceived knowledge of the observer, consideration of performance obstacles, and perceptions of fairness/accuracy all contributed directly to acceptance of the appraisal system.

Supervisory impact directly influenced feedback, post-observation discussion, and perceptions of fairness/accuracy. Consideration of performance obstacles also directly influenced perceptions of fairness/accuracy.

As shown in Figure 10, how teachers felt about the post-observation discussion, and indirectly about the knowledge of the observer and the feedback, helped determine perceptions of fairness/accuracy. Perceptions of fairness/accuracy were then based in part on how teachers felt about these activities. Because fairness/accuracy greatly impacted acceptance, teachers' perceptions of feedback and post-observation discussion ultimately also impacted acceptance.

Knowledge did not impact fairness/accuracy directly, perhaps because, in the 2 + 2 model, teachers were free to disregard any 2 + 2 feedback they did not consider useful. Knowledge did, however, directly influence acceptance of the 2 + 2 program, and post-observation discussion. Therefore, it appeared that the more knowledgeable an observer was perceived to be, the greater the acceptance of 2 + 2 as an alternative appraisal program. In addition, when an observer is perceived to be knowledgeable, post-observation discussion was considered more desirable.

Achievement orientation was very high among all respondents. The mean for each scale item was no lower than 4.6 on a five point Likert scale. Individuals who are high-achievement oriented have a need to perform well and to be recognized for their accomplishments. If these achievement needs are met by an appraisal process, that process will be more easily accepted. The model revealed a strong connection between achievement orientation and desire for feedback. In the 2 + 2 model, teachers' desire to achieve is tied most directly to obtaining feedback, rather than to issues of fairness/accuracy or acceptance of the 2 + 2 program itself.

Performance obstacles refer to aspects of performance that are reviewed but are not under the teacher's control. This construct appeared to be of great importance to teachers. It directly impacted fairness/accuracy in both the NPS appraisal system and the 2 + 2 appraisal system models, as well as acceptance in the 2 + 2 system model. These findings are consistent with teachers' comments in focus groups that expressed irritation at suggestions which focused on something outside of their control such as the number

of textbooks available. Teachers also had some difficulty with suggestions addressing strategies which had, in fact, been employed before or after the observer's visit. Such an occurrence was also outside of a teacher's control, in that it was not possible to be recognized for employing the strategy, and, in addition, it may be assumed by the observer that the teacher had never used it. Further study may be useful in determining the best way to minimize teachers' frustration with the review of performance obstacles in appraisal systems. When teachers are distracted by feelings of frustration and unfairness, other more valid evaluation feedback may be less acceptable to them as well.

Goodness-of-fit indices supported the post-2+2 model. Although the chi-square statistic was significant ($X^2 = 72.36$, $p = .01$), the other goodness-of-fit indicators cited showed a basic consistency supporting the fit. The root mean square residual (RMR) was .06 (<.10 is desirable), and the comparative fit (CFI) index was .93 (>.90 is desirable). The complexity of the model relative to the pre-2+2 model may have contributed to the significant chi-square statistic..

The path analyses were exploratory and, due largely to the small sample size, parameter estimates were unreliable, especially for the post-2+2 model. But the models showed an interesting preliminary view of how the two appraisal systems differ, and which variables influence acceptance of each system. The findings suggest that further research is warranted.

Summary

The questionnaire analysis provided preliminary evidence that teachers' attitudes toward the 2 + 2 appraisal process are fundamentally different than those toward the NPS appraisal system in place from 1983-1995. The NPS appraisal system path model showed two disconnected sections, where teachers' perceptions of performance feedback were unrelated to perceptions of fairness/accuracy of the system, or to its acceptance. The 2 + 2 appraisal system model was an integrated whole, where positive attitudes toward feedback were directly influenced by achievement orientation. Achievement orientation did not appear in the NPS appraisal system model.

The achievement orientation of teachers was very high in both questionnaires. Its absence in the NPS appraisal model is significant in that high-achievement oriented

individuals have a need to perform well, and expect recognition for their accomplishments. An appraisal process is typically the arena where recognition and evaluation of an individual's efforts is carried out, providing affirmation for high-achievement oriented teachers. Theoretically, the validation of their performance would lead to perceptions of fairness/accuracy and to acceptance of the appraisal system. The absence of the achievement orientation variable in the model may signify that the recognition needs of teachers have not been met by the NPS appraisal system.

In the 2 + 2 appraisal system model, achievement orientation does not contribute to fairness/accuracy or to acceptance. However, it does highly impact feedback. This is consistent with the view that high-achievement oriented individuals have a need for recognition, and want their teaching performance to be examined. Because achievement orientation did not contribute to fairness/accuracy or acceptance of the 2 + 2 system, it might be concluded that in the 2 + 2 model, perceptions of fairness/accuracy, and even acceptability of the appraisal system, were secondary to a need and desire for performance feedback.

In the 2 + 2 model, feedback impacts fairness/accuracy secondarily via the post-observation discussion. Because of the position of the post-observation discussion in the model, further study of 2 + 2 should be undertaken to ascertain in more detail how much interaction among teachers is taking place as a direct result of 2 + 2 observations. This would help clarify its importance to the 2 + 2 appraisal process.

Some similarities existed between the two models. For example, the section of the NPS model which addressed feedback showed similarities with the 2 + 2 model. In both cases, supervisory impact contributed to feedback and post-observation discussion. Knowledge also contributed to post-observation discussion in both models. However, whereas in the NPS model, the post-observation discussion impacted how teachers felt about the feedback, the reverse was true of the 2 + 2 model. Teachers' feelings about the feedback contributed to a desire for post-observation discussion. The difference is interesting because teachers may be more independent in evaluating feedback under the 2 + 2 model.

Other similarities included performance obstacles as a contributing factor to

perceptions of fairness/accuracy in both models. This corroborates teachers' feedback in focus group meetings that suggestions referring to situations outside of their control were met with frustration and irritation. The implications for appraisal systems merit further study, including examination of the range of issues considered by teachers to be not under their control.

Overall, the questionnaire provided strong preliminary evidence that the 2 + 2 appraisal system was perceived by teachers much as it was intended to be. The strong influence of the achievement orientation construct on feedback lends support to the assumption that 2 + 2 serves the needs of teachers by providing opportunities for recognition, as well as for improvement through both the suggestion component and post-observation discussions. The questionnaire enlightened the evaluation of the 2 + 2 program by highlighting constructs teachers consider important in the acceptance of the system. The observer's knowledge level, supervisory impact, performance obstacles, and post-observation discussion in particular are key areas of anticipated discussion within the 2 + 2 program.

Limitations. The objective in analyzing the questionnaire was to obtain a model that would help to describe the 2 + 2 program in terms of an alternative to traditional teacher appraisal programs. The relatively small number of responses made parameter estimates unreliable, and precluded a more rigorous structural equation modeling. In addition, a biased model of the pre-2+2 questionnaire, which referred to the NPS appraisal system in effect from 1983-1995, may have resulted from the respondent sample, all of whom were planning to opt for the 2 + 2 program. As shown in the interview data above, however, there was no evidence that a disproportionate number of teachers had negative views toward the NPS appraisal system. The study provided an interesting exploratory model, and opens possibilities for future research into the area of teacher acceptance of performance appraisal systems. Ideally, teachers would complete the questionnaire again in a year or two to confirm current findings.

In further study already in progress, the model will be compared with a larger sample of non-2 + 2 teachers from eleven PRIME and non-PRIME schools who also completed the pre-2+2 questionnaire. A similar sample of teachers also completed a

questionnaire asking about attitudes toward the new NPS teacher appraisal system initiated during the 1996-97 school year. The results of the larger study will aid in better understanding how the 2 + 2 program is perceived in relation to traditional teacher performance appraisal programs.

Summary

In this chapter, data collected during the implementation of the 2 + 2 for Teachers: Alternative Performance Appraisal Program was analyzed. Issues which arose as the implementation proceeded were discussed utilizing a case study approach. Preliminary outcomes of the 2 + 2 program implementation based on interviews, focus group meetings, 2 + 2 observation forms, self-reflection reports, a survey, a questionnaire were also analyzed and discussed. In Chapter V, the implications of the data analysis will be further explored.

CHAPTER V

CONCLUSIONS AND IMPLICATIONS

Introduction

The focus of the current evaluation study of the 2 + 2 for Teachers: Alternative Performance Appraisal program implementation centered on three questions:

1. How was the 2 + 2 program implemented?

2. What difference did the implementation of the 2 + 2 program make?

3. What were teachers' perceptions of the benefits and drawbacks of the program?

The current evaluation focused on the time frame of August 1996 to June 1997. Additionally, historical information from the program's inception in 1995 was presented and discussed. In Chapter IV, data which had been collected during the 1996-97 implementation year were analyzed, and the implementation process, including the program's history, was described. Data triangulation was achieved through varied data sources including interviews, focus group meetings, 2 + 2 observations forms, self-reflection reports, a survey, and questionnaires. In this chapter, implications and conclusions drawn from the data analysis will be examined, and recommendations for future 2 + 2 implementation offered.

Evidence documented and analyzed in Chapter IV points to several conclusions:

1. The 2 + 2 program made a positive difference for teachers.

2. The 2 + 2 program was a viable and attractive alternative performance appraisal program for teachers.

3. The 2 + 2 program received limited administrative support.

4. Systemic issues will be important to consider in future implementations.

The 2 + 2 program implementation achieved virtually unanimous acceptance among participating teachers. Despite minor implementation concerns and minimal active administrative support, teachers' enthusiasm and numbers of 2 + 2 observations grew over the course of the school year. Evaluation data indicated that 2 + 2 made a significant difference in many positive ways. Teachers felt overwhelmingly that 2 + 2 helped them share expertise, overcome isolation and expand their organizational

perspective, and that it led to professional growth. Teachers preferred participation in the 2 + 2 program to evaluation under the traditional Norfolk Public School (NPS) appraisal system. Most teachers also thought the 2 + 2 appraisal program was superior to the NPS appraisal system.

To adequately understand the implications of the 2 + 2 program implementation, it must be discussed in the context of the PRIME project, a systemic reform effort. The 2 + 2 program was implemented, not in isolation, but in a complex environment of educational change. Finally, suggestions for the success of future 2 + 2 program implementations are integrated into the chapter throughout. Recommendations proposed by the PRIME Steering Committee 2 + 2 sub-committee and supported by the researcher are also described near the close of the chapter.

Outcomes Discussion: The 2 + 2 Program Made a Positive Difference for Teachers

Some of the major ways that 2 + 2 made a difference are discussed below. The outcomes discussion is organized around major objectives of the 2 + 2 program, including improvement of instructional practice, alleviation of teacher isolation, reduction of teacher anxiety while being observed, and implementation of an alternative appraisal system. The discussion of each outcome is prefaced by a brief synopsis of results.

Improvement of Instructional Practice

Synopsis

As stated in the evaluability assessment, improvement of instruction is one major objective of the 2 + 2 program, important because, as hypothesized in the 2 + 2 model, it leads to increased student achievement. Neither improvement of instruction nor increased student achievement could be evaluated here in a quantifiable way. However, teachers cited improvement of instruction many times as evidence of the value of 2 + 2 and were overwhelmingly positive about the usefulness of 2 + 2 feedback in relation to improvement of instruction. Teachers did perceive that gaining multiple perspectives on instruction led to improvement in their own teaching. Teachers not only valued their

colleagues' suggestions, but also perceived great benefit from seeing other teachers model different instructional strategies.

Discussion

Teachers' expectations about the 2 + 2 program focused on improvement of instruction from the outset. In interviews conducted in the fall of 1996, a total of 80% of teachers cited improvement of instruction, ways to grow professionally, or peers helping in ways administrators cannot, as reasons for participating in 2+ 2. Nearly 90% of teachers' responses to the question, "what are your expectations of the 2 + 2 program?", indicated that new ideas to improve teaching, and honest and mutual feedback were primary expectations of 2 + 2. Only 12% mentioned 2 + 2 in the context of an alternative evaluation system, and some of these responses focused on the expectation that 2 + 2 would be a "more helpful" alternative, that is, would provide better feedback for growth. Asked if they expected to develop new instructional strategies as a result of 2 + 2, 91% indicated "yes", "hopefully", or "probably". Clearly, teachers first and foremost wanted and expected to learn and grow as professionals.

Did 2+2 fulfill these optimistic expectations? The self-reflection reports written by high school teachers at the end of the first semester, the focus groups, and the survey all indicated a high value was placed on the 2 + 2 experience. Twenty-three of the 31 (74%) high school teachers who submitted a self-reflection report mentioned receiving specific suggestions to change instructional practice, while 78% reported they had received insights and perspectives from observing other teachers. Many specifically mentioned receiving "fresh ideas", and adapting what they observed for their own classrooms. Several teachers also mentioned the value of 2 + 2 as a means of "teachers making teachers better".

In a recent study of ten high schools engaged in restructuring efforts, "teachers commonly report[ed] that they [had] received the most valuable 'professional development' by working in teams with their colleagues and thus having the chance to regularly observe strategies used by their fellow teachers" (Cawelti, 1997, p. 2). 2 + 2 participants strongly corroborated these findings. Likewise, the first three major themes of the self reflection reports all pertain to professional growth. Teachers commented that

observing other teachers provided insight and perspective, and being observed provided specific suggestions to improve instructional practice. Teachers were also encouraged by praise and compliments. These positive comments indicate the value teachers placed on mutual feedback.

As indicated in Chapter IV, the high school focus groups were enthusiastically positive about the value of 2 + 2 as a source of learning and feedback. Many responses lauded the program for its potential to "help teachers grow" and lamented that it was not in existence at the beginning of their careers.

The survey data also show strong affirmation that 2 + 2 was instrumental in helping teachers experiment with different instructional strategies. Ninety-six percent of respondents reported having implemented at least one 2 + 2 suggestion, while 85% reported implementing strategies observed in other classrooms. Nearly 90% of those surveyed disagreed that 2 + 2 does not lead to professional growth.

Thus, the data triangulation about 2 + 2 is highly suggestive that the 2 + 2 program met, if not exceeded, teachers' expectations concerning improvement of instruction and professional growth. Teachers perceived the 2 + 2 program as a vehicle for improving instruction. The sample of interviewees and the sample who submitted the self-reflection reports and attended focus group meetings overlapped but were not identical. However, responses from both groups were similar and consistent.

Further study is needed to determine whether experimentation with new classroom strategies, or the integration of new ideas as a result of 2 + 2 observations, leads to improvement of instruction.

Teacher Isolation

Synopsis

As indicated by the evaluation data, the 2 + 2 program provided a powerful and effective mechanism to alleviate teacher isolation. The reduction of teacher isolation, one of the stated objectives of the 2 + 2 program, is critical to teacher learning and collaboration. As teacher isolation dissipates, more and different kinds of dialogue among teachers can emerge and lead to increased teacher collaboration. Teachers' sense of certainty about their own professional competence, and about teaching norms in

general grows. An unprecedented number of teacher observations took place during the fall semester, nearly 400 in the high school alone. By the year's end, most teachers (72%) agreed that the quality of their interaction with colleagues improved as a result of 2 + 2.

Positive feedback in the form of praise and encouragement from other teachers was especially valued by 2 + 2 teachers. During the 2 + 2 implementation, discussion often focused on the suggestions component of the 2 + 2 program in addressing teachers' concerns. In addition, rationale for 2 + 2 as an appraisal alternative tended to hinge on the ability of teachers to offer suggestions. But the power of positive feedback to motivate and encourage change was underscored by teachers' reactions. Indeed, an absence of positive feedback can seriously impact a teacher's sense of efficacy, reducing the likelihood of growth in student achievement (Ashton and Webb, 1986; Rosenholtz, 1989).

Discussion

In the interviews, most teachers (81%) said they discussed instructional strategies with colleagues at least several times per week, with nearly 60% saying they discussed them daily, or almost daily. Yet 81% of teachers also wished for more opportunity to interact with colleagues professionally. Perhaps the reason for this was that, as teachers indicated, much existing interaction took place within departments or clusters only. The isolation among teachers was aptly demonstrated in each of the high school focus group sessions, where introductions among teachers were necessary.

There is a need to differentiate between departmental collegiality and collaboration on a school-wide or interdepartmental, interdisciplinary basis. While being part of a reference group is positive and promotes interdependence among teachers, collaboration restricted only to departments can make school wide improvement difficult. Departmental decisions often impact teachers greatly, and can lead to competition among departments, creating a win-lose situation for resources (Johnson, 1990). Teachers themselves were cognizant of the need to expand their circle of referents. Some of the most positive comments made by teachers have been about "getting off my floor" or "getting out of the math and science department". Several teachers spoke

enthusiastically about visiting other school levels.

Did 2 + 2 alleviate teacher isolation? Nearly 400 2 + 2 observations took place at the high school from October 1996 through January 1997. Almost 100 times per month, on average, teachers visited a colleague's classroom, an unprecedented occurrence. In addition to classroom visitations, collegial interaction among teachers increased. As several teachers pointed out in the focus group meetings, conversations or discussions about 2 + 2 observations were a natural outgrowth of the program. In the survey, 72% of the respondents agreed or strongly agreed that the quality of interaction among their colleagues had improved as a result of 2 + 2.

Over 50% of high school teachers submitting a self-reflection report included comments that lauded a sense of decreased isolation and greater sharing of teacher expertise. Over 60% of the teachers indicated that they valued praise and encouragement from other teachers. The need for positive feedback is vividly demonstrated here. Although any type of feedback decreases feelings of isolation, frequent positive feedback is expected to have a salutary effect on teachers' sense of certainty (efficacy) as well (Rosenholtz, 1989).

Dearth of feedback about teacher performance reinforces a teachers' sense of isolation, and is a source of teachers' uncertainty about their own performance and their capacity to help students. Teachers who are uncertain are less likely to risk asking for help, or offering help, for fear of revealing shortcomings or being rebuffed. The more positive feedback teachers receive, the more confident they will be in their own abilities, and in their ability to influence outcomes for students (Rosenholtz, 1989). Thus, the praise and encouragement cited by 2 + 2 teachers, along with the opportunity for teachers to see colleagues struggling with similar problems, are as important an outcome of 2 + 2 as specific suggestions to improve instruction.

A path model developed by researcher Susan Rosenholtz (1989) shows a direct contribution of teacher collaboration, the antithesis of teacher isolation, to greater teacher certainty, or efficacy. At the same time, teacher collaboration influences teachers' perceptions of positive feedback, which also directly affects teacher efficacy. In Rosenholtz's full structural model, "the data indicate a reciprocal relationship between

teacher collaboration and their certainty" (p. 114). In other words, even as collaboration influences teacher efficacy, teachers with greater efficacy tended to engage in more collaborative efforts, in a kind of mutually reinforcing cycle. Teachers who participated in the 2 + 2 program experienced a reduction of uncertainty about their teaching, and about their colleagues' teaching. If Rosenholtz's model is reflective of relationships among 2 + 2 teachers, 2 + 2 teachers will be more likely to collaborate with colleagues, and further increase their sense of certainty.

In their teacher efficacy research, Ashton and Webb (1986) state that "our study of teachers' sense of efficacy leads us to conclude that *the central social-psychological problem facing teachers today is how they can maintain a sense of satisfaction and accomplishment in a profession that offers so few supports for, and myriad threats to, their sense of professional self-respect*" (p. 66). The power of positive 2 + 2 feedback, as well as the satisfaction achieved by collaborative efforts to learn and grow through suggestions and observations, is likely a major reason for the unparalled success of the 2 + 2 implementation. More importantly, if greater satisfaction can lead to greater teacher efficacy, increased student achievement may be all the more likely.

Reduction in Anxiety

Synopsis

Anxiety about being observed is exacerbated by teacher isolation and the infrequency of summative evaluation observations. Isolation can lead to greater uncertainty, and hence anxiety, but anxiety can reinforce isolation as well. Therefore, a reduction in anxiety is an important objective of the 2 + 2 program.

Data in Chapter IV, in particular the questionnaire, indicated that 2 + 2 teachers experienced less observation anxiety under the 2 + 2 appraisal system than under the NPS appraisal system. Another indicator was the gradual diffusion of the 2 + 2 program to non-participating teachers who began to participate in an informal way.

Discussion

Teachers typically experienced some anxiety over the performance appraisal process. Because of the culture of teacher isolation and infrequency of appraisal, an observation from an administrator is an unusual and potentially threatening occurrence.

Contributing to the discomfort is the nature of the traditional appraisal system which is perceived less as a helpful feedback process and more as a mechanism to rate performance. The interview data showed that some 2 + 2 participants (29%) were also anxious about peer observations, though nearly 70% of those interviewed reported feeling very comfortable about the process. The data in Chapter IV suggested that teachers experienced less anxiety under the 2 + 2 appraisal system than under the NPS appraisal system.

As reported in Chapter IV, the questionnaire path model indicated that anxiety contributed in a minimal way to the acceptance of the traditional NPS appraisal system, but did not appear in the 2 + 2 appraisal system model. Perhaps teachers felt that in the NPS appraisal system they were under pressure to perform and, as is true in many types of performance situations, a degree of anxiety is beneficial to success. In any event, levels of anxiety were lower in the 2 + 2 appraisal model. Means on the anxiety scale were 20.8 for the NPS appraisal system compared with 14.6 for the 2 + 2 model. The maximum score recorded for anxiety was 39 out of a possible 40 in the NPS appraisal system questionnaire, and 25 of a possible 40, in the 2 + 2 questionnaire.

Reduction of anxiety is important because it can be a barrier to interaction with colleagues (see Teacher Isolation above). Anxiety about one's own teaching performance may inhibit a teacher from asking for suggestions or collaborating with other teachers. Evidence that 2 + 2 worked to quell anxious feelings about opening classrooms to observers went beyond questionnaire data. Teachers reported approaching non-2 + 2 teachers and asking if they could observe their classrooms. For example, one teacher wanted to see a veteran teacher who had a reputation of being good at teaching writing. The veteran teacher was reluctant, but the 2 + 2 teacher persisted, saying she had seen her bulletin boards and heard good reports and would just observe without writing compliments and suggestions. The teacher relented, and was later invited to observe the 2 + 2 teacher's class. Since that time, the veteran teacher has asked for feedback, and has also taken over the 2 + 2 teacher's class at different times to demonstrate strategies.

Examples such as this one illustrate how a process of diffusion of the 2 + 2

program aided the reduction of anxiety in three of the four participating schools. As mentioned earlier, the number of teachers participating in 2 + 2 at the high school eventually reached 90 teachers and interns, largely through a process of informal requests by 2 + 2 teachers to observe and be observed. At the elementary school, teachers in the first focus group were positive, but somewhat uncomfortable about observers in their classrooms. In the second focus group session near the end of the school year, 2 + 2 teachers enthusiastically brainstormed ways to involve a greater number of faculty in the 2 + 2 program in 1997-98. During the second semester, six teachers joined the program informally, in addition to seven teachers already participating in lieu of the NPS summative evaluation.

Anxiety reveals a basic discomfort with or distrust of a situation. When anxiety is diminished, it follows that trust will grow. The data indicated that teachers experience of the 2 + 2 process was non-threatening and positive, even when it presented a challenge. Several teachers mentioned that 2 + 2 kept them "on their toes" because they never knew when someone might come in and observe them. The comment was positive in that the teachers felt it helped them focus and improve their teaching performance. It also showed a basic respect for the opinions of their colleagues, and trust that their feedback is not only valuable but well-intentioned. This is an encouraging sign that distrust of collegial feedback is dissipating.

Alternative Performance Appraisal

Synopsis

One of the objectives of the 2 + 2 program implementation was to provide a viable alternative to the traditional NPS appraisal system. Teachers have verified that this objective has been met. In survey responses, and in focus groups, teachers overwhelmingly preferred the 2 + 2 program as an evaluation system, and found it a superior vehicle for professional growth as well. Analysis of the questionnaire responses showed different path models of teacher attitudes toward each appraisal system.

It was not the purpose of the 2 + 2 program evaluation to conduct a comparative study of the two systems. However, inasmuch as 2 + 2 is a program which is an alternative to the traditional NPS appraisal system, data was collected on how teachers

reacted to 2 + 2 as an alternative performance appraisal system. Two aspects of teacher evaluation were addressed: 1) the performance evaluation (assessment) aspect, and 2) the evaluation as professional growth aspect.

In the interviews at the beginning of the school year, teachers were asked to voice their opinions about the traditional system. Likewise, the survey near the end of the year contained statements about the relative merits of each system. The survey revealed that, in fact, teachers overwhelmingly preferred the 2 + 2 program to the traditional NPS evaluation system. The interviewees and survey respondents represented a cross section of 2 + 2 participants, including those not participating in lieu of the NPS appraisal system.

Discussion

Interviews conducted early in the school year revealed that many teachers were less than positive, or ambivalent about the NPS system of evaluation. Teachers' comments often centered on the lack of "any constructive feedback", or the fact that the NPS appraisal system "rated teachers, but didn't improve performance". Even many positive comments had qualifiers such as "It was fine, but I like this [2 + 2] better". Only 25% of responses were classified as positive. Just 40% of respondents offered negative comments (the remainder were neutral), indicating that most teachers were not participating in 2 + 2 because they were disgruntled with the traditional NPS appraisal system. Comments did indicate that teachers were participating in the 2 + 2 program to grow professionally and improve instruction.

Comparative teacher attitudes. Near the end of the school year, survey results in May 1997 (Table 7) showed strong teacher support for the 2 + 2 program as an alternative to the traditional appraisal system. The majority (67%) of respondents agreed or strongly agreed that "2 + 2 is a better appraisal alternative than the traditional teacher evaluation system". An additional 13% of teachers (for a total of 80% of teachers) preferred the 2 + 2 program to the traditional appraisal system. Only 4% agreed that "the traditional evaluation system is the best system for evaluating teacher performance".

Teachers were also asked about the 2 + 2 program and professional growth. In every forum, including focus groups, survey, and self-reflection reports, teachers were in

agreement that 2 + 2 led to professional growth. Of more than 70 survey responses, for example, only one teacher agreed that "the 2 + 2 system does not lead to professional growth". In the self-reflection reports, teachers focused heavily on professional growth issues, consistent with expectations voiced in the interviews that 2 + 2 would help them achieve new instructional perspectives and improve professionally. Only one comment in the 2 + 2 self-reflection reports concerned the evaluative aspect of the program. One teacher wrote, "The 2 + 2 process is an excellent evaluative tool for assessing teacher performance.

Teachers were more enthusiastic about the 2 + 2 program as a source of professional development than the traditional Norfolk Public School appraisal system. An excerpt of the survey results is found in Table 24. The strongly agree/agree responses have been combined, as have the strongly disagree/ disagree responses, so that results are condensed into three columns: disagree, neutral, and agree.

Table 24

Excerpted Survey Results Pertaining to Professional Growth Statements

Survey Statements	% Disagree	% Neutral	% Agree
(7.) Improvement of instruction is more likely to occur as a result of 2 + 2 than the traditional system.	3.9	23.7	72.4
(8.) The 2 + 2 system does not lead to professional growth.	89.5	9.2	1.3
(9.) The traditional Norfolk Public School evaluation system does not lead to professional growth.	46	39.5	14.5

Clearly teachers are ambivalent about the NPS evaluation system as a source of professional growth. In a direct comparison statement (7.) above, a large majority (72%) believe that improvement of instruction is more likely to occur as a result of 2 + 2 than the traditional NPS evaluation system.

Feedback and appraisal systems. The data from the interviews and focus groups show that teachers seldom felt they received helpful feedback to improve their teaching from the traditional NPS appraisal system. Even those rated as excellent teachers commented that, although the recognition was nice, meaningful feedback would have been welcome.

Perhaps most dramatically, the questionnaire path analysis results depict a major contrast in teacher attitudes toward the 2 + 2 program, and the traditional NPS appraisal system (Figures 9 and 10). The NPS model is comprised of two disconnected sections. Teachers' acceptance of the NPS evaluation system is based largely on how fair/accurate the system is perceived to be. This perception appears to be influenced by two factors: 1) the performance rating, and 2) whether performance obstacles were considered in the evaluation. Feedback does not factor into a teacher's acceptance of the system, and is part of the separate section of the model. One could conclude that feedback plays a peripheral role in the acceptance of the NPS appraisal model, lending support to teachers' comments that they receive only negligible amounts of feedback anyway. Teachers' achievement orientation does not factor into the model at all.

In the 2 + 2 model, achievement orientation is tied to teachers' desire for feedback. The more teachers wish to perform at their best or be recognized for a job well done, the more they are likely to want 2 + 2 feedback. Again, the acceptance of the 2 + 2 appraisal system was directly influenced by perceptions of fairness/accuracy. Here, however, the fairness/accuracy perceptions were influenced directly by post-observation discussion, and indirectly by the feedback and the perceived knowledge of the observer. Fairness/accuracy was also influenced by whether performance obstacles were considered in the observation, and supervisory impact, that is, the quality of the interaction with the observer in post-observation discussion.

Significant is that teachers' desire to perform at their best (achievement orientation) is directly tied to a need for feedback. Teachers who responded to the first questionnaire on attitudes toward the NPS appraisal system, were no less achievement oriented than those responding to the second questionnaire on attitudes toward 2 + 2. Indeed, many of them were the same teachers. But in these preliminary findings at least,

the traditional NPS system of evaluation was not found to support or promote teachers' achievement orientation in any way.

Based on the questionnaire results, one recommendation for the 2 + 2 program is to encourage post-observation discussions, whether they occur informally in the hall, or by appointment after school. In fact, one modification of the program in the second semester of the 1996-97 school year was to enable teachers to document up to five post-observation discussions and count them toward the required number of observations. However, more could be done to encourage post-observation discussion as part of the 2 + 2 culture.

Appraisal system purpose. The contrast between the 2 + 2 approach to teacher performance appraisal and the NPS evaluation system seems to be characterized by divergent purposes. The 2 + 2 approach is based on frequent, mutual feedback, with the objective of supporting professional growth and improvement of instruction. It is also a system of appraisal, where colleagues' observations, and a teacher's own reflection on those observations, are a means of assessing performance. Central to the NPS appraisal system, and most traditional evaluation systems, is the concept of assessment, where assessment connotes a rating continuum. Most evaluation systems have the written objective of improving teacher performance as well. But the rating, and the judgment whether a teacher is deemed competent or better, remains the major, if not exclusive, focus of most typical evaluation systems. In other words, the difference in approaches is a normative vs. a formative perspective.

Interestingly, while only 4% of teachers agreed that the traditional evaluation system was best, 28% agreed or strongly agreed that "I like to have my teaching performance ranked by an administrator". Only 12% disagreed or strongly disagreed with this statement, with 60% remaining neutral. Why do some teachers who do not believe in the evaluation system still like to have their performance ranked? One possible explanation is that, in a type of "grade inflation", teachers typically receive above average ratings on their evaluations. The ranking provides teachers with rare positive feedback on their teaching performance. Understandably, if a rating is the only visible validation of a teacher's performance, a teacher may wish to be ranked by an

administrator in order to fulfill a need for positive feedback, even if he/she dislikes the whole process.

The enduring attraction of a rating process is one indication of the general dearth of performance feedback in a typical school. At the same time, it distracts from the ostensible purpose of evaluation, the improvement of teaching. Once every third or fourth year, a teacher has three administrator visits, three post-observation conferences, and a summative conference. Yet much of the value of feedback lies in receiving it regularly, if not continuously. Timing, content, frequency, and follow-up are all important. Under the traditional appraisal system, years are spent teaching with no feedback from adult referents at all. No feedback is generally worse than negative feedback at all.

Teachers are not opposed to evaluation, but are not happy with inappropriate evaluation. Disenfranchisement occurs when teachers accurately perceive the traditional evaluation process to be geared toward identifying incompetent staff, even when their own teaching performance is rated as satisfactory or better. Administrators do have important roles in identifying marginal teachers, organizing professional assistance (possibly from colleagues), and, if necessary, arranging plans of action. They also have important contributions to make when offering feedback to competent teachers. But when teachers are to be rated in a summative assessment, administrators are expected to be experts in every subject area, experts in pedagogy, and sensitive to all performance obstacles. They cannot, however, be all things to all people, and are set up to fail when their main purpose, performing a task perceived as disagreeable by many administrators and teachers alike, becomes to give a rating once every several years.

Reframing performance appraisal. Is it possible for 2 + 2 to serve both the purpose of assessment and that of continuous professional improvement through frequent feedback? For this to occur to the satisfaction of teachers and administrators alike, the system/concept of evaluation must be reframed. First, assessment and professional development cannot be separated. Any type of professional growth involves critical judgment of some aspect(s) of the act of teaching. At the same time, assessment that neglects clear, specific, and ongoing feedback on both positive accomplishments and

areas for improvement is relatively meaningless. Summative evaluations themselves are not at issue here. They can serve a useful purpose and contribute to professional development. But all too often, summative evaluation, in practice, emphasizes ratings and offers little feedback to improve teaching practice.

Second, the goal of evaluation for teachers already in the acceptable range should be to improve teaching. In traditional evaluation systems, ratings, not feedback become the main objective. Furthermore, years pass between evaluations. In the environment of summary assessments every few years, teachers are understandably cautious. Often situations which further limit opportunities for learning may be created. For example, where teachers set goals as part of the appraisal process, they are likely to identify an area of strength rather than weakness to minimize risk. Yet growth depends on risk taking behavior, trying new approaches, and experimentation, behavior not encouraged by the traditional appraisal approach. But what is the purpose of assessment where it does not lead to growth or improvement?

Third, as evidence from the 2 + 2 program implementation indicates, teachers get powerful feedback about teaching performance by acting as observers as well as by being observed. The experience of observing another professional model their own strategies is rare in the teaching world, but enlarges perspectives. At least one 2 + 2 teacher spoke of the 2 + 2 program literally in terms of "tearing down walls". At the same time, teachers must engage in critical judgment when writing compliments and suggestions, an exercise that can lead to greater awareness of their own strengths and weaknesses, that is, self-assessment.

Complex issues are involved in changing an evaluation system. If administrators and teachers do not frame the concept of evaluation differently than in the past, responses to the new model can undermine its effectiveness. Administrators may perceive participation in the new model as relinquishing control of a process that has been traditionally firmly rooted in their domain. Both teachers and administrators may consider a departure from the old system something other than "real" evaluation. When operating in an environment of isolation, trust may be difficult to muster for some teachers. Trust in a traditionally hierarchical environment may not come easily for some

administrators. These and other concerns will be discussed further, below.

Other Outcomes

Student 2 + 2s. Teachers had mixed reactions to student 2 + 2s. Although the interview data showed that many teachers seemed willing to try soliciting student 2 + 2s, little interest was generated in the focus group meetings. An exception was at Lake Taylor Middle School (LTMS) where teachers enthusiastically reported soliciting student feedback. Teachers there recommended framing the feedback not as a teacher evaluation, but as feedback about a particular lesson itself. Teachers recommended implementing some of the students' suggestions, after discovering this positively impacted students' levels of trust. Elementary school teachers reported occasionally asking their students what they liked or didn't like about a particular lesson.

In contrast, teachers at Azalea Garden Middle School (AGMS) were opposed to asking for student feedback, even when the strategy employed at the other middle school was suggested. Teachers nearly became hostile when the idea was pursued by the researcher. One reason given by the teachers was the "type of student" at AGMS, which has a large percentage of minority and at-risk students from housing projects in the city. This line of reasoning is in essence no different from high school teachers who thought they might ask only their more advanced or more mature students for feedback.

Fear of negative or inappropriate feedback appeared to override the benefits of knowing how students perceived an instructional strategy, project, or class. Yet, as in LTMS, feedback empowers both students and teachers. When 2 + 2 student feedback is routinely collected and consulted, it may prove both an indicator of and contributor to greater teacher efficacy. Further study is needed in the area of student 2 + 2 observations to help determine how all types of students can offer meaningful feedback, and how the opportunity to give feedback can help students feel empowered.

Participation. In the high school, the majority of teachers participating formally in the 2 + 2 program, that is, in lieu of the NPS summative evaluation, were in their summative evaluation year. Those participating formally were, however, a minority of the total staff. The distinction between formal and informal participation is important because formal participation required adherence to guidelines regarding numbers of

observations and the filing of the self-reflective report. Informal participation could mean opening the classroom to observers only, or the completion of just a few observations. In the remaining three schools, 2 + 2 teachers comprised quite a small minority. Who were the 2 + 2 teachers?

Susan M. Johnson, in her book, Teachers at Work (1990), identifies five factors which permitted and promoted greater interaction and interdependence among teachers. Those factors are good teachers, supportive organizational norms, reference groups, sufficient time, and administrators who provided encouragement and accommodation. What are "good teachers"? Good or outstanding teachers were defined by other teachers in the study as being "committed, generous, open to change, eager to learn, and able to see beyond private successes and failures" (p. 167).

The 2 + 2 program encourages greater interaction and interdependence among teachers. Therefore, these five factors are likely to be among those which promote involvement in the 2 + 2 program as well. The 2 + 2 program might look for ways to help other teachers become more open and less defensive, to be "good teachers". Related to this factor are organizational norms which must be developed to support a culture of openness and trust. The other factors are possibly even more important, especially time to conduct observations, and administrative support, because they are likely to influence organizational norms and provide encouragement for teachers to become "eager to learn". These factors will be examined in the following discussion.

Process Implementation Discussion

The promising outcomes of the 2 + 2 program become even more remarkable when the many pitfalls of the implementation are considered. Two issues in particular will be considered here: 1) administrative support and leadership, and 2) time.

Building Level Leadership

In his book, Utilization-Focused Evaluation, Patton (1997) references George Odiorne, saying he " dissected 'the anatomy of poor performance' in managing change and found gargantuan human obstacles including staff who give up when they encounter

trivial obstacles, people who hang onto obsolete ideas and outmoded ways of doing things, emotional outbursts when asked to perform new tasks, muddled communications, poor anticipation of problems, and delayed action when problems arise so that once manageable problems become major management crises." The intention here is to identify barriers that might be obstacles to the growth and success of future 2 + 2 programs. Data from the current study indicates that a combination of excessive demands on administrators, and varying levels of commitment to the 2 + 2 program contributed to uneven administrative support.

The 2 + 2 implementation was in no way immune to the challenges of change management, especially since schools were expected to make simultaneous progress on other PRIME project initiatives. Internship programs, curriculum compression, alternative scheduling including ninth grade clusters at the high school, and parental involvement programs,, were all PRIME initiatives competing for the attention of administrators and teachers. These were external environmental factors, in addition to the usual stresses of the school year, impacting the 2 + 2 program implementation. This made adequate administrative support difficult, yet, at the same time, especially critical.

Once initiated, any program requires continued, active, highly visible administrative support. A number of factors were responsible for a general absence of continued, active support beyond the external factors already mentioned. These include an apparent belief among administrators that passive support was entirely adequate for the task. An unintended effect of the external support role performed by the researcher was to buttress this attitude. In other words, except for the elementary school, administrators were not proactive in their leadership of the implementation; the researcher was relied upon to do whatever needed to be done.

In the middle schools, administrative support for 2 + 2 was particularly vague. The implementation of 2 + 2 was, at best, tolerated by the principals. At Azalea Gardens Middle School, for example, only two teachers participated in 2 + 2 in lieu of the NPS evaluation system. No visible monitoring system existed at Azalea Gardens Middle School, and teachers there had the lowest number of 2 + 2 observations. Yet the researcher was assured by the principal more than once that 2 + 2 was "wonderful" and

flourishing. At Lake Taylor Middle School, observations conducted and received were documented in a database file by the administration. In spite of documentation, however, little action was taken by the administrations to remind teachers to complete their required number of 2 + 2 observations. The result was minimal monitoring without any consequences or feedback. Middle school administrators also aired philosophical differences with 2 + 2 at PRIME Steering Committee meetings near the end of the school year, including the need for administrators to be in charge of the evaluation process. Yet, administrators, with few exceptions, did not conduct 2 + 2 observations themselves. Most "forgot" that administrator observations are part of the 2 + 2 program guidelines.

The assistant principal at the high school was the coordinator of PRIME initiatives, including the 2 + 2 program. It was understood from the beginning that the researcher would help support the 2 + 2 implementation, especially at the high school with its large staff. In the end, though, almost all monitoring was accomplished by the researcher. Supportive action was not initiated by the administration, except at the beginning of the school year when teachers were strongly encouraged to participate in the 2 + 2 program in lieu of the new NPS appraisal system. Promises to schedule full 2 + 2 participant meetings were not fulfilled.

In contrast to the middle schools, however, 2 + 2 was considered a potentially valuable program at the high school. More than at the other schools, external factors seemed to overwhelm the capacity of the school leadership to be proactive in the 2 + 2 implementation and sustain support and visibility for the program. The many conversations with the assistant principal regarding the future viability of the 2 + 2 program at the high school led to the proposal of a building coordinator in the 2 + 2 sub-committee's proposed guidelines for 1997-98. Many other issues needing resolution in the implementation process such as accountability for numbers of observations, were first raised by the assistant principal and discussed with the researcher. Later in the year, the high school principal became a strong advocate for the 2 + 2 program. The 2 + 2 Sub-committee was co-chaired by the high school principal, who demonstrated a striking level of support for the 2 + 2 program as an alternative appraisal and professional

development system at a time when the middle school principals were most vocal in their opposition.

The elementary school, the most proactive of the 2 + 2 program schools, appointed a teacher representative to monitor 2 + 2 observations, and hired a substitute teacher for several days to enable 2 + 2 teachers to visit their colleagues. Both the principal and the assistant principal were known within the school as advocates of the 2 + 2 program, and both accepted 2 + 2 as an alternative appraisal system. Ironically, the teachers seemed slower than at the high school level to accept the authenticity of the program. Again, competing agendas, particularly the curriculum compression initiative, compromised the time and energy spent developing interest and support among the faculty for the 2 + 2 program. Nevertheless, as mentioned above, the number of teachers conducting observations doubled to 12 during the course of the year, and teachers meeting in May were very positive, despite continuing problems with time, about engaging the interest of other faculty members in the program.

In general, 2 + 2 received renewed attention as the end of the school year neared. Some defensiveness was in evidence due to the belated recognition that numbers of observations were fewer than required. The challenges to the viability of 2 + 2 as an evaluation program, especially by both middle school principals, is documented in Chapter IV. Middle school administrators also aired the criticism that 2 + 2 lacked focus. As evidenced in the program description in Chapter I, which was circulated among all PRIME Steering Committee members and 2 + 2 participants on many occasions, a staff development or grade level or cluster focus has always been an integral part of the 2 + 2 program. Initiative is required on the part of schools, however, to reach this level of sophistication. In fact, each school was asked in early 1997 to develop a list of instructional categories, best suited to their own instructional emphasis, which teachers could use to help focus their 2 + 2 observations. A sample list of categories was given to each school to assist them in getting started. No school, however, responded by developing such a focus.

Perhaps it would be too much to expect 2 + 2 to progress to another level of sophistication, including a staff development focus unique to each school, during its

initial year of implementation. Initiating the basic 2 + 2 model was a challenge. Teachers found that visiting colleagues in other departments and grade levels was rewarding without a school wide, or departmental staff development focus. By the end of the year, though, a core of teachers were ready to take the next step in making 2 + 2 yet more useful and focused. To take this step, against a background of an overfilled school calendar, is a function of the school leadership.

Beyond specific concerns about a staff development focus or evaluation considerations, the real issue is leadership. Are PRIME administrators committed to the concept of 2 + 2 as both an alternative appraisal and a professional development system? Are administrators unwilling to fully support its implementation, or are they simply overwhelmed by the range of administrative tasks already required of them? Do administrators share teachers' enthusiasm for sharing knowledge? Do administrators seek solutions to perceived weaknesses of the program? Or, not understanding the power of feedback, do they raise multiple objections and resist further participation? Do administrators resist 2 + 2 as a performance appraisal program? How convinced are administrators that 2 + 2 offers a feasible mechanism for improving teaching? Are administrators prepared to offer the level of support and visibility necessary to lead to an institutionalization of the program? Are administrators committed to the vision of teachers helping teachers? One consultant to the PRIME project, Dr. Walter Heinecke of the University of Virginia, expressed his concern in July, 1997, that if there is no vision at the top, that is, at the level of the principalship and PRIME Steering Committee, there will be little likelihood of change.

Yet, leadership at the school level is critical to the success of new initiatives. Hall and Hord (1987) cite research by Teacher Corps that examined the role of the principal in the change process. When implementing a new, externally funded program, specific contributions by the principal were critical to the program's success. A principal's enthusiasm during the program initiation and communication of support, his/her "active, positive role in the project" (p. 46), and, during implementation, a continuing, demonstrable show of interest and willingness to help solve problems were important behaviors for program success. Another researcher, in a study of principals from sixty

schools, concluded that "although many factors affect implementation, the leadership of the principal appears to be one of the most important factors in the success of demise of an alternative program" (Thomas, as cited in Hall and Hord, 1978, p. 47). Clearly, commitment to an initiative moves beyond compliance with an implementation, and must be communicated unequivocally by the school principal.

Time

Contrary to what might have been expected, teachers carried on with the 2 + 2 implementation process, and showed great interest despite a frequent absence of administrative feedback. Teachers readily attended focus group meetings, and some were in favor of 2 + 2 feedback sessions among participating teachers. Many teachers, though, found it difficult to summon the self-discipline required to organize regular 2 + 2 observations. For most, short term concerns related to their students needs and the day's classes crowded out thoughts of a semester commitment to 2 + 2 observations. Preparation for class, grading papers, meetings with students, or photocopying can easily occupy every free moment. Despite the difficulty establishing a 2 + 2 observation routine, the trend was an increase in 2 + 2 observations. A review of the second semester self-reflection reports showed that most teachers significantly increased their number of 2 + 2 visits in the spring.

The number of observations was not, in most cases, solely a function of self-discipline. Teachers at different school levels differed in the amount of discretionary time available to them. The high school teachers had the most discretion over their time, while elementary teachers and teachers at Azalea Gardens Middle School reported they could find time only with the help of substitute teachers. The elementary school 2 + 2 program was implemented with the assistance of substitute teachers. Other elementary school observations occurred on the teachers' own time. Most elementary school participants were unable to meet the stated number of required observations, even with the extra substitute time, largely because of the isolating nature of elementary schedules. If teachers are so isolated they have no time to interact with colleagues even if they want to, 2 + 2 observations will require so much extra effort and related stress that the costs will outweigh the many benefits.

In fact, without adequate time, 2 + 2 may wither away and meet the fate of countless other school programs. Time remains the greatest factor working against 2 + 2 and other forms of teacher collaboration. Teaching is one of the few professions where services must be delivered continuously, with little or no time for discussion, collaborative decision making, and coordination of instruction. Consequently, teachers at both middle school and at Lake Taylor High School were sometimes frustrated by administrative policies such as scheduling over which teachers have no control, and which thwart collegial interaction. Institutionalized isolation becomes the norm, where the thought of actively encouraging collegial contact through specific school practices rarely occurs to those exercising control over time. Research supports the premise that the principal is seen as the catalyst in supporting and encouraging school policies of shared decision making, and a collaborative culture (Johnson, 1990).

Time alone, however, does not guarantee collegial teacher interaction or more teacher 2 + 2 observations. Amidst calls for teamwork and group problem solving among students, teacher autonomy stands in odd contrast. Teacher autonomy and acceptance of isolation may be, as discussed earlier, a function of anxiety and uncertainty. Another perspective is advanced by Flinders (1988), who found that some teachers utilized isolation as an adaptive strategy to responsibilities which, in an open-ended profession, never end. Their choice not to interact with colleagues came from a professional desire to be the best possible provider of instruction. Ironically, the excellent instructional quality these teachers strive for is undermined in the long term by their isolation.

Teachers and administrators need to develop a school culture which values and promotes the exchange of ideas and collaborative efforts. The gradual diffusion of the 2 + 2 process witnessed in the high school implementation illustrates but one approach to breaking isolating barriers. A model of group work and team effort would then exist for the student body, a powerful lesson of collaborative learning that demonstrates the value of working together in solving problems beyond the school building.

Future of the 2 + 2 Program

The future of 2 + 2 depends on its active promotion by a school's administration, and on provisions in the schedule for time to accomplish observations. The resistance on

the part of school administrators discussed earlier becomes a key factor. If resistance is not addressed, the effort required on the part of administrators to institutionalize the 2 + 2 program will probably not be forthcoming. Peter Senge, in his book The Fifth Discipline (1990), discusses the qualities of systems as they undergo change. All systems naturally seek to preserve themselves, and, therefore, have limits to growth unless underlying root causes of resistance are removed. Limits to growth may be a resource, an implicit goal or norm. Reform efforts seem to succeed until they push against political power bases, threaten someone's authority or control, or exhaust a needed resource. At this point, they may continue to exist, but will not grow because underlying structures and power bases remain the same. Though the 2 + 2 program may continue to exist, its potential will be strictly limited if administrators consider it a stepchild to evaluation, a nice activity teachers can do on their own time, or a throwaway, and do not internalize the implications of collegial professional development.

The implication is that a professional staff will be developed which helps determine its own professional development agenda, engages in true self-reflection and shared decision making, and is supported in a culture of collaboration. In the 2 + 2 program, teacher interest is so strong that the program may continue to grow at first. At some point, however, without administrative commitment, the program will gradually lose ground.

Another limit to growth is a lack of necessary resources. Time, for example, poses a limit against which teachers push. Participation in the 2 + 2 program will eventually be discouraged if time for observations causes other obligations to be neglected. Organization of staff development based on 2 + 2 foci requires time. When time is an issue, exhortations to try harder to find time will not solve the problem. Senge suggests that "to change the behavior of the system, you must identify and change the limiting factor" (p. 101). The limiting factors, the root causes, are not always obvious, but ignoring them will affect the entire reform effort.

As defined by middle school researchers, "the school is a *system* of interactions involving structures, materials, pedagogy, professional development opportunities, and approaches to leadership and management...that *together* produce outcomes...Not until a

critical mass of reforms is in place and operating together in an integrated manner do significant positive changes in student outcomes occur" (Lipsitz, Jackson, Austin, 1997). Cawelti (1997) echoes these findings. He has identified seven elements through the research literature which are critical to school restructuring. He has postulated that in order to significantly affect positive changes in student achievement, and move the concept of school from "seat time" to performance, all seven elements should function together in a synergistic manner, creating a different, more productive system. In a study of ten high schools that have been engaged in restructuring activities for a number of years, results were highly suggestive that the more elements a school had successfully implemented, the higher the gains in student achievement. The elements replicate many PRIME initiatives such as a focus on teaching and learning, and curriculum integration. But Cawelti's emphasis is on the necessity for these elements to interact synergistically as a system of changes.

This is why resistance to the underlying implications of the 2 + 2 program will also affect the PRIME initiative of alternative scheduling, the call for site based management and shared decision making, and ultimately the whole PRIME reform effort. The 2 + 2 program should interact with these and other PRIME initiatives to bring about real change in student outcomes. Time allocation reflecting a belief in a collaborative management culture, where teachers and administrators together coordinate an educational program, will require changes in the power structure and conceptions of authority.

In the long term, commitment to a different system of beliefs requires more than "buy-in". Senge (1990) describes a continuum of attitudes toward a vision where buy-in suggests a process of being sold something, but "'Committed' describes a state of being not only enrolled but feeling fully responsible for making the vision happen" (p. 218). In this implementation of the 2 + 2 program, various stages of "compliance" were evidenced by administrators, including grudging compliance: "Does not see the benefits of the vision. But, also, does not want to lose job. Does enough of what's expected because he has to, but also lets it be known that he is not really on board" (Senge, 1990, p. 219). Genuine compliance was also represented: "Sees the benefits of the vision. Does

everything expected and more. Follows the 'letter of the law'" (Senge, 1990, p. 219).
But commitment reflects another level of responsibility: "Wants it. Will make it happen.
Creates whatever 'laws' (structures) are needed" (Senge, 1990, p. 219). Creating
structures to support, optimize, and institutionalize the 2 + 2 program, as well as the
entire PRIME project, will require commitment to a vision of systemic change over the
long term.

One of the most important prerequisites on the road to shared vision and
commitment is an honest view of the current state of reality. Problems should not be
denied, or current achievements inflated. At the same time, clear signals from
organizational leaders about expected levels of compliance are necessary in an
atmosphere of honest discourse.

A first step toward gaining a sense of current reality is to realize that every
stakeholder is, in one way or another, a part of the problem. "At the heart of the learning
organization is a shift of mind - from seeing ourselves as separate from the world to
connected to the world, from seeing problems as caused by someone or something 'out
there' to seeing how our own actions create the problems we experience. A learning
organization is a place where people are continually discovering how they create their
reality. And how they can change it" (Senge, 1990, pp. 12-13). Such a realization
highlights the ineffectiveness of blame placing actions and denials of responsibility,
typical reactions when organizational crises arise. The PRIME project, along with other
educational reform projects, might do well to seek to achieve the mindset of a learning
organization. The alternative is to attempt reform while mired in a traditional paradigm,
with familiar results: Circumstances or individuals targeted for blame when progress
lags, participant feelings of victimization, and/or lowered goals are common outcomes of
reform efforts.

Commitment to a vision can never be forced on another person, but conditions
can be created that support open discussion. The PRIME project, like other
organizational change efforts, requires an environment which encourages ongoing
challenges to the status quo. To make progress toward becoming a learning
organization, a decision from upper management to support individuals in the process is

needed. Trust is essential for breaking out of the dependent, victim status experienced by many participants in schools undergoing restructuring, and gaining a sense of control over change. District leaders can assist school level administrators to relinquish the role of victim of change by redefining the traditional patriarchal relationship. In that model, upper level managers are protective care of those they supervise, as long as the subordinates do as they are told (Patterson, 1997). The same relationship often exists between administrators and teachers at the school level. This is not a model which supports personal growth or risk-taking.

Personal concerns of teachers and administrators must be addressed and resolved for learning to take place, much the same as connections to students' frames of reference must be made for learning in the classroom to occur. Mixed messages, where the district rewards compliance with traditional roles and expectations, but hopes for innovation and improvement, effectively squelches commitment to radically different visions of the educational process, and thwarts honest dialogue about how to integrate innovation with organization guidelines. With more clearly articulated expectations and better defined levels of empowerment, the PRIME Project will be in a position to move its participants beyond traditional roles, with the 2 + 2 program a powerful mechanism to support greater trust and openness among all stakeholders.

Recommendations

In the shorter term, a number of proposed modifications to the 2 + 2 program have been suggested by teachers, administrators, and the researcher alike. Some of these have already been implemented. Others are being discussed for possible adoption.

Implemented Recommendations

In response to teacher suggestions, the number of observations per semester was reduced from 20 to 15 during the spring semester. Due to the value teachers placed on post-observation discussion, the option to document up to five such discussions and count them toward the total of 15 observations was also approved by the PRIME Steering Committee. The PRIME Steering Committee also agreed that the observation

deadlines for those teachers participating in the 2 + 2 program in lieu of the NPS evaluation system should follow those of the NPS system.

<div align="center">Recommendations Pending</div>

The PRIME Steering Committee established a sub-committee to develop guidelines for the 1997-98 implementation of the 2 + 2 program in May 1997. The new guidelines address many of the concerns which surfaced during the 1996-97 implementation of 2 + 2, making the document a very positive outcome of this 2 + 2 evaluation. The draft proposal may be found in Appendix D. Among the recommendations was a reinstatement of the guideline that all teachers, regardless of years of experience, be eligible for participation in 2 + 2 in lieu of the Norfolk Public Schools (NPS) evaluation. Also significant was the recognition of the need for support by school administrators and department heads working, together with a 2 + 2 building coordinator. Figure 11 summarizes the recommendations. The proposed recommendations represent a major commitment to the 2 + 2 program, and reflect the belief of the high school principal that 2 + 2 is a valuable program. The adoption of this proposal would be a significant step toward the institutionalization of the 2 + 2 program if school administrations view it as a priority.

Beyond Logistics

The draft proposal provides a blueprint for program implementation that will significantly aid in standardizing the 2 + 2 process. Beyond the logistical and programmatic processes, however, a focus on progressive improvement of program substance must be maintained. The current 2 + 2 implementation stands on its own merits. Teachers have enthusiastically embraced 2 + 2 observations, and experimented with ideas which were suggested or which they observed in other classrooms. It has made a positive difference to teachers. But to realize continuous gains in creating a collaborative culture, and, especially, to significantly impact improvement of instruction, it will be necessary for schools to further develop the potential of 2 + 2 program. Observation foci may be identified and related to staff development initiatives, such as alternative assessment, or interdisciplinary curriculum. Among options teachers might consider are paired 2 + 2 observations, with mutual coaching, or other groupings such as

Major Recommendations of the PRIME Steering Committee Sub-Committee for the
Future Implementation of the 2 + 2 Program

- 2 + 2 will remain voluntary

- 2 + 2 will follow the observation and time line schedule of the NPS Teacher
  Appraisal Handbook for summative years.

- A teacher, administrator, or intern will be appointed building coordinator of the 2
  + 2 program to arrange training, organize/schedule/monitor observations, share
  information, and problem solve.

- Numbers of observations to be required were based on the survey distributed in
  May 1997: summative year teachers will complete 10 observations including 3
  conferences per semester and professional development year teachers will
  complete 5 observations per semester including 1 conference.

- At least 90% of the required first semester observations must be completed by
  the end of January and the teacher must demonstrate an acceptable level of
  proficiency. Otherwise removal from the 2 + 2 program will be mandatory.

- Training for the 2 + 2 program will be part of the pre-service staff development
  annually.

- A school administrator must make at least three, and department heads at least
  four 2 + 2 visits. There will be a summative evaluation conference.

- School administrators and department heads must work with the building
  coordinator to support the 2 + 2 program.

- The self-reflection report form was redesigned as the 2 + 2 Teacher Observation
  Report and includes a section for the administrator to confirm that the teacher is
  "proficient" in each domain addressed by Norfolk Public School evaluation
  system.

Figure 11. Summary of the Draft Guidelines for 2 + 2 Observation and Evaluation

grade level, cluster, or subject area concentration.

Most teachers also need ongoing support to feel competent making suggestions for improvement. Teachers took first, tentative steps in critiquing their colleagues. While many of their comments were insightful, others did not address substantive issues of instruction. Even lower level comments, for example, to use the overhead projector more, can be very useful. Too often, though, suggestions were non-productive, or not provided at all. The type of training teachers have said would be most useful is that provided by peers in informal group settings, where teachers can discuss instructional issues and observation techniques. Such meetings would have a secondary effect of further increasing teacher interdependence.

Substantive inquiry into the nature of teaching, and learning also requires deep self-reflection. Logistical issues are important, but educational change is brought about by inquiry and reflection on difficult, fundamental issues affecting student learning. If schools view 2 + 2 as a tool to support open inquiry, honest self-analysis, and trust among teachers and administrators alike, greater likelihood for real change in student outcomes will exist. Improved student performance is the ultimate goal.

The Current 2 + 2 Program Evaluation

The current study had a beneficiary effect on the 2 +2 program implementation. Feedback from teachers influenced program development and implementation. For example, an observation chart was posted and schedules were distributed at the high school. Teachers enjoyed opportunities to meet and give feedback, especially the chance to discuss their 2 + 2 experiences with participating colleagues. One teacher even mentioned in the self-reflection report that the interview was a good idea. Throughout the year, teachers' suggestions and concerns were brought to the PRIME Steering Committee, where changes were made in the number of observations required, and accountability issues were discussed. For administrators, especially the high school principal, survey results provided quantitative data that were easily accessible and served user needs to understand teachers' attitudes.

The evaluation also provided a means to identify program ambiguities and implementation issues that had not been explicitly addressed prior to the 1996-97 implementation. These included, for example, the need for a designated building coordinator of the program, and clarification of the role of the administrator in the 2 + 2 summative year evaluations. The issues were brought to the attention of individual principals and the PRIME Steering Committee and became the basis for the comprehensive proposed guidelines for 1997-98.

Limiting factors included time and resources for further examination of the 2 + 2 program. An analysis of individual teachers' 2 + 2 observation forms over time could have indicated how critical judgement skills were developing, and help define where coaching or training in performing observations might be useful. Individual interviews later in the year might have yielded more specific information on how often and with what success teachers implemented suggestions.

In particular, a higher level of user participation in the evaluation would have been desirable. This, however, assumes an interest in the process, and a willingness to devote time and energy to exploring evaluation uses and questions. Both these commodities were in short supply during the 1996-97 school year. It is hoped that the current evaluation spurs interest in long term assessment of the 2 + 2 program.

Inferences made in this evaluation were based on preliminary data. Further evaluation of the program is necessary to confirm the impact of the 2 + 2 program on instructional practice and student achievement.

Implications for Further Evaluation Research

There are rich opportunities for further research in the 2 + 2 program. Study of the implementation process in 1997-98 may suggest additional program modifications, for example. Many questions could be asked regarding program components. For example, which types of suggestions, post-observation contacts, goal-setting, training, coaching relationships or combinations of these elements best encourage modification of teaching behaviors? How do variations in the 2 + 2 process, that is, pairing of teachers,

grouping in clusters, departments, or grade level, affect teacher learning? To what extent are teachers attitudes, knowledge, and skills affected by the 2 + 2 process?

These questions lead inevitably to a very important avenue of future evaluation, the assessment of how the 2 + 2 program affects improvement of instruction. Teachers indicated in the current study that they had implemented suggestions made by their peers, or adapted ideas observed in other classrooms for their own teaching. Further research is needed to determine if increased feedback and multiple perspectives leads to improvement of instruction. As mentioned above, specific components or combinations of 2 + 2 components might be evaluated insofar as each contributes to improvement of instruction. A time series design might be employed to observe and document teachers' improvement over time. Measurement criteria could be adopted by PRIME schools, based on, for example, teaching standards developed by the Center on Organization and Restructuring of Schools at the Wisconsin Center for Education Research. The Center has developed "specific teaching standards that measure the extent to which students are challenged to think, to develop in-depth understanding, and to apply academic learning to important, real-world problems" (Newmann & Wehlage, 1995, p. 2). Alternatively, establishing staff development foci for 2 + 2 observations could also help focus an evaluation of teaching improvement by limiting the scope of the observational research.

Another possibility could be to compare instructional improvement in two middle schools, one where 2 + 2 is being implemented and the other, for example, Azalea Gardens Middle School (AGMS), where 2 + 2 is not being promoted. Such a comparison would be difficult to design. However, the improvement of selected 2 + 2 teachers might be compared to selected non-2+ 2 teachers over time, controlling as much as possible for extraneous variables. The progress of teachers participating in the 2 + 2 program might also be compared with that of non-participating teachers in a single school.

A separate line of inquiry for further study might be based on findings from the General Linear Model Procedures in the survey and questionnaires which are not adequately explained in the literature. Two findings pointed to possible gender differences in attitudes toward appraisal, specifically that male teachers were both more

likely to accept the 2 + 2 appraisal system, and to consider the Norfolk Public Schools appraisal system fair and accurate, than were female teachers. If these findings are confirmed, implications for future development of appraisal systems include consideration of male/female differences.

Two additional findings indicated that a difference may exist in attitudes toward performance appraisal between elementary school teachers and high school teachers. Elementary school teachers were more likely to consider their supervisors knowledgeable about teacher performance than high school teachers. They also were more likely than high school teachers to view the Norfolk Public Schools appraisal system favorably. High school teachers were more likely than elementary school teachers to view the 2 + 2 appraisal system favorably. Possible differences in perceptions of autonomy, or locus of control, may also hold implications for differentiated appraisal systems by school level. Alternatively, further research may suggest a greater need to encourage shared decision making and other empowerment measures at the elementary level.

Future evaluation research should make every effort to involve stakeholders in designing the evaluation. Administrators, teachers, department chairs, teacher union representatives, and district level personnel should be involved at a minimum. As rich an area for evaluation as the 2 + 2 program presents, time and resources for gathering and analyzing data remains at a premium. An evaluation task force would be able to best identify what information would be most useful. Together with an evaluator, the task force would determine which data might be obtained systematically by instituting regular data collection and other non-intrusive measures. This might include such data as numbers of observations, and analyzing 2 + 2 observation forms over time. Other data would need to be obtained using trained observers to assess instructional improvement. A task force would also serve the function of processing the hidden assumptions, for example, evaluation and control, which threaten continued growth of the 2 + 2 program.

Conclusion

The evaluation of the 2 + 2 for Teachers: Alternative Performance Appraisal program has confirmed the potential of 2 + 2 to significantly impact teacher isolation and teacher learning. Teachers were enthusiastic in their support of the program, and recognized that it represented a fundamental change in the performance appraisal process, and in the way they interacted with colleagues. Teachers themselves were the most vocal advocates of the 2 + 2 program. This, together with the simplicity of 2 + 2 accounted for the gradual diffusion of the program throughout the high school, and to a lesser extent in one middle school and the elementary school. Its most basic component, the classroom observation, is an activity any teacher can understand. Teachers found that, without in service or special training, without special permission or forms to sign, they, too, could have a colleague observe them, or step next door and "do a 2 + 2" on another teacher. Although the involvement of numerous teachers on an informal basis led, in some cases, to fewer observations performed on summative evaluation year teachers, informal participation prepared the way for large scale formal participation in 2 + 2 in the long term. It also allowed the seeds of a culture of cooperation and collaboration to be sown.

The integrity of the program was not threatened by informal participation, since specific guidelines existed for participation in the 2 + 2 program as an alternative to the NPS appraisal system. New guidelines proposed above further delineate expectations for those participating in both summative and non-summative evaluation years. An integrated staff development focus is expected to evolve gradually.

Administrators must perform a crucial role in the future development of the 2 + 2 program. Preoccupation with other reform initiatives, external pressures, and demands by the district, distracted administrators from the 1996-97 implementation of the 2 + 2 program. Discomfiture with the 2 + 2 program, and its departure from the familiar, traditional evaluation model, were additional causes for the lack of attention given to the 2 + 2 implementation in some schools. In the current 1996-97 implementation year, minimal harm was incurred by this approach. Teachers supported each other and

accomplished a surprising amount of program diffusion as well. Without strong administrative commitment in the future, however, danger exists that the 2 + 2 program will go the way of many programs that work, but are not actively supported by organizational leaders.

It is difficult to address the implementation of the 2 + 2 program without looking at the larger system into which it was introduced. New programs do not stand on their own merits alone and are not implemented in a vacuum. Systemic forces shape both implementation of the program and its potential sustainability. Where a series of initiatives are implemented with the intention of achieving substantial systemic change, resistance is a natural reaction. In the PRIME project, changes in curriculum, teaching, staffing, scheduling, in addition to teacher appraisal, deeply affect assumptions of power, control, time, teaching, and learning held by the typical school system. Commitment to educational change, then, also means deep reflection on these assumptions and active pursuit of changes which are believed to make a positive difference for students. For initiatives to move beyond appendages to the "regular" system, and begin to positively influence student performance, they must be value driven (Fullan, 1993; Patterson, 1997, Senge, 1990).

Teachers must be supported in risk-taking behaviors, in creating change, and in accepting increased accountability. Administrators, though, also need support and encouragement. Implicit in many school systems is a patriarchal hierarchy in which an administrator or teacher is protected as long as they do what is expected. In such an environment, it would be unusual for administrators to choose ambiguity, risk, and accountability for shaping change, and more natural to perpetuate what brought rewards in the past (Patterson, 1997). The question must be asked what rewards exist for administrators to take risks, learn by trial and error, and embrace uncertainty.

Moral purpose, defined by Fullan (1993) as "making a difference in the lives of more and more individual students" (p. 111) is at the heart of change. Many, if not most, teachers and administrators enter the profession because they genuinely wish to make a difference in students' lives. Encouragement to pursue a moral purpose in creating conditions which support continuous systemic change is necessary, as are the knowledge

and skills to understand and shape systemic forces in complex environments. But most teachers and administrators also need information and reassurances about how changes will affect them personally. Understanding the implications for the larger system, and committing time and energy to changes based on personal and organizational values comes after personal concerns have been identified and addressed (Hall and Hord, 1987).

The new guidelines propose by the 2 + 2 sub-committee are a positive evolution of the 2 + 2 program in this direction. The guidelines reduce uncertainty and are, in part, a result of some trial and error during the 1996-97 implementation. Eligibility for all teachers was reinstated, and the role of the administrator was more clearly defined. Significantly, the guidelines were drafted by PRIME participants themselves. It remains for the schools to secure the 2 + 2 program as part of the educational culture.

REFERENCES

Ackland, R. (1991). A review of the peer coaching literature. Journal of Staff Development, 12, 22-27.

Allen, D.W. (1992). Schools for a new century: A conservative approach to radical school reform. New York: Praeger Publishers.

Allen, D. W., Nichols, R. D., Jr., & LeBlanc, A. C. (1997, June/July).The prime teacher appraisal program: 2 + 2 for teachers. The High School Magazine, 4, 30-35.

Anderson, R. D. (1995, September). Curriculum reform, dilemmas and promise. Phi Delta Kappan, 77, 33-36.

Ascher, C. (1991). Retaining good teachers in urban schools. ERIC/CUE Digest, 77. (ERIC Document Reproduction Service No. ED 341 762)

Ashton, P. T. & Webb, R. B. (1986). Making a difference: Teachers' sense of efficacy and student achievement. New York: Longman.

Askins, B. E. (1994, February 12-16). Collegial support teams replace teacher evaluations in a partnership school. Atlanta, GA: Association of Teacher Educators. (ERIC Document Reproduction Service No. ED 367 654)

Barber, L. W. (1990). Self-assessment. In J. Millman & L. Darling-Hammond (Eds.), The new handbook of teacher evaluation assessing elementary and secondary teachers (pp. 216-228). Newbury, CA: Sage Publications, Inc.

Berndt, A. E. (1990). Employee attitudes about the performance appraisal process: Structural equation modeling of the performance appraisal attitude questionnaire. Unpublished master's thesis, Old Dominion University, Norfolk, VA.

Berube, M. R. (1994). American school reform: Progressive, equity, and excellence movements, 1883-1993. Westport, CT: Greenwood Press.

Brophy, J.E. & Good, T.L. (1986). Teacher behavior and student achievement. In M.Wittrock (Ed.) Handbook of research on teaching (pp. 328-375). New York: Macmillan.

Canady, R. L. & Rettig, M. D. (1995). Block scheduling: A catalyst for change in high schools. Princeton, NJ: Eye on Education.

Cawelti, G. (1997). Does high school restructuring produce results? Paper presented at the 1997 conference of the Association for Supervision and Curriculum Development in Baltimore, MD.

Centra, J. A. (Ed.). (1977). Renewing and evaluation teaching: New directions for higher education. San Francisco: Jossey-Bass, Inc., Publishers.

Conley, D. T. (1993). Road map to restructuring policies, practices and emerging visions of schooling. Eugene, OR: ERIC Clearinghouse on Educational Management. (ERIC Document Reproduction Service No. ED 359 593)

Eisner, E. W. (1992, Summer). Educational reform and the ecology of schooling. Teachers College Record, 93, 610-627.

Elmore, R. F. (1995). Structural reform and educational practice. Educational Researcher, 24, 23-26.

Flinders, David J. (1988, fall). Teacher isolation and the new reform. Journal of Curriculum and Supervision, 4, 17-29.

Fullan, M. G. (1993). Change forces: Probing the depths of educational reform. Bristol, PA: The Falmer Press, Taylor & Francis Inc.

Fullan, M. G. (1996, February). Turning systemic thinking on its head. Phi Delta Kappan, 77, 420-423.

Fullan, M. G. & Hargreaves, A. (1991). What's worth fighting for? Working together for your school. (Report No. ISBN-1-878234-02-1). Andover, MA: Regional Laboratory for Educational Improvement of the Northeast and Islands. (ERIC Document Reproduction Service No. ED 342 128)

Goldberg, M. F. (1996, June). An interview with Theodore Sizer: Here for the long haul. Phi Delta Kappan, 77, 685-687.

Good, T. L. (1980). Classroom observations: Potential and problems. In W. R. Duckett (Ed.). Observation and the evaluation of teaching. (A CEDR Monograph). (pp. 2-45). Bloomington, IN: Phi Delta Kappa.

Good, T.L. & Mulryan, C. (1990). Teacher ratings: A call for teacher control and self-evaluation. In J. Millman & L. Darling-Hammond (Eds.). The new handbook of teacher evaluation assessing elementary and secondary school teachers.

(pp. 191-215) Newbury Park, CA: Sage Publications, Inc.

Goodlad, J.I. (1996, November). Sustaining and extending educational renewal. Phi
Delta Kappan, 78, 228-234.

Gottesman, B. L. & Jennings, J.O. (1994). Peer coaching for educators. Lancaster, PA:
Technomic Publishing company.

Gredler, M.E. (1996). Program evaluation. Englewood cliffs, NJ: Prentice-Hall, Inc.

Hall, G. E. & Hord, S. M. (1987). Change in schools: Facilitating the process. Albany,
NY: State University of New York Press.

Hargreaves, A. & Earl, L. (1990). Rights of passage: A review of selected research
about schooling in the transition years. Toronto, Ontario: MGS Publications
Services.

Hoover-Dempsey, K., Bassler, O. & Brissie, J. (1987). Parent involvement:
Contributions of teacher efficacy, school socioeconomic status, and other school
characteristics. American Educational Research Journal, 24, 417-435.

Johnson, J. (1995). Assignment incomplete: The unfinished business of educational
reform. A report from public agenda. Washington, D.C.: Institute for Educational
Leadership. (ERIC Document Reproduction Service No. ED 387 921)

Johnson, S.M. (1990). Teachers at work: Achieving success in our schools.New York:
Basic Books.

Kahne, J. & Westheimer, J. (1993). Building community: A model for teacher
education and staff development. San Francisco, CA: The American Educational
Research Association. (ERIC Document Reproduction Service No. ED 346 074)

Karlsen, J.I. (1991). Action research as method: Reflections from a program for
developing methods and competence. In W.F. Whyte (Ed.). Participatory action
research (pp. 143-158). Newbury, CA: Sage Publications, Inc.

Lee, V.E., Dedrick, R.F., & Smith, J.B. (1991, July). The effect of the social
organization of schools on teachers' efficacy and satisfaction. Sociology of
Education,64, 190-208.

Lipsitz, J., Jackson, A. W. & Austin, L. M. (1997, March). What works in middle-
grades school reform. Phi Delta Kappan, 78, 517-519.

Little, J.W. (1990). The persistence of privacy: Autonomy and initiative in teachers' professional relations. Teachers College Record. 91. 509-536.

McCarthey, S. J. & Peterson, K.D. (1988). Peer review of materials in public school teacher evaluation. Journal of Personnel Evaluation in Education. 1. 259-267.

Moore, W.P. & Esselman, M.E. (1994). Exploring the context of teacher efficacy: The role of achievement and climate. New Orleans, LA: American Research Association. (ERIC Document Reproduction Service No. ED 370 919)

National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform. Washington, DC: U. S. Department of Education.

Newmann, F. M., & Wehlage, G.G. (1995). Successful school restructuring: A report to the public and educators. [On-line]. Available: http://144.92.174.191/Completed_Projects/Center_on_Organization/Successful_Restruct.html

Norfolk Public Schools. (1996). Teacher Appraisal System & Professional Growth Plan. Norfolk, VA: Author.

Patterson, J. L. (1997). Coming clean about organizational change: Leadership in the real world. Arlington, VA: American Association of School Administrators.

Patton, M. Q. (1978). Utilization-focused evaluation. Beverly Hills, CA: Sage Publications, Inc.

Patton, M. Q. (1990). Qualitative Evaluation and Research Methods. Newbury Park, CA: Sage Publications, Inc.

Patton, M. Q. (1997). Utilization-focused evaluation. (3rd ed.). Newbury Park, CA: Sage Publications, Inc.

Peterson, K. D. (1984). Methodological problems in teacher evaluation. Journal of Research and Development in Education. 17. 62-70.

Peterson, K.D. (1995). Teacher evaluation: A comprehensive guide to new directions and practices. Thousand Oaks, CA: Corwin Press.

Popham, W. J. (1987). The shortcomings of champagne teacher evaluations. Journal of Personnel Evaluations. 1. 25-28.

Practice into theory: Teachers coaching teachers. (1990, November). Horace, 7.

Raudenbush, S., Rowan, B. & Cheong, Y. (1992). Contextual effects on the self-perceived efficacy of high school teachers. Sociology of Education, 65, 150-167.

Rebell, M. A. (1990). Legal issues concerning teacher evaluation. In J.Millman & L. Darling-Hammond (Eds.). The new handbook of teacher evaluation assessing elementary and secondary teachers. (pp 337-355). Newbury Park, CA: Sage Publications, Inc.

Reyes, P. (1992, May). Preliminary models of teacher organizational commitment: Implications for restructuring the workplace. Austin, TX: Center on Organization and Restructuring of Schools. (ERIC Document Reproduction Service No. ED 349 680)

Rosenholtz, S. J. (1987). Education reform strategies: Will they increase teacher commitment? American Journal of Education, 95, 534-562.

Rosenholtz, S. J. (1989). Teachers' workplace: The social organization of schools. New York: Longman.

Ross, J.A. (1992). Teacher efficacy and the effect of coaching on student achievement. Canadian Journal of Education, 17, 51-65.

Ross, J.A. (1994). Beliefs that make a difference: The origins and impacts of teacher efficacy. Peterborough, Ontario: Ontario Institute for Studies in Education Trent Valley Centre. (ERIC Document Reproduction Service No. ED 379 216)

Rossi, P. H. & Freeman, H. E. (1993). Evaluation: A systematic approach. Newbury Park CA: Sage Publications, Inc.

Sarason, S. B. (1990). The predictable failure of educational reform. San Francisco, CA: Jossey-Bass.

Sarason, S. B. (1995, September). Some reactions to what we have learned. Phi Delta Kappan, 77, 84-85.

Sarason, S. B. (1996). Revisiting The culture of the school and the problem of change. New York: Teachers College Press.

Sashkin, M. & Egermeier, J. (1993). School change models and processes. A review and synthesis of research and practice. Washington, DC: U.S. Department of

Education. Office of Educational Research and Improvement.

Scriven, M. (1981). Politics of teacher evaluation. In J. Millman (Ed.), Handbook of teacher evaluation. (pp. 248-261). Beverly Hills, CA: Sage Publications, Inc.

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

Shanker, A. (1990, January). The end of the traditional model of schooling - and a proposal for using incentives to restructure our public schools. Phi Delta Kappan, 71, 345-357.

Steven, H. & Stigler, J. (1992). The learning gap: Why our schools are failing and what we can learn from Japanese and Chinese education. New York: Summit Books.

U.S. Department of Education, National Center for Education Statistics. (1991). Trends in academic progress. Washington, DC: U.S. Government Printing Office.

U.S. Department of Education, National Center for Education Statistics. (1994). Digest of education statistics. Washington, D.C.: U.S. Government Printing Office

Whitaker, K. S. & Moses, M. C. (1994). The restructuring handbook: A guide to school revitalization. Needham Heights, MA: Allyn and Bacon.

White, P. A. (1992, Spring). Teacher empowerment under *ideal* school-site autonomy. Educational Evaluation and Policy Analysis, 14, 69-82.

Wholey, J.S. (1979). Evaluation: Promise and performance. Washington, DC: The Urban Institute

Whyte, W. F. (Ed.). (1991). Participatory action research. Newbury Park, CA: Sage Publications, Inc.

Whyte, W. F., Greenwood, & Lazes. Participatory action research: Through practice to science in school research. In W.F. Whyte (Ed.), Participatory action resear

APPENDIX A

2 + 2 ALTERNATIVE PERFORMANCE APPRAISAL SURVEY

JUNE 1996

2 + 2 Alternative Performance Appraisal Evaluation Survey

Dear PRIME teachers and interns: Please take a few minutes of your very busy schedule to complete this survey. A valid and useful evaluation is dependent on a high percentage of returned surveys. Feedback from each of you (even if you did not participate) is important to improve 2 + 2. Thanks for your help.

P.S. Please return to the 2 + 2 box in your school's main office by **June 15.**

I am a(n): _____ intern _____ teacher at (school name):_____

1. When were you introduced to the 2 + 2 concept?_____

2. Did you participate in 2 + 2? ___Yes ___No

   If not, why? (Check all that apply)
   ___No time to perform observations ___Do not like the 2 + 2 idea
   ___Administration did not assign me ___Other (please specify)_____
      to program
   ___Will/would like to participate next year

3. Did you receive training on 2 + 2? ___Yes ___No
   If yes, was the training helpful? ___Yes ___No
   Was it about the right amount of training? ___Yes ___No

Comments:_____

4. About how many times did you observe another teacher or intern using 2 + 2 (per semester)?
   ___1-5 times   ___6-10 times ___more than 10 times

5. About how many times were you observed (per semester)?
   ___1-5 times   ___6-10 times ___more than 10 times

6. Please comment on the compliments you received (were they unexpected, helpful, brightened your day, provided motivation, etc):

7. Please comment on the suggestions you received (were they irrelevant, useful, did you implement some of them, irritating, all of the above, etc.):

8. As a result of my participation in 2 + 2, I am ___more ___less interested in participating in 2 + 2 in the future.
   Why?_____

9. What suggestions do you have to make 2 + 2 more useful or meaningful to you?

10. What compliments do you have for the 2 + 2 program?

APPENDIX B

2 + 2 FOR TEACHERS APPLICATION FORM

2 + 2 FOR TEACHERS

APPLICATION FORM

Please submit this form to the principal in duplicate.

I would like to participate in the 2 + 2 for Teachers program. I understand I will be given one period of release time every two weeks to accomplish the 2 + 2 observations.

Name_____   Dept_____

As a 2 + 2 participant I agree to:

* observe at least two colleagues every other week using the 2 + 2 observation form
* give one copy of the triplicate 2 + 2 form to the teacher I observe, keep one copy for my observer portfolio, and submit one copy to the administration
* open my classroom to teachers, interns, and administrators without prior notice
* keep a supply of 2 + 2 observation forms in my room
* maintain an observation portfolio of 2 + 2 compliments and suggestions I have received
* document which suggestions have been most useful: at the end of the fall semester I will complete a summative evaluation form which lists the 10 most useful compliments and suggestions
* participate in surveys and/or interviews to help evaluate 2 + 2 and shape its implementation
* be observed at least 20 times by teachers, interns, administrators, students, etc. each semester. If this goal looks as though it will not be met, I will actively solicit observers and/or alert PRIME staff. 2 + 2 will replace the Norfolk Public School teacher evaluation program for participating teachers in PRIME schools. Optionally, the 2 + 2 Program encourages the use of the 2 + 2 format to solicit student compliments and suggestions. This is not a requirement of the 2 + 2 Program.

I understand it is my choice whether to share my portfolio with observing teachers.

I understand that the administration retains the right to terminate my participation in the 2 + 2 Program should an administrator have concerns about my teacher performance.

_____   _____
                signature   date

```
For Administrative Use Only
Application for participation in the 2 + 2 program has


_____been approved


_____not been approved at this time



_____          _____
            administrative signature                            date
```

APPENDIX C

2 + 2 IMPLEMENTATION ISSUES

ISSUE: THE IMPLEMENTATION AND INSTITUTIONALIZATION OF 2 + 2
(to be discussed with PRIME principals and the PRIME Steering Committee)

General Questions:
How is 2 + 2 doing?
How is 2 + 2 making a difference in PRIME schools?
How is 2+2 different than you thought it would be?
How is 2 + 2 impacting instruction? Professional collegiality?

WHAT IS THE BEST DESIGN FOR NEXT YEAR?

Logistics:
Setting deadlines for submitting 2+2 pink forms to the administration; when?
Quarterly?
Monitoring
Follow up
Sessions for feedback
Development of a database?
How many observations should be required in summative years? Non-summative years?

Summative Evaluation Report:
Changes in report form to reflect value of observation to the observer and well as the teacher being observed
Criteria for...
Importance of...
Value of...
Should there be a summative report in non-summative years?
Superintendent letter attached to form explaining his support of program (for permanent record)

Summative Conferences
What form should the summative conference take?
What should be discussed at the summative?
How should a summary report be different in non-summative years?

2 + 2 Options:
Paired 2+2s
Group 2+2s: grade levels, subject area dept., clusters, coaching optional
School: a staff development focus is identified for the school
Addition of staff development foci to 2+2 forms
Identify other options

Training
Is there a need for more training?
What should the training look like?

Role of Administrator
>What are the responsibilities of the administrator in 2 + 2?
>How many 2 + 2 observations should administrators make?
>Do administrators have a role in training teachers for 2+2?
>Do the department heads have a role in training teachers for 2 + 2?

Goals
>Are the formal goal setting sessions at the beginning of the year part of 2+2?
>How do they fit in with the summative evaluation? With 2+2 forms?
>Should goals be eliminated?

Domains not addressed by 2+2
>How can issues such as "professionalism" be addressed.

2+2 as a mandatory vs. voluntary program
>What are the pros and cons of each?
>If mandatory, how should it be scheduled?
>What are the pros and cons?
>What can be done to encourage the perception of 2+2 as a professional growth issue, rather than an evaluation issue?
>How can we broaden the appeal and participation in 2+2?

Conferencing and/or coaching?
>Should coaching be an optional component of 2 + 2 ?
>Are informal followup conferences occuring?

Are student 2+2s being solicited? How is it important?
Do you have suggestions for appropriate ways to collect 2+2s from students?

Accountability
What kind of accountability should there be with regard to completion of of 2+2 observations?
What kinds of measures should we take to spread out the observations better?

2 + 2 forms
Any changes to be made here?

Program History and Development
What is your perception as to how 2+2 got established ?
Who should have the major responsibility for decision making about the 2+2 program?
Where have the decisions come from in the past?

APPENDIX D

2 + 2 SUB-COMMITTEE DRAFT PROPOSAL
AS SUBMITTED TO THE PRIME STEERING COMMITTEE

draft

PRIME   **FAIRLAWN LITTLE CREEK ST. HELENA**
2+2   **AZALEA   LAKE TAYLOR MIDDLE   HIGH**

GUIDELINES FOR 2 + 2 OBSERVATION AND EVALUATION

1. OBSERVATION - EVALUATION

2 + 2 will remain a voluntary component of the PRIME initiatives. Teachers may opt to substitute 2 + 2 for the NPS appraisal system during summative years (the first three years of employment, every fourth year, first year in new assignment).

2 + 2 will be both a means for professional growth, sharing of techniques and a means to achieve the professional development hours required by NPS.

Pages 1-11 of the NPS Teacher Appraisal Handbook will be followed for summative years.

2. LOGISTICS OF 2 +2

A teacher, intern or administrator will be the **Building Coordinator** for 2 + 2. The building coordinator will arrange training, organize and monitor observations, follow-up observation problems, sharing observation information.

Each school will develop a system where structure and/or scheduling of observations is established. There may be a sign up sheet with blanks under each teacher's name. Each observer signs up under another teacher's name. When the blanks are full, observers must select another teacher to observe.

Building Coordinator will complete a memorandum to teachers listing each teacher and the names of observers that will visit. Each quarter the Building Coordinator will collect data and report by teacher the number of observations completed, the number of visits made to each teacher's class and suggestions for next quarter.

For **summative years** each teacher must complete 10 observations per semester including 3 conferences. For **professional development** each teacher is asked to complete 5 observations per semester including 1 conference. 2 + 2 observation forms are to be submitted to the building coordinator the day of the observation. The Building coordinator will file each observation in a separate file folder for each teacher. Administrators will have access to the file. Each teacher will have access to their file folder.

Summative Evaluation and the **2 + 2 Teacher Report** (attached) will be completed by May 30 of each year. For Professional Development, only the **teacher report** must be completed and signed by the teacher. For Summative Evaluation, both the **teacher report** and the **Principal's Summation** must be completed and signed by the teacher and the principal. If 2 + 2 has been chosen by a teacher for summative evaluation, at least 90% of the required first semester observations must be completed by the end of January and the teacher must have demonstrated acceptable level of proficiency or removal from 2 + 2 is mandatory. **Summative** conference with the principal or designee will be completed in summative years by the May 30.

2 + 2 OPTIONS, ROLES AND GOALS

Options for Sharing results of observations
- paired 2 + 2
- group 2+2 by grade level, subject area, clusters, coaching
- a staff development focus
- small groups rotating during a session

Training for 2 + 2
- part of preservice staff development annually
- Building coordinator will provide coaching in techniques upon request by teacher
- options for sharing listed above may provide peer training

A School Administrator must make at least 3 visits, complete 2+2 form, and discuss observation with each teacher on summative evaluation. Department heads must complete at least 4 observations, complete 2+2 form and discuss observation with each teacher on summative evaluation. The school administrators and department heads must work with the building coordinator to support 2+2.

PRIME
2+2

2 + 2 TEACHER OBSERVATION REPORT

NORFOLK PUBLIC SCHOOLS

Teacher_____   Grade level/Subject_____ June, ____

School_____   Professional Development___ Summative ___

Professional Goals for the school year

_____

_____

_____

_____

2 + 2 Observations

Observations completed_____   Observations received_____   Portfolio completed___

Most Significant Compliments Received:

1.

2.

3.

4.

5.

Most significant suggestions received:

1.

2.

3.

4.

5.

2 + 2 Analysis

Explain which of the above compliments and suggestions were especially useful, and how they have led to reinforcement of your classroom teaching strategies.

Explain the most helpful ideas gleaned from observing colleagues.

Explain how observing your colleagues led to changes in your classroom teaching strategies.

Discuss how observing your colleagues enhanced/did not enhance your ability to grow professionally.

Teacher's Signature_____ date_____

PRINCIPAL'S SUMMATION
(Mandatory for first 3 years of teaching, every 4th year of continuing contract, new assignment to building)

_____(Teacher)

_____ (has/has not) demonstrated proficiency in the 2 + 2 observation process.
_____ (has/has not) demonstrated proficiency in knowledge of subject curriculum content.
_____ (has/has not) demonstrated proficiency in human relations and communication skills.
_____ (has/has not) demonstrated proficiency in professional responsibilities.
_____ (has/has not) demonstrated proficiency in planning for instruction.
_____ (has/has not) demonstrated proficiency in the management of student behavior.
_____ (has/has not) demonstrated proficiency in the delivery of instruction.
_____ (has/has not) demonstrated proficiency in student classroom achievement.

I agree _____ I do not agree _____ with the Principal's Summation

Teacher Signature	Date	Principal's Signature	Date

Table___

2 + 2 Program Survey Results: Percentage of Teachers who "Agreed" or "Strongly Agreed"* with the following statements:

Statements	Percent

Evaluation Alternatives
1. I prefer the 2+2 system to the traditional teacher evaluation system.	79%
2. The traditional evaluation system gave me the best feedback for my growth as a teacher.	13%
3. The 2+2 system is a better appraisal alternative than the traditional teacher evaluation system.	67%
4. I prefer the self evaluation summative form to the traditional teacher evaluation form.	61%
5. I like to have my teaching performance ranked by an administrator.	28%
6. The traditional evaluation system is the best system for evaluating teacher performance.	5%
7. Improvement of instruction is more likely to occur as a result of 2+2 than the traditional system.	72%
8. The 2+2 system does not lead to professional growth.	~~11%~~ 1.3%
9. The traditional Norfolk Public School evaluation system does not lead to professional growth.	14%

Outcomes of the 2 + 2 Program
1. As a result of 2+2 participation, I have experimented with new instructional strategies.	84%
2. During 2+2 observations, I have seen strategies I wanted to try with my own students.	87%
3. More training would make 2+2 feedback more useful.	55%
4. I have implemented at least one 2+2 suggestion.	96%
5. I have implemented strategies I observed in other classrooms.	85%
6. 2+2 observations have helped me gain perspective on my own teaching abilities.	83%
7. The 2+2 program has helped me gain confidence as a teacher.	65%
8. The quality of interaction among my colleagues has improved as a result of 2+2.	72%
9. 2+2 has had no impact on my professional growth.	7%

Programmatic Issues
1. Participation in 2+2 should remain voluntary.	78%
2. Only a few observations should be required during a teacher's non-summative years.	66%
3. The 2+2 system should be mandatory for all PRIME teachers during summative evaluation years.	30%
4. Student feedback is an important component of 2+2.	29%
5. Regular administrator observations are an important component of 2+2.	43%
6. The 2+2 system should be mandatory for all PRIME teachers, every year.	15%

Which specific changes would you make in the 2 + 2 program?
No comment =51.3%

Summary of responses	Number of observations	Percentage of respondents
Optimal number of observations during summative evaluation year per semester:	10	37%
	5	14%
	15	13%
Optimal number of observations during non-summative evaluation year per semester:	5	31%
	10	17%
	2	16%

* The Likert scale utilized in the survey: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree

April 29, 1997

Dear 2+2 Participant,

Please complete the following survey to provide feedback for the 2 + 2 Program for the next school year. Your views are very important for future program development, whether you were on a summative evaluation this year, or if you participated informally. The PRIME Steering Committee will be reviewing the 2 + 2 Program in May. Guidelines, based largely on your experiences and feedback, will be formally established in making progress toward institutionalization of the program. Thank you.

Is 1996-1997 your summative evaluation year?___yes   ___no
IF YES, did you opt for 2 + 2 in lieu of the Norfolk Public Schools Teacher Appraisal System?___yes___no
Approximate number of times you were observed during 1996-97 school year so far:____
Approximate number of observations you made 1996-1997 school year so far:_____

EVALUATION ALTERNATIVES - refers to the relative merits of the traditional Norfolk Public Schools (NPS) teacher evaluation system from 1983-1995/96, and the 2+2 Alternative Appraisal System. (If new to NPS, please respond based on your experience with the evaluation system in your former district.)

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
I prefer the 2+2 system to the traditional NPS teacher evaluation system.	____	____	____	____	____
The traditional NPS evaluation system gave me the best feedback for my growth as a teacher.	____	____	____	____	____
The 2+2 system is a better appraisal alternative than the traditional NPS teacher evaluation system.	____	____	____	____	____
I prefer the self evaluation summative form to the traditional NPS teacher evaluation form.	____	____	____	____	____
I like to have my teaching performance ranked by an administrator.	____	____	____	____	____
The traditional NPS evaluation system is the best system for evaluating teacher performance.	____	____	____	____	____
Improvement of instruction is more likely to occur as a result of 2+2 than the traditional system.	____	____	____	____	____
The 2+2 system does not lead to professional growth.	____	____	____	____	____
The Norfolk Public School evaluation system does not lead to professional growth.	____	____	____	____	____

OUTCOMES OF THE 2+2 PROGRAM - refers to the outcomes of 2+2 participation.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
As a result of 2+2 participation, I have experimented with new instructional strategies.	____	____	____	____	____
During 2+2 observations, I have seen strategies I wanted to try with my own students.	____	____	____	____	____

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
	1	2	3	4	5
More training would make 2+2 feedback more useful.	___	___	___	___	___
I have implemented at least one 2+2 suggestion.	___	___	___	___	___
I have implemented strategies I observed in other classrooms.	___	___	___	___	___
2+2 observations have helped me gain perspective on my own teaching abilities.	___	___	___	___	___
The 2+2 program has helped me gain confidence as a teacher.	___	___	___	___	___
The quality of interaction among my colleagues has improved as a result of 2+2.	___	___	___	___	___
2+2 has had no impact on my professional growth.	___	___	___	___	___

PROGRAMMATIC -refers to ways the 2+2 program might be modified to both institutionalize 2+2 and to broaden participation.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
	1	2	3	4	5
Participation in 2+2 should remain voluntary.	___	___	___	___	___
Only a few observations should be required during a teacher's non-summative years.	___	___	___	___	___
The 2+2 system should be mandatory for all PRIME teachers during summative evaluation years.	___	___	___	___	___
Student feedback is an important component of 2+2.	___	___	___	___	___
Regular administrator observations are an important component of 2+2.	___	___	___	___	___
The 2+2 system should be mandatory for all PRIME teachers, every year.	___	___	___	___	___

Which specific changes would you make in the 2 +2 program?

_____

_____

_____

_____

I believe the optimum number of observations in the summative year per semester should be: Per semester:_____

I believe the optimum number of observations in non-summative years should be: Per semester_____

APPENDIX E

INDIVIDUAL AND AGGREGATE RESULTS (N=50)
OF 2 + 2 OBSERVATION FORM ANALYSIS

Table E1. Content Categories for 2 + 2 Observation Forms

C = Compliments / S = Suggestions	1 C	1 S	2 C	2 S	3 C	3 S	4 C	4 S	5 C	5 S	6 C	6 S	7 C	7 S
OBJECTIVES	0	0	0	0	2	1	3	0	0	0	3	0	8	1
Learning objectives general					2	1	3				3		8	1
CURRICULUM	0	0	2	0	4	10	1	0	0	0	0	0	0	0
Curriculum general														
Builds on previous lessons			1		2									
Warm up			1		2	9	1							
Quality of examples					1									
Interdisciplinary study														
INSTRUCTIONAL STRAT.	8	5	3	2	25	13	21	10	14	7	18	5	19	7
Instructional strategies general	1		1		3	4	1		3	2	3	3	1	1
Relate to real world/Applications			1		1	1	1	1						
Clarity of instructions					1								1	1
Classroom displays						1					3			
Group study/Cooperative learning	1	1			4		1	1	3	1				
Classroom discipline/Strategies				1	3	3	1					1	7	3
Class participation	1	1			1	1	3	1	1			2		1
Planning and preparation					2		1		1		3		2	
Use students names/Rapport	1				2	1	5		3		2		3	
Positive feedback/Reinforcement				1	3		1				2			
Move around the room					1								1	
Use games				1										
Questioning	3	2	1		1	3	3	3	3					1
Peer helping														
Student engagement/On task	1				6		5	1	2	1	5		2	
PRESENTATION VARIABLES	0	0	0	0	4	3	6	0	2	0	0	0	0	2
Vivid/Captivating					1	2	1		1					
Presentation variables general					1									
Time management/Flow					3		5		1					2
Slow down														
Speak up														
MATERIALS	1	0	1	4	2	3	4	5	0	0	3	0	0	1
Materials general				3		2		4			1			
Visual aids	1			1	1	1	1	1						
Handouts				1										1
Technology					1		2							
Material preparation							1				2			
ASSESSMENT	3	2	1	0	2	2	4	0	0	3	2	0	4	1
Assessment general			1							1				
Student involvement				1						2	1		1	
Monitor student progress	3	1			2	2	4				1		3	1
NON PRODUCTIVE	0	3	1	2	0	6	0	20	1	4	1	17	0	12
Blank		3		2		6		11		4		11		6
Continue teaching style								9				6		6
Good job (blowing sunshine)			1							1	1			
TOTAL Compliments/Suggestions	12	10	8	8	39	38	39	35	17	14	27	22	31	24
NUMBER OF FORMS	5	5	4	4	18	18	16	16	7	7	12	12	12	12

Note. * Numbers have been substituted for the names of individual observers

Table E1. (Continued)

C = Compliments / S = Suggestions	8		9		10		11		12		13		14	
	C	S	C	S	C	S	C	S	C	S	C	S	C	S
OBJECTIVES	0	1	4	0	0	1	1	0	0	0	0	0	4	0
Learning objectives general		1	4				1	1					4	
CURRICULUM	0	1	0	1	0	0	0	2	1	0	0	0	0	2
Curriculum general														
Builds on previous lessons														
Warm up			1				1							1
Quality of examples		1					1							1
Interdisciplinary study									1					
INSTRUCTIONAL STRAT.	8	7	15	9	3	3	15	11	2	1	6	1	32	23
Instructional strategies general	1	1					1		5				4	5
Relate to real world/Applications	1			1			1						3	
Clarity of instructions				5										
Classroom displays			1								1			
Group study/Cooperative learning			1	1			1		2			1		2
Classroom discipline/Strategies	2	5	2	1			1	3			1		5	9
Class participation			1	2			1	1			1		3	4
Planning and preparation					1		2						2	
Use students names/Rapport	1						3						2	
Positive feedback/Reinforcement	2				1				2				4	1
Move around the room													2	
Use games			1				1						1	
Questioning			1	2			1	1	2		1	1	1	1
Peer helping			1				1						1	1
Student engagement/On task	1	1	4			1	1	1			2		4	
PRESENTATION VARIABLES	0	4	0	3	1	0	1	0	1	1	0	0	2	2
Vivid/Captivating			1										1	
Presentation variables general		2					1							
Time management/Flow		1	2	1					1	1			1	2
Slow down														
Speak up		1												
MATERIALS	1	4	0	5	1	0	0	4	0	0	0	2	1	0
Materials general				1				4				2		
Visual aids	1	3	2											
Handouts		1			1									
Technology			2											
Material preparation													1	
ASSESSMENT	3	1	2	1	1	0	0	0	0	0	0	1	0	0
Assessment general	1											1		
Student involvement														
Monitor student progress	2	1	2	1	1									
NON PRODUCTIVE	2	0	1	2	0	2	0	1	0	2	0	2	1	14
Blank						2								4
Continue teaching style			2											
Good job (blowing sunshine)	2		1					1		2		2	1	10
TOTAL Compliments/Suggestions	14	18	22	21	6	6	17	18	4	4	6	6	40	41
NUMBER OF FORMS	7	7	11	11	3	3	9	9	2	2	3	3	20	20

Table E1. (Continued)

C = Compliments / S = Suggestions	15 C	15 S	16 C	16 S	17 C	17 S	18 C	18 S	19 C	19 S	20 C	20 S	21 C	21 S
OBJECTIVES	1	0	2	1	0	0	0	0	0	0	1	0	0	0
Learning objectives general	1		2	1							1			
CURRICULUM	1	0	1	0	0	1	1	0	5	0	0	0	2	0
Curriculum general						1			2					
Builds on previous lessons													1	
Warm up	1		1				1							
Quality of examples									3				1	
Interdisciplinary study														
INSTRUCTIONAL STRAT.	7	10	14	9	5	5	6	9	31	23	2	3	10	6
Instructional strategies general		1	1	2		3		2	2	5	2	1	4	2
Relate to real world/Applications									1				1	
Clarity of instructions		1	4				1		3	1				
Classroom displays		2											1	1
Group study/Cooperative learning			1	1	1				1	4				
Classroom discipline/Strategies		2	1	4			1	2	4	6			1	1
Class participation	2			1			1	2		4	1			1
Planning and preparation	2					1			1	1	1	2		
Use students names/Rapport	1	1	1							1				
Positive feedback/Reinforcement		2							1	2				
Move around the room									1					1
Use games				1										
Questioning		1					1	1	7	2				
Peer helping						2				1				
Student engagement/On task	2		6		3		2		8				1	
PRESENTATION VARIABLES	0	0	0	0	0	0	1	1	2	2	0	0	2	1
Vivid/Captivating													1	
Presentation variables general														
Time management/Flow							1	1	2	2			1	
Slow down														1
Speak up														
MATERIALS	0	0	2	2	0	0	1	0	1	1	0	1	1	0
Materials general			1	2										
Visual aids							1		1	1			1	1
Handouts			1											
Technology														
Material preparation														
ASSESSMENT	1	0	0	3	3	0	2	0	4	1	0	0	3	2
Assessment general					1		1			1			3	2
Student involvement							1		1					
Monitor student progress	1			3	2				3					
NON PRODUCTIVE	0	0	2	6	0	2	0	2	0	12	1	0	0	9
Blank				6		1		2		1				7
Continue teaching style										11				1
Good job (blowing sunshine)			2			1					1			1
TOTAL Compliments/Suggestions	10	10	21	21	8	8	11	12	43	39	4	4	18	18
NUMBER OF FORMS	5	5	10	10	4	4	5	5	19	19	2	2	9	9

Table E1. (Continued)

C = Compliments / S = Suggestions	22		23		24		25		26		27		28	
	C	S	C	S	C	S	C	S	C	S	C	S	C	S
OBJECTIVES	0	0	0	0	0	0	0	0	0	1	0	1	0	0
Learning objectives general										1		1		
CURRICULUM	0	1	2	0	1	1	0	1	2	0	1	0	0	0
Curriculum general									1					
Builds on previous lessons				1										
Warm up			2		1				1		1			
Quality of examples									1					
Interdisciplinary study		1												
INSTRUCTIONAL STRAT.	7	5	2	6	10	14	8	6	9	4	5	1	3	1
Instructional strategies general				2	2	4	1	2						
Relate to real world/Applications	1					1			1	1	1			
Clarity of instructions									1					1
Classroom displays									1					
Group study/Cooperative learning		1							1					
Classroom discipline/Strategies	1	1	1	2	2	3		1	3	3				
Class participation		1			1	2		1	1					
Planning and preparation				1		1	1				1			
Use students names/Rapport	3				1	1	1		1			1		
Positive feedback/Reinforcement	1				1						1		1	
Move around the room					1	1					1			
Use games														
Questioning		1			2	1	1				1	1		
Peer helping		1												
Student engagement/On task	1				1	1	4		2			1		
PRESENTATION VARIABLES	0	1	4	5	1	0	1	0	0	0	3	2	0	0
Vivid/Captivating											1			
Presentation variables general				1	1									
Time management/Flow		1	3	1							2	1		
Slow down				3								1		
Speak up				1				1						
MATERIALS	2	1	2	3	1	1	1	2	0	2	0	0	0	0
Materials general	1				1									
Visual aids		1	1	3	1		1	2	1					
Handouts	1													
Technology			1						1					
Material preparation														
ASSESSMENT	1	0	0	1	1	2	0	0	0	0	3	0	0	1
Assessment general				1		1								
Student involvement														
Monitor student progress	1				1	1					3			1
NON PRODUCTIVE	0	1	1	0	2	0	4	5	4	9	1	8	1	2
Blank							4	5	1	9		8		
Continue teaching style														2
Good job (blowing sunshine)		1	1		2					3	1		1	
TOTAL Compliments/Suggestions	10	9	11	15	16	18	14	14	15	16	13	12	4	4
NUMBER OF FORMS	4	4	6	6	8	8	7	7	8	8	6	6	2	2

218

Table E1. (Continued)

C = Compliments / S = Suggestions	29		30		31		32		33		34		35	
	C	S	C	S	C	S	C	S	C	S	C	S	C	S
OBJECTIVES	0	0	0	0	0	0	0	0	0	0	0	0	3	2
Learning objectives general													3	2
CURRICULUM	0	0	1	0	0	0	1	1	0	0	1	1	1	1
Curriculum general														
Builds on previous lessons							1							
Warm up						1					1			1
Quality of examples			1									1	1	
Interdisciplinary study														
INSTRUCTIONAL STRAT.	2	1	4	5	9	2	12	13	10	7	2	2	11	4
Instructional strategies general			1	1	1		1	4	2	2		2	2	1
Relate to real world/Applications									1				1	
Clarity of instructions				1						1		1	2	
Classroom displays										1				
Group study/Cooperative learning			1		1					1	1		1	
Classroom discipline/Strategies				2	3	2	4	3	1					1
Class participation							1	1	1					
Planning and preparation	1						1							
Use students names/Rapport	1				1	2	1	1						1
Positive feedback/Reinforcement							1					1		
Move around the room							1	3	1					
Use games														
Questioning			1						1		1	1	2	1
Peer helping				1							1			
Student engagement/On task			1		2		2		3				2	
PRESENTATION VARIABLES	1	0	1	0	0	0	3	4	1	1	0	0	1	0
Vivid/Captivating	1		1				1		1				1	
Presentation variables general														
Time management/Flow							1	4		1				
Slow down														
Speak up							1							
MATERIALS	0	1	2	0	0	1	2	1	0	4	0	1	1	1
Materials general								1		1	3			
Visual aids			1				1			1		1	1	1
Handouts			1	1										
Technology							1							
Material preparation														
ASSESSMENT	0	0	2	0	2	0	1	1	1	0	1	0	3	2
Assessment general			1											
Student involvement										1		1	1	
Monitor student progress			1		2		1	1					2	2
NON PRODUCTIVE	1	2	0	6	3	11	3	2	0	0	0	0	0	11
Blank		2		6	1	11		2						10
Continue teaching style														1
Good job (blowing sunshine)	1				2		3							
TOTAL Compliments/Suggestions	4	4	10	11	14	14	22	22	12	12	4	4	20	21
NUMBER OF FORMS	2	2	5	5	7	7	11	11	6	6	2	2	10	10

Table E1. (Continued)

C = Compliments / S = Suggestions	36 C	36 S	37 C	37 S	38 C	38 S	39 C	39 S	40 C	40 S	41 C	41 S	42 C	42 S
OBJECTIVES	0	0	1	1	0	0	0	0	0	0	0	0	1	1
Learning objectives general			1	1									1	1
CURRICULUM	6	0	1	0	0	1	0	0	0	0	0	1	3	0
Curriculum general			1											
Builds on previous lessons	2										1			
Warm up													2	
Quality of examples	3					1							1	
Interdisciplinary study	1													
INSTRUCTIONAL STRAT.	17	14	1	3	10	10	5	5	1	1	20	5	6	6
Instructional strategies general	2	4				1	1	1		1	2	1	1	2
Relate to real world/Applications											1			
Clarity of instructions								2						2
Classroom displays						1			1					
Group study/Cooperative learning	1	2			2	2	1	1			2	2		
Classroom discipline/Strategies	4	4			3	2	2							1
Class participation	1	1					1		1		5	1		
Planning and preparation	2												1	
Use students names/Rapport	2					2	1				3	2		
Positive feedback/Reinforcement	2													
Move around the room					1									
Use games									1		2			
Questioning		2				1					1		1	1
Peer helping		1				1	2				1	1		
Student engagement/On task	3			1		1	1			1	3		1	
PRESENTATION VARIABLES	2	5	1	0	0	1	0	2	1	0	1	0	0	2
Vivid/Captivating														
Presentation variables general														1
Time management/Flow	2	3						1	1	1	1			1
Slow down				1					1					
Speak up				1	1									
MATERIALS	4	7	0	0	1	1	2	0	0	0	0	2	0	2
Materials general	2					1						1		
Visual aids	2	5				1	1							2
Handouts		2					1							
Technology														
Material preparation											1			
ASSESSMENT	2	1	0	0	0	0	1	0	1	0	0	0	1	1
Assessment general														1
Student involvement							1							
Monitor student progress	2	1							1				1	
NON PRODUCTIVE	2	5	0	0	1	0	0	1	1	3	1	14	0	1
Blank		5								1	3	12		1
Continue teaching style	1											1		
Good job (blowing sunshine)	1					1				1	1	1		
TOTAL Compliments/Suggestions	33	32	4	4	12	13	8	8	4	4	22	22	11	13
NUMBER OF FORMS	16	16	2	2	6	6	4	4	2	2	11	11	5	5

Table E1. (Continued)

C = Compliments / S = Suggestions	43		44		45		46		47		48	
	C	S	C	S	C	S	C	S	C	S	C	S
OBJECTIVES	0	1	0	0	0	0	0	0	0	0	2	0
Learning objectives general		1									2	
CURRICULUM	1	2	2	1	0	0	0	1	1	0	0	0
Curriculum general												
Builds on previous lessons		1	1									
Warm up	1			2					1	1		
Quality of examples		1										
Interdisciplinary study												
INSTRUCTIONAL STRAT.	5	6	20	14	4	0	10	8	6	1	9	11
Instructional strategies general	3	3	1	4			2	1	1		2	4
Relate to real world/Applications			1	2			1					
Clarity of instructions				3			2		2			
Classroom displays				1								
Group study/Cooperative learning								2			1	
Classroom discipline/Strategies	2	5	2	1			1	1		1	2	5
Class participation		1		2				1				1
Planning and preparation				1					1		1	
Use students names/Rapport	1			1			1					
Positive feedback/Reinforcement				1	1							
Move around the room	1											
Use games												
Questioning			1	2					2	1		1
Peer helping												
Student engagement/On task			5	2	2		4			1	3	
PRESENTATION VARIABLES	1	2	0	3	0	0	2	1	0	0	4	1
Vivid/Captivating												
Presentation variables general			1									
Time management/Flow							2	1			2	
Slow down				1								1
Speak up	1	1		2							2	
MATERIALS	4	1	1	4	1	0	0	1	0	1	1	1
Materials general				1					1			1
Visual aids	3	1	1	3				1			1	
Handouts	1			1								
Technology												
Material preparation												
ASSESSMENT	5	2	1	1	1	0	0	0	0	0	0	1
Assessment general				1								
Student involvement												
Monitor student progress	5	2	1	1								1
NON PRODUCTIVE	0	0	2	3	0	4	0	1	0	4	1	2
Blank				1		2				4		2
Continue teaching style				1		2						
Good job (blowing sunshine)			2	1					1		1	
TOTAL Compliments/Suggestions	16	14	26	26	6	4	12	12	7	6	17	16
NUMBER OF FORMS	7	7	13	13	2	2	6	6	3	3	8	8

Table E1. (Continued)

C = Compliments S = Suggestions	49		50	
	C	S	C	S
OBJECTIVES	2	0	2	0
Learning objectives general	2		2	
CURRICULUM	2	0	2	0
Curriculum general				
Builds on previous lessons			1	
Warm up	1		1	
Quality of examples	1			
Interdisciplinary study				
INSTRUCTIONAL STRAT.	7	7	4	0
Instructional strategies general				
Relate to real world/Applications				
Clarity of instructions		1		
Classroom displays				
Group study/Cooperative learning		2		
Classroom discipline/Strategies				
Class participation	1			
Planning and preparation	1			
Use students names/Rapport				
Positive feedback/Reinforcement			1	
Move around the room			2	
Use games				
Questioning	2	3		
Peer helping		1		
Student engagement/On task	3		1	
PRESENTATION VARIABLES	0	1	0	0
Vivid/Captivating				
Presentation variables general				
Time management/Flow		1		
Slow down				
Speak up				
MATERIALS	0	1	0	0
Materials general		1		
Visual aids				
Handouts				
Technology				
Material preparation				
ASSESSMENT	0	1	0	0
Assessment general				
Student involvement		1		
Monitor student progress				
NON PRODUCTIVE	0	1	1	8
Blank		1		8
Continue teaching style				
Good job (blowing sunshine)			1	
TOTAL Compliments/Suggestions	11	11	9	8
NUMBER OF FORMS	6	6	4	4

Compliments by Individual Category
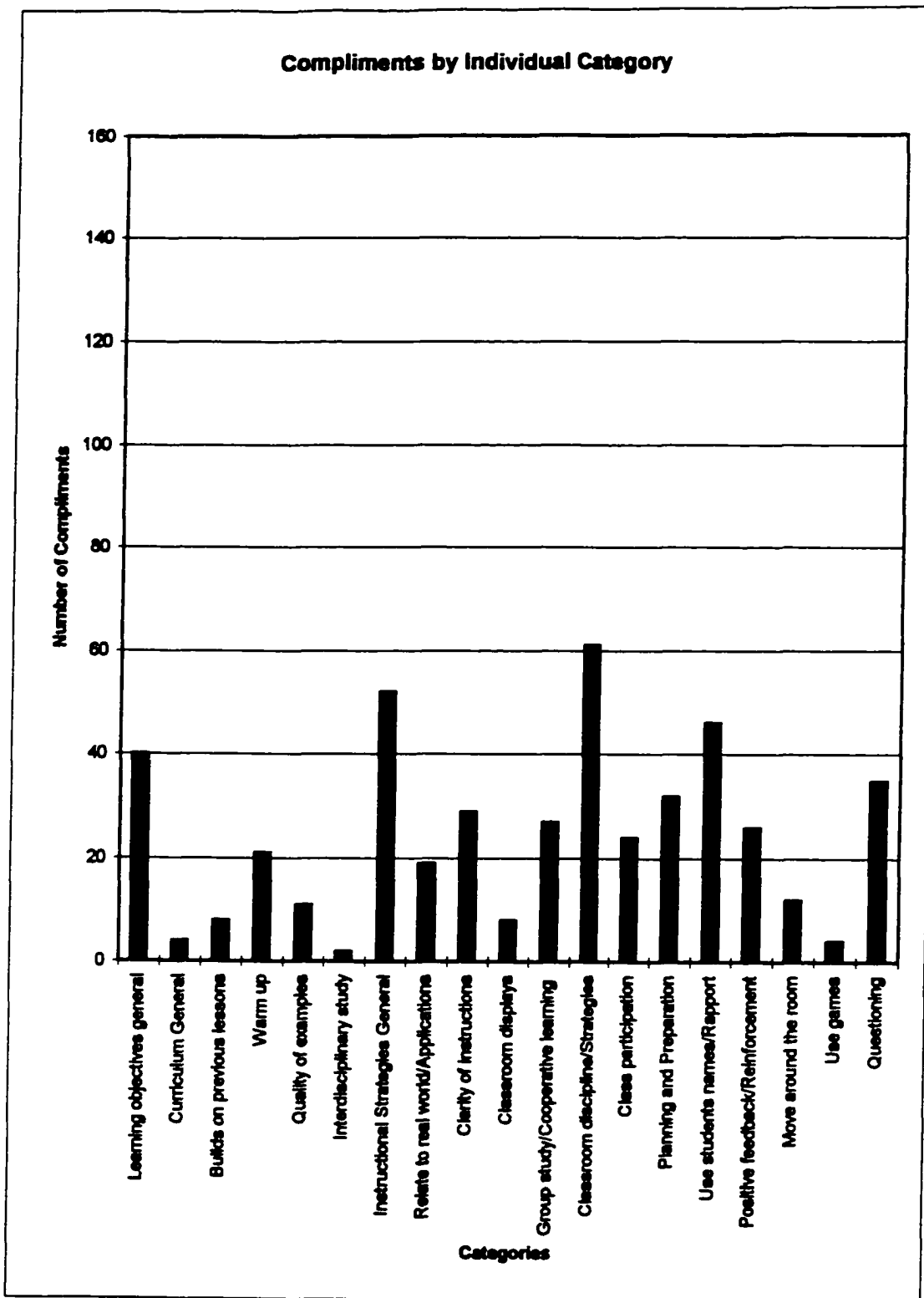
Figure E1. Compliments by individual categories

Compliments by Individual Category (Continued)
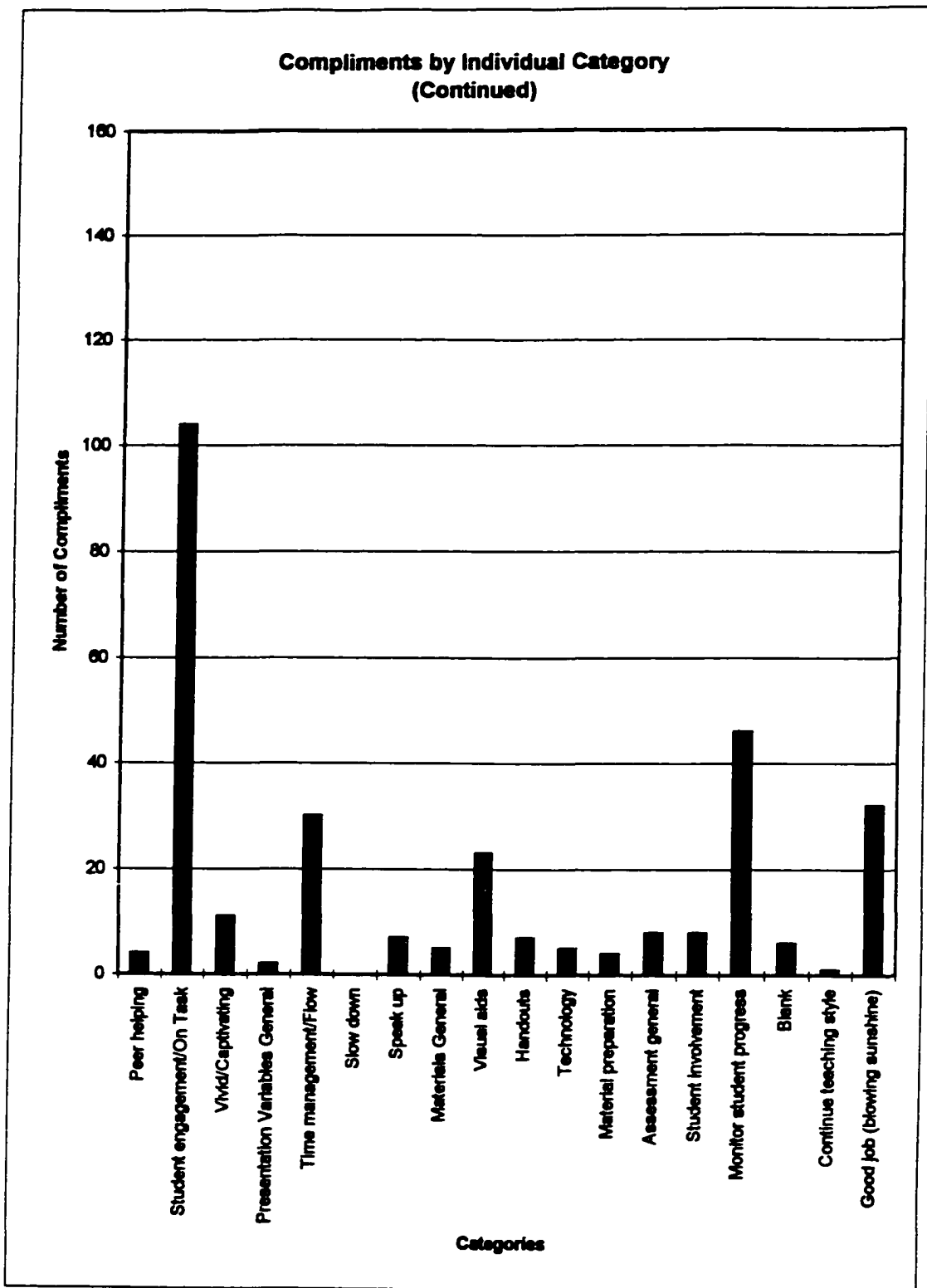
Figure E1. Continued
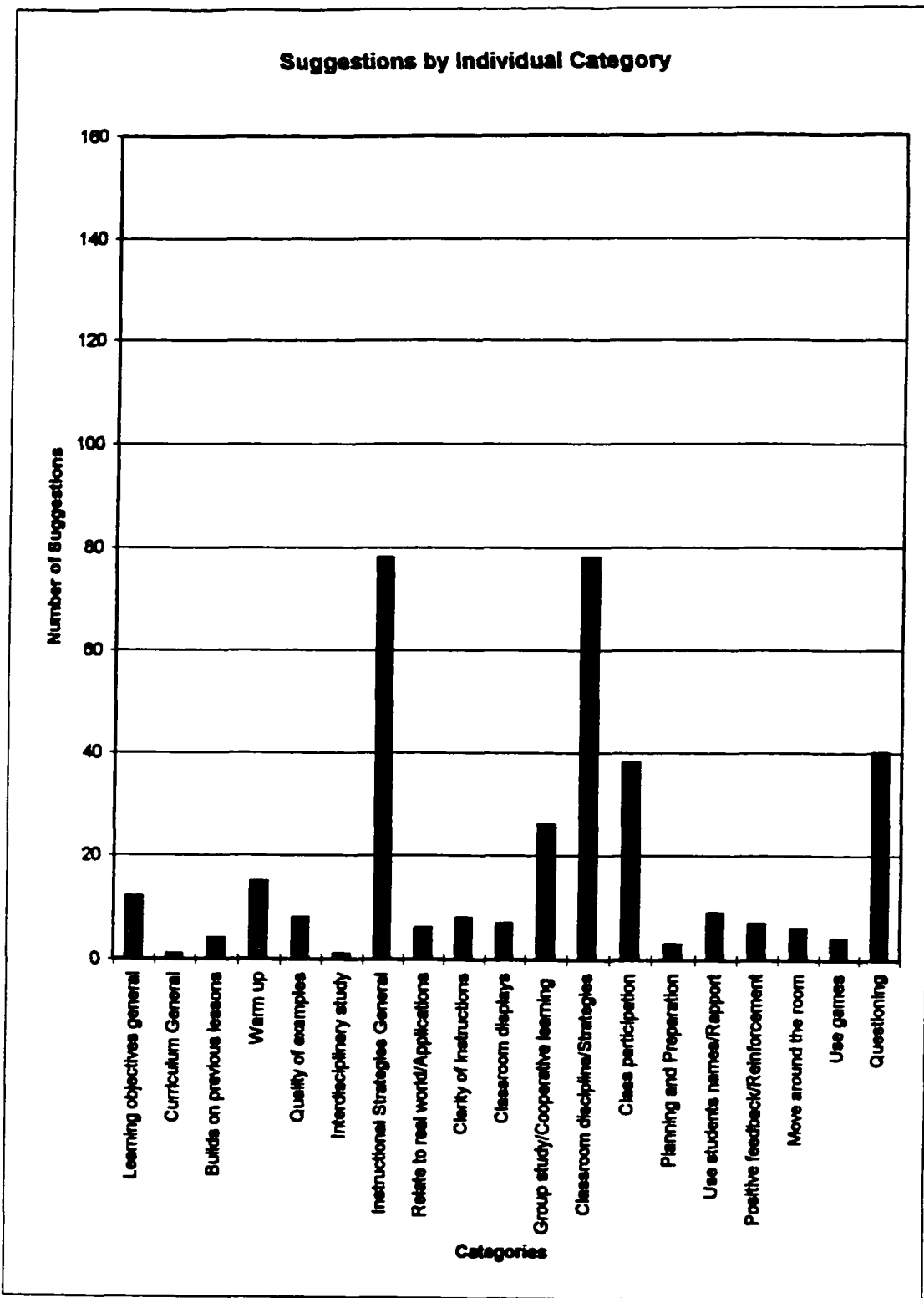
Figure E2. Suggestions by individual categories

Suggestions by Individual Category (Continued)
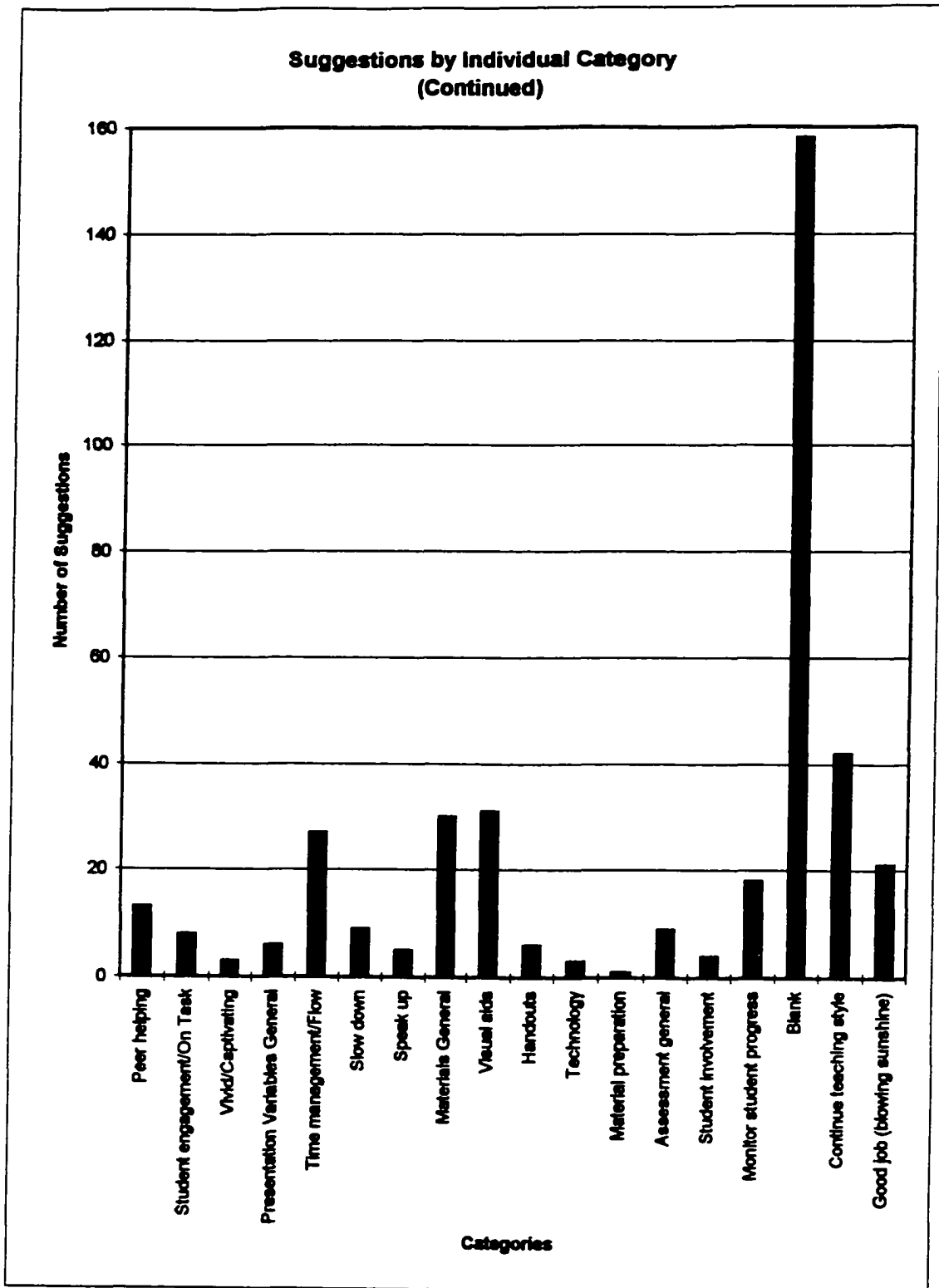
Figure E2. Continued
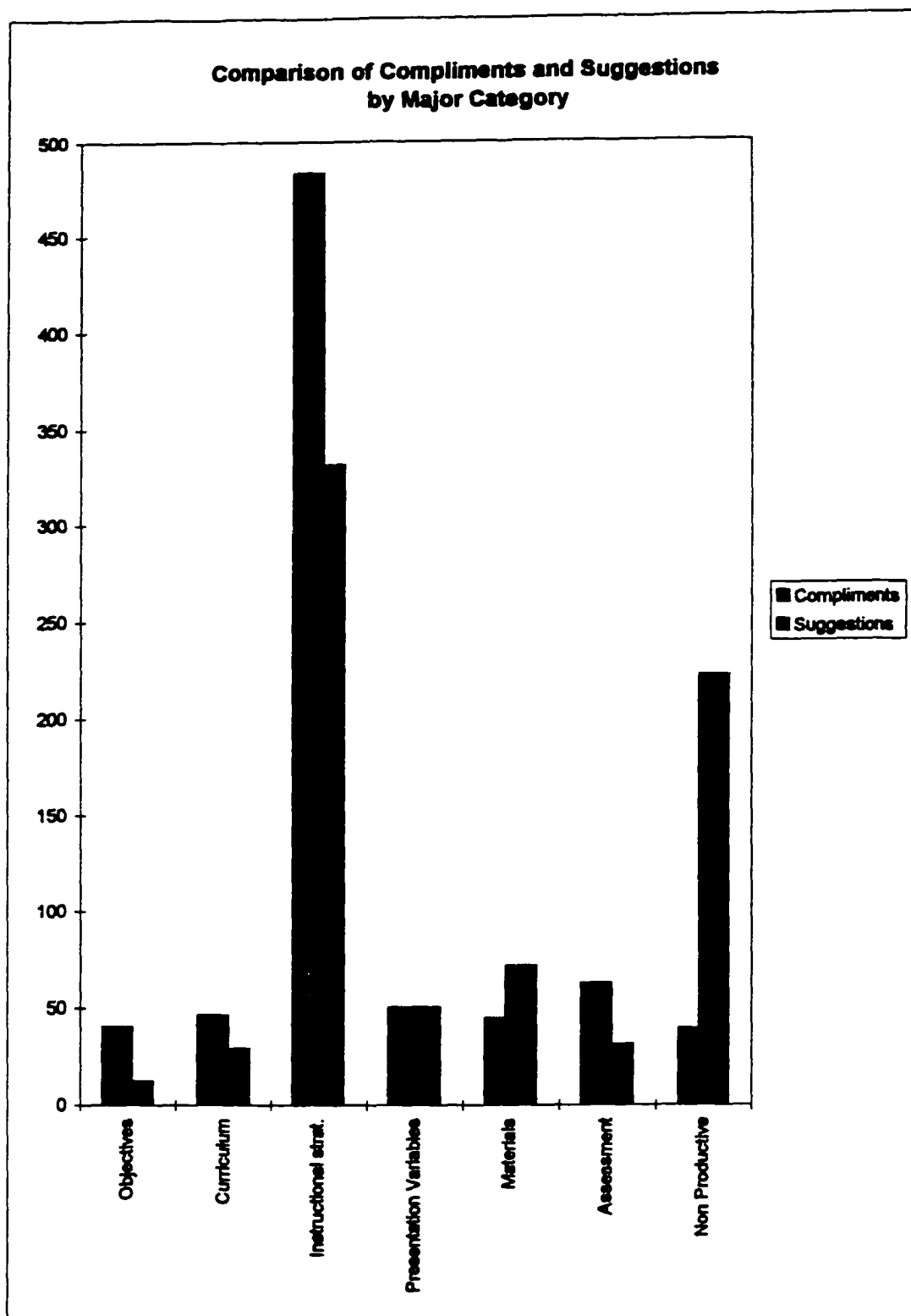
Figure E3. Comparison of compliments and suggestions by major category

APPENDIX F

SAMPLE SELF-REFLECTION REPORT FORM
AND COMPREHENSIVE RESULTS OF SELF-REFLECTION REPORT FORMS
ANALYSIS

2+ 2 Summative Teacher Report Form

Norfolk Public Schools
Fall Semester 1996

Teacher Name:_____

School:_____Grade Level/Subject Area:_____

Most Useful/Significant Compliments Received:

1.

2.

3.

4.

5.

6.

7.

8.

9.

10.

2 + 2 Summative Report (page 2)

Most Useful/Significant Suggestions Received:

1.

2.

3.

4.

5.

6.

7.

8.

9.

10.

2 + 2 Summative Report (page 3)
(Use reverse side if necessary.)

Analysis:

Explain which of the above compliments and suggestions were especially useful, and how they have led to reinforcement or improvement of your classroom teaching routines.

List any future agenda items which have emerged as a result of the 2 + 2 process. Please add any reflections about the entire 2 + 2 observation process, including your assessment of the value of observing others.

Teacher's Signature_____ Date_____

Principal's Signature_____ Date_____

(To be maintained in employee's permanent personnel file.)

Self Reflection Reports

Combined Middle School Responses (N = 10)

Analysis, Reflections and Observations

10 OBSERVING OTHER TEACHERS PROVIDES INSIGHT AND PERSPECTIVE: Found that teachers observed had excellent strengths. Enjoyed seeing others at work. It brought me great pleasure to contribute to the growth of my colleagues. Valuable to see how others handle situations. I have learned many techniques from my observations of other teachers - and try to use them in a variety of ways. It is valuaable to "assess" fellow teachers to gain new ideas and helpful hints. Observing others allows you to take the best and use it. When you observe someone else, you always learn something

9 2+2 PROVIDES SPECIFIC SUGGESTIONS TO CHANGE INSTRUCTIONAL PRACTICE: I'm now using the blackboard more effectively. I think more about how I talk with my students. Realization of weaknesses in questioning techniques. I began posting charts that kept track of students' success and giving rewards and reinforcement. When the teacher who made the suggestion complimented my changes I was extremely pleased. Acting on a 2+2 suggestion, I allowed more time for students to come to the board and display their responses. This helped to increase class participation, especially from those students who do not usually raise their hands. I gained a potpourri of ideas and strategies to expand my instructional program. Appreciate the specific methods gained from 2+2 observations I can use to improve my teaching. Assigning readers before the reading lesson is a wonderful suggestion. I do this now and the reading flows much more smoothly. I have students critique my model paper. It helps the editing step in writing. Incorporating more writing in LPT preparation has been very effective as students were able to create their own stories to share with their classmates

7 ENCOURAGED BY COMPLIMENTS AND PRAISE FROM 2+2: I appreciated the compliments and suggestions made by my peers. I was pleased to see how my colleagues noticed the effective learning environment in my classroom. Suggestions have made me a better teacher. I was encouraged when observers remarked very positively on my efforts. Each evaluation showed me that I am doing my job well. I enjoy the pat on the shoulder

4 2+2 HELPED MAKE LESSONS MORE STUDENT CENTERED: Now that I'm more aware, I feel I can make my lessons more student centered. Though I'm 100% dedicated to student led lessons, I realized that my

lessons remain too teacher centered. I have built an effective team monitoring system as a result of suggestions

4 2+2 ENCOURAGES TEAMWORK AMONG TEACHERS: Improving instruction is a team effort, and to work as a team in order to improve instruction. 2+2 has impacted the professional growth of teachers

2 2+2 HAS HELPED ME WITH THE APPLICATION OF KNOWLEDGE: I make more effort to make sure my students realize the significane of what they are learning and how it applies to life

2 SOME 2+2 SUGGESTIONS AREN'T USEFUL, BUT THEY ARE EASILY DISREGARDED: The suggestion to use tables is not feasible. One of the suggestions I had already used the previous day

2 2+2 observations reinforced my belief in a positive instructional climate. I find 2+2 to be both comfortable and informative

Difficult to make "accurate observations" when few weaknesses are noted

Time for observations was a major problem

I want to do more 2+2s

Perhaps add to the form: "One thing I have learned from observing your class: "

I like 2+2 observations - they are non-threatening

Students are now more relaxed about taking notes in my class

The pressure of being observed is not as burdensome and consequently allows for a more accurate observation

I plan to observe more teachers - and those not involved in the program

Perhaps next year more teachers will be involved

Good reminder to call on more students

Compliments and suggestions have reminded me what teachers and administrators expect in a classroom

Useful to learn others expectations - they want more quiet - a wake up call for my planning

I wish more people could observe me. I would love to get new ideas from other teachers

Unfortunately many of the teachers considered "master teachers" are not in 2+2 so I wasn't able to observe them

I need to prioritize getting my observations done. Some release time would be really great

Most significant compliments

8 Keeping students on task, actively involved

4 excellent rapport with students

4 effective student participation - all students involved

4 excellent examples, creative examples

3 good atmosphere for learning

3 clear objectives, expectations clear and effective

3 good cooperative learning, group skills, good peer assistance in learning

3 effective monitoring of student progress
2 TESA skills
2 good questioning skills
2 use of visual aids
2 clear directions
2 positive reinforcement, upbeat atmosphere

excellent classroom management
lively
good accountability for students
effective use of simple activity - making it different and enjoyable
use of graphics organizer
extra credit chart
multiple modalities
effective ruberics
Students were allowed to give input
Lesson plans were well organilzed and clear
Good transitions between lessons
Modeled instructions while giving oral directions
Complimented brainstorming technique
Calling on students to get their attention is effective
Effective individualized instruction
Assisted students in sounding out difficult words
Allowed students to display prior knowledge during introduction
Effective movement around classroom to offer assistance
Dynamic presentations
Compliments on techniques used to practice for the LPT

Most significant suggestions

5 better use of av aids, use black markers and print bigger, invite students to
    come to the board use full space of blackboard
4 increase use of small groups - suggestions for improving them, make non-
    participating students more accountable in team work
3 more student involvement in instruction
2 better questions, stop and insert questions in presentation
2 procedures for processing hand/outs and hand/ins
2 make sure directions are understood
2 suggestions for handling class projects
2 more wait-time
2 More oral questioning

Cut out additional newspaper articles in advance to avoid wasting time
Move students to reduce talking
Have students write answers on board or overhead
Provide warm-up activity to get students settled
More discussion about stories
Incorporate a writing process

Have students critique your writing model
Be more specific in setting group goals
Have a more detailed lesson
Model instruction while giving directions
Assign readers in the beginning of the reading lesson
Clarify the homework assignment - be more specific
Good specific suggestion to color-code a paragraph
better storage of book bags etc.
suggestions how to involve uninvolved students
devise penalties for inattention
each student could be given an organizer to avoid visual transerence from board
     to paper
provide handout for reference until objectives are learned
suggestions for improved physical arrangements

Lake Taylor Middle (N = 5)

Most significant compliments

4 excellent rapport with students
3 effective student participation - all students involved
3 good atmosphere for learning
2 excellent time on task
2 creative examples - good examples
2 use of visual aids
2 TESA skills - (one "by accident - didn't know I was using")
2 good questioning skills
excellent classroom management
positive reinforcement
lively
clear objectives
expectations clear and effective
good accountability for students
effective use of simple activity - making it different and enjoyable
use of graphics organizer
extra credit chart
multiple modalities
good cooperative learning, group skills
effective ruberics
clear directions

Most significant suggestions

2 better use of av aids
2 stop and insert questions in presentation - better questions
2 more student involvement in instruction
2 increase use of small groups - suggestions for improving them
2 procedures for processing hand/outs and hand/ins

black markers and print bigger
invite students to come to the board
query students' understanding of directions
better storage of book bags etc.
make sure directions are understood
suggestions how to involve uninvolved students
devise penalties for inattention
each student could be given an organizer to avoid visual transerence from board
     to paper
provide handout for reference until objectives are learned
several suggestions for handling class projects
suggestions for more wait-time
use full space of blackboard
make non-participating students more accountable in team work
suggestions for improved physical arrangements

Analysis:

I think more about how I talk with my students
Realization of weaknesses in questioning techniques
I have built an effective
 team monitoring system as a result of suggestions
2 I appreciated the compliments and suggestions made by my peers
Though I'm 100% dedicated to student led lessons, I realized that my lessons
     remain too teacher centered
Now that I'm ore aware, I feel I can make my lessons more student centered
I'm now using the blackboard more effectively
Students are now more relaxed about taking notes in my class
I was pleased to see how my colleagues noticed the effective learning
     environment in my classroom
2+2 observations reinforced my belief in a positive instructional climate
I gained a potpourri of ideas and strategies to expand my instructional program
I bega posting charts that kept track of students' success and giving rewards
     and reinforcement. When the teacher who made the suggestion
     complimented my changes I was extremely pleased
I make more effort to make sure my students realize the significane of what they
     are learning and how it applies to life
I was encouraged when observers remarked very positively on my efforts
Acting on a 2+2 suggestion, I allowed more time for students to come to the
     board and display their responses. This helped to increase class
     participation, especially from those students who do not usually raise their
     hands

Reflections on observations:

Found that teachers observed had excellent strengths
Enjoyed seeing others at work
Difficult to make "accurate observations" when few weaknesses are noted

Time for observations was a major problem
I want to do more 2+2s
Perhaps add to the form: "One thing I have learned from observing your class: "
I like 2+2 observations - they are non-threatening
Suggestions have made me a better teacher
2+2 has impacted the professional growth of teachers at LTMS
It brought me great pleasure to contribute to the growth of my colleagues
Improving instruction is a team effort, and to work as a team in order to improve
      instruction
I find 2+2 to be both comfortable and informative
The pressure of being observed is not as burdensome and consequently allows
      for a more accurate observation
I enjoy the pat on the shoulder
Appreciate the specific methods gained from 2+2 observations I can use to
      improve my teaching
Valuable to see how others handle situations
I plan to observe more teachers - and those not involved in the program
Perhaps next year more teachers will be involved

Azalea Gardens (N = 5)

Most significant compliments

4 Students actively involved
2 Keeping students on task
2 Excellent examples
2 Effective monitoring of student progress
Students were allowed to give input
Lesson plans were well organilzed and clear
Effective monitoring of students
Good transitions between lessons
Clear objectives
Good specific suggestion to color-code a paragraph
Modeled instructions while giving oral directions
Explicit directions
Complimented brainstorming technique
Calling on students to get their attention is effective
Effective individualized instruction
Good group learning
Good peer assistance in learning
Assisted students in sounding out difficult words
Allowed students to display prior knowledge during introduction
Everyone was involved
Effective movement around classroom to offer assistance
Upbeat atmosphere
Dynamic presentations
Compliments on techniques used to practice for the LPT

Most significant suggestions

2 More oral questioning
Cut out additional newspaper articles in advance to avoid wasting time
Move students to reduce talking
Have students write answers on board or overhead
Provide warm-up activity to get students settled
Involve all students in some facit of the lesson
More discussion about stories
Incorporate a writing process
Have students critique your writing model
Be more specific in setting group goals
Have a more detailed lesson
Use more cooperative learning groups
Model instruction while giving directions
Assign readers in the beginning of the reading lesson
Clarify the homework assignment - be more specific

Analysis:

Assigning readers before the reading lesson is a wonderful suggestion. I do this
    now and the reading flows much more smoothly
I have students critique my model paper. It helps the editing step in writing
Good reminder to call on more students
Incorporating more writing in LPT preparation haws been very effective as
    students were able to create their own stories to share with their
    classmates
Compliments and suggestions have reminded me what teachers and
    administrators expect in a classroom
Useful to learn others expectations - they want more quiet - a wake up call for my
    planning
Each evaluation showed me that I am doing my job well
The suggestion to use tables is not feasible
One of the suggestions I had already used the previous day

Reflections on observations:

I have learned many techniques from my observations of other teachers - and try
    to use them in a variety of ways
It is valuaable to "assess" fellow teachers to gain new ideas and helpful hints
Observing others allows you to take the best and use it

I wish more people could observe me. I would love to get new ideas from other
teachers

Unfortunately many of the teachers considered "master teachers" are not in 2+2
so I wasn't able to observe them

I need to prioritize getting my observations done. Some release time would be
really great

When you observe someone else, you always learn something

Lake Taylor High School Interns   (N = 18)

Analysis:

18 2+2 OBSERVATIONS OF MY CLASS HELP ME IMPROVE From 2+2
evaluations of my class I've been able to come up with ideas for my
classroom, to change my behavior, to know what works. When you
use something learned in 2+2 there is an awakening. Compliments
and suggestions have led to reinforcement/improvement of my
classroom teaching routines. I feel more confident about my
teaching ability. I've gathered many excellent ideas from 2+2.
Specific practices have changed: making students raise hands, listen
to each other, avoid choral responses, revised questioning
procedures, developed peer tutoring, put in more "fun activities."
At the beginning, discipline was my biggest fear. Observing and
being observed I have been able to put that worry at ease. 2+2
helped me revamp and revise classroom management. I am
pleasantly surprised by the evaluations of others, an essential tool in
helping an educator evaluate his/her strengths and weaknesses.
The professional yet personal touch of the observation process is a
welcome event to an educator. Makes you aware of where you
need improvement.

12 2+2 OBSERVATIONS OF OTHER TEACHERS ARE HELPFUL AND
ENJOYABLE. I've enjoyed observing other teachers. Observing
others helps you think of many things you would like to try.

8 2+2 LETS ME KNOW WHAT I'M DOING RIGHT.

8 OBSERVING OTHERS GIVES ME MORE CONFIDENCE TO EVALUATE
MYSELF. Builds confidence I know I'm using the right methods
for my particular students

4 I'VE LEARNED THAT THERE ARE MANY DIFFERENT WAYS TO
HANDLE A SITUATION. Variety keeps students actively involved.

4 2+2 REMINDS ME TO REINFORCE/PRAISE MY STUDENTS
CONTINUOUSLY The compliments remind me to continue to focus
on praising students and giving reasons for particular activities.
Students need/enjoy continuous praise

3 2+2 EXPOSES ME TO OTHER DISCIPLINES AND GIVES PERSPECTIVE
Afforded me the opportunity to visit teachers in various
departments thereby exposing me to various disciplines and
services. Afforded me the opportunity to get to know more staff

3 2+2 HELPS ME SEE THINGS I COULDN'T SEE BY MYSELF It's good to
have someone in the class monitoring my activity, as I can't see
myself and may fall into bad habits. Others observations provide a
"window"

3 I am not observed as much as I'd like to be

3 2+2 has caused me to continue my development of student centered learning

2 2+2 has helped me become more assertive

2+2 will help keep teachers from becoming stagnant in their jobs

Frequent 2+2 evaluations have helped me develop patience with my progress

Methods can rarely carry over to other disciplines. That's why methods are
taught by subject area at the university level

Some difficult problems require more visits for effective help

I need all the suggestions I can get.

Research of the contemporary examples/applications of teaching topics pays off.
Kids get most of their current events information from school.

I enjoy the teacher to teacher interaction after 2+2.

I have developed a "withitness" to observe the class as a whole and monitor
behavior without sacrificing the learning environment.

Particularly would appreciate 2+2 visits from experienced teachers in September

Interns must be encouraged to do their 2+2 observations early in the year as well

I particularly learn from being in the classroom with teachers who have an
excellent rapport with their students

Some 2+2 visits to my classroom have been superficial. I don't think observers
have taken the necessary time for "deep" suggestions

2+2 is the perfect reason to talk in on otherwise a perfect stranger and
legitimately evaluate performance

Expectations of observers can color their suggestions Make it manditory for
interns and voluntary for teachers

Most significant compliments

8 Good discipline, human relations skills, timely, good proximity control

8 Effective positive reinforcement, applauding students' achievements

6 Good rapport , caring, respectful, empathetic relationship with students,
students not afraid to make mistakes, giving feedback to allow for
remediation

9 Effective monitoring of students, checking for understanding using examples,
monitoring of warm-up writing activity BEFORE going on

5 Good use of questions, prompting questions to promote discussion, and
answering student questions, at different levels

7 Students on task - presentations started immediately

8 Effective use of warm-ups

9 Effective rapport with students at beginning of class, reinforce appropriate
behavior, constant praise of students, comfortable atmosphere, call
students by name

5 Move around the class well

5 Good knowledge of subject matter, the writing process

4 Student centered learning - student presentations

4 Effective planning

5 Clear directions, directions were posted for students to read - established
schedule

3 Relaxed atmosphere, good atmosphere

3 Good voice projection - effective voice

3 Good review of previous day's lesson, good repetitiveness

2 Everything explained, why is it important

2 Good use of analogies

3 Good use of technology, audio-visual aids

2 Effective to take attendance at the end, use of note cards for attendance
3 Effective group work, great cooperative learning
2 Present material with confidence
3 Effective application of knowledge, relating knowledge to students' lives
2 Provides immediate feedback, effective feedback
Good classroom participation
Effective individual work
You get your students to class on time
Clear answers to student questions
Involved passive students
Good use of time - pass back papers at end of bell
Effective routine
Appropriate expectations
Acknowledge problems without making a big deal out of them
Good explanations
Nice sense of humor
Student activities to reinforce learning
You praise publicly, criticize privately
Actively engaged students in problem solving
Excellent use of interpreter
Effective to allow students to choose vocabulary words from story
Good summary and closure, questions at the end of the week
Allow non readers to be tested orally
Effective use of peer tutoring
Effective classroom management skills
Good variety
Display student work well
Effective pacing - slower for objectives
Good example of students working at different levels
Objectives clearly stated
Complimented on how study hall was handled at the same time other instruction
            was going on
Students are "hard working"
Good brainstorming activity
Students are encouraged to attempt higher level skills
Assignments are fun
Encourages kids to rely on themselves and each other
Orderly class
Teacher reading aloud with students following in their texts
Thorough coverage of material
Good prompting of students to achieve success
Use of games in reviewing
Very observant, knows what is going on
Good strategy to give students an example of your own writing
Uses "target language" effectively
Good flexibility - noted when students weren't prepared and changed
            assignment
Compliments from students:
        Appreciate group work

Enjoy discussions
Teacher treats like adults
The class is easy but I'm learning a lot

Most significant suggestions

5 Give clear directions, explain assignments better, give model answers and examples when explaining an assignment, ask students to repeat the directions to make sure they understand, write assignments on board as well as giving them orally

3 Make sure students are quiet before you start talking

3 Involve all students in classroom participation, call on students who don't have their hand up

3 Make a sign in sheet for tardy students

3 Display student work - for parent conferences, career week

3 Separate "buddies" to help offset off-task behavior

3 Ask more probing questions, work on questioning techniques

2 Put the name of the class on the board to orient observers

2 Ask for a peer review of each other's work and the presentation topics

2 Have students present projects to class, teach the class

2 Keep the class moving faster

2 Use overheads rather than the chalkboard, reduce handouts

Avoid choral answers to questions

Give prompts to students who need help so they can succeed in answering questions

Be firm, give detention if misbehavior continues

Develop a "formal recognition" for student participation to decrease multiple responses

Refuse choral answers, call on students individually

More group activities

Model group activity in advance

Visual projects great for group activity

Groups are too noisy - don't allow group work to continue if noise continues

Require students to follow class rules of behavior

Collect warm ups now and then

When it's time for the bell make sure homework is on the board

Have students get books after journal entry - less clutter

Give students who finish journal entry first their choice of characters in the play

Require students to "sit up"

Write page numbers on board with objectives and goals.

Post daily objectives

Check to make sure students really understand what you say/ask

Speak more clearly

Make groups more nearly even in number of participants

Get more student involvement by asking one of them to hold the cards,

Make sure that learning groups are heterogeneously assigned

Assign group membership rather than always allowing them to be with friends

Reduce group size to a maximum of four students per gbroup

Give students more time to work in groups before being called on

More time for journal writing
Focus on one item at a time
Think of ways to motivate students for homework
Watch out for cheating
Ask students if they have finished copying before moving on
Use cassettes to emphasize the lesson
Move the desks further apart so you can walk between them
Use different colors of chalk for variety and emphasis
2 Monitor students during quizes by making eye contact with "eye roamers,"
        watch some students more closely tos reduce cheating
Be careful about talking to the board
Leave information on the overhead projector for a longer period of time
Break objectives down into smaller units of content
Write objectives on the board
Reinforce objectives with concept maps
Provide reinforcement (enrichment?) activity if students finish early
Recognize students with their hands up before calling on others
Consider inter-class competion for good behavior
Encourage individual students
Watch out for unauthorized book bags
Have each student write a question about each topic
Require students to keep notebooks which become a part of their grade
Try team study before big quiz or test
Take attendance after activity has begun to save class time
2 Move around the class more to help keep students on task
Ignore students who blurt out unrelated comments
In a co-teaching situation, make sure one teacher is always monitoring the class
Enlist student help to pass out and collect papers
Look into reading to learn activities
Praise students for their responses
Inform students privately about the work they have missed
Grading suggestions - for warm ups +1 for correct, 0 if wrong, -1 if not done
Print notes to make them easier to copy
Slide page of notes up so students who get ahead can remain involved
Find something to keep the hyper student busy
Demonstrate procedures
Reexplain key points
Devote more time to closure activities
Repeat questions which students ask so others can hear
Give one example, ask students to think of others
Make graph paper worksheets available for student to use
Have students volunteer for warm ups
Shorten warm up
Distribute hand outs before explaining them
Use body language to convey authority and confidence
Deal with late students after class, not during class
Students are getting out of their seats too much
If a confrontation develops, take the student aside or outside the class
More organization will help your presentation go more smoothly

Make sure that students follow thru on what you have asked them for

Don't allow students to write on tests so they can be reused

Teach students to block in drawings and then add refinements

Provide alternatives for slower students

Build interest by having a discussion

Pose specific questions like "do we really pay for the president's vacation."

Bring in specific artifacts (W-4 forms)

A horizontal line in fractions rather than a slanted one avoids confusion

Be at the door to invite students in

Don't allow students to leave, once they are in

Have AP students help others

Consider permanent seating arrangements

Suggestions from students:

    Too many notes

    Tests seem a little difficult

ANALYSIS OF 31 FIRST SEMESTER
SELF-REFLECTION REPORTS AT
LAKE TAYLOR HIGH SCHOOL, FEBRUARY 1997

PARTS THREE AND FOUR

103 statements were gleaned from parts three and four of the self-reflection reports which
contained responses to the following two prompts:

Part 3. Explain which of the above compliments and suggestions were especially useful,
and how they have led to reinforcement or improvement of your classroom
teaching routines.

Part 4. List any future agenda items which have emerged as a result of the 2 + 2 process.

Statements were categorized under one of five themes:

1. Teachers encouraged by praise and compliments from other teachers (26 statements)
2. Teachers provided specific suggestions to change instructional practice (24 statements)
3. Observing other teachers provided insight, perspective, and ideas (22 statements)
4. 2 + 2 encouraged collegiality and collaboration among teachers (17 statements)
5. Suggestions and comments about the 2 + 2 program (14 statements)

Theme 1: TEACHERS ENCOURAGED BY PRAISE AND COMPLIMENTS FROM
OTHER TEACHERS

Validates use of "visual aids and warm up activities."

"I know that using the Socratic method...frequently irritates some students initially. I also know
that using this method along with other teaching techniques is the best for teaching math
in general and algebra in particular. My task is to reach the right balance."

"I enjoy the comments of my fellow teachers. Many of them experience the same problems in
their classes that I have in mine. They understand that kids will not always behave properly."

"...comments have reinforced me to continue my current teaching style which includes using
various methods such as oral practice, workbook, cooperative learning and pairs."

"The compliments have reinforced by belief that the way I am working with students meets their
needs and help to achieve their academic goals.."

"It was good to see that other teachers appreciated my work and saw it as worthy."

"I appreciated reading the "reflections" - teachers need positive feedback."

"Several observations have shown me that my directions almost always clear, concise, and
complete."

"Teachers were consistent in making similar complimentary remarks relating to classroom
management, time on task, teaching style and human relations."

"As a result of the 2+2 process, I have determined to continue to give my complete and best efforts
to providing a positive educational experience for my students."

"I feel that the compliments listed gave me a lot of positive reinforcement. They also gave me
encouragement to try additional innovative and creative activities to help my students grasp
the concepts..."

"All compliments, if the teacher truly know he/she is doing the cited behavior/technique, reinforce
instruction by causing you to want to keep on doing what you are doing well."

"After reviewing the compliments and suggestions, I feel that I will continue teaching in the same

positive manner"

"The compliment of [well-paced] instruction is significant. I never want to run through information just to say it has been taught."

"Comments have reinforced my belief that I am able to effectively communicate with the students...as a guidance counselor."

"It is very good to know that my [colleagues] observed this program [good classroom management program] being implemented in my classroom."

"...reinforced the fact that I do most things right - like planning, implementing, and human relationships."

"The compliments have encouraged me to continue to provide these instructional skills on a daily basis."

"It is great to have colleagues let you know what they feel is good about your teaching/classes, and some of their suggestions I have already put into practice or increased the practice."

"I will continue to respect students and try to teach students to respect others."

"Respect of the students determine whether they focus on the assignment or not, this is a strong characteristic that teachers can possess."

"I believe I do have good rapport with my students and will work to maintain high level of respect."

"Knowing the climate of a classroom...is used in help students develop strategies so that they can be successful in their classes" (counselor)

"the most commonly occurring compliment dealt with my rapport with students which, I believe, leads to good classroom management."

"I use my sense of humor to "humanize" myself to students"

"Encouragement and reinforcement are very important in teaching a skill subject."

Theme 2: TEACHES PROVIDED SPECIFIC SUGGESTIONS TO CHANGE INSTRUCTIONAL PRACTICE:

"I have used small group activity before, but I will increase that. Also I want the students to question each other and ask why instead of turning to me for the answer."

"The suggestion to reprimand [tardy] students at the end of the class is something I will try.

"My classroom routine has improved to become more efficient as I have several students who have volunteered to collect homework and classwork, leaving me more time to prepare the class for the next activity. Also, during oral drills or reviewing classwork, I have insisted students speak in a loud, clear voice as his classmates are being held responsible for his information."

To signal readiness to begin class, "I am there to greet [students]. It also allows peer tutoring when I am unavailable to help right away."

"The 2+2 observation process is effective in the improvement of instruction."

"I certainly will attempt to increase my use of transparencies."

"The compliments and suggestions have helped me assess the extent to which I am able to engage my students in meaningful learning experiences."

"I have enlarged and am very conscious when I do questioning techniques-Socratic-to give the student ample time to answer, and to give more time to questions in time allocated in lesson plan."

"I have incorporated more warm-up questions at the appropriate time for students."

"All suggestions were considered and weighed in reference to how I could use them to improve my planning, delivery and assessment."

"I'm trying to become a better questioner during instruction."

"I survey class after timed writings for speed and accuracy."

"I used the strategy of keying lines backwards to reinforce 'eyes on copy' and keying letter by letter."

"[2+2] has made a difference in my method of instruction"

"I've used the idea concerning day to day student responsibilities. This eliminated some of the make-up work problems."

"I will make specific point to get students in front of class to present material."

"I intend to increase their explanation time with the overhead graphing calculator. They love this activity."

"...I need to stay alert to such things as questioning skills, demanding good behavior, and group work divisions."

"The suggestion has made me aware...that positioning students when working in cooperative learning teams is certainly significant."

"I have tried to involve the students more in the learning process by using more group and cooperative learning activities and by using some practical problem solving techniques."

"I have...established a written tardy policy and behavior policy that will be used second semester."

"I will work on providing work space for students in my office" (counselor)

"I do need to utilize my students more as monitors/technicians. They know a lot more than I do on occasion."

"I changed my policy on allowing kids out to the clinic. If they repeat their request two or more times, I take them seriously."

Theme 3: OBSERVING OTHER TEACHERS PROVIDED INSIGHT AND PERSPECTIVE:

"This process has given me an even greater appreciation of the work my colleagues do."

"...observing other teachers could have the outcome of improving the observer."

"The 2+2 process is good in that it gives teachers the opportunity to visit other classes to observe what other teachers are doing."

" I discovered new ways of presenting vocabulary and grammar...I feel I benefited most by observing others' methods of discipline management and classroom routine."

"I have seen some excellent teaching: relevant, interesting and student-centered."

"It has been a fascinating, informative, exciting, enlightening, and, on rare occasions, shocking experience to view other teachers in their element."

"I developed more confidence in observing my peers."

"It is helpful to see methods used in other academic areas."

"Observing others provides the observer with fresh ideas."

"The 2+2 program has enabled me as a seasoned teacher to gain meaningful insights and new ideas as I visit my peers"

"...I found there to be similarities in teaching styles, classroom management skills, human relations skills, and curriculum."

"Observing teachers allows you to evaluate and improve your instruction."

"In all of the visits I made I found that I learned something and came away with ideas which I could adapt for implementation in my own classes."

"By observing others, I get reinforcement that what happens in my classes is normal and that I'm on target in my planning and implementing strategies that result in success."

"Observing others gives a person confidence in his or her teaching duties."

"The differences in the approach to classroom management and introduction of new topics gives me food for thought."

"Classroom observations...gives me a greater appreciation for the work classroom teachers do." (Counselor)

"It is always nice to get out and observe other teachers presenting the same or different disciplines of the educational field."

"I enjoyed...seeing how other teachers handled classroom management problems, and how they organized their classes. For me, this is of high value and help me make changes in my own management routine."

"It was helpful observing a "'troubled student' in another classroom situation...to see what works or does not work for that particular individual'"

"It is very informative to see how other classrooms function...especially...anticipatory sets, varied learning styles, warm-up activities, closures, and learning resources."

"The compliments and suggestions I noted when observing other teachers will be the most help to me as I change classroom teaching routines."

Theme 4: 2+2 ENCOURAGED COLLEGIALITY AND COLLABORATION AMONG TEACHERS

"Many times, we have been doing something a certain way for so long, it is not until someone else points it out, that we can begin to correct it."

"The more 2+2s you do, the more relaxed you are in doing them."

"During one observation I found out that several of my students had just finished a unit on Mexico. This, in turn, helped me to plan my lesson on food in Mexico."

"Observing other teachers has led me to see how my department [media center] might better interact and enhance instruction in other departments"

"The compliments and suggestions also remind me that while I am teaching, I cannot see or hear everything."

"Interdisciplinary lessons in culinary art and art have been planned."

"The 2+2 concept decreases my isolation and makes me feel more a part of a team effort. It's an excellent tool for teacher growth."

"This process has led me to a better understanding of the entire working of our faculty as a whole."

"I would like to try more interdisciplinary projects with other teachers...2+2 has helped me to understand better how we can interrelate some of the disciplines."

"I look forward to more visits from my colleagues...I also look forward to visiting my colleagues."

" Peer observers have an understanding of students, many of whom they may have taught, and take into account their behavior in giving suggestions in your instructions of a lesson."

"This program gives teachers a choice which was not possible before. Choice promotes empowerment, creativity, and good morale...the 2+2 program opens lines of communication among teachers which could possibly lead to cooperative teaching endeavors."

"It is an opportunity for teachers to help teachers by sharing their expertise."

"I have had the opportunity to share teaching ideas and become acquainted with other disciplines..."

"I plan to 'get out of my little box' more often to visit classes. The 2+2 process has allowed me to share as well as obtain so many new ideas and teaching strategies."

"I like to think of 2+2 as a means for teachers to help teachers. If done thoughtfully and reflectively, it does this!!"

" 2+2 has got to be the very best way of teachers making teachers better.

Theme 5: 2 + 2 COMMENTS/SUGGESTIONS

"I think the 2+2 process is helpful to all teachers in many ways and will become a very useful evaluation instrument."

"The 2+2 process is an excellent evaluative tool for assessing teacher performance."

"I am still not being observed as much as I wish to be, will there be a list going around of teachers that need to be observed more often?"

"The number of required observations seem to be excessive."

"Schedule small group sessions during the semester and update current 2+2 schedule."

"Schedule a meeting 2+2 participants mid-semester for the purpose of expressing progress or concerns."

" Stress the importance of teachers observing colleagues across all disciplines."

"I cannot but help wonder if the actual amount of observation per semester is realistic."

"Teachers need practice in observing and additional in service (especially suggestions section.)

"It is hard to find time to do 2+2s with my busy schedule."

"I feel limited in access to teachers I would like to observe" due to schedules

"...time may not always be available to effectively observe in order to write an accurate evaluation. Fewer observations per semester (one per month) is a suggestion for next year."

"I thought the taped interviews were a good idea"

"The section on suggestions gives the teacher the opportunity to see himself as other see him."

APPENDIX G

TEACHER PERFORMANCE EVALUATION ATTITUDE QUESTIONNAIRES

Thank you for taking time out of your busy schedule to respond to the Teacher Performance Evaluation Attitude Questionnaire. The school district will use this information to evaluate teacher perceptions and attitudes about the 2+2 program, the new Teacher Evaluation System, and the Teacher Evaluation System that was in place before September of 1996. Today's questionnaire will be used to gather your attitudes about the Teacher Evaluation system that was used before September of 1996. We will return to your school in April or May to get your perceptions and attitudes about the 2+2 program or the new Teacher Evaluation Program.

To help us match your fall responses with your spring responses and to preserve your confidentiality, we need you to provide us with a four digit subject identification number. This number can be any number you would like, but we suggest that you select a number that will be easy to remember such as the first or last four digits of your social security number, a birth date, a house number, or license plate. You will be asked to fill in your subject identification number on the next page of the survey which asks for demographic information.

We will NOT reveal your subject identification number to any school personnel. Additionally, once we have matched your responses, we will discard any subject identification numbers you have given us. We will bring a list of numbers that teachers have chosen when we collect the second wave of surveys to help you remember your number. Also, we have provided you with a tear-off section to record your subject number that can be placed in your wallet.

If you have any questions, concerns, or comments regarding the study, feel free to contact:

     Alyce LeBlanc (ODU - Urban Services)
     Andrea Berndt (ODU - Psychology)

■ ■

Tear off this section to record your subject identification number and put it in a safe place.

MY SUBJECT IDENTIFICATION NUMBER FOR THE TEACHER EVALUATION STUDY CONDUCTED BY OLD DOMINION UNIVERSITY IS:

    ___ ___ ___ ___

IF I HAVE ANY QUESTIONS, I CAN CONTACT
    ALYCE LEBLANC
    ANDREA BERNDT

BACKGROUND INFORMATION

The background information will help us to compare your responses with those provided by other individuals. This information is confidential and will NEVER be used to identify any individual. All analyses conducted on the survey will be reported to individual schools and districts as group information. Individual schools and the district will NOT receive the raw data (i.e., the original surveys). Analyses conducted on the data will be performed by non-school personnel (i.e., doctoral candidates at Old Dominion University), who will present the final analyses to the schools as aggregated information.

SUBJECT IDENTIFICATION NUMBER   _____

For the following questions, check the appropriate response.

Gender _____ Male _____ Female

Age
_____ 21-26
_____ 27-31
_____ 32-40
_____ 41-49
_____ 50 +

Educational Background
_____ Bachelor's
_____ Master's
_____ Ph. D / Ed.D
_____ other_____

Ethnicity
_____ African-American
_____ Asian-American
_____ Caucasian
_____ Hispanic
_____ Native American
_____ other_____

School Level
_____ Elementary
_____ Middle
_____ High

Years of Teaching
_____ 0-3 years
_____ 4-10 years
_____ 11 or more

(If Secondary)
Subject Area
_____ English
_____ Math
_____ Science
_____ Social Studies
_____ Electives (e.g., Language, Music, Arts, PE)

Years of Teaching
in Norfolk Public Schools
_____ 0-3 years
_____ 4-10 years
_____ 11 or more

Length of time since your last summative evaluation
_____ Within last 12 months
_____ Within last 13-24 months
_____ Between 2-3 years ago
_____ 4 years ago
_____ More than 4 years ago

Are you up for summative evaluation during 96-97? _____Yes _____ No

Have you opted to participate in 2+2 for 96-97? _____Yes _____No

REFER TO THE TEACHER EVALUATION SYSTEM IN USE BEFORE SEPTEMBER 1996

TEACHER PERFORMANCE EVALUATION ATTITUDE QUESTIONNAIRE

This questionnaire has been developed to collect information concerning your experiences as a teacher with the traditional teacher evaluation process. This questionnaire also investigates your attitudes about the people and the methods that are used for evaluating your teaching performance. You learn about your teaching performance in many ways. Your responses will enable us to gain a better understanding of the appraisal process.

Listed on the following pages are statements about the **teacher evaluation process** that was used in the Norfolk Public School system **before September of 1996.** The statements are contained in sections that refer to a particular aspect of the teacher evaluation process (e.g., ACCURACY). When you respond to each section, make sure that you consider the traditional teacher evaluation system that was in place before September of 1996.

Read each statement and decide to what degree you **agree** or **disagree** with it, using the scale given below. For analysis purposes, it is critical that you respond to all items even if you are not sure about a particular response. Please indicate how strongly you agree or disagree with the statements below by checking the corresponding blank.

ACHIEVEMENT ORIENTATION -
refers to desires for achievement and performing well.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
When being evaluated, I want to do well on my formal classroom observations.	___	___	___	___	___
I always try to perform at my best.	___	___	___	___	___
I want to be rated in the top category on all categories of the summative evaluation	___	___	___	___	___
I take pride in my summative evaluation.	___	___	___	___	___
The quality of my work is important to me.	___	___	___	___	___
I like to be recognized for a job well done.	___	___	___	___	___
I want my teaching to be evaluated.	___	___	___	___	___
I want my teaching evaluation to recognize my goal accomplishments.	___	___	___	___	___
I want my teaching evaluation to recognize my work efforts.	___	___	___	___	___

ACCEPTANCE - refers to approval of the entire teacher evaluation process (as you have experienced it),
including the specific forms, objectives, goals, methods, and purposes that are part of that process.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
The teacher evaluation system is acceptable for rating job performance.	___	___	___	___	___
Our teacher evaluation system can review my teaching abilities satisfactorily.	___	___	___	___	___
The step-by-step process used to evaluate performance is acceptable to me.	___	___	___	___	___
Strategies for improvement based on observation are appropriately determined by evaluation information.	___	___	___	___	___

1

REFER TO THE TEACHER EVALUATION SYSTEM IN USE BEFORE SEPTEMBER 1996

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
Generally, teachers agree that the traditional teacher evaluation process is the best way to measure performance.	___	___	___	___	___
Information from teacher evaluations is the most appropriate basis for making decisions about teacher in-service sessions.	___	___	___	___	___
The summative evaluation provides an acceptable description of my work efforts.	___	___	___	___	___

ACCURACY - refers to comparison between actual performance, and evaluation observations and summative evaluation ratings.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
Teacher evaluations do not provide accurate descriptions of teacher work efforts.	___	___	___	___	___
I feel that my teaching is evaluated accurately by the traditional teacher evaluation process.	___	___	___	___	___
At my last summative evaluation, I received the ratings I expected.	___	___	___	___	___
Overall, the teacher evaluation system accurately reflects how well teachers perform on the job.	___	___	___	___	___
Our teacher evaluation system cannot provide accurate evaluations for all specializations.	___	___	___	___	___
The summative evaluation accurately identified areas where I could improve my teaching performance.	___	___	___	___	___
Classroom observations can pinpoint unique areas of teacher performance strengths and weaknesses.	___	___	___	___	___
I don't believe the teacher evaluation system is a real reflection of teacher performance.	___	___	___	___	___
Teacher evaluations provide accurate descriptions of teacher performance.	___	___	___	___	___

ANXIETY - refers to concerns about the teacher evaluation system.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
During my observations, I found myself preoccupied with the results of a poor evaluation.	___	___	___	___	___

2

REFER TO THE TEACHER EVALUATION SYSTEM IN USE BEFORE SEPTEMBER 1996

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
When I know my teaching performance is being observed, I worry about making errors.	___	___	___	___	___
Thinking about my summative evaluation conference makes me feel sick to my stomach.	___	___	___	___	___
I worry that my teaching will be compared to more capable teachers.	___	___	___	___	___
I get disturbed by my summative evaluation.	___	___	___	___	___
I worry that my teaching evaluations will limit my opportunities.	___	___	___	___	___
Discussing my observations in post-observation conferences makes me nervous.	___	___	___	___	___
My summative evaluation conference makes me nervous.	___	___	___	___	___

FAIRNESS - refers to perceptions of how fairly performance is reviewed.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
I feel that my teaching is evaluated fairly.	___	___	___	___	___
The step-by-step process used to make summative evaluation ratings is fair.	___	___	___	___	___
Summative evaluations only consider my actual job performance.	___	___	___	___	___
My evaluation rating is unfairly influenced by opinions of others.	___	___	___	___	___
Unfair information is often used when evaluating teacher performance.	___	___	___	___	___
The evaluation process does not justly assess all teachers.	___	___	___	___	___
My evaluation rating is an honest representation of my teaching activities.	___	___	___	___	___
Teacher evaluations are fair because they are based on teacher performance information.	___	___	___	___	___
The teacher evaluation system seemed fair to me.	___	___	___	___	___

3

REFER TO THE TEACHER EVALUATION SYSTEM IN USE BEFORE SEPTEMBER 1996

FEEDBACK - refers to the information that is given to the teacher about performance throughout the year and at the post-observation and summative evaluation conferences.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
My summative evaluation conference gives me useful information about my teaching performance.	___	___	___	___	___
Feedback from my post-observation/evaluation conferences leads to improvements in my teaching.	___	___	___	___	___
The way to improve my teaching performance goals has been clearly explained to me.	___	___	___	___	___
I receive feedback from administrative observations throughout the year.	___	___	___	___	___
Most of the feedback I receive about my teaching performance is positive.	___	___	___	___	___
Appropriate information about my teaching performance has been shared with me.	___	___	___	___	___
Negative aspects of my teaching performance are discussed with me.	___	___	___	___	___
I receive specific information about my teaching performance.	___	___	___	___	___
I receive more negative than positive feedback about my teaching performance.	___	___	___	___	___

KNOWLEDGE - refers to the rater's understanding of the teacher's work performance, job duties, and standards for performance.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
The person who conducts my teaching evaluation has a complete understanding of my teaching performance.	___	___	___	___	___
My evaluation is conducted by someone who understands my job duties.	___	___	___	___	___
The person who rates my work performance knows less about my job duties than I do.	___	___	___	___	___
In our evaluation system, performance evaluations are conducted by individuals who have extensive knowledge of teacher performance.	___	___	___	___	___
Summative evaluations are conducted by individuals who understand teachers' job duties.	___	___	___	___	___

4

REFER TO THE TEACHER EVALUATION SYSTEM IN USE BEFORE SEPTEMBER 1996

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
The standards for performance are well known by persons conducting teacher evaluations.	___	___	___	___	___
The individual(s) who conduct my evaluation have enough time to become familiar with my teaching performance.	___	___	___	___	___
My administrative reviewer sees enough of my teaching performance to rate me effectively.	___	___	___	___	___
The individual who evaluates my performance considers whether my teaching performance differs from teaching requirements.	___	___	___	___	___

TEACHER EVALUATION FORM - refers to the particular rating form or instrument that is used in evaluations.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
I am familiar with all forms used in the teacher evaluation process.	___	___	___	___	___
My evaluation form identifies specific areas where I should make improvements.	___	___	___	___	___
My evaluation form accurately describes specific areas where I perform at recommended levels.	___	___	___	___	___
My opinion about my summative ratings can be included on my evaluation.	___	___	___	___	___
The evaluation form can distinguish performance differences among teachers.	___	___	___	___	___
I understand the evaluation form which is used to appraise my teaching performance.	___	___	___	___	___
The summative evaluation form has space to include written comments explaining why ratings were chosen.	___	___	___	___	___
The classroom observation forms have space to describe specific examples of positive and negative teaching performance.	___	___	___	___	___
All classroom teachers, specialists, and licensed instructional support personnel can be evaluated with the same evaluation form.	___	___	___	___	___

5

REFER TO THE TEACHER EVALUATION SYSTEM IN USE BEFORE SEPTEMBER 1996

PERFORMANCE OBSTACLES - refer to aspects of performance that are reviewed but are not under the teacher's control.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
	1	2	3	4	5
My teaching evaluations are restricted to items under my control.	___	___	___	___	___
My teaching evaluations directly measure my performance.	___	___	___	___	___
Factors that interfere with my performance are not considered in my evaluation.	___	___	___	___	___
My teaching evaluation only considers activities over which I have control.	___	___	___	___	___
Restrictions to my performance are considered in my summative evaluation.	___	___	___	___	___
My evaluations fail to consider the availability of resources (e.g., materials, supplies, copiers) my job requires.	___	___	___	___	___
My evaluations do not consider obstacles that restrict my teaching performance.	___	___	___	___	___
I am evaluated on teaching performance factors that are not under my control.	___	___	___	___	___

POST-OBSERVATION AND SUMMATIVE EVALUATION CONFERENCES - refers to how post observation / evaluation information is shared with teachers during conferences. This includes the usefulness and clarity of the information shared, discussion of goals, and goal setting.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
	1	2	3	4	5
During summative evaluation conferences, specific goals are identified.	___	___	___	___	___
I generally agree with the goals that are set during my evaluation conferences.	___	___	___	___	___
During conferences, I often receive useful suggestions for instructional performance.	___	___	___	___	___
In my summative evaluation conference, there is a discussion about how to achieve performance goals.	___	___	___	___	___
My post-observation conference includes useful information.	___	___	___	___	___
My summative evaluation conferences do not have a clear purpose.	___	___	___	___	___
I participate in setting my performance goals during my evaluation conferences.	___	___	___	___	___

6

REFER TO THE TEACHER EVALUATION SYSTEM IN USE BEFORE SEPTEMBER 1996

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
My summative evaluation conferences reconsider the goals that were set during the previous conference.	——	——	——	——	——
My post-observation conferences are a waste of time.	——	——	——	——	——

POLICIES - refer to the district's principles for the teacher evaluation process, including the assumptions, values, and rules about that process.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
My teaching is observed three times a year.	——	——	——	——	——
The person conducting my evaluation and I must agree on my teaching duties before observations are conducted.	——	——	——	——	——
Teacher evaluations follow a standardized approach.	——	——	——	——	——
Individuals who conduct teacher evaluations are rewarded for their efforts.	——	——	——	——	——
The school system gives individuals who rate teachers enough time to observe and evaluate teachers' performance.	——	——	——	——	——
I have been told about school policies concerning teacher evaluation.	——	——	——	——	——
I am unclear why teacher evaluations are conducted.	——	——	——	——	——
Open discussion of the teacher evaluation process is encouraged throughout the year by the administration.	——	——	——	——	——
My school guarantees that teacher evaluation information is confidential.	——	——	——	——	——
My school district has a policy that outlines the consequences of a negative teacher evaluation.	——	——	——	——	——

PROCEDURES - refers to methods for developing and changing the evaluation system, administrator training in evaluation, and other formal activities related to the teacher evaluation system.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
Teachers helped to develop our teacher evaluation system.	——	——	——	——	——

7

REFER TO THE TEACHER EVALUATION SYSTEM IN USE BEFORE SEPTEMBER 1996

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
Teachers are asked to suggest changes to our teacher evaluation system.	—	—	—	—	—
Teacher are encouraged to share their opinions about the teacher evaluation system.	—	—	—	—	—
Teachers are informed of their expected role in the evaluation process.	—	—	—	—	—
Administrators are trained in how to conduct classroom observations and summative evaluations.	—	—	—	—	—
Administrators periodically receive refresher training in conducting observations/summative evaluations.	—	—	—	—	—
There is a procedure to use if a teacher wants to rebut a summative evaluation rating.	—	—	—	—	—
Teachers participated in defining the standards for effective teacher performance.	—	—	—	—	—
Teachers receive information that emphasizes the importance of the teacher evaluation system.	—	—	—	—	—

PURPOSES - refers to the functions served by the teacher evaluation system.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
The teacher evaluation system is used for teacher development and growth.	—	—	—	—	—
I know why my school district has a teacher evaluation system.	—	—	—	—	—
My school district is clear about the reasons for conducting teacher evaluations.	—	—	—	—	—
My summative evaluation will not affect my chances for future promotions.	—	—	—	—	—
The purposes of teacher evaluation have been explained to me.	—	—	—	—	—
Teacher evaluations are used to determine who receives additional training opportunities.	—	—	—	—	—
Teacher evaluations serve as a basis for salary decisions.	—	—	—	—	—
Teacher evaluations are used to make decisions about firing.	—	—	—	—	—

8

REFER TO THE TEACHER EVALUATION SYSTEM IN USE BEFORE SEPTEMBER 1996

SUPERVISORY IMPACT - refers to the reactions and actions of the administrator as they take part in the observation/summative evaluation conference.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
After my classroom observation, the administrator discusses the information with me.	——	——	——	——	——
When the administrator shares evaluation information with me, I can express my views on it.	——	——	——	——	——
The administrator reviews my evaluation form with me.	——	——	——	——	——
During my post-observation/summative evaluation conferences the administrator gives me plenty of time to express my opinions.	——	——	——	——	——
After the administrator discussed my evaluation information with me, I knew exactly what I could do to improve my performance.	——	——	——	——	——
During my evaluation conference, the administrator took enough time to discuss the results of my performance rating with me.	——	——	——	——	——
The individual who observed me generates my summative evaluation rating.	——	——	——	——	——
My post-observation/summative evaluation conferences serve as an opportunity for the administrator and I to discuss my goals.	——	——	——	——	——
The administrator spends most of the summative evaluation conference criticizing my performance.	——	——	——	——	——

SUPERVISORY TRUST - refers to help and support given by the administrator.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
My administrator listens to my opinions about my teaching performance.	——	——	——	——	——
My administrator is credible and trustworthy.	——	——	——	——	——
My administrator has a helping attitude in evaluating my teaching performance.	——	——	——	——	——
My administrator maintains confidentiality of my evaluation information.	——	——	——	——	——
My administrator and I have a positive relationship.	——	——	——	——	——
I can have honest and open communication with my administrator about my teaching performance.	——	——	——	——	——

9

REFER TO THE TEACHER EVALUATION SYSTEM IN USE BEFORE SEPTEMBER 1996

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
My administrator encourages me to share differences of opinion about my teaching performance.	___	___	___	___	___
My administrator helps me to solve problems associated with my teaching.	___	___	___	___	___
My administrator cares about my well being.	___	___	___	___	___

10

REFER TO THE TEACHER EVALUATION SYSTEM IN USE BEFORE SEPTEMBER 1996

Summative Evaluation Rating

For the next 8 items, please circle your response.

1. At your last summative evaluation, did you receive the ratings you expected?

| Much less than expected | Less than expected | Exactly what was expected | Slightly more than expected | Surpassed expectations |

2. I was satisfied with my last teacher evaluation ratings.

| Not at all satisfied | Somewhat unsatisfied | Neutral | Moderately satisfied | Very satisfied |

3. Were you satisfied that you received accurate teacher evaluation ratings at your last summative evaluation?

| Not at all satisfied | Somewhat unsatisfied | Neutral | Moderately satisfied | Very satisfied |

4. My most recent summative evaluation ratings were fair.

| Not at all true | Somewhat not true | Unsure | Somewhat true | Very true |

5. My most recent summative evaluation ratings were just.

| Not at all true | Somewhat not true | Unsure | Somewhat true | Very true |

6. My last summative evaluation ratings were firmly based on my teaching performance.

| Not at all true | Somewhat not true | Unsure | Somewhat true | Very true |

7. My last evaluation rating is a good representation of my actual teaching performance.

| Not at all true | Somewhat not true | Unsure | Somewhat true | Very true |

8. How well did your last summative evaluation ratings compare to your expectations?

| Much less than expected | Less than expected | Exactly what was expected | Slightly more than expected | Surpassed expectations |

9. On your last summative evaluation ratings, check whether you had:

NEEDS IMPROVEMENT	COMPETENT	PROFICIENT (+2)
____0 categories	____0 categories	____0 categories
____1-2 categories	____1-2 categories	____1-2 categories
____3-5 categories	____3-5 categories	____3-5 categories
____6-10 categories	____6-10 categories	____6-10 categories

11

TO: ALL SUMMATIVE YEAR 2 + 2 PROGRAM TEACHERS

RE: FOLLOW-UP TEACHER PERFORMANCE EVALUATION ATTITUDE SURVEY

Thank you for your responses to the survey about teacher attitudes toward performance appraisal which you filled out last September. At that time, you were asked to respond to the items by referring to the teacher evaluation system in place in the Norfolk Public Schools from 1983-1995. This follow-up survey is intended to assess your feelings about the value of the 2 + 2 Alternative Performance Appraisal system. **This survey should be completed even if you did not complete the survey last fall.** (Teachers not participating in 2+2 will be asked to fill out a follow-up survey later in the year about the new NPS appraisal system.)

Below is a list of subject ID numbers from last September's survey. The list is to assist you in remembering the number you chose. If a number appears twice, it is because two people chose it. Please use the same number you chose last fall to identify this survey. In a few cases, your number may not appear below due to error as to who is participating in 2+2. If your number does not appear, please fill out the survey anyway and provide us with your subject ID number if you remember it. Otherwise leave the ID number space blank. Also, leave the subject ID number blank if you did not participate in the fall survey. **Remember to answer each question even if you are unsure of a particular response.** Your cooperation is critical, and will provide maximum understanding of teacher reactions to the 2+2 program, as well as provide insights for the future direction of 2+2.

ID NUMBERS OF 2+2 PARTICIPANTS AS IDENTIFIED ON FIRST SURVEY:

0005	1207	1970	5612
0009	1226	1998	6139
0136	1227	2175	6367
0278	1234	2275	6648
0320	1234	2279	7729
0405	1317	2599	8148
0411	1420	2607	8513
0442	1492	3288	8548
0475	1505	3632	9-14-90
0828	15-0-15-0	4040	9664
0997	1515	4419	825F
1023	1549	4444	BEME
1026	1602	4638	HATS
1029	1613	4659	
1032	1671	4711	
1033	1717	5001	
1085	1908	5100	
1088	1919	5317	
1122	1943	5412	

If you have any questions, concerns, or comments regarding the study, feel free to contact:
Alyce LeBlanc (ODU - Urban Services)
Andrea Berndt (ODU - Psychology)

BACKGROUND INFORMATION

The background information will help us to compare your responses with those provided by other individuals. This information is confidential and will NEVER be used to identify any individual. All analyses conducted on the survey will be reported to individual schools and districts as group information. Individual schools and the district will NOT receive the raw data (i.e., the original surveys). Analyses conducted on the data will be performed by non-school personnel (i.e., doctoral candidates at Old Dominion University), who will present the final analyses to the schools as aggregated information.

For the following questions, check the appropriate response:

I ___ DID/ ___ DID NOT participate in the survey on teacher attitudes towards performance appraisal system in place in the Norfolk Public Schools from 1983-1995/96.

SUBJECT IDENTIFICATION (ID) NUMBER: _____
(To be filled in only if you participated in the survey last fall; otherwise leave blank and continue with the following questions.)

Gender ____ Male ____ Female

Age
____ 21-26
____ 27-31
____ 32-40
____ 41-49
____ 50+

School Level
____ Elementary
____ Middle
____ High

(If Secondary)
Subject Area
____ English
____ Math
____ Science
____ Social Studies
____ Electives
    (e.g., Language, Music, Arts, PE)
____ Special Education

Educational Background
____ Bachelor's
____ Master's
____ Ph.D/Ed.D
____ other_____

Ethnicity
____ African-American
____ Asian-American
____ Caucasian
____ Hispanic
____ Native American
____ other_____

Years of Teaching
____ 0-3 years
____ 4-10 years
____ 11 or more

Years of Teaching
in Norfolk Public Schools
____ 0-3 years
____ 4-10 years
____ 11 or more

Length of time since your last summative evaluation
____ Within last 12 months
____ Within last 13-24 months
____ Between 2-3 years ago
____ 4 years ago
____ More than 4 years ago

Were you up for summative evaluation during 96-97? ____ Yes ____ No

If yes, did you participate in 2 + 2 in lieu of the Norfolk Public Schools appraisal system? ____ Yes ____ No

REFER ONLY TO THE 2+2 ALTERNATIVE PERFORMANCE APPRAISAL SYSTEM IN PRIME SCHOOLS

TEACHER PERFORMANCE EVALUATION ATTITUDE QUESTIONNAIRE

This questionnaire has been developed to collect information concerning your experiences as a teacher with the 2+2 Appraisal Process. This questionnaire also investigates your attitudes about the people and the methods that are used in the 2+2 Appraisal Process. You learn about your teaching performance in many ways. Your responses will enable us to gain a better understanding of the 2+2 appraisal process.

Listed on the following pages are statements about the 2+2 Appraisal Process that is being used in the Norfolk Prime Schools. The statements are contained in sections that refer to a particular aspect of the teacher evaluation process (e.g., ACCURACY). When you respond to each section, make sure that you consider the 2+2 Alternative Performance Appraisal System.

Read each statement and decide to what degree you **agree** or **disagree** with it, using the scale given below. For analysis purposes, it is critical that you respond to all items even if you are not sure about a particular response. Please indicate how strongly you agree or disagree with the statements below by checking the corresponding blank.

ACHIEVEMENT ORIENTATION -
refers to desires for achievement and performing well.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
When being observed, I want to do well on my 2+2 classroom observations.	—	—	—	—	—
I always try to perform at my best.	—	—	—	—	—
I take pride in my 2+2 summative evaluation.	—	—	—	—	—
The quality of my work is important to me.	—	—	—	—	—
I like to be recognized for a job well done.	—	—	—	—	—
I want my teaching to be observed.	—	—	—	—	—
I want my 2+2 evaluations to recognize my goal accomplishments.	—	—	—	—	—
I want my summative 2+2 evaluation to recognize my work efforts.	—	—	—	—	—

ACCEPTANCE - refers to approval of the entire 2+2 Alternative Performance Appraisal (as you have experienced it), including the specific forms, objectives, goals, methods, and purposes that are part of that process.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
The 2+2 Performance Appraisal system is acceptable for giving feedback on teacher performance.	—	—	—	—	—
Our 2+2 system can review my teaching abilities satisfactorily.	—	—	—	—	—
The 2+2 summative report used to evaluate performance is acceptable to me.	—	—	—	—	—
Strategies for improvement based on observation are appropriately determined by 2+2 compliments and suggestions.	—	—	—	—	—

3

REFER ONLY TO THE 2+2 ALTERNATIVE PERFORMANCE APPRAISAL SYSTEM IN PRIME SCHOOLS

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
Generally, teachers agree that the traditional teacher evaluation process is the best way to improve performance.	——	——	——	——	——
Information from 2+2 observations is the most appropriate basis for making decisions about teacher in-service sessions.	——	——	——	——	——
The 2+2 summative evaluation provides an acceptable description of my work efforts.	——	——	——	——	——
Generally, teachers agree that the 2+2 Alternative Performance Appraisal process is the best way to improve performance.	——	——	——	——	——

ACCURACY - refers to comparison between actual performance, and 2+2 observations.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
2+2 observations do not provide accurate descriptions of teacher work efforts.	——	——	——	——	——
I feel that my teaching is observed accurately by 2+2 feedback.	——	——	——	——	——
At my last 2+2 observation, I received the compliments and suggestions I expected.	——	——	——	——	——
Overall, the 2+2 system accurately reflects how well teachers perform on the job.	——	——	——	——	——
The 2+2 system cannot provide accurate feedback for all specializations.	——	——	——	——	——
2+2 observations have accurately identified areas where I could improve my teaching performance.	——	——	——	——	——
2+2 observations can pinpoint unique areas of teacher performance strengths and weaknesses.	——	——	——	——	——
I don't believe the 2+2 system is a real reflection of teacher performance.	——	——	——	——	——
2+2 observations provide accurate descriptions of teacher performance.	——	——	——	——	——

4

REFER ONLY TO THE 2+2 ALTERNATIVE PERFORMANCE APPRAISAL SYSTEM IN PRIME SCHOOLS

ANXIETY - refers to concerns about the 2+2 teacher evaluation system.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
While being observed, I found myself preoccupied with the possibility of poor 2+2 feedback.	——	——	——	——	——
When I know my teaching performance is being observed, I worry about making errors.	——	——	——	——	——
Thinking about formal and informal comments after 2+2 observations makes me feel sick to my stomach.	——	——	——	——	——
I worry that my teaching will be compared to more capable teachers.	——	——	——	——	——
I get disturbed by my 2+2 feedback	——	——	——	——	——
I worry that my 2+2 observations will limit my opportunities.	——	——	——	——	——
Discussing my 2+2 observations in post-observation conferences makes me nervous.	——	——	——	——	——
Discussion following a 2+2 observation makes me nervous.	——	——	——	——	——

FAIRNESS - refers to perceptions of how fairly performance is reviewed.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
I feel that my teaching is observed fairly.	——	——	——	——	——
The step-by-step process leading to my summative evaluation is fair.	——	——	——	——	——
2+2 observations only consider my actual job performance.	——	——	——	——	——
My 2+2 feedback is unfairly influenced by opinions of others.	——	——	——	——	——
Unfair information is often used when giving feedback in 2+2 observations.	——	——	——	——	——
The 2+2 system does not justly assess all teachers.	——	——	——	——	——
My 2+2 feedback is an honest representation of my teaching activities.	——	——	——	——	——
2+2 observations are fair because they are based on teacher performance information.	——	——	——	——	——
The 2+2 system seems fair to me.	——	——	——	——	——

5

REFER ONLY TO THE 2+2 ALTERNATIVE PERFORMANCE APPRAISAL SYSTEM IN PRIME SCHOOLS

FEEDBACK - refers to the information that is given to the teacher about performance through 2+2 observations throughout the year and during informal followup discussion.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
My 2+2 feedback gives me useful information about my teaching performance.	___	___	___	___	___
Feedback from any followup conferences leads to improvements in my teaching.	___	___	___	___	___
The way to improve my teaching performance goals has been clearly explained to me.	___	___	___	___	___
I receive 2+2 feedback from administrative observations throughout the year.	___	___	___	___	___
Most of the 2+2 feedback I receive about my teaching performance is positive.	___	___	___	___	___
Appropriate information about my teaching performance has been shared with me.through 2+2 feedback	___	___	___	___	___
Negative aspects of my teaching performance are addressed through 2+2 feedback	___	___	___	___	___
I receive specific information about my teaching performance through 2+2 feedback.	___	___	___	___	___
I receive more negative than positive feedback about my teaching performance in 2+2 feedback.	___	___	___	___	___

KNOWLEDGE - refers to the observer's understanding of the teacher's work performance, job duties, and standards for performance.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
The individuals who observe my teaching have a complete understanding of my teaching performance.	___	___	___	___	___
2+2 observations are conducted by persons who understand my job duties.	___	___	___	___	___
The persons who observe my work performance know less about my job duties than I do.	___	___	___	___	___
In the 2+2 system, observations are conducted by individuals who have extensive knowledge of teacher performance.	___	___	___	___	___
The standards for performance are well known by persons conducting teacher observations.	___	___	___	___	___

6

REFER ONLY TO THE 2+2 ALTERNATIVE PERFORMANCE APPRAISAL SYSTEM IN PRIME SCHOOLS

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
The teachers who observe me have enough time to become familiar with my teaching performance.	___	___	___	___	___
Administrators in my school see enough of my teaching to provide effective feedback.	___	___	___	___	___
The teachers who observe my performance consider whether my teaching performance differs from teaching requirements.	___	___	___	___	___
The administrators who observe my performance consider whether my teaching performance differs from teaching requirements.	___	___	___	___	___

2+2 OBSERVATION FORMS - refers to the particular 2+2 forms or instruments used in generating feedback and evaluations.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
I am familiar with all forms used in the 2+2 Appraisal process.	___	___	___	___	___
My 2+2 feedback form identifies specific areas where I should make improvements.	___	___	___	___	___
My 2+2 feedback form accurately describes specific areas where I perform at recommended levels.	___	___	___	___	___
My opinion about my 2+2 feedback can be included on my summative evaluation report.	___	___	___	___	___
The 2+2 feedback and summative evaluation form can distinguish performance differences among teachers.	___	___	___	___	___
I understand the 2+2 feedback form which is used when others observe my teaching performance.	___	___	___	___	___
The 2+2 observation forms have space to describe specific examples of positive and negative teaching performance.	___	___	___	___	___
All classroom teachers, specialists, and licensed instructional support personnel can be evaluated with the same 2+2 form.	___	___	___	___	___
The purpose of the 2+2 summative evaluation form is clear to me.	___	___	___	___	___

7

REFER ONLY TO THE 2+2 ALTERNATIVE PERFORMANCE APPRAISAL SYSTEM IN PRIME SCHOOLS

PERFORMANCE OBSTACLES - refer to aspects of performance that are reviewed but are not under the teacher's control.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
My 2+2 feedback I receive is restricted to items under my control.	___	___	___	___	___
My 2+2 feedback directly measures my performance.	___	___	___	___	___
Factors that interfere with my performance are not considered in the 2+2 feedback I receive.	___	___	___	___	___
2+2 feedback I receive only considers activities over which I have control.	___	___	___	___	___
Restrictions to my performance are considered in the 2+2 feedback I receive.	___	___	___	___	___
2+2 feedback fails to consider the availability of resources (e.g., materials, supplies, copiers) my job requires.	___	___	___	___	___
2+2 feedback does not consider obstacles that restrict my teaching performance.	___	___	___	___	___
My 2+2 feedback is based on teaching performance factors that are not under my control.	___	___	___	___	___

POST-OBSERVATION DISCUSSION - refers to how post observation information is shared with teachers, either formally or informally. This includes the usefulness and clarity of the information shared, discussion of goals, and goal setting.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
During post-observation discussions, specific teaching goals are identified.	___	___	___	___	___
I generally agree with the goals that are discussed.	___	___	___	___	___
During discussions, I often receive useful suggestions for instructional performance.	___	___	___	___	___
During post-observation discussions, I receive information about how to achieve teaching goals.	___	___	___	___	___
My post-observation discussions include useful information.	___	___	___	___	___
My post-observation discussions do not have a clear purpose.	___	___	___	___	___

8

REFER ONLY TO THE 2+2 ALTERNATIVE PERFORMANCE APPRAISAL SYSTEM IN PRIME SCHOOLS

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
I have rarely had a post-observation discussion after my teaching has been observed.	____	____	____	____	____
To be effective, all 2+2 observations should be followed by post-observation discussions.	____	____	____	____	____
My post-observation discussions are a waste of time.	____	____	____	____	____

POLICIES - refer to the PRIME school's principles for the 2+2 process, including the assumptions, values, and rules about that process.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
My teaching is observed more than three times a year.	____	____	____	____	____
The person observing my classroom and I must agree on my teaching duties before observations are conducted.	____	____	____	____	____
2+2 observations follow a standardized approach.	____	____	____	____	____
Individuals who conduct 2+2 observations are rewarded for their efforts.	____	____	____	____	____
The PRIME schools give individuals who observe teachers enough time to observe and give feedback on teacher performance.	____	____	____	____	____
I have been told about school policies concerning the 2+2 process.	____	____	____	____	____
I am unclear why 2+2 observations are conducted.	____	____	____	____	____
Open discussion of the 2+2 appraisal process is encouraged throughout the year by the administration.	____	____	____	____	____
My school guarantees that 2+2 observations and evaluations are confidential.	____	____	____	____	____
My school has criteria for an unsuccessful 2+2 appraisal.	____	____	____	____	____

9

REFER ONLY TO THE 2+2 ALTERNATIVE PERFORMANCE APPRAISAL SYSTEM IN PRIME SCHOOLS

PROCEDURES - refers to methods for developing and changing the evaluation system, administrator training in evaluation, and other formal activities related to the 2+2 system.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
Teachers helped to develop our 2+2 system.	___	___	___	___	___
Teachers are asked to suggest changes to our 2+2 system.	___	___	___	___	___
Teacher are encouraged to share their opinions about the 2+2 appraisal system.	___	___	___	___	___
Teachers are informed of their expected role in the 2+2 program.	___	___	___	___	___
Administrators are trained in 2+2 classroom observations and summative evaluations.	___	___	___	___	___
Administrators periodically receive refresher training in conducting 2+2 observations/summative evaluations.	___	___	___	___	___
There is a procedure to use if a teacher wants to rebutt a 2+2 feedback or summative evaluation.	___	___	___	___	___
Teachers participated in defining the standards for effective teacher performance.	___	___	___	___	___
Teachers receive information that emphasizes the importance of the 2+2 appraisal system.	___	___	___	___	___

PURPOSES - refers to the functions served by the 2+2 appraisal system.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
The 2+2 system is used for teacher development and growth.	___	___	___	___	___
I know why my school has implemented the 2+2 system.	___	___	___	___	___
My school is clear about reasons for conducting the 2+2 program.	___	___	___	___	___
My 2+2 summative evaluation report will not affect my chances for future promotions.	___	___	___	___	___
The purposes of peer observation have been explained to me.	___	___	___	___	___
2+2 observations are used to determine who receives additional training opportunities.	___	___	___	___	___
2+2 feedback serves as a basis for salary decisions.	___	___	___	___	___
2+2 feedback is used to make decisions about firing.	___	___	___	___	___

10

REFER ONLY TO THE 2+2 ALTERNATIVE PERFORMANCE APPRAISAL SYSTEM IN PRIME SCHOOLS

OBSERVER IMPACT - refers to the reactions and actions of the observer as they take part in the optional observation discussions.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
After my classroom observation, the observer discusses the feedback with me.	___	___	___	___	___
When the observer shares feedback information with me, I can express my views on it.	___	___	___	___	___
The observer reviews my feedback form with me.	___	___	___	___	___
During my post-observation discussions, the observer gives me plenty of time to express my opinions.	___	___	___	___	___
When the observer discusses my feedback information with me, I know exactly what I could do to improve my teaching.	___	___	___	___	___
My post-observation discussions serve as an opportunity for the observer and I to discuss my teaching goals.	___	___	___	___	___
My observers spend most of the post observation discussions criticizing my teaching.	___	___	___	___	___

SUPERVISORY TRUST - refers to help and support given by the administrator.

	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
My administrator listens to my opinions about my teaching performance.	___	___	___	___	___
My administrator is credible and trustworthy.	___	___	___	___	___
My administrator has a helping attitude in evaluating my teaching performance.	___	___	___	___	___
My administrator and I have a positive relationship.	___	___	___	___	___
I can have honest and open communication with my administrator about my teaching performance.	___	___	___	___	___
My administrator encourages me to share differences of opinion about my teaching performance.	___	___	___	___	___
My administrator helps me to solve problems associated with my teaching.	___	___	___	___	___
My administrator cares about my well being.	___	___	___	___	___

11

VITA

ALYCE C. LEBLANC

EDUCATION

OLD DOMINION UNIVERSITY
Ph.D. Urban Services/Education, 1997
    College of Education
    Educational Curriculum and Instruction
    Old Dominion University
    Norfolk, VA 23529

UNIVERSITY OF SOUTHERN CALIFORNIA
M.S. Systems Management, 1993

STAATLICHE HOCHSCHULE FÜR MUSIK
Konzertexam, mit Auszeichnung, 1976

OBERLIN COLLEGE
B.M. Music, 1972

EXPERIENCE

OLD DOMINION UNIVERSITY
Liaison for the PRIME Project, Adjunct Professor, Research Assistant   1994-1997
    Administrator, co-author, National Science Foundation Planning Grant ($85,000)
    Administrator, co-author, USIA grant for Czech educators ($63,000)
    Classroom and distance education teaching experience

CHINA July 1996
    Presenter at conferences in Wutai Shan and Guilin on educational reform on behalf of the
    Shanxi Institute and the United Nations Development Programme

CZECH REPUBLIC May 1995
    University of Prague, seminar on systemic change

Prior experience as assistant professor of piano, Oberlin College and Conservatory of Music
Instructor of piano, Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany
Concerts in Germany, France, England, Switzerland, Scandinavia; radio and television recordings.
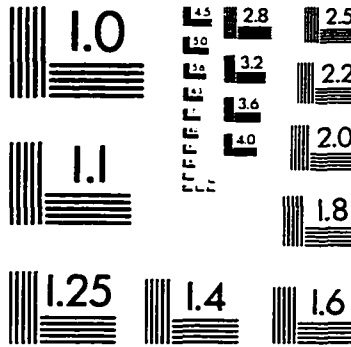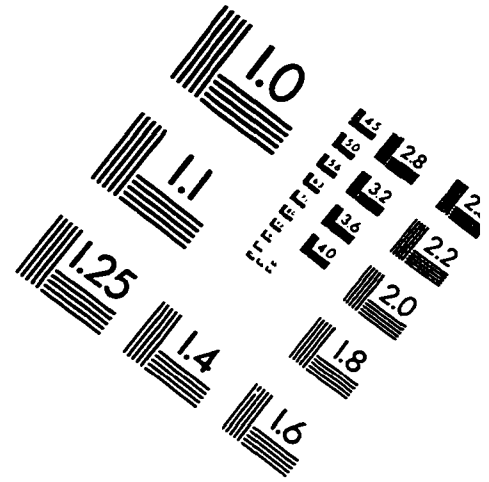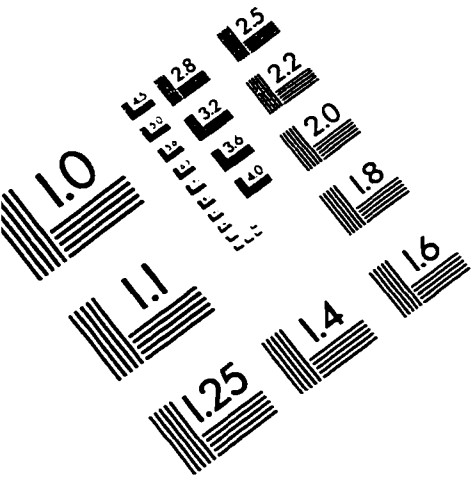
AWARDS

Numerous prizes at national and international piano competitions
Fulbright Scholarship, 1972-1974, music studies in Germany

PUBLICATION

Allen, Dwight W. , Nichols, Roy D., LeBlanc, Alyce C. (1997, June/July). The PRIME Teacher
Appraisal Program: 2 + 2 for Teachers. The High School Magazine, 4, 30-35.

IMAGE EVALUATION
TEST TARGET (QA-3)

150mm

6"
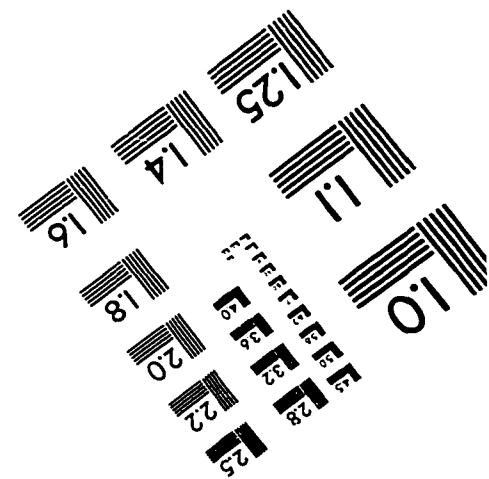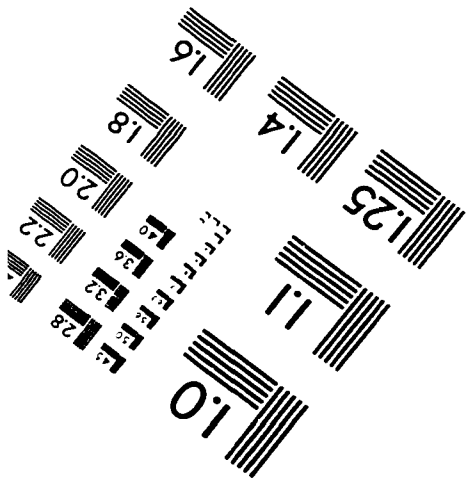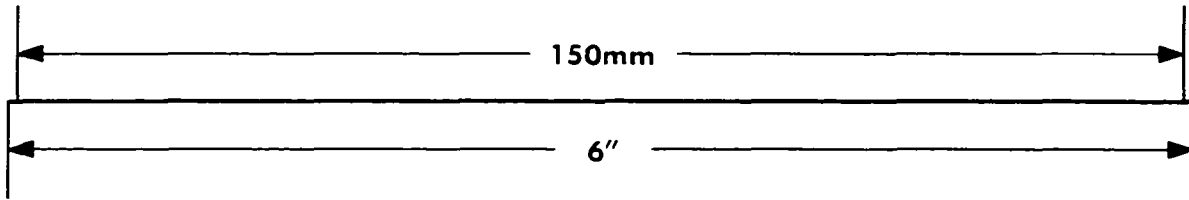
APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989