THE MEDIATING EFFECT OF PARTICIPATION AND PROCESS OUTCOMES ON EVALUATION USE IN BRITISH COLUMBIA SCHOOL ACCREDITATION

by

BARBARA TURNBULL

B.A., The University of British Columbia, 1985
M.A., The University of British Columbia, 1992

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Department of Educational Psychology and Special Education)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

October 1997

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Educational Psychology + Special Education

The University of British Columbia
Vancouver, Canada

Date October 21, 1997

ABSTRACT

The term participatory evaluation is commonly understood as stakeholder involvement in evaluation decision making and is generally accepted as a means of increasing the use of evaluation information. In spite of the popularity of participatory evaluation, there are few empirical studies which explain the casual processes of the participation-use relationship and few theories of participatory evaluation. Furthermore, it is not yet known what variables mediate participation and use, or what evaluation methodology best identifies these variables.

This dissertation was designed to test causal relations between participation and use in a proposed model of participatory evaluation. The constructs in the model were Participative Climate, Level of Participation in Decision Making, Influence in Participative Decision Making, Process Outcomes, and Instrumental and Symbolic Use. An intervening mechanism design (Chen, 1990) was used to test the hypotheses that (a) Participative Climate, Level of Participation in Decision Making, and Influence in Participative Decision Making predict Process Outcomes and (b) Process Outcomes predicts Instrumental Use and Symbolic Use.

The sample included 315 elementary and secondary teachers who participated in the 1995/1996 British Columbia (B.C.) School Accreditation Program, which is a participative school evaluation program sponsored by the B.C. Ministry of Education. Structural equation modeling was used to test the fit of the model. Overall, the analysis indicated that both hypotheses were tenable and the model was a plausible representation of the data. Furthermore, cross-validation strategies indicated that the model would likely replicate in

other independent samples. Specifically, the findings indicated that (a) teacher participation in pre-evaluation decisions, (b) influence in decision making, and (c) teacher perception of the participative process mediated the relationship between Participative Climate and Use of evaluation information. Moreover, teacher perceptions of Process Outcomes is a key factor in understanding the nature and function of participatory evaluation. The model tested in this study provides an empirically based explanation of how participatory evaluation can be expected to work and thereby provides a basis for further development of a theory of participatory evaluation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

It is with pleasure that I would like to thank Dr. Nand Kishor, who guided me through my doctoral program with expertise and precision. He has been an exemplary supervisor. I would also like to thank my committee members, Dr. Don Allison and Dr. Judith Ottoson, for their guidance, insights, and sense of humor throughout the process of researching and writing this dissertation. I would like to thank Dr. Larry Gray for sharing his time as well as his resources on the British Columbia School Accreditation Program. In addition, I would like to thank Donna Fischbacher, my mother, who was an immense help through all stages of the writing process, and Rina Bonanno and Irene Wanless for the time they took to proof-read my work. Also, I would like to thank my friends Jody, Rina, and Andrea for their patience and interest in the study.

Finally, this work is dedicated to Gregory, my husband, who has always believed in my ability to succeed.

CHAPTER ONE:
INTRODUCTION

Program evaluation researchers and theorists have long been concerned with devising ways to increase the use of evaluation information (e.g., Barrick & Cogliano, 1993; Brunner & Guzman, 1989; Fetterman, Kaftarian, & Wandersman, 1996; Greene, 1987, 1988a; Mark & Shortland, 1985; Patton, 1986; Weiss, 1983). A common method to increase use is to involve program stakeholders in evaluation decision making. This form of evaluation is called stakeholder and participatory evaluation. The terms participatory evaluation (Garaway, 1995) and stakeholder evaluation (Patton, 1986) are generally used to describe evaluation models that focus on the link between stakeholder participation in evaluation decision making and the use of evaluation information.

In stakeholder-based evaluation, people directly linked to a program, such as program staff, provide input into decision making that occurs during the evaluation, but the evaluation is primarily the responsibility of an external evaluator. Participatory evaluation is an extension of stakeholder evaluation. In participatory evaluation, stakeholders are involved in evaluation decision making as well as share joint responsibility for the evaluation report with an external evaluator (Cousins & Earl, 1992; Garaway, 1995). Stakeholders can be involved in decisions related to all phases of the evaluation process such as clarifying program objectives, designing the evaluation, collecting data, and interpreting and reporting results. It is generally believed that stakeholders who participate in evaluation decision making will have greater understanding of their program and greater investment and motivation to use the evaluation information (e.g., Chen, 1990; Patton,

1986; Weiss, 1983; Wholey, 1981). In current literature, the distinction between stakeholder and participatory evaluation is becoming less defined, and it is now common to see the term participatory evaluation used to characterize any evaluation that involves stakeholders as a means of increasing use (Papineau & Keily, 1996).

The widespread use of participatory evaluation is a response to the much earlier and more deeply rooted issue of *evaluation use.* Evaluation use has traditionally been described as: (a) instrumental, (b) symbolic, or (c) conceptual (Shadish, Cook, & Leviton, 1991). Instrumental use occurs when direct and specific changes are made to a program based on evaluation information; symbolic use occurs when evaluation information is used to persuade others of a predetermined position; and conceptual use occurs with increased understanding and learning about a program as a result of an evaluation. Participatory evaluation is generally accepted as a means of facilitating all three types of evaluation use (Garaway, 1995; Greene, 1988a, 1988b; Papineau & Keily, 1996).

Literature on evaluation use is extensive (e.g., Chen, 1990; Patton, 1986; Scriven, 1967; Weiss, 1983; Wholey, 1981) and it is probable that the study of evaluation use will continue to be of interest to researchers in that it is one of the four evaluation standards proposed by the Joint Committee on Standards for Educational Evaluation (Joint Committee, 1994). However, in spite of the concern for evaluation use, the literature on participatory evaluation consists mostly of project descriptions and position statements. There are relatively few empirical studies which use analysis of original data and formal theories to derive research questions and interpret findings. As a result, there are no empirically based theories of participatory evaluation. Considering the widespread use of participatory program evaluation, it seemed problematic that the relationship between

participation and use had not been substantiated by empirical research. Therefore, the question of what variable(s) mediate participation and use and what evaluation methodology best identifies these variables(s) was addressed in the current study.

In an effort to find empirical research on participation and use and possible mediator variables, organizational behavior literature on causal linkages between participation and employee performance was examined. The findings indicated repeatedly that *employee participation in decision making may not be a sufficient nor a necessary condition to motivate increased employee performance* (cf. Locke & Latham, 1990; Locke & Schweiger, 1979; Miller & Monge, 1986; Schweiger & Leana, 1986; Shalley, Oldham, & Porac, 1987; Tubbs, 1986; Vanderslice, Rice, & Julian, 1987; Wagner & Gooding, 1987a, 1987b). Therefore, in spite of the widespread use of participatory strategies, they appear to have an equivocal effect on job performance. In response to the lack of evidence for the motivational link between employee participation and performance (cf. Locke & Schweiger, 1979; Miller & Monge, 1986; Wagner & Gooding, 1987b), researchers have begun to investigate constructs called cognitive outcomes of participation (i.e., Latham, Winters, & Locke, 1994; Locke & Latham, 1990; Locke & Schweiger, 1979).

Locke and Schweiger (1979) defined cognitive outcomes of participation as (a) the attainment of new and useful information, (b) the generation of new ideas, and (c) the development of strategies and tactics to solve problems. In a recent study, Latham et al. (1994) found that cognitive outcomes mediated the relationship between participation in decision making and employee performance. They argued that participants who acquired task relevant knowledge as a result of participation were more likely to use this knowledge, which in turn positively affected performance. Moreover, in a related set of studies,

Tjosvold and colleagues (i.e., Tjosvold, 1985, 1986, 1987; Tjosvold, Andrews, & Jones, 1983) found that cognitive outcomes such as effective information sharing, open expression of ideas, and decisions that reflected a variety of opinions were more likely to occur in participative workplace climates characterized by cooperative goals and constructive controversy. Therefore, findings from organizational behavior research indicated that participative climate and cognitive outcomes were possible mediator variables in the participation-use relationship.

The purpose of the current study was to assimilate possible mediator constructs, which had been substantiated by prior research, into a theory-based model of participatory evaluation. It was intended that the model be used as a framework on which new developments could be related to previous work and thus, decrease the prevailing assumptions regarding participatory evaluation. The constructs in the model were Participative Climate, Level of Participation in Decision Making (Level of PDM), Influence in Participative Decision Making (Influence in PDM), Process Outcomes, Instrumental Use, and Symbolic Use.

The model was developed using an intervening mechanism design (Chen, 1990) which involved testing an action theory and a conceptual theory (see Figure 1). The action and conceptual theories of the intervening mechanism design were the primary hypotheses in the study. The action theory tested the effect of the antecedent program variables (Participative Climate, Level of PDM, and Influence in PDM) on the intervening variable (Process Outcomes). The conceptual theory tested the effect of the intervening variable (Process Outcomes) on the outcome variables (Instrumental and Symbolic Use). Together,

the action and conceptual theories were a means of testing the underlying causal processes of participatory evaluation.

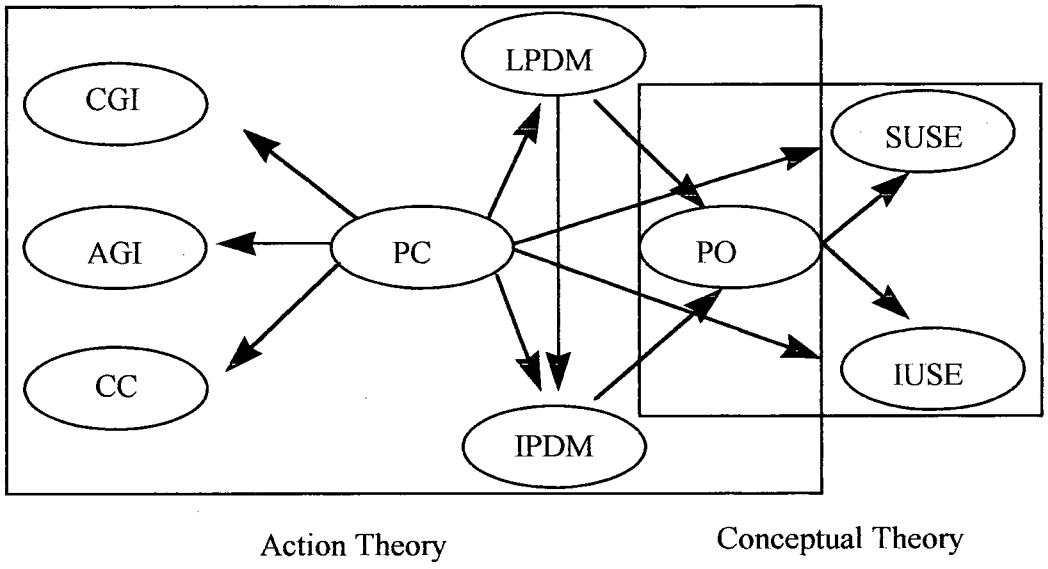


Figure 1. Intervening mechanism design.

CGI = Cooperative Goal Interdependence, AGI = Autonomous Goal Interdependence, CC = Constructive Controversy, PC = Participative Climate, LPDM = Level of Participation in Decision Making, IPDM = Influence in Participative Decision Making, IUSE = Instrumental Use, PO = Process Outcomes, SUSE = Symbolic Use.

The participants in this study were 315 elementary and secondary teachers involved in the 1995/1996 B.C. School Accreditation, which is a mandatory school evaluation program sponsored by the provincial government. The accreditation process requires that staff progress through a series of steps which involve collecting evidence to demonstrate their performance on a series of Ministry set criteria. The evaluation evidence is organized into areas of school strengths and weaknesses and prioritized into a school growth plan, which is intended to guide the school development over a period of approximately six years. The accreditation program is a good example of participatory evaluation in that all decision making is participative and reflects the opinions of the entire staff. Teachers organize and

participate in all phases of the evaluation process from defining school objectives to reporting the evaluation results. The entire accreditation process usually requires a full school year (September to June) to complete.

Definition of Terms

The following definitions are provided to explain how specific terms are intended to be understood within the context of the following Chapters:

1. The term participatory evaluation has been given many different definitions by a variety of evaluation researchers. In the context of the current study, it is intended to mean "...an educational process through which groups produce action-oriented knowledge about their reality, clarify and articulate their norms and values, and reach consensus about further action" (Brunner & Guzman, 1989, p. 21). In participatory evaluation, program stakeholders may work with an external evaluator, but it is the stakeholders not the external evaluator, who are ultimately responsible for carrying out the evaluation and interpreting and reporting the results.
2. Evaluation decision making refers to all the types of decisions that must be made *while* conducting an evaluation. For example, what data will be collected, how the data will be collected, and how the data will be analyzed, interpreted, and reported. These tasks involve logistical as well as content related decision making. Evaluation decision making is not intended to refer to decisions about whether or not to conduct an evaluation.
3. In a seminal review on participation literature, Locke and Schweiger (1979) discussed various definitions for the term participation in decision making (PDM). They pointed

out that although there is little consensus as to its exact meaning, there are several commonalities among the many definitions. For example, they argued that most definitions of participatory decision making include the following criteria: (a) refers specifically to participation in the processes or activities involved in reaching decisions, (b) generally requires that at least two people be involved and that something in common must be shared by these two people, (c) can involve both subordinates and superiors, (d) involves some degree of sharing, but the sharing may not necessarily be equal, (e) can be forced or voluntary, (f) can be informal or formal such as participation between unions, committees, boards (formal) or between managers and subordinates (informal), (g) can be direct (an employee takes the opportunity to assert opinions) or indirect (where elected representatives speak for a larger body of participants), and (h) can also vary in degree (from no-participation to situations where employees and superiors have equal voting power) and context (such as personnel functions, working conditions, or policy).

Participation in decision making within the current study is intended to mean mandatory participation in group decision making related to program evaluation. It is equal in that all those who participate have equal voting power and decision authority. It is informal in that it is based on relationships among program staff, and it is direct in that all participants have the opportunity to provide input.

4. The terms Level of Participation in Decision Making (Level of PDM) and Influence in Participative Decision Making (Influence in PDM) refer to one's perception of the nature and function of their participation. Level of PDM refers to perceptions regarding

degree of participation in pre-evaluation decisions and Influence in PDM refers to perceptions of influence in evaluation content decisions.

5. Cooperative Goal Interdependence refers to employee perceptions of whether their workplace has common goals which all employees strive to achieve. Autonomous Goal Interdependence refers to employee perceptions of whether their workplace goals have little or no effect on other employees. Constructive Controversy refers to perceptions of whether workplace controversy is considered to be positive in that it is used to raise alternate points of view which are incorporated in resulting decisions and actions.
6. The term Process Outcomes is defined as one's perceptions regarding outcomes of participative evaluation processes. As previously stated, participative evaluation processes can involve clarifying program objectives, designing the evaluation, collecting data, and interpreting and reporting results. In participatory evaluation, it is assumed that employee participation in these processes will result in certain outcomes. For example, it is assumed that employees will learn about their program and thereby be more likely to use the new knowledge to make changes to and improve their program. In the context of the current study, the term Process Outcomes refers specifically to teacher perceptions of the Process Outcomes involved in the B.C. School Accreditation Program. The expected Process Outcomes of the accreditation program are that teachers (a) learn useful things about their school, (b) acquire the necessary knowledge to contribute to the development of the school growth plan, (c) acquire a better understanding of the points of view of other staff members, and (d) become better advocates of their school. The concept of Process Outcomes was based on the

assumption that teachers would be more likely to use evaluation information if they believed the participative processes were effective.

7. Evaluation Use and utilization are intended to be interchangeable. The definition of evaluation use in this study is based on the traditional categorization of instrumental (action-oriented use), conceptual (education-oriented use), and symbolic (persuasive-orientated use)(e.g., Greene, 1988b; Shadish et al., 1991).

CHAPTER TWO:
REVIEW OF THE LITERATURE

Introduction

The purpose of this literature review was to provide a theory-based rationale for the three principal relationships in the current study: (a) the relationship between participation in evaluation decision making and workplace climate, (b) the relationship between participation in evaluation decision making and Process Outcomes, and (c) the relationship between Process Outcomes and evaluation use. The review begins with a discussion of organizational research on participative decision making and employee performance, followed by a discussion of workplace climate factors believed to affect participation. Next, I discuss education research on the effects of teacher participation in school decision making, which includes a discussion of teacher participation in evaluation decision making. Finally, a theory-based rationale is provided for the variable called Process Outcomes and its relationship to evaluation use. The chapter concludes with a summary of the literature review.

Organizational Research on Participative Decision Making

Miller and Monge (1986) conducted a meta-analytic review of research on the effects of participation in decision making (PDM) on satisfaction and productivity in journals and books (published in English) in the areas of social psychology, management, organizational behavior, and communication. They excluded dissertations and other

unpublished research and identified 106 articles and book chapters on PDM of which 47 contained quantifiable estimates of the relationship between PDM and satisfaction or productivity (41 correlations between participation and satisfaction and 25 correlations between participation and productivity). The 47 studies were categorized into three explanatory mechanisms through which participation is believed to affect employee satisfaction and productivity: (1) cognitive models, (2) affective models, and (3) contingency models.

Cognitive Models of employee participation are based on the assumption that PDM results in exchange and use of information. Furthermore, cognitive models of participation posit that employees have extensive knowledge of their work, and in many instances more extensive knowledge than management. Thus, if employees participate in company decision making, the resulting decisions will be based on a broader and more accurate pool of information. Also, cognitive models assume employees who participate in decision making will have a better understanding of how to implement decisions once they are made. The knowledge they acquire through PDM will teach them more about their jobs and the company in general.

Affective Models of participation are based on assumed links between worker satisfaction and increased performance. According to Miller and Monge (1986), the causal chain between PDM and productivity in affective models could be explained as: (a) participation fulfills needs, (b) fulfilled needs lead to satisfaction, (c) satisfaction strengthens motivation, and (d) increased motivation improves workers' productivity.

Contingency Models of participation are based on the assumption that PDM affects productivity and satisfaction differently for different people in different situations.

Contingency models focus on personality issues, context specific decision situations, manager and employee relationships, job levels, and values. No single model is believed to be appropriate for all employees in all organizations.

The meta-analytic results of Miller and Monge (1986) showed a mean correlation of .34 between participation and satisfaction and a mean correlation of .15 between participation and productivity. The results indicated only moderate correlational support for cognitive models ($\underline{r}$ = .27) and moderate to strong support for affective models ($\underline{r}$ = .46). The results for contingency models were not significant. Miller and Monge argued the significance of the affective models may have been due in part to methodological moderators such as laboratory versus field studies and type of participant. For example, laboratory studies generally involved simple and well-defined tasks such as question games or device manipulation. Field studies, however, involved subjects in naturally occurring participative activities such as job design or pay incentive plans.

Another methodological moderator believed to affect PDM research is the percept-percept problem. Campbell (1982) described the percept-percept problem as the collection of both PDM and outcome data in a questionnaire administered to a particular group in a single sitting. When the two types of questions were asked in the same questionnaire, respondents typically responded the same way to questions about attitudes toward participation as they did to questions about degrees of participation. Wagner and Gooding (1987a) reported that the percept-percept technique often resulted in, "...spurious similarities....[and] inflated measured effects" (p.245). They concluded that, "...significant participation-outcome research between 1950 and 1985 in America is mainly the product of percept-percept artifacts" (p. 257).

In a similar but later study, Wagner and Gooding (1987b) performed another meta-analytic review of participation literature to provide further evidence of the impact of percept-percept technique. This study involved the effects of four situational moderators on five types of outcomes. The four moderators were (a) group size, (b) task interdependence, (c) task complexity, and (d) performance standards. The outcomes were (a) task performance, (b) decision performance, (c) motivation, (d) satisfaction, and (e) acceptance. The situational moderator subgroups were nested within percept-percept and multi-source groups. Multi-source correlations were defined as, "those that researchers had gathered using at least one objective measure or assigned condition, different respondents for data on participation and outcome variables, or a longitudinal break between the collection of data on both participation and outcome variables for the same respondents" (Wagner & Gooding, 1987b, p. 153). The results were similar to their previous study in that percept-percept correlations were typically larger than multi-source correlations.

Overall, the meta-analytic research by Miller and Monge (1986) and Wagner and Gooding (1987a, 1987b) tended to support affective models of participation to a greater extent than contingency or cognitive models. Locke and Latham (1990) pointed-out, however, that the greater evidence for affective models may be due in part to the greater number of studies on motivational links between PDM and performance as compared to the relatively small number of experimental and quasi-experimental studies on cognitive effects of participation. In response to the vast and inconclusive literature on affective models of PDM, some researches have begun to re-examine the cognitive benefits of PDM (i.e., Latham et al., 1994). Based on the cognitive studies in the meta-analytic reviews discussed in previous sections and on further review of the most current research, four studies were

found to be most representative of the current understanding of cognitive outcomes of PDM: Latham and Saari (1979), Erez and Arad (1986), Campbell and Gingrich (1986), and Latham et al. (1994).

Latham and Saari (1979) have often been cited (e.g., Campbell & Gingrich, 1986; Erez & Arad, 1986; Latham et al., 1994; Locke & Latham, 1990) as one of the first studies to provide evidence that participation increases performance through increased information sharing. The purpose of the study was to test the importance of a supportive climate in participative and non-participative goal setting when difficulty of goals was held constant. The authors hypothesized that (a) supportive management style would lead to higher job performance than non-supportive management style, (b) setting specific goals would lead to better performance than assigning goals in a "do your best" manner, and (c) participative goal setting would result in higher performance than assigned goals.

The findings indicated supportive climates resulted in higher goal setting than non-supportive climates and participatively set goals led to better performance than assigned goals. These findings were contradictory to the majority of research, which suggested no significant difference between assigned and participatively set goals when goal difficulty was controlled (cf. Ivancevich, 1977; Locke & Latham, 1990; Shalley, Oldham, & Porac, 1987; Vanderslice, Rice, & Julian, 1987). The authors attributed the contradictory findings to the impact of supportive participation on task understanding. They discovered more questions were asked about the task requirement in the participative categories than in the assigned categories. The question dialogue was believed to be responsible for the increased understanding and in turn significant results in the participative climate in general and the supportive participative climate in particular. The nature of participation was believed to

have affected the level and quality of group interaction, which in turn was thought to have affected the degree of information sharing and utilization of information in the resulting decisions. They argued that a supportive climate, which fostered question dialogue, appeared to increase understanding and performance.

In a similar study, Erez and Arad (1986) examined the effects of group discussion, motivation, and information on PDM and performance in a simulated setting. The independent variables were (a) perceived social interaction, (b) perceived involvement in goal-setting, (c) perceived amount of relevant information, (d) ability, (e) performance quantity and quality, (f) attitudes (goal acceptance), (g) satisfaction with job and satisfaction with co-workers, (h) group commitment, and (i) incidental learning. The hypotheses concerning attitudes (i.e., goal acceptance, group commitment, co-workers satisfaction, and work satisfaction) and incidental learning were tested with a series of three-way ANOVAs. Some of the results of these analyses were: (a) group discussion and involvement in goal setting had a significant positive effect on performance quantity, (b) participants in the conditions of low-group discussion, low-involvement, and low-information performed significantly worse than participants in high-group discussion, high-involvement, and high-information, (c) goal acceptance was higher in the high-involvement and group discussion conditions than the low-involvement and low-group discussion conditions, (d) group discussion had a positive and significant main effect on goal commitment, and (e) there was no significant effect of information on co-worker or worker satisfaction.

The findings listed above support Latham and Saari's (1979) hypothesis that discussion and question dialogue promote task understanding and in turn performance. Although Erez and Arad (1986) were not specifically examining the impact of a supportive

climate, the treatment conditions of high-discussion and high-involvement were much the same as the conditions in Latham and Saari. This similarity suggested that the concepts of supportive climate used in Latham and Saari and high-involvement used in Erez and Arad had comparable effects on information sharing and performance.

In another commonly cited study on the cognitive benefits of PDM (e.g., Latham et al., 1994; Locke & Latham, 1990), Campbell and Gingrich (1986) examined the interactive effects of task complexity, participation, and performance. The study involved forty computer programmers who were required to perform the task of writing either a simple or a complex program. Half of the programmers were able to participate in the decision regarding a task completion date as well as to discuss the project. The other half of the programmers were assigned completion dates equal to the first group but were not permitted discussion regarding a completion date for the task. Campbell and Gingrich hypothesized that for complex tasks, those who used PDM would perform better than those who did not. They also hypothesized participation would have no effect on simple tasks.

The results indicated complex tasks involving PDM resulted in significantly higher performance than the other groups. The difference in performance was attributed to information sharing. The authors posited that benefits for simple tasks involving participation were not readily apparent and that there may be little difference in outcomes between assigned or participative decisions. However, when tasks were complex, the cognitive outcomes of participation such as strategy development, information processing, and problem solving resulted in more effective task strategies and subsequent decision outcomes than if participation was not employed.

A recent study by Latham et al. (1994) provided further evidence for a cognitive link between PDM and performance in organizational contexts. The intent of the study was to determine whether (a) participation in strategy development, task strategy quality, which was operationalized as the effectiveness of the task strategy, and (b) self-efficacy mediated the relationship between PDM and performance. The findings suggested learning which occurred as a result of participation, and the self-efficacy associated with this learning, had a strong mediating effect on the participation-performance relationship. The results supported their hypothesis that participation in strategy development and task strategy quality had a significant effect on performance. Figure 2 depicts a diagram of their findings. The results indicated that self-efficacy and performance increased when participants were able to share and implement task relevant knowledge.

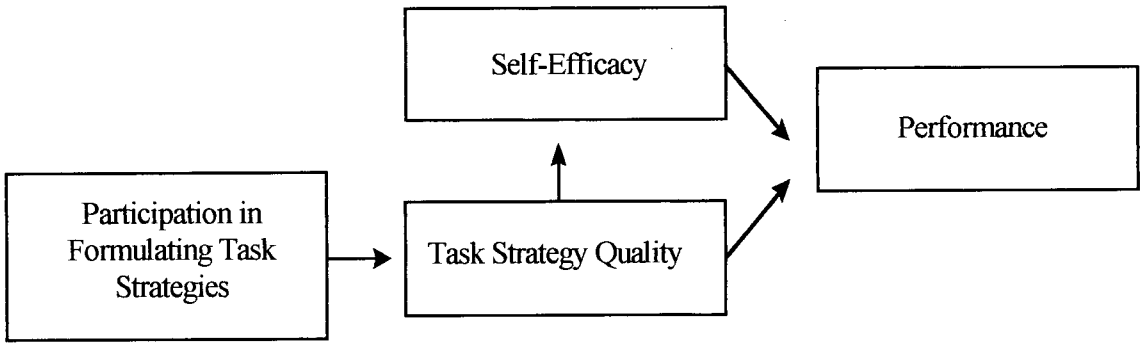


Figure 2. Relationship between participation, self-efficacy, and performance as reported in Latham et al. (1994).

Summary

The studies included in this discussion represent the most recent citations available through the on-line computer data-bases (e.g., CARL; Dissertation Abstracts; Psych Info;

Reveal Uncover; RIE; Social Sciences Citation Index CD-ROM)[1]. The findings provided evidence for a cognitive link between PDM and employee job performance. Collectively, the research suggested that in participatory decision making contexts, the following six factors were found to have positive effects on employee performance: (a) group discussion involving question dialogue and information sharing (Latham & Saari, 1979), (b) participation in goal setting (Erez & Arad, 1986), (c) participation in complex versus simple tasks (Campbell & Gingrich, 1986), (d) participation in the development of task strategies (Latham et al., 1994), (d) task strategy quality (Latham et al., 1994), and (e) self-efficacy (Latham et al., 1994).

Organizational Research on Participative Climate

Concomitant with the growing body of literature on PDM has been an expanding research and theory base derived from studies on participative climate. A discussion of PDM would, therefore, be incomplete without addressing effects of contextual factors on PDM. A commonly cited series of studies by Tjosvold and colleagues (i.e., Tjosvold, 1985, 1986, 1987; Tjosvold, Andrews, & Jones, 1983; Tjosvold & Deemer, 1980; Tjosvold & Johnson, 1978; Tjosvold & McNeely, 1988; Tjosvold Wedley, & Field, 1986) suggested that (a) organizational goal orientation and (b) employee controversy affected PDM.

---

[1] The descriptors used for the current literature review were (a) participative decision making, (b) participation, (c) employee participation, (d) participative goal setting theory, (e) employee performance and employee motivation, (f) employee productivity, (g) management theory and styles, (h) models of organizational management (I) participatory program evaluation, (j) evaluation use, (k) self-efficacy, (l) regression analysis, (m) LISREL, (n) latent-variable modeling.

Organizational Goal Orientation

Deutsch (1980) theorized that when people work in cooperation, they believe their goals to be positively related. When one person achieves a goal, it is more likely others will also reach their goals; one person's success helps others to be successful. However, when people work in competition, they believe their goals to be negatively related. When one person reaches a goal, it makes it less likely that others will reach their own goals; one person's success hinders others' success. Deutsch hypothesized that in cooperation people are more likely to encourage others to be effective because their success is interconnected. In competition, however, people may actually work against one another as a means of ensuring their own success.

Based on the assumptions of Deutsch's theory of cooperation and competition (also called goal linkage theory), Tjosvold and colleagues (i.e., Tjosvold, 1985, 1986; Tjosvold et al., 1983; Tjosvold & McNeely, 1988) conducted a series of studies on goal interdependence in organizations. The findings provided evidence to substantiate the hypothesis that goal linkage theory has implications for work relationships. For example, Tjosvold et al. (1983) found that employees who perceived their leaders' goals to be similar to their own, believed "goal-linkage" had a positive effect on their job performance. Conversely, employees believed their job performance was affected negatively when they thought their leaders had competitive or individualistic goals.

In a more recent study, Tjosvold (1986) tested the concepts of goal linkage theory in a study to determine whether (a) employees who perceived their goals as positively linked helped others to succeed by sharing information and exchanging resources and (b) if

employees who perceived their goals as negatively linked saw others' success as a threat and subsequently did not share information or help others to succeed. It was hypothesized that positive goal linkage and independent goal linkage (employees perceive their goals as unrelated to those of other workers so their success doesn't hinder or assist the success of others) would facilitate higher work performance than negative goal linkage.

Participants included 50 employees of a social service organization for the mentally handicapped. An interview was complemented with a questionnaire designed to elicit information on participants' interactions with co-workers and supervisors. In the interview, participants were asked to explain in detail a recent effective interaction with a co-worker and an ineffective interaction with a supervisor. The interviewer asked questions pertaining to participants' objectives and goals and their co-workers' objectives and goals. The interviewer defined the three goal linkages: (a) cooperative, (b) competitive, and (c) autonomous. Participants were asked to distribute ten points over the three goal linkages and give the most points to the one that best represented the situation. The goal linkage given the most points was used to classify the situation.

The results showed large differences among the three goal linkages for all of the interview scales, which suggested that goal orientation affected participation in decision making. Tjosvold (1986) concluded that

> in positive linkage organizational members exchanged valuable resources, negotiated without affronting social face, felt successful and satisfied during interaction, and thought the interaction was productive and strengthened their relationship. Independently linked interactions had dynamics characterized by less exchange and less constructive conflict management than positively

> linked situations, but not less than negatively linked interactions and was associated with productivity. (p. 525)

Controversy

Controversy is conflict that occurs when one person's ideas, information, and opinions are incompatible with those of another when they discuss problems and make decisions. When controversy is productive, people are more likely to understand and assimilate opposing points of view and arguments, which results in high-quality decisions (Tjosvold, 1987). When controversy is unproductive, people express their opinions but in a closed-minded manner. For example, Tjosvold (1985) found in non-constructive controversy, employees often tried to find

> weaknesses in opposing arguments rather than trying to incorporate them into their own. Then they are better able to counterattack, undercut other positions, and make their own views dominate, relying on superior authority or other means to try to impose their solution. In this way, controversy creates polarization and results in low-quality decisions which only the winners are committed to implementing. (p. 41)

Productive controversy is more likely to occur in a climate with cooperative goals than in a climate with competitive goals (Tjosvold & Deemer, 1980). Tjosvold (1987) stated, "Decision makers in cooperation have been found to express controversy openly, consider opposing points of views without bias, and arrive at effective, integrated solutions" (p. 741). Furthermore, Tjosvold (1985) found that in competitive climates, there was more rejection of opposing points of view and less incorporation of others' ideas.

Tjosvold (1987) linked constructive controversy to cognitive outcomes through cognitive development theory (Berlyne, 1963, 1965). Berlyne hypothesized if people encounter ideas that contradict their own, they experience conceptual conflict which in turn leads to uncertainty as to which idea is correct. To alleviate this uncertainty, they ask questions and search for additional information. The process of trying to understand both points of view can lead to the attainment of additional or new information, which is hypothesized to result in greater acceptance and assimilation of contradictory ideas.

Tjosvold and Johnson (1978) posited that a participative climate with cooperative goals and constructive controversy was the most effective in facilitating the dialogue necessary to understand conflicting points of view. They noted, however, that both cooperative and competitive climates involving controversy were effective in producing the level of uncertainty necessary to initiate increased question dialogue and increased understanding of conflicting points of view. This does not necessarily mean controversy will be more effective in a competitive climate. Tjosvold and Johnson claimed the high degree of information acquired in competitive climates may not be understood or assimilated into the decision outcomes due to closed-minded attitudes and disregard for the opponent and the opponent's ideas. However, in a cooperative climate, the regard for group members and their ideas is much greater, and although conflict will occur, varying ideas are more likely to be utilized.

Summary

Research on participative climate suggests that goal linkage and controversy have implications for the cognitive outcomes of PDM. Participative climates which fostered cooperative goals and constructive controversy were found to promote effective information

sharing, open expression of conflicting ideas, and integration of solutions (Tjosvold, 1986; Tjosvold et al., 1983). Moreover, participants who experienced opposing points of view in a climate with cooperative goals and constructive controversy, as opposed to competitive goals and non-constructive controversy, were more likely to assimilate various perspectives into decision outcomes, which was assumed to result in high-quality decisions (Tjosvold, 1985, 1987).

Based on the organizational and participative climate research discussed thus far, the following eight factors are believed to have positive effects on employee job performance in participatory decision making contexts:

1) group discussion involving question dialogue and information sharing (Latham & Saari, 1979)
2) participation in goal setting (Erez & Arad, 1986)
3) participation in complex versus simple tasks (Campbell & Gingrich, 1986)
4) participation in the development of task strategies (Latham et al., 1994)
5) task strategy quality (Latham et al., 1994)
6) employee self-efficacy (Latham et al., 1994)
7) participation involving constructive versus non-constructive controversy (Tjosvold, 1985, 1987)
8) participation in cooperative versus competitive climates (Tjosvold, 1986; Tjosvold et al., 1983)

Education Research on Participation in Decision Making

Teacher participation in general school decision making

Literature on teacher participation in school decision making is vast and has been perpetuated in part by several inherent assumptions. Estler (1988) stated, "Perhaps more than any other tradition within the decision-making literature, the participatory model is rooted more in values and beliefs than in empiricism" (p. 309). Policy makers and educational planners have intuitively justified the participatory approach based on (a) the utilization of teachers' first-hand knowledge of students and (b) valuing participation as a humanistic, democratic, and ethical management approach (Conway, 1984; Estler, 1988). Consequently, educational researchers have devoted considerable energy to investigating teacher satisfaction with PDM.

It is generally believed if teachers are involved in PDM the assumed links between PDM and increased school effectiveness will occur. This logic is evident in effective schools literature (e.g., Carnegie Commission, 1986; Carnegie Foundation, 1988). Educational policy makers have promoted management and decentralization strategies such as site-based management and PDM as a means of giving teachers more decision making authority (Ambrosie & Haley, 1988). It is believed the distribution of power will increase teacher professionalism and ultimately empower teachers, which in turn will result in more effective teachers and schools (Cohen, 1982; Goodlad, 1984).

The mass of literature on teacher participation has focused on teacher satisfaction with PDM. The findings suggested that teachers are more satisfied with participation when they have influence in the decision process. Duke, Showers, and Imber (1980) found that

teachers distinguished between participation and influence and that in order to understand PDM, it was first necessary to distinguish among factors that make teachers feel their involvement has a bearing on the decision process. In a similar study, Imber, Neidt, and Reyes (1990) examined factors that contributed to teacher satisfaction with participative decision making. Their findings were similar to Duke et al. (1980) and they concluded that, "Involvement alone such as attending meetings, expressing an opinion, casting a vote and so forth does not necessarily make PDM a satisfactory experience" (p. 223). They argued that the PDM processes must also give teachers influence, which according to Imber et al. (1990), could be achieved by implementing the decisions that result from PDM in that implementation would be seen as verification of teacher influence.

In spite of the large literature on teacher satisfaction with participation, few empirical studies have tested the assumed causal link between PDM and school performance. Furthermore, findings from this research are equivocal. For example, a recent study by Taylor and Bogotch (1994) suggested the effects of PDM on teacher and school-level outcomes were unclear. The intent of the study was to determine whether increasing teacher involvement in decision making had an effect at the school level. Specifically, the authors questioned whether teacher participation in school decision making could be organized into categories and whether the categories would correlate with teacher job satisfaction, teacher and student attendance, student achievement, and student behavior.

The participants were 1654 teachers in 33 schools in a large US school district. A matched-pair design was used with schools that did and did not participate in a decentralization pilot project. Teachers from schools involved in the decentralization project had greater exposure and experiences with PDM in various dimensions of school

decision making than did teachers in non-participating schools. Exploratory factor analysis was used to identify dimensions of teacher participation in decision making. These dimensions were correlated with job satisfaction and the school-level outcomes. The results indicated that (a) although several different dimensions of teacher participation in school decision making were identified, teacher participation did not have a significant effect on student or teacher outcomes and (b) teachers felt they were deprived of decision making power. Furthermore, teachers felt they had less opportunity to participate in decision making, less influence in the decision making process than they desired, and felt frustrated by the lack of administrative support in that many administrators were not able to give up decision making control. Taylor and Bogotch (1994) concluded that the decentralization project had little impact on school performance.

In a recent study on the effects of PDM, Smylie, Lazarus, and Brownlee-Conyers (1996) attributed the lack of empirical research on the effects of PDM in education contexts to the nature of the scholarly literature. They argued that the literature on PDM in schools

> consists mostly of position statements, essays, projects descriptions, and status reports. A relatively small proportion consists of systematic empirical investigations-both qualitative and quantitative-with identifiable questions for inquiry, specified methodologies, and collection and analysis of original data [and] relatively few studies are designed to explain the nature and function of participation. (p. 182)

In a an effort to examine the nature and function of participation using sound empirical methods, Smylie et al. (1996) examined the relationship between PDM and instructional outcomes using a longitudinal design. Specifically, they examined the

relationship among (a) participative decision making, (b) three change mechanisms: individual autonomy, individual accountability, and organizational learning, (c) instructional improvement, and (d) three student outcomes (teacher reports, change in reading achievement, change in math achievement). The design involved three data collection periods over six years: 1990 (n = 116), 1992 (n = 162), and 1994 (n = 174). The participants were elementary teachers in seven schools from one district. The school rather than the teacher was used as the unit of analysis. There were six schools in the final analysis. Correlation analysis with the six schools indicated a significant relationship between PDM and change in individual accountability ($\underline{r} = .96, p < .01$). The correlation between change in reading achievement and instructional improvement was also significant ($\underline{r} = .92, p < .01$). However, the other correlations, although reported as meaningful were not reported with probability levels.

I question whether Smylie et al. (1996) or Taylor and Bogotch (1994) addressed what Smylie et al. called the nature and function of PDM. In order to address the nature and function of participation, further research is needed which incorporates variables used in organizational research. For example, education PDM research has yet to test the effects of variables such as supportive versus non-supportive climates, cooperative versus competitive climates, controversy, goal setting, complex versus simple tasks, participation in decision strategies, task strategy quality, and self-efficacy on teacher and school performance. The reasons why education researchers have not used such variables could in part be due to the political realities of collecting data in unionized settings where it is often prohibited to comment on other teachers or issues related to other staff members.

In spite of the issues of collecting data in union settings, there are PDM studies which include school climate variables. As explained by Smylie et al. (1996), subsumed in the education research on PDM and school performance is a large literature based on position statements and project descriptions associated with PDM. One such example is a commentary article by Weiss, Cambone, and Wyeth (1992), which examined teacher conflicts in participative decision making. This article is described as a commentary because the authors did not explain their methods or analysis strategies. The article is an interesting discussion of the authors' perceptions of approximately 180 interviews conducted with public high school teachers in 45 schools across 15 states in a two year period. Based on the interview data, the authors posited that issues such as constructive controversy and understanding colleagues' points of view were crucial to effective PDM in that teachers were generally unable to use conflict constructively and had difficulty integrating various points of view into decisions. They believed training in the process of decision making was crucial for effective PDM. The points raised by Weiss et al. (1992) are very much like those tested in organizational PDM research (i.e., Tjosvold, 1985, 1987). However, educational researchers have yet to test the effects of these variables on PDM and school and teacher performance.

Teacher participation in evaluation decision making

Although the empirical literature on causal links between PDM and school performance is limited, there are even fewer empirical studies on outcomes of PDM in school evaluation contexts. The only empirical study I found on the effects of teacher participation in evaluation decision making was by Brandon, Wang, and Heck (1994). In a

school needs assessment context, the authors employed structural equation modeling to test the effects of (a) teacher participation in pre-evaluation decisions such as how decisions would be made and (b) evaluation content decisions on teacher agreement with assessed needs. The authors posited that teacher involvement in pre-evaluation decisions would predict teacher participation in evaluation content decisions and, in turn, agreement with the assessed needs. Brandon et al. (1994) concluded that the model was a plausible representation of the data. The direct effect between input into content decisions and agreement was .73 and .49 between participation in pre-evaluation decision and participation in content decisions. There was a relatively small (.17) direct effect between input into process and agreement. However, the indirect effect through input into content was moderate (.49 x .73 = .36). The findings of Brandon et al. are significant in that they emphasize the importance of differentiating between involvement in pre-evaluation and evaluation content decisions, which suggests that each has a separate effect on the outcomes of participation. Teachers who had involvement in pre-evaluation as well as evaluation content decisions were more likely to agree with the decisions that were made.

Further evidence of the shortage of empirical research on teacher participation in school evaluation is in a recent and thorough review article by Cousins and Earl (1992), who summarized 31 empirical studies on organizational learning and use in educational settings dating from 1980 to 1992. The studies were chosen on the criteria that (a) original data were collected, (b) the use of information was the focus of the study, and (c) evidence was provided to support the linkage between collecting and using information. Of the 31 studies, there were only five survey designs and none was a school evaluation study.

Based on the findings of the 31 studies, Cousins and Earl (1992) proposed a theoretical model of participatory evaluation intended for educational settings. The model included five organizational elements and six evaluator requirements. The organizational components were that (a) the organization must see value in conducting an evaluation, (b) the time and resources necessary to conduct an evaluation must be made available by the organization, (c) the organization must be committed to learning and improvement, (d) evaluation participants must be willing and motivated to participate, and (e) some of the participants must have research skills or have the potential to learn them quickly. The six evaluator requirements were that (a) the evaluator must have the skills to conduct the evaluation, (b) be accessible to the organization, (c) have resources to conduct the evaluation appropriately, (e) be willing to take on a pedagogical role, (f) be motivated to participate, and (g) have intolerance for imperfection. Based on these criteria, Cousins and Earl believed, "...participatory evaluation offers a powerful approach to the improvement of educational organizations by creating learning systems that enhance organizational learning and, consequently, lead to better informed decisions" (p. 412).

The model proposed by Cousins and Earl (1992) did not address the nature and function of participation. Instead, it reflected the assumption that teacher satisfaction with PDM would result in favorable outcomes. Hence, the model incorporated elements that were believed to contribute to teacher satisfaction. Factors that would have addressed the nature and function of teacher participation such as controversy, cooperation, competition, task complexity, and efficacy were not included. Furthermore, although the model was based on a thorough review of previous literature, its usefulness is limited until tested empirically.

Summary

Empirical research on teacher participation in school and evaluation decision making has resulted in equivocal findings. In a recent study, Smylie et al. (1996) found several significant correlations among PDM and teacher and school performance. Taylor and Bogotch (1994); however, found that PDM had little effect on student, teacher, or school performance and argued that PDM was less effective than anticipated because teachers had less influence in decision making than they desired. Our understanding of PDM in education evaluation contexts has been hindered by a lack of empirical research. In a recent study on PDM in education evaluation contexts, Brandon et al. (1994) found that teacher participation in pre-evaluation and evaluation content decisions was positively related to agreement with the decisions made. In a discussion paper, Cousins and Earl (1992) proposed a theory-based model of PDM in school settings. However, the model did not include variables which were found to affect PDM and performance in organizational research. Education evaluation research has yet to examine factors similar to those used in organizational research such as the impact of controversy, conflict, and efficacy on PDM and use of evaluation information.

Process Outcomes

The concept of Process Outcomes is a key component and the primary mediator variable in the current study. The concept was derived from two theoretical perspectives: (a) Bandura's (1986) theory of perceived self-efficacy and (b) the theory of conceptual use in program evaluation literature (Greene, 1988a, 1988b; Shadish et al., 1991).

Self-efficacy as a theoretical basis for Process Outcomes

Bandura's (1986) theory of perceived self-efficacy has been applied in many areas of study such as (a) psychological research on fear arousal and phobics (e.g., Bandura, 1983; Bandura, Reese, & Adams, 1982), (b) organizational theory research on employee performance (e.g., Gist, 1987, 1992; Latham, et al., 1994), and (c) education research on causal connections between self-efficacy and school related variables such as achievement, motivation, career choices, and teacher change (e.g., Ambrosie & Haley, 1988; Carnegie Commission, 1986; Carnegie Foundation, 1988; Dembo & Gibson, 1984; Leithwood, 1992; Owen & Froman, 1992; Stein & Wang, 1988). Moreover, there is a vast literature on measuring self-efficacy (e.g., Gable & Wolf, 1993) as well as many published measures of self-efficacy (e.g., Dembo & Gibson, 1984). Based on the breadth of the self-efficacy construct, it is reasonable to assume that it would be relevant in a participatory evaluation context.

Bandura (1986) defined perceived self-efficacy as "people's judgments of their capabilities to organize and exercise courses of action required to attain designated types of performance" (p. 391). Bandura's theory was based on the assumption that someone who has the skills needed to perform a particular act does not necessarily have the belief that they can or are able to perform the act. Therefore, self-efficacy has been described as a "...cognition that mediates between knowledge and action" (Raudenbush, Rowan, & Cheong, 1992, p. 150). Self-efficacy is context specific. Those who have high self-efficacy in one domain may have low self-efficacy in other domains. According to Bandura, people develop self-efficacy perceptions through (a) physiological cues such as anxiety, nervousness, or sweating during a particular task, (b) verbal persuasion from others, and (c)

positive previous experience with specific behaviors in that successful experiences result in higher self-efficacy than failures.

Like self-efficacy, Process Outcomes is believed to be a cognition that mediates between participation and use of evaluation information. It is defined as teacher perceptions of participative Process Outcomes and is measured by determining whether participative processes produced desired outcomes (explained further in Chapter Four). Research on school improvement suggests that self-efficacy variables similar to the concept of Process Outcomes are related to program implementation and thereby provides a theoretical basis for the function of Process Outcomes in the current study. For example, Stein and Wang (1988) conducted a study on teacher implementation of innovative education programs in which they examined the relationship among (a) perception of self-efficacy for implementing the program, (b) perceived value of the program, and (c) teacher perceptions of success in implementing the program. One of the predictor variables for the perceived value construct was called "teacher perception of program efficacy." They defined this variable as teacher perceptions of program outcomes and general assessment of the program. The authors found that teacher perception of program efficacy was positively related to perceived program value and teacher success in implementing the program.

There is a parallel between the Stein and Wang (1988) "teacher perception of program efficacy" variable and the Process Outcomes variable in the current study in that the former is concerned with teacher perceptions of the efficacy of an education program and the latter is concerned with teacher perceptions of the effectiveness of a participative process. Stein and Wang reported a significant relationship between program efficacy and perceived implementation success, which formed the basis for assuming a significant

relationship between Process Outcomes and Instrumental and Symbolic Use in the current study.

Conceptual use as a theoretical basis for Process Outcomes

As previously stated, conceptual use is commonly described as increased understanding and learning about a program as a result of an evaluation (Greene, 1988a, 1988b; Shadish et al., 1991). The Process Outcomes construct as defined in the current study embodies the notion of conceptual use in that it incorporates the so called cognitive outcomes or learning related outcomes of conceptual use. However, Process Outcomes is not considered to be use. Rather, it is one's perception of the participatory process outcomes. For example, one expected outcome of participatory evaluation processes is that participants learn something about their program, which is central to the idea of conceptual use. There are, however, additional and more specific outcomes which could be expected to result from participatory processes. For example, it would be expected that participants (a) acquire a better understanding of colleague's point of view, (b) be able to identify strengths and weaknesses of the program, and (c) be able to apply what they have learned in terms of developing implementation strategies. Therefore, within the current study, conceptual use is considered to be an element or component of the Process Outcomes construct.

As part of a study on participation and evaluation use, Greene (1988a) developed a scale called Stakeholder Perceptions of the Usefulness of the Evaluation Process and Results. Three of the ten items on the questionnaire were listed as either conceptual or symbolic "Uses of the Process." The three items were as follows, (a) I've learned about some youth employment/day care problems I wasn't aware of (conceptual use), (b) The process has helped me stop and think about what's most important for the YE/DC program

(conceptual use), and (c) The process of participating has made me a better advocate for the YE/DC program (symbolic use). Greene argued that responses to these three items reflected participant perceptions of the "Uses of the Process".

I disagree with Greene's (1988a) interpretation of these items as process uses. I believe they describe Process Outcomes, not process uses. As will be discussed in Chapter Four, I adapted Greene's three items and included them in the various scales developed for the current study. However, the adapted items were intended to reflect Process Outcomes rather than process uses. I believe the items (and other structurally similar items) are easier to understand when defined from a self-efficacy perspective in that each one addresses perceptions of process effectiveness in achieving particular outcomes. It must be noted, however, that although Greene used the items to describe process uses and I used them to describe process outcomes, we both posit that the concepts precede or predict Instrumental and Symbolic Use of evaluation information.

Testing Mediation

Due to the reference to Process Outcomes as a mediator variable, further explanation of the mediation concept is appropriate. Statistical mediation is generally described as either complete or partial (James & Brett, 1984). A complete mediation model is such that $A \rightarrow B \rightarrow C$, where A is an antecedent, B is the mediator, and C is the consequence. The relationship between the antecedent and the consequence is indirect in that all of the effect of A on C is transmitted by B. This means the effect of A on C is *completely* mediated by B. A partial mediation model (in a recursive relationship) is such that $A \rightarrow B \rightarrow C$. In a partial mediation model, the antecedent A has both a direct and indirect effect on the

consequence C, which means the total effect of A on C is only due in part to the mediation by B.

Structural equation modeling has traditionally been accepted as the most effective means of testing mediation (Fiske, Kenny, & Taylor, 1982; James & Brett, 1984). However, due to relatively recent accessibility to modeling programs such as LISREL, mediation has most commonly been tested by regression models (Baron & Kenny, 1986). Testing mediation with regression involves estimation of three regression equations: (a) regress the independent variable on the mediator, (b) regress the independent variable on the dependent variable, and (c) regress the mediator on the dependent variable. Mediation is established through the following conditions: (a) in the first regression equation, the independent variable must predict the mediator, (b) the independent variable must predict the dependent variable in the second regression equation, and (c) the mediator must predict the dependent variable in the third equation. If all these conditions are established, then the effect of the independent variable on the dependent variable will be greater in the second equation than in the third (Baron & Kenny, 1986).

Multiple regression analysis assumes there is no measurement error in the independent variables (that the observed variables are 100% reliable). Traditionally, the issue of unreliability has been addressed through latent-variable modeling and the use of multiple indicators. Subsequently, latent variable modeling (also called structural equation modeling) has become the preferred method to test mediation (Fiske et al., 1982; James & Brett, 1984). A more detailed description of structural equation modeling is provided in the following chapters.

Chapter Summary

A series of review articles (Miller & Monge, 1986; Wagner & Gooding, 1987a, 1987b) showed little empirical support for the PDM-performance relationship. Gooding and Wagner (1987a, 1987b) suggested most significant findings linking PDM to job satisfaction and performance were due to methodological artifacts such as the percept-percept techniques, inferences made from laboratory to field settings, and type of subjects involved. In response to the equivocal nature of the PDM-performance relationship, researchers have begun to study PDM from a cognitive perspective. Models predicting cognitive outcomes of PDM assume learning that results from PDM leads to better informed decision makers and thereby increased performance (Locke & Latham, 1990). A summary of key findings on cognitive outcomes of PDM suggested that (a) PDM in supportive climates was related to increased information sharing and employee performance (Latham & Saari, 1979), (b) group discussion, information, and involvement in goal setting affected employee performance (Erez & Arad, 1986), (c) PDM with complex rather than simple tasks was positively related to employee performance (Campbell & Gingrich, 1986), (d) employee participation in the development of task strategies was positively related to performance (Latham et al., 1994), and (e) task strategy quality and self-efficacy mediated the relationship between participation and performance (Latham et al., 1994).

Research on participative climate has implications for PDM. For example, work environments with cooperative goals and constructive controversy were found to promote effective information sharing and more open expression of ideas (Tjosvold, 1986; Tjosvold et al., 1983). Furthermore, participants who experienced conflict in a climate with

cooperative rather than competitive goals were more likely to assimilate opposing points of view in decision outcomes, which resulted in more effective decision making (Tjosvold, 1985, 1987).

Although there is a substantial literature on participation in educational settings, few researchers have tested causal links between PDM and performance. Moreover, the findings reported in this small literature were equivocal. Taylor and Bogotch (1994) correlated dimensions of teacher participation in school decision making with student, teacher, and school outcomes and found that PDM had little relationship with teacher or school performance. However, Smylie et al. (1996) examined PDM and teacher and school performance and found significant relationships between (a) teacher participation and change in individual accountability and (b) reading achievement and instructional improvement.

Brandon et al. (1994) was the only study found which tested the effects of teacher participation in an evaluation context. The results indicated that teacher participation in pre-evaluation decisions affected participation in content decisions and both variables affected agreement with decisions in a school needs assessment context. The findings emphasized the importance of differentiating between participation at various stages of an evaluation process.

Overall, education research has demonstrated little evidence to substantiate a causal relationship between PDM and school or teacher performance. This may be due in part to the observation made by Smylie et al. (1996) that education PDM literature has not addressed the nature or function of participation. Education researchers have yet to examine

the effects of variables such as controversy, competition, and efficacy on PDM and performance in PDM in general or in evaluation decision making in particular.

CHAPTER THREE:

HYPOTHESES

There are two primary hypotheses tested in the current study. Hypothesis one tests the action theory and hypothesis two tests the conceptual theory for the intervening mechanism design (see Figure 3). Hypotheses one and two are each comprised of several individual path hypotheses and are explained in the following sections. The intervening mechanism design is explained further in Chapter Four.

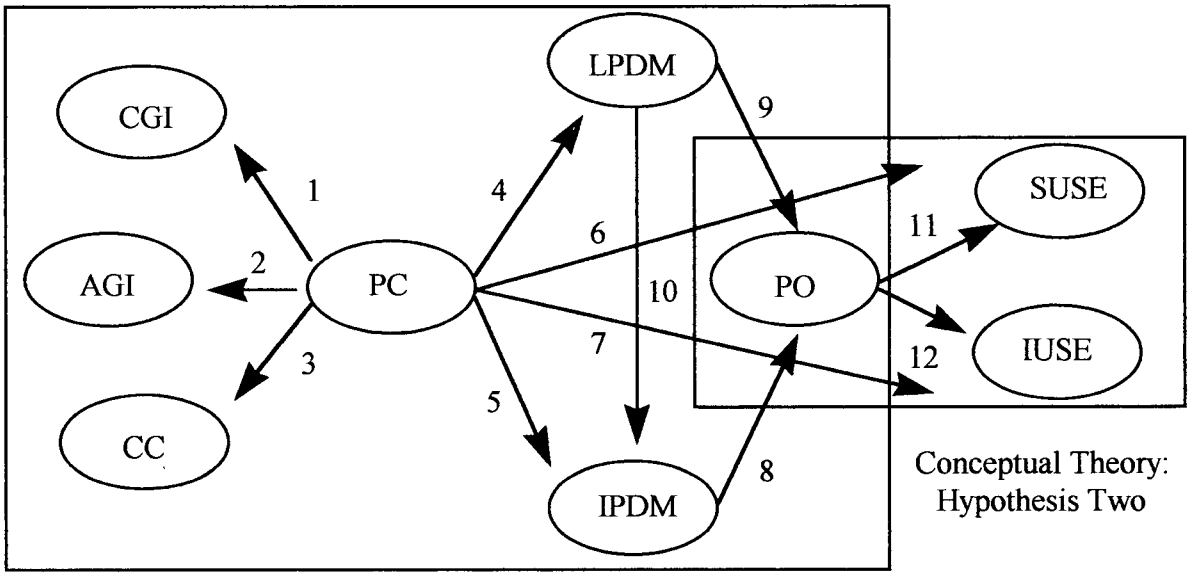


Figure 3. Hypothesized structural model.

CGI = Cooperative Goal Interdependence, AGI = Autonomous Goal Interdependence, CC = Constructive Controversy, PC = Participative Climate, LPDM = Level of Participation in Decision Making, IPDM = Influence in Participative Decision Making, IUSE = Instrumental Use, PO = Process Outcomes, SUSE = Symbolic Use.

Hypothesis One (Action Theory): The three antecedent program variables, Participative Climate, Level of Participation in Decision Making (Level of PDM), and Influence in Participative Decision Making (Influence in PDM), predict Process Outcomes (see Figure 3).

Paths 1, 2, and 3: These paths are not considered to be hypotheses. Instead, they are latent indicator variables for the Participative Climate construct. Their primary function in the model is through Participative Climate and they are discussed further in Chapter Five.

Paths 4 and 5: Paths 4 and 5 illustrate the hypothesis that Participative Climate predicts Influence in PDM and Level of PDM. This hypothesis was based on research (i.e., Erez & Arad, 1986; Latham & Saari, 1979; Tjosvold, 1986; Tjosvold et al., 1983) which indicated that supportive work climates characterized by cooperative or autonomous goals and constructive controversy led to greater information sharing and integration of opposing points of view than work climates with competitive goals and non-constructive controversy. Therefore, in the current study, it was hypothesized that participative climates characterized by cooperative or autonomous goal interdependence and construct controversy would also predict desired levels of employee participation and greater perceived influence in decision making.

Paths 6 and 7: Paths 6 and 7 illustrate the hypothesis that Participative Climate predicts Instrumental and Symbolic Use. The antecedent variable, Participative Climate, is hypothesized to have both direct and indirect effects on Instrumental and Symbolic Use. The total effect of Participative Climate on Instrumental and Symbolic Use is only in part due to the mediation of Level of PDM, Influence in PDM, and Process Outcomes. The hypothesis of a partial rather than full mediation between Participative Climate and

Instrumental and Symbolic Use is based on organizational research on controversy (Latham & Saari, 1979; Tjosvold, 1985, 1987) and goal theory (Tjosvold, 1986; Tjosvold et al., 1983), which repeatedly indicated the importance of workplace climate on participation and employee performance. Therefore, the importance of Participative Climate was emphasized by hypothesizing direct as well as indirect effects between Participative Climate and Instrumental and Symbolic Use.

Paths 8 and 9: Paths 8 and 9 illustrate the hypothesis that Level of PDM and Influence in PDM predict Process Outcomes. This hypothesis was based on the findings of Brandon et al. (1994) who posited that teacher participation in pre-evaluation decisions and evaluation content decisions predicted teacher agreement with decision outcomes. The Level of PDM and Influence in PDM variables used in the current study are much the same as the variables used by Brandon et al. except that in the current study the focus was on teacher perceptions of Process Outcomes whereas Brandon et al. focused on teacher perceptions of decision outcomes.

As discussed in Chapter Two, Brandon et al. (1994) reported the strongest direct effect between participation in content decisions and agreement with decision outcomes (.73). They reported a relatively small but significant direct effect between participation in pre-evaluation decisions and agreement with decision outcomes (.17). It was hypothesized that a similar relationship would occur in the current study in that the direct effect (path 8) between Influence in PDM and Process Outcomes would be greater than the direct effect (path 9) between Level of PDM and Process Outcomes.

Path 10: Path 10 illustrates the hypothesis that Level of PDM predicts Influence in PDM. This hypothesis was based on prior research which suggested teachers differentiate

between participation and influence in decision making and that the nature of participation affects teachers' perceived influence in decision making (i.e., Duke et al., 1980; Imber et al., 1990; Taylor & Bogotch, 1994). Furthermore, Brandon et al. (1994) found that teacher involvement in pre-evaluation decisions predicted teacher participation in evaluation content decisions. Therefore, it was assumed that the relationship between participation in pre-evaluation and content decisions would also be present in the current study.

Hypothesis Two (Conceptual Theory): Process Outcomes predict Instrumental and Symbolic Use.

Paths 11 and 12: Paths 11 and 12 illustrate the hypothesis that the latent variable Process Outcomes predicts Instrumental and Symbolic Use. Because the Process Outcomes construct used in the current study was unique (I was unable to find the same construct used in prior evaluation research) the hypotheses for paths 11 and 12 were based on research with variables related to but not exactly the same as the Process Outcomes variable. As discussed in Chapter Two, Bandura's (1986) theory of perceived self-efficacy has been related to performance. For example, Latham et al. (1994) found that participation in decision making predicted increased self-efficacy, which in turn predicted increased performance. Furthermore, Stein and Wang (1988) found that teacher perceptions of program efficacy were related to perceptions of program value, which in turn were related to perceptions of success in program implementation. These findings suggested that teacher perceptions of whether a program was capable of achieving its intended outcomes affected teacher perceptions of how well they were able to implement a program. This finding indicated that teachers base their perception of program efficacy on their perceptions of success in implementing the program. In the current study, it was hypothesized that teacher

perceptions of participative Process Outcomes affect perceptions of Instrumental and Symbolic Use. That is, if teachers perceive the participative process to be effective, they are more likely to believe that the evaluation information has been used in instrumental and symbolic ways. In both the current study and the study by Stein and Wang, perceptions of process are hypothesized to predict perceptions of implementation. Stein and Wang predicted perceptions of program implementation, and in the current study, I predict perceptions of use of evaluation information.

CHAPTER FOUR:
METHOD

Intervening Mechanism Evaluation Design

The design for the current study involved two components: (a) an intervening mechanism evaluation design and (b) a research design. The intervening mechanism design explains the relationships modeled in the current study. The research design explains the methodological and statistical strategies used. Figure 4 depicts Chen's (1990) conception of an intervening mechanism design.

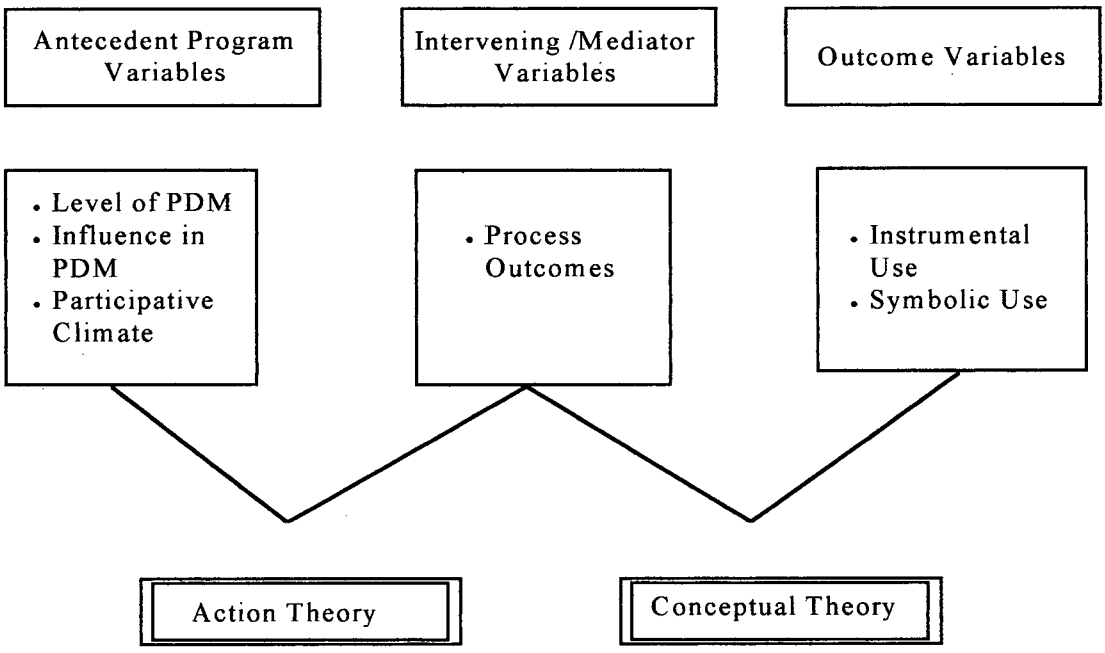


Figure 4. Intervening mechanism evaluation design (Chen, 1990).

According to Chen (1990), the purpose of the intervening mechanism design is to explain the causal mechanisms between the antecedent and the outcome variables in an

evaluation context. Chen described intervening mechanism designs as having two theories to test: (a) a test of "conceptual theory," which determines whether mediator variable(s) affect outcome variables and (b) a test of "action theory," which determines whether antecedent program variables, manipulated through program processes, influence the mediator variable(s) of the conceptual theory (p. 199). Not only must the antecedent program variables succeed in activating the intervening variable, as specified in the action theory, but the intervening variables must also affect the outcome variables, as specified in the conceptual model.

Figure 5 shows Chen's (1990) intervening mechanism design applied to this study.

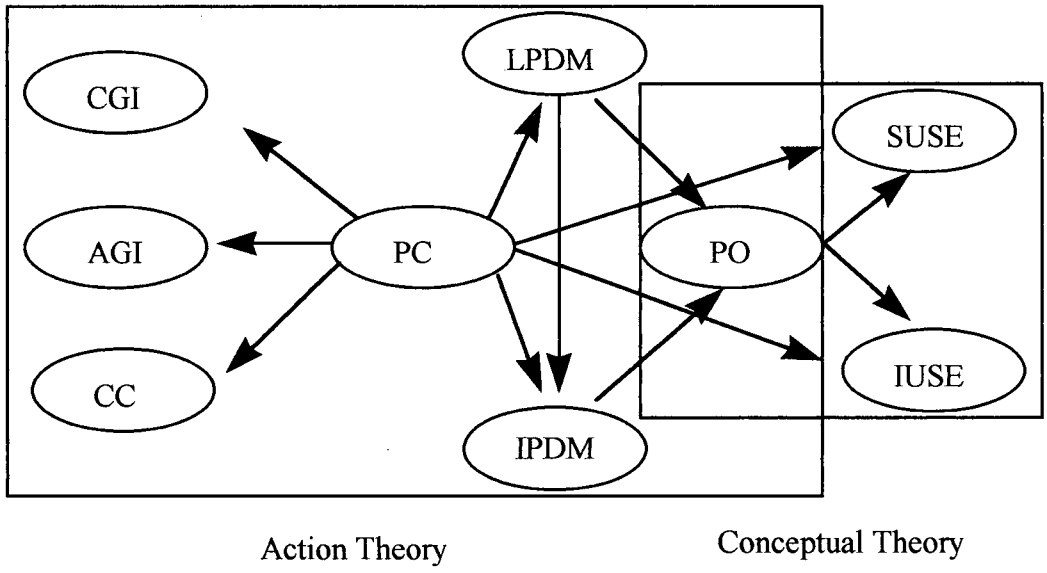


Figure 5. The action and conceptual theories of the hypothesized causal model.

CGI = Cooperative Goal Interdependence, AGI = Autonomous Goal Interdependence, CC = Constructive Controversy, PC = Participative Climate, LPDM = Level of Participation in Decision Making, IPDM = Influence in Participative Decision Making, IUSE = Instrumental Use, PO = Process Outcomes, SUSE = Symbolic Use.

In the current study, Participative Climate, Level of PDM, and Influence in PDM were the antecedent program variables; Instrumental and Symbolic Use were the outcome

variables; and Process Outcomes was the mediator variable. The *conceptual theory* in the current study was that Process Outcomes predicted Instrumental and Symbolic Use. The *action theory* was that Participative Climate, Level of PDM, and Influence in PDM predicted Process Outcomes. The action theory was tested through hypothesis one and the conceptual theory was tested through hypothesis two (see Figure 3).

The intervening mechanism design requires success in both the action theory and the conceptual theory. That is, if the antecedent program variables do not affect the intervening variable, the intervening mechanism design is considered ineffective. If, in this case, the conceptual theory was considered sound, then the problem would be the manner in which the program was implemented, and the antecedent program variables would have to be reconsidered in terms of their relationship to the intervening variable. This scenario, applied to the context of B.C. School Accreditation, would mean the accreditation processes were not creating desired levels of teacher participation and influence (antecedent program variables), which in turn would affect Process Outcomes (mediating variable) and use of evaluation information (outcome variables). If, in the current study, the intervening variable (Process Outcomes) did not activate the outcome variables (Symbolic and Instrumental Use) thereby making the conceptual theory flawed, then the study would have to be redirected with an alternate conceptual theory.

Research Design

Context of the Study

The British Columbia (B.C.) School Accreditation Program was the context for the current study. It was chosen because it was a good example of participative program evaluation and it met the criteria necessary to test the hypotheses of this study. The accreditation program was such that it enabled the measurement of all the constructs. Furthermore, the literature review indicated cognitive benefits of PDM were most likely to occur when (a) tasks were complex versus simple and (b) when participants were involved in developing task strategies to implement decisions. The B.C. School Accreditation provided for these conditions by requiring teachers to organize and carry out the entire accreditation process. This involved complex tasks of comprehending, interpreting, evaluating, and synthesizing evaluation information. It also required teachers to participate in decision making as well as develop ways to implement the decisions.

The accreditation model of school evaluation was first introduced in British Columbia Schools in the 1920s (Ministry of Education, 1996). Since then, it has been revised several times and is continually updated to meet needs of teachers, administrators, and government standards. The purpose of the accreditation program was defined as follows in the May, 1995 British Columbia School Accreditation Resource Guide:

> Accreditation is a process for school improvement and accountability with emphasis on school improvement. Such a process fosters continual positive growth, assists with setting direction, staff development and vision formulation and confirmation. Accreditation incorporates accountability whereby student outcomes (knowledge, skills and attitudes), parent/teacher satisfaction, and community satisfaction are assessed and reported. (p. 3)

The Ministry of Education finances the accreditation program which is now compulsory in the British Columbia school system. The accreditation process has six sequential components: (1) notification that school accreditation is required, (2) school staff undertake self-evaluation process, (3) external team visit and reports are drafted, (4) external teams report to school staff, superintendent, and school trustees, (5) assistant director of accreditation awards final status, and (6) follow-up on recommendations.

The staff self-assessment is of most concern to this study. Self-assessment is a process of collecting various forms of evidence to show how students and the school as a whole performed on a series of eight provincial education standards: (1) intellectual development, (2) human and social development, (3) career development, (4) attributes of quality, (5) leadership, (6) professional and staff development, (7) school culture, and (8) the school and its community. Based on the evidence gathered for these eight standards, the teachers rate their performance, and the results are examined and validated by an external assessment team. The specific steps of the self-assessment as outlined by the Ministry of Education Accreditation Manual (1995) are explained in Table 1.

The accreditation guidelines require that the entire school staff participate. The accreditation process begins in April when school administrators begin meeting with

Ministry accreditation personnel. Staff and administration training occurs in May and June and the self-assessment steps begin in September. The external teams usually visit during April and May and the entire process is generally completed by the end of June.

The self-assessment process is conducted primarily by the teaching staff and is headed by a self-assessment committee. The committee chair leads the accreditation process and is responsible for most of the organizing necessary to complete the self-assessment. Some of the responsibilities of the assessment chair include generating staff support, acting as manager of the assessment process, and enlisting staff participation in open, frank, and well managed sessions.

Table 1

Overview of the Self-Assessment Process of the B.C. School Accreditation

Step One	Clarifying of criteria statements: Purpose of this step is to ensure all participants understand the meaning of each criterion statement.
Step Two	Listing items of evidence in light of knowledge, skills, and attitudes: • evidence is gathered to show strengths, strengths which are building, and areas of needed improvement. • evidence can be qualitative, quantitative, and/or demonstrative.
Step Three	Determining levels of staff satisfaction: • staff reaches consensus on how well the criterion statement has been met. • it is a staff judgment on how well students and school are performing. • each criterion statement is judged using a 4-point Likert-type scale.
Step Four	Completing internal comments: Information that will help the external team understand the reality of the school situation.
Step Five	Completing summary for each section: Usually done by the self-assessment chair.
Step Six	Categorizing school priorities: Categorize information from summary statements into 4 categories: areas of strength, areas of strength but need support, areas needing change, areas of change that need further attention.
Step Seven	Building the preliminary school growth plan with the description of the "Educated Citizen" as filter, and develop action plans: Components of an action plan are: strategies, staff development, needed resources, person responsible, target dates, evaluation, cost analysis.
Step Eight	Building direct involvement of all shareholders in the growth plan: Involve students, parents, the community, and the school district.
Step Nine	Reviewing and building mission statement in light of the school growth plan.

The means of appointing a self-assessment chair varies from school to school. However, it is commonly on a volunteer basis with the final decision made by an administrator. The remainder of the staff is organized into sub-committees headed by a sub-committee chair. The sub-committees have varying responsibilities throughout the accreditation processes. The means of appointing sub-committee chairs also varies from school to school. Decision making is based on consensus and decisions are made either in small or whole staff groupings. Decisions made during the accreditation steps are supposed to reflect the opinions of the majority of the staff.

Sample

Of the 315 elementary and secondary school teachers who participated in this study, approximately 173 were elementary teachers and approximately 110 were secondary teachers. Due to missing categorical data on the school and the grade(s) taught, it was not possible to determine whether some of the teachers were from secondary or elementary schools. This, however, had no effect on the analysis. Categorical data were collected for descriptive purposes only. The 315 questionnaires represented teachers from 47% of the schools (91 of 194) in the 1995/1996 B.C. School Accreditation Program.

Procedures

School Districts in the province of British Columbia have different regulations and procedures to guide research within their schools. Because of this, the methods used to solicit participation varied slightly from district to district and from school to school. There were two means of asking teachers to participate in the study. The first was by direct contact with 351 teachers (from 27 schools) through a cover letter and attached package

with six questionnaires and a postage-paid return envelope. This first method was used in school districts that supplied names of individual teachers. The second method was through an introductory letter to 167 principals followed by a cover letter and attached package of six questionnaires. This second method was used in school districts that would not supply names of individual teachers because they preferred that teacher participation be solicited through the administration. All correspondence to principals is included in Appendix A.

The introduction letter, a postage-paid return envelope as well as a phone, fax, or e-mail reminder to return the questionnaires were strategies used to increase the return rate. It must also be noted that initial contact with principals either by phone or letter was only possible once school district approval had been acquired. Means of acquiring district approval varied, but it was generally required that I complete a request-to-conduct research form and submit either a summary or the entire research proposal. Once approval was granted the district explained the processes for collecting the data. In some cases, the district contacted the principals with my request to conduct research, waited for a response, and then contacted me as to the principal's decision. Other districts allowed me to send the questionnaire packages immediately after approval had been granted. Upon receipt of the questionnaire package, principals were asked to distribute the questionnaires to the staff. Upon completion, staff members who chose to participate returned their questionnaires to the office. Questionnaires were returned in a postage-paid envelope that was included in the package.

The data collection procedure was in part constrained by the B.C. School Accreditation timeline. Data collection did not begin until the schools had completed or nearly completed the entire accreditation process, which was typically between March and

May. In addition to the accreditation timeline, the data collection procedures were in part guided by the constraints imposed by the necessary avoidance of the percept-percept problem. The percept-percept problem refers to the collection of both participation and outcome data in a questionnaire administered to a particular group in a single sitting (Campbell, 1982). Wagner and Gooding (1987a) suggested this form of data collection could inflate correlations and can be avoided by using at least one objective or assigned condition, different respondents for data on participation and outcome variables, or leaving a time lag between the collection of the participation and outcome data.

Participatory literature, which has addressed the percept-percept problem, has not specified a length of time required for the lag between the collection of the participation and outcome data (i.e., Campbell, 1982; Wagner & Gooding, 1987a, 1987b). In the current study, the percept-percept problem was addressed by using two time lags between completion of the questionnaires: (a) one-day and (b) one-week. However, due to the lack of research on the optimal time lag, the choice of one-day and one-week time lags was somewhat arbitrary. It was rationalized that the one-day time lag was long enough to avoid the percept-percept problem and the one-week time lag was not too long in that teachers would forget about or lose the questionnaires.

The questionnaires sent directly to the 351 teachers were based on the one-day time lag. This meant participants had to complete four questionnaires, then wait a day; complete another questionnaire, wait a day and then complete the final questionnaire. The questionnaires pertaining to Level of PDM, Influence in PDM, and Participative Climate were completed first; the questionnaire pertaining to Process Outcomes was completed second; and the questionnaire on Instrumental and Symbolic Use was completed last. The

questionnaires sent to principals required participants to follow the same procedure except instead of waiting one-day, they were asked to wait one-week between completion of the three sets of questionnaires. Of the 315 useable questionnaires returned, there were 102 completed with a one-day time lag and 213 completed with a one-week time lag. The reason for the proportionately high return rate for the one-week time lag group is not known. It is possible, however, that teachers needed the extra time to complete the questionnaire and the one-week time lag enabled them to work it into their schedule more easily than did the one-day time lag.

Measures

There were six measures used in this study. Four were developed for this study and two were previously published. All six measures are included in Appendix B.

Developed Measures

1. Level of Participation in Decision Making
2. Influence in Participative Decision Making
3. Process Outcomes
4. Use of Evaluation Information

Published Measures

5. Goal Interdependence (Tjosvold et al., 1983)
6. Constructive Controversy (Tjosvold et al., 1986)

Analysis of the developed questionnaires included several common features. Internal consistency reliability was determined by Cronbach's (1951) coefficient alpha (r_α). Construct validity was established through exploratory factor analysis. The

component structures were examined through principal components analysis with orthogonal (varimax) and oblique rotations and a loading of .40 or greater for an item was viewed as meaningful (cf. Gable & Wolf, 1993). Based on a recommendation presented by Gable and Wolf (1993), oblique rather than orthogonal (varimax) solutions are reported. Gable and Wolf posited that in most cases of factor analysis in educational psychology research, it is likely that item clusters used to define factors will not be orthogonal (varimax). Therefore, they advised that oblique rotation be used because it allows the axis system to be less than 90 degrees. The following sections explain the analysis for the developed and existing measures.

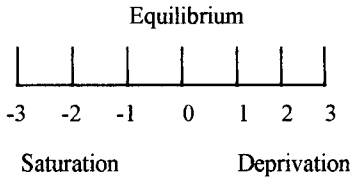
Developed Questionnaires

Level of Participation in Decision Making (Level of PDM) Questionnaire

I designed the Level of PDM questionnaire to measure teachers' perceptions of level of participation in decision making in a program evaluation context. It consisted of two scales: (1) contribution to pre-evaluation decision making and (2) contribution to evaluation content decisions. There were 12 questions, six for each scale. Participants were required to indicate the degree to which they actually contributed and the degree to which they would have preferred to contribute. A four point Likert scale was used for *actual* and *preferred* levels of contribution. The *actual* score was subtracted from the *preferred* score and the resulting score was called either a deprivation, saturation, or equilibrium score.

The difference score method used for the Level of PDM was based on previously published measures of participative decision making (cf. Alutto & Belasco, 1972; Bacharach, Bamberger, Conley, & Bauer, 1990; Conway, 1976). For example, Alutto and

Belasco (1972) used a difference score strategy as a means of distinguishing between actual and desired levels of participation. They defined five levels of participation: actual, preferred, deprivation, saturation, and equilibrium. The deprivation, saturation, and equilibrium scores are generally used as indicators for Level of PDM and are acquired by subtracting the actual participation level from the desired participation level. For example, if actual participation was described as 3 on a Likert item and desired participation was described as 2, then the deprivation score would be -1. If on the other hand, the actual score was 3 and the desired score was 2, then the saturation score would be 1. And, if the actual and desired scores were the same, then the equilibrium score would be 0. The range of possible scores is from positive three through to negative three. Deprivation scores range from one to three, saturation scores range from negative three to negative one, and the equilibrium score is zero:



This method of acquiring deprivation scores is based on cognitive dissonance theory (Festinger, 1957) and has implications for the cognitive outcomes of participation in decision making. Measuring participation in decision making in terms of discrepancies between actual and desired participation is analogous to actual and desired participation being viewed as two conflicting cognitions (Miller, Ismail, Giesen, Adams-Price, & Topping, 1993). For example, the discrepancy state between actual and desired participation causes dissonance. The reduction of dissonance between an act and an attitude (the act being the actual state of participation, and the attitude being the desired state of

participation) could manifest itself in a number of ways. For example, tension could be lessened by derogating the source. That is, the participant could derogate the act, which in the case of participation in decision making would be the actual level of participation. Derogating the actual level of participation would be the same as invalidating the participation processes. If participants become frustrated by the discrepancy between how they would like to participate and how they are participating, they could discredit the participative process as being ineffective. This would lessen the dissonance, but it would also impact their effectiveness as participative decision makers. Therefore, in the current study, it was assumed that participants who were satisfied with their level of involvement were more likely to be effective decision makers than those who felt their participation was saturated or deprived.

Reliability and Content Validity

Reliability of the difference scores between actual and preferred contribution for the full scale (all 12 items) was $r_{\alpha} = .92$ (n = 315). Reliability was also calculated using a difference score formula suggested by Crocker and Algina (1986), which resulted in a coefficient of $r_{\alpha} = .92$ (n = 315). Reliability of differences scores for scale one (items 1 - 6) was $r_{\alpha} = .89$ and $r_{\alpha} = .87$ for scale two (items 7 - 12). Mean inter-item correlations computed for the total measure and each of the scales are presented in Table 2. Item-total scale correlation coefficients were high; all 12 items had coefficients of .54 or higher.

Table 2

Internal Consistency Reliability Estimates for Level of Participation in Decision Making Questionnaire

	n	r_α	r_{ii}	*Range* of r_{it}
Scale 1				
Pre-evaluation decisions	315	.89	.59	.61 - .79
Scale 2				
Evaluation content decisions	315	.87	.53	.54 - .80

Note. r_α = coefficient alpha reliability; r_{ii} = mean inter-item correlation; r_{it} = item-total scale correlations.

Items for the initial draft of the Level of PDM questionnaire were derived from the B.C. Accreditation Manual and represent each of the major steps in the accreditation process. Content validity was established by reviewing the initial draft for ambiguity and wording by persons who had not previously experienced the accreditation program. It was then reviewed by several teachers (elementary and secondary) who had recently experienced the accreditation program. Revisions were made before the questionnaire was pilot-tested on 28 teachers in one elementary and two secondary schools in the greater Vancouver area.

Construct Validity

The results of the oblique principal components solution are presented in Table 3.

Table 3

Rotated (oblique) Component Loadings of Level of Participation in Decision Making Items

No.	Item Description	Components 1	2
1.	deciding who would be self-assessment chair	**.631**	.109
2.	deciding who would be the sub-committee leaders	**.667**	.141
3.	deciding how final decisions would be made	**.794**	.061
4.	deciding when meetings would be held	**.849**	-.006
5.	deciding how meetings would be conducted	**.817**	.073
6.	deciding how long meetings would be	**.910**	-.174
7.	clarifying criteria statements	-.000	**.778**
8.	gathering evidence	-.007	**.737**
9.	determining satisfaction levels	.109	**.804**
10.	adding internal comments	.248	**.565**
11.	categorizing school priorities	.125	**.714**
12.	building school growth plan	-.070	**.831**
Eigenvalue		6.31	1.45
Variance Explained		51.10%	12.10%

Two components with eigenvalues (λ) of 6.13 and 1.45 accounted for 63.20% of the total variance. A Scree test (Cattell, 1966) also showed a strong two component solution. Items one to six loaded on the first component and items 7 to 12 loaded on the second component. The first rotated component was characterized as teacher perceptions of decision making in pre-evaluation decisions. All items on this component demonstrated strong loadings with minimal shared variance. The second component was characterized as teacher perceptions of decision making in evaluation content decisions. This component, as well, was clearly defined with minimal shared variance.

The component score correlation was $\underline{r} = .55$, which suggested a relatively high degree of interrelationship and made collapsing the scales an option (cf. Gable & Wolf,

1993). However, the reliability for each scale was high ($\underline{r}$ =. 89 and $\underline{r}$ = .87) and the scales were meaningful so they were not collapsed.

Influence in Participative Decision Making (Influence in PDM) Questionnaire

I designed the Influence in PDM questionnaire to measure teachers' perceptions of influence in participative decision making in a program evaluation context. There were 12 items each with a 6 point Likert scale. Like the Level of PDM questionnaire, the items for the Influence in PDM questionnaire were derived from the B.C. School Accreditation Manual and represent each of the major steps in the accreditation process. The questionnaire format was similar to the Level of PDM measure in that items addressed participation in pre-evaluation and evaluation content decision making. The theoretical basis for the Influence in PDM questionnaire was the same literature used for the Level of PDM (i.e., Alutto and Belasco, 1972; Bacharach et al., 1990; Conway, 1976). It was decided, however, that the difference score strategy used with the Level of PDM questionnaire was not appropriate because it would be difficult for respondents to know how much influence they *actually* had versus what they *preferred* to have had. Therefore, a six point Likert scale was used instead.

Reliability and Content Validity

The reliability for the 12 items on the Influence in Participative Decision Making questionnaire was r_{α} =.92 (n = 315). Reliability for scale one (items 1 - 6) was r_{α} =.92 and r_{α} =.91 for scale two (items 7 - 12). Reliability and mean inter-item correlations were computed for the total measure and each of the scales and are presented in Table 4. Item-total scale correlation coefficients were high with all items demonstrating coefficients of .69

or higher. The process used to establish content validity for the Level of PDM questionnaire was also used for the Influence in Participative Decision Making questionnaire.

Table 4

Internal Consistency Reliability Estimates for Influence in Participatory Decision Making Questionnaire

	n	r_α	r_{ii}	*Range* of r_{it}
Scale 1				
Pre-evaluation decisions	315	.92	.67	.70 - .86
Scale 2				
Evaluation content decisions	315	.91	.63	.69 - .79

Note. r_α = coefficient alpha reliability; r_{ii} = mean inter-item correlation; r_{it} = item-total scale correlations.

Construct Validity

The results of the oblique solution are presented in Table 5. Two components with eigenvalues (λ) of 6.74 and 1.80 accounted for 71.20% of the total variance. Items one to six loaded on the first component and items seven to 12 loaded on the second component. A Scree test (Cattell, 1966) showed a strong two component solution. The first rotated component was characterized as teacher perceptions of influence in pre-evaluation decision making. All items on this component demonstrated strong loadings with minimal shared variance with the other component. The second component was characterized as teacher perceptions of influence in evaluation decision making. This component, as well, was clearly defined with minimal shared variance. The component score correlation was $\underline{r} = .56$, which indicated a relatively high degree of interrelationship. However, the scales were not collapsed because each was independently meaningful and the reliability for each was high ($\underline{r} = .92$, $\underline{r} = .91$) (cf. Gable & Wolf, 1993).

Table 5

Rotated (oblique) Component Loadings of Influence in Participative Decision Making Items

No.	Item Description	Components 1	2
I had influence on:			
1.	deciding who would be chosen as the self-assessment chair	**.837**	-.071
2.	deciding who the subcommittee leaders would be	**.762**	.004
3.	deciding how final decisions would be made	**.785**	.074
4.	deciding when meetings would be held	**.884**	.028
5.	deciding how meetings would be conducted	**.921**	-.006
6.	deciding how long meetings would be	**.878**	.034
7.	decisions about clarifying criteria statements	.151	**.703**
8.	decisions about gathering evidence	.090	**.717**
9.	decisions regarding satisfaction levels	-.057	**.892**
10.	decisions regarding the internal comments	.070	**.814**
11.	decisions about categorizing school priorities	-.064	**.903**
12.	decisions about building the school growth plan	-.067	**.879**
Eigenvalue		6.74	1.80
Variance Explained		56.20%	15.00%

Process Outcomes Questionnaire

I designed the Process Outcomes questionnaire to measure Process Outcomes in an evaluation context. The questionnaire had 16 items each with a six point Likert scale. Items five, nine, and 14 were reverse scored. Two items (number 1 and 3) on the questionnaire were adapted from a scale developed by Greene (1988a). The items on the Process Outcomes questionnaire were intended to assess teacher perceptions of the outcomes of the participatory processes used in the accreditation program. For example, it was assumed that the participative nature of the accreditation process would require teachers to examine what was important for their school and in the process acquire new, useful, and necessary information which they would be able to apply to the development of

the school growth plan. Furthermore, it was assumed that the participative nature of the accreditation process would result in a better understanding of the points of view and needs of other staff as well as increased school advocacy.

Reliability and Content Validity

The reliability for items one to 16 on the Process Outcomes questionnaire was r α = .96, the mean inter-item correlation was .61, and the range of item-total scale correlations was .64 to .85. The procedure used to establish the content validity for the Level of PDM and Influence in PDM questionnaires was also used for the Process Outcomes questionnaire. The results of the oblique solution are presented in Table 6. One component, with an eigenvalue (λ) of 10.16, accounted for 63.50% of the total variance. The items on the Process Outcomes questionnaire are structurally similar to efficacy measures such as the Teacher Perception of Program Efficacy Scale (TPPO) (Stein & Wang, 1988). Examples of TPPO items are (a) Generally, students tend to spend more time on task under this program and (b) Students feel better about themselves because of this program. The TPPO items were intended to address teacher perceptions of expected program outcomes. Likewise, items on the Process Outcomes questionnaire were intended to address teacher perceptions of expected participation outcomes. Two examples of Process Outcomes items are (a) Through the accreditation process I've learned about problems in my school that I wasn't aware of and (b) The accreditation activities provided me with the necessary knowledge to contribute to the development of our school growth plan.

Table 6

Rotated (oblique) Component Loadings of Process Outcomes Items

No.	Item Description	Component
1.	Through the accreditation process, I've learned about problems in my school that I wasn't aware of.	**.717**
2.	I learned new things about my school because of the accreditation program.	**.814**
3.	The accreditation process has made me think about what's important for my school.	**.834**
4.	Because of the accreditation program, I learned useful things about my school.	**.889**
5.	The accreditation program did not teach me anything new about my school.	**.797**
6.	The accreditation gave me a better understanding of the strengths of our school.	**.792**
7.	The accreditation activities provided me with the necessary knowledge to contribute to the development of our school growth plan.	**.853**
8.	The accreditation provided me with a better overall understanding of our school practices.	**.856**
9.	The steps in the accreditation were not useful in giving me ideas for the development of our school growth plan.	**.743**
10.	The accreditation increased my understanding of my school's role in the community.	**.806**
11.	The accreditation gave me a better understanding of the weaknesses of our school.	**.801**
12.	I did not learn anything useful about my school as a result of going through the accreditation program.	**.870**
13.	The accreditation gave me a better understanding of other staff members' needs.	**.759**
14.	The accreditation provided no new details about our school.	**.668**
15.	The accreditation gave me a better understanding of the points of view of other staff members.	**.752**
16.	The accreditation activities gave me ideas for developing the various components of the school growth plan.	**.756**
	Eigenvalue	10.16
	Variance Explained	63.50%

The items on the Process Outcomes questionnaire are also similar in structure to items developed by Greene (1988a) for a measure of "Stakeholder Perceptions of the Usefulness of the Evaluation Process and Results." As discussed in Chapter One,

items from Greene's questionnaire classified as "Uses of the Process" were adapted and used in the current study. However, the items as used in the current study were intended to represent participatory process outcomes, not uses of participatory process as intended by Greene. It was rationalized that the Process Outcomes items were essentially the same in terms of structure and content as the TPPO items developed by Stein and Wang (1988). Therefore, they were more closely related to the notion of process efficacy than process use. In either case, the content validity of the Process Outcomes items was substantiated by both the TPPO items (Stein & Wang, 1988) and the "Uses of Process" items (Greene, 1988a).

Use of Evaluation Information Questionnaire

I developed the 14 item Use of Evaluation Information questionnaire to measure perceived occurrence of Symbolic and Instrumental Use of evaluation information. The Symbolic Use scale had six items and the Instrumental Use scale had eight items. A six point Likert scale was used. Six items (5, 7, 8, 10, 13, and 14) were adapted from a previous scale by Rinnie (1993). Also, item three was adapted from Greene (1988a).

Reliability

The reliability and mean inter-item correlations computed for the total measure and each of the scales are presented in Table 7. The reliability for the 14 items was r α = .80 (n = 315). The reliability for scale one (items 1 - 6) was r α = .88 and r α = .79 for scale two (items 7 - 12). Item-total scale correlation coefficients were moderate with all of the 14 items having coefficients of .21 or higher.

Table 7

Internal Consistency Reliability Estimates for Use of Evaluation Information Questionnaire

	n	r_{α}	r_{ii}	*Range* of r_{it}
Scale 1				
Symbolic Use	315	.88	.54	.52 - .73
Scale 2				
Instrumental Use	315	.79	.33	.21 - .77

Note. r_{α} = coefficient alpha reliability; r_{ii} = mean inter-item correlation; r_{it} = item-total scale correlations.

Content and Construct Validity

The theoretical basis for the content validity of the Use of Evaluation Information questionnaire was derived from a review of the literature on previous measures of evaluation use (Newman & Cai, 1995; Ramirez, 1985; Rinnie, 1993). The questionnaire was pilot-tested in the same manner as the other questionnaires developed for this study. An initial draft was developed and items were reviewed by content area specialists and edited before being pilot-tested.

The results of the oblique solution did not correspond to the originally predicted two component structure with items one to six on scale one and items seven to 14 on scale two. Instead, three components with eigenvalues (λ) of 5.98, 1.88, and 1.05 accounted for 63.70% of the total variance (see Table 8). The first rotated component was characterized as teacher perceptions of Symbolic Use. All items on this component, with the exception of item nine, demonstrated strong loadings with minimal shared variance with the other components. The second and third components were both considered to reflect teacher perceptions of Instrumental Use. With the exception of item 12, these components were

also defined with minimal overlap and minimal shared variance. Components one and three had a moderate to high correlation (r = -.45). Therefore, they were collapsed into a single component (cf. Gable & Wolf, 1993), which corresponded to the originally predicted two-component structure.

Table 8

Rotated (oblique) Component Loadings of Use of Evaluation Information Items

No.	Item Description	Components 1	2	3
The accreditation information:				
1.	raised discussion about overall school achievement.	**.427**	.060	-.322
2.	is useful for persuading parents to support what we are doing in our school.	**.799**	.043	.007
3.	has made me a better advocate of my school.	**.769**	.283	-.014
4.	resulted in a feeling of satisfaction with overall school achievement.	**.797**	-.097	-.075
5.	convinced staff members of need for maintaining existing school practices.	**.845**	-.103	.060
6.	is useful in gaining community support for what we are doing in our school.	**.830**	.065	.033
9.	confirmed the effectiveness of the existing school practices.	**.517**	-.471	-.250
8.	resulted in some existing school practices being terminated.	-.081	**.731**	-.035
12.	was used as a starting point for further evaluation of specific school practices.	.205	**.464**	-.431
14.	resulted in new school policy being implemented.	.301	**.678**	-.111
7.	resulted in no action to make changes in our school practices.	.091	-.138	**-.737**
10.	resulted in modifications of existing school practices.	-.238	.091	**-.785**
11.	resulted in suggestions for how to improve the school.	.142	-.001	**-.812**
13.	resulted in implementation of new ideas on how to improve school practices.	.292	.198	**-.634**
Eigenvalue		5.98	1.88	1.05
Variance Explained		42.70%	13.40%	7.60%

Existing Questionnaires

Goal Interdependence

The Goal Interdependence questionnaire (Tjosvold, et al., 1983) had 24 items and three scales all with acceptable internal consistency reliability: (a) cooperative goal orientation (8 items; r α = .91), (b) competitive orientation (9 items; r α = .90), and (c) individualistic goal orientation (8 items; r α = .76). The reliability coefficients were derived from 310 medical laboratory technicians from ten different Western Canadian hospitals (Tjosvold et al., 1983).

In personal correspondence with Dean Tjosvold (October 10, 1995), it was recommended that the Goal Interdependence questionnaire, as published in 1983, should be reduced to 23 items. Based on the analysis for the current study using 23 items, the internal consistency reliability was comparable to the 1983 study: (a) cooperative goal orientation (7 items; r α = .89), (b) competitive orientation (8 items; r α = .92), and (c) individualistic goal orientation (8 items; r α = .84). Items 2-4, 6-7,9-10,12-14,16,18-19, and 21-23 were reverse scored.

Constructive Controversy

The Constructive Controversy questionnaire (Tjosvold et al., 1986) had 15 items and was reported to have internal consistency reliability of r α = .88. The questionnaire was administered to 58 managers enrolled in an MBA program at Simon Fraser University in British Columbia, Canada as well as work teams from business organizations in Hong Kong

and China with similar results (personal correspondence with Dean Tjosvold, October 10, 1995). The internal consistency reliability for the sample in the current study was also $r_{\alpha} = .88$.

Issues in Model Testing

The following methodological issues were considered prior to using structural equation modeling (SEM): (a) sample size, (b) model identification, (d) model estimation, and (e) fit indices. The following section describes how these issues were addressed within the current study.

Sample Size

The issue of sample size has received considerable attention in SEM literature (e.g., Bollen, 1990; Cudeck & Henly, 1991; Tanaka, 1987). This is due primarily to the effects of sample size on goodness-of-fit indices. Bollen (1990, p. 256) argued that much of the contradictory information on the effects of sample size (i.e., Bentler & Bonett, 1980; Joreskog & Sorbom, 1988; Marsh, Balla, & McDonald, 1988) can be reconciled by considering two types of sample size influences: (a) whether *N* is used directly in the calculation of the fit index or (b) "... whether the means of the sampling distributions of the fit indices are associated with *N*." Bollen posited that *N* is positively associated with the means of the sampling distribution for the Bentler and Bonett (1980) index, Bollen (1986) index, Goodness-of-Fit Index (GFI), and Adjusted Goodness-of-Fit Index (AGFI) (Joreskog & Sorbom, 1988), however, their calculation is not influenced by *N*. On the other hand, the Tucker and Lewis (1973) and Bollen (1989) indices, which have no normed maximum

value, are the opposite in that the calculation is affected by *N* but not by the means of the sampling distributions. Therefore, the advantage of the Bentler and Bonett (1980), Bollen (1986), GFI, and AGFI indices is that they all have a normed maximum or minimum which makes them easy to interpret. The advantage of the Bollen (1989) and Tucker and Lewis (1973) indices is that they have means that are stable across *N*'s, however, they are not normed so they are more difficult to interpret. Based on the differences between fit indices, Bollen (1990) recommended that more than one fit index be reported. Therefore, as will be discussed in the following sections, four fit indices were used in the current study.

The critical *N* (CN) statistic (Hoelter, 1983) can be used to determine acceptable sample size. The CN statistic provides sample size at which the *F* value would reject the null hypothesis $H_o: S = \Sigma$ for a given alpha level. The CN statistic is available in LISREL 8 (Joreskog & Sorbom, 1993) or can be calculated by $CN = (\chi^2/F) + 1$. This is, however, a post hoc way of approaching the sample size question. In a review of sample size literature, Schumacker and Lomax (1996) reported that larger samples are better and that most published studies had from 250 to 500 cases. However, they also cited several studies which posited that smaller sample sizes ranging from 100 to 400 were satisfactory for SEM analysis (i.e., Boomsma, 1982, 1983; Ding, Velicer, & Harlow, 1995). Furthermore, Bentler and Chou (1987) argued that five subjects per observed variable would be sufficient if the sampling distribution was normal and if there were multiple indicators per latent variable. Moreover, Hayduk (1987) posited 100 to 500 subjects should be adequate, but fewer than 100 often led to problems of nonconvergence and negative error variances. Based on this information, the sample size for the current study was considered satisfactory with 315

cases, a cases to latent-variable ratio of 35:1 (9 latent variables), and a cases to observed-variables ratio of 14:1 (23 observed variables).

Model Identification

The underlying premise of model identification is to determine whether constraints placed on the model are sufficient to determine unique estimates of the structural coefficients. If it can be proven that the unknown parameters are functions only of the identified parameters and these functions lead to unique solutions, then the unknown parameters are identified. If it can not be shown that the unknown parameters have unique solutions, then one or more parameters are under identified and the entire model is considered to be under identified. A model is over identified if one or more parameters are over identified (Bollen, 1989).

Traditional methods of calculating identification involve solving for the parameters of the model algebraically in terms of the variance and covariance of observed variables. This is done by first showing that the parameters of the measurement model are identified, including the covariance and variances of the factors. Once the covariances of the factors are identified, the structural parameters can be identified by determining the covariances among the factors, rather than trying to solve them directly in terms of the covariances among the observed variables (Bollen, 1989).

Unfortunately, with complex models the algebraic calculation can be very difficult and time consuming. Fortunately, rules exist that enable identification using less complex methods. The Two-Step Rule was used in the present study (Bollen, 1989). The Two-Step rule incorporates the *t*-Rule, a necessary but not sufficient condition of identification. If the

condition is not met, it means only that the model is not unidentified. The first step in the Two-Step Rule requires that measurement parameters be identified. This can be done by rearranging the structural model into a confirmatory factor analysis model so the original x and y variables become x variables and the ξ and η factors become ξ variables. The *t*-Rule can be used to determine if this new model is identified. If this condition is met, the second step requires the latent variable model parameters be identified. This can be done by rearranging the latent variable equation so that it is a structural equation made up of observed variables. This means that each latent variable is assumed to be perfectly measured. This new model can also be identified using the *t*-rule. If the second step is met, then the model is identified. The Two-Step Rule was applied to the proposed covariance structure model and both steps were identified.

Model Estimation

The fundamental aim of SEM is to reduce the difference between the covariance matrix of observed variables S and the estimated/implied population covariance matrix Σ. Model estimation is the process of obtaining an estimated or implied covariance matrix Σ that is as close as possible to the observed matrix S. If there is little difference between the observed S and implied Σ matrices, the residual matrix will be near zero and the goodness-of-fit indices will indicate a good fit between the hypothesized model and the data. After the hypothesized model has been specified (all parameters have been set as either fixed, free, equal, or constrained), estimates of the free parameters are derived from the matrix of observed variables. That is, the estimates for the implied covariance matrix Σ are obtained from the observed matrix S. Generalized least (GLS) squares, unweighted least squares

(ULS), and maximum likelihood (ML) are commonly used estimation methods. These estimation methods are iterative fitting functions designed to minimize differences between S and Σ matrices. Each estimation method produces parameter estimates, standard error of estimates for the free parameters, and a chi-square model test (Hoyle, 1995). All estimation methods involve a set of starting values, which are used to begin estimation of the free parameters of the Σ matrix. After each iteration, the Σ matrix is compared to the S matrix and a residual matrix is obtained. Iterations continue until the residual matrix can not be further minimized. At this point, the model is said to have converged and a fitting function value is given, which is then used as a basis to test the fit of the model. Of the commonly used estimation methods, ML has been found to be the most consistent, unbiased, scale-invariant, scale-free, and normally distributed if the observed variables are multivariate normal and independent (Schumacker & Lomax, 1990). Therefore, because the observed variables in the current study were multivariate normal (as described in Chapter Four) and independent, ML estimation was used.

Goodness-of-Fit Indices

Model goodness-of-fit is determined simultaneously with estimation. The maximum likelihood (ML) estimation procedure previously described, produces a fitting function value that indicates the similarity between the observed S and implied Σ covariance matrices. The goodness-of-fit index tests the statistical significance of the fitting function. There are many published goodness-of-fit indices for structural equation models. As well, there are many opinions as to the best measures of model fit, comparison, and parsimony (cf. Bentler & Bonett, 1980; Bollen, 1990; Cudeck & Henly, 1991; Marsh et al., 1988;

McDonald & Marsh, 1990; Mulaik, James, Van Alstine, Bennett, Lind, & Stilwell, 1989). It is common, however, for the Tucker-Lewis Index (TLI) to be recommended over other fit indices when using ML estimation. For example, Bollen (1990) recommended the non-normed TLI in that it had advantages over other normed fit indices. And, Hoyle (1995) and Schumacker and Lomax (1996) also recommended the TLI over other comparative fit indices when using ML estimation.

Based on the literature and various recommendations, the chi-square goodness-of-fit index (χ^2), Goodness-of-Fit Index (GFI), Adjusted Goodness-of-Fit Index (AFGI), Root Mean Square Residual (RMSR), and the Tucker-Lewis Index (TLI) were chosen for use in the current study. The fit indices and the associated critical values for the indices are listed in Table 9.

Table 9

Indices of Model Fit and Comparison

A. Indices of Model Fit	Critical Values
Chi-square goodness-of-fit index (χ^2)	• Chi-square table value ($p > \alpha$)
Goodness-of-fit index (GFI)	• 0 (no fit) to 1 (perfect fit)
Adjusted goodness-of-fit index (AGFI)	• 0 (no fit) to 1 (perfect fit)
Root mean-square residual (RMSR)	• less than .05
B. Index of Model Comparison	
Tucker-Lewis index (TLI)	• Not normed so values need not fall between 0 and 1 (the greater the value the better the fit)

Chi-Square Index

The chi-square goodness-of-fit index tests the fit between the specified model and the sample data by the function:

$$\chi^2 = (n-1)F_{ML}, \quad \text{where } F_{ML} = tr(S\Sigma^{-1}) - (p+q) + In|\Sigma| - In|S|$$

An obtained chi-square value that is greater than chi-square critical is considered significant, which means the estimated covariance matrix Σ is significantly different from the observed covariance matrix S. A non-significant chi-square value means the difference between Σ and S is due to sampling fluctuations and the two matrices are not significantly different. A non-significant difference is the desired outcome in that it is preferred that the S and Σ matrices are as similar as possible in order to claim the hypothesized model fits the data well. Although the chi-square is commonly used as an index of model fit in SEM research, it is generally used in conjunction with other fit indices because of its sensitivity to sample size (Joreskog & Sorbom, 1988). As sample size increases, so does chi-square, which tends to result in greater probability of a significant chi-square (a poor fitting model). When sample size decreases, it tends to indicate nonsignificant probability levels (Hayduk, 1987). Therefore, if the sample size is small enough (generally below 100), a good fit is very likely, whereas if the sample size is large enough (generally above 200), a good fit is difficult to achieve (Schumacker & Lomax, 1996).

Goodness-of-Fit Index (GFI) and Adjusted Goodness-of-Fit Index (AGFI)

The GFI is an indicator of the amount of variance and covariance in the observed variable matrix S that is predicted by the implied variable matrix Σ. The GFI function is:

$$GFI = 1 - tr[(\hat{\sum}^{-1}S - I)^2] / tr[(\hat{\sum}^{-1}S)^2]$$

The AGFI differs from the GFI in that it adjusts for the degrees of freedom in the model. The AGFI function is:

$$AGFI = 1 - [q(q+1)/(2df_m)][1 - GFI],$$

where S is the sample covariance matrix of q observed variables,
Σ is the sample implied covariance matrix, and
I is an identity matrix.

The AFGI adjusts for models with fewer degrees of freedom in that models with many free parameters are likely to have better fit than more parsimonious models, which makes it possible to use the AGFI as an index of model parsimony. Both the AGFI and the GFI have values ranging from zero to 1.00 with a value of .90 or greater indicating a good fit.

Root Mean Square Residual (RMSR)

The RMSR index is the square root of mean squared differences between the elements of S and Σ. The RMSR ranges from zero to 1.00 and good fit is indicated by values smaller than .05.

Indices of Model Comparison

Indices of model comparison compare the fit of a hypothesized model to the null model. The null model could be any model that is used as a basis for comparison of other models.

Tucker-Lewis Fit Index (TLI)

The TLI (which is the same as the Bentler and Bonett (1980) Non-normed Fit Index) is calculated using chi-square:

$$[(\chi^2_{null} / df_{null}) - (\chi^2_{proposed} / df_{proposed})] / [(\chi^2_{null} / df_{null}) - 1]$$

The TLI is non-normed which can make it difficult to interpret because it does not have a maximum and minimum value. In most cases, values will fall between 0 and 1 with values of .90 or greater reflecting good fit. Bollen (1990) found the calculation of the TLI is affected by sample size, however, the means of the sampling distributions are mostly unaffected by *N*. This means that although TLI values are not normed, they are more stable when comparing samples with different *N*'s.

Chapter Summary

An intervening mechanism evaluation design (Chen, 1990) was used to test the action and conceptual theories of the proposed model of participatory program evaluation. The conceptual theory was based on the hypothesis that Process Outcomes would predict Instrumental and Symbolic Use. The action theory was based on the hypothesis that Participative Climate, Level of PDM, and Influence in PDM would predict Process Outcomes. The intervening mechanism design as applied in the current study, required that both the action and conceptual theories be successful in order to determine the underlying causal processes of the B.C. School Accreditation Program in particular and participatory program evaluation in general.

The sample included 315 elementary and secondary teachers, from 91 elementary and secondary schools, who participated in the 1995/1996 B.C. School Accreditation Program. Participants were asked to complete three sets of questionnaires with either a one-day or a one-week time lag between each set. The time lags were incorporated into the design as a means of avoiding the percept-percept problem, which can cause inflated correlations due to collection of participation and outcome data in a questionnaire

administered in a single sitting (Campbell, 1982). Two of the six questionnaires used in the study were existing measures with acceptable internal consistency reliability. The remaining four questionnaires, which were developed specifically for this study, demonstrated internal consistency reliability ranging from .79 to .96 as well as content and construct validity. Construct validity was established through the use of principal components analysis.

The research design employed structural equation modeling to test the theory-based causal assumptions included in the intervening mechanism design. The structural model was identified using the two-step rule and maximum likelihood estimation. Model fit was determined through the use of the chi-square statistic, Goodness-of-Fit Index, Adjusted Goodness-of-Fit Index, Root Mean Square Residual, and the Tucker-Lewis Index.

CHAPTER FIVE:

RESULTS

Preliminary Analysis

Sample Size and Missing Data

The total number of cases for the current study was 315. There were 23 observed variables and nine latent variables. The ratio of cases to observed variables was 14:1, the ratio of cases to estimated parameters was 5:1, and the ratio of cases to latent variables was 35 :1. Based on the recommendations of Schumacker and Lomax (1996) and the other authors discussed previously in Chapter Four, a sample size of 315 was considered sufficiently large for the design. Missing data for each observed variable were replaced with the variable mean.

Missing observations ranged from 8.0% to 0.6%, which is considered low enough to have very minimal effects on the results (Little, 1987) (see Appendix C). If variable nine (the item with the highest percentage of missing data) were to be excluded, the range of missing data would have been considerably less (0.6% to 4.0%). Variable nine was a question on perceptions of actual versus preferred participation in determining subcommittee leaders. Missing data on this item may have been due to cases in which staff had little or no say in this pre-evaluation decision.

Missing data were tested for random patterns using a strategy suggested by Tabachnick and Fidell (1989). A variable with two groups was constructed from variable

nine. Group one consisted of non-missing data and group two consisted of missing data. Tests of mean differences were performed on Process Outcomes, Instrumental Use, and Symbolic Use between groups. None of the tests were significant, which indicated there was likely no difference between participants who did and did not respond to variable nine. See Appendix D for tests of mean differences between groups.

Outliers and Normality

The data were screened for multivariate outliers using SPSS to calculate the Mahalanobis distance (D^2) for each case. An outlier was indicated by a D^2 that was significant at $p < .001$ level. Mahalanobis distance was calculated as chi-square with degrees of freedom equal to the number of observed variables in the hypothesized model (Tabachnick & Fidell, 1989). For the current study, which had 23 observed variables, the chi-square critical value at $p < .001$ was approximately 49.7. If a D^2 for a single case exceeded 49.7 it was considered an outlier. Ten cases were identified as outliers. Table 10 includes the univariate summary statistics for the ten outlier cases as well as the total sample with the outlier cases removed (n = 305). A comparison indicated that skewness, kurtosis, and means were comparable between the outlier cases and the total sample. Therefore, it was decided all were accurate observations and none would be deleted from the analysis.

Multivariate normality was examined through PRELIS[2], which provided Mardia's (1970) measure of multivariate kurtosis. The relative multivariate kurtosis for the current study was considered to be satisfactory at .117. Table 10 lists the univariate mean, standard

---

[2] PRELIS is a data screening program included with LISREL8 (Joreskog & Sorbom, 1993).

deviation, skewness, and kurtosis for the 23 variables included in the study. None of the univariate analyses indicated severe skewness or kurtosis.

Main Analysis: The Two Step Approach to Modeling

Step One: Fitting the Measurement Model

Anderson and Gerbing (1988) described the measurement model as the hypothesized model with the least number of constraints and thereby fewest degrees of freedom. In a measurement model, all latent variables are related to all other latent variables and all parameters are estimated. To fit the measurement model in the current study a process was used in which each set of observed variables was fit to its original latent construct. That is, the items for each questionnaire were fit to their original hypothesized latent construct. For example, there were originally 12 items on the Level of PDM questionnaire. The latent construct called Level of PDM was, therefore, hypothesized to predict these 12 items. The items that had the highest loadings and the best fit with the Level of PDM construct were chosen as the set of marker variables for that particular latent construct (cf. Gable & Wolf, 1993). The process of determining marker variables was repeated for each latent variable. The latent variables and corresponding marker variables were then combined into a measurement model and re-tested for overall fit. When fitting the measurement model, I attempted to retain at least three marker variables per latent construct[3].

---

[3] The measurement model in Figure 4 shows that Ksi 2 has only two marker variables whereas the other latent variables all have at least three. Ksi 2 was originally fitted with eight observed variables. However, six of the eight variables correlated highly with marker variables from other constructs and caused severe positive residuals. Therefore, I decided to retain only 2 marker variables in that the model was still identified.

Table 10

Univariate Summary Statistics for Observed Variables

Total sample minus outlier cases (n = 305)					Outlier cases (n = 10)				
Variable	Mean	SD	Skewness	Kurtosis	Variable	Mean	SD	Skewness	Kurtosis
Y1	3.32	1.02	-.54	-.18	Y1	3.53	.96	.36	-.41
Y2	3.18	.94	-.35	-.19	Y2	3.00	1.25	-.43	.17
Y3	3.01	.88	-.14	-.46	Y3	3.10	.57	.08	1.51
Y4	3.32	.72	.22	-.07	Y4	3.10	1.20	-.23	-.37
Y5	3.11	.97	-.03	-.60	Y5	2.40	1.17	.04	-1.46
Y6	3.61	.83	-.35	-.39	Y6	3.50	.71	-1.18	.57
Y7	3.78	.84	-.34	-.25	Y7	3.90	.74	.17	-.73
Y8	3.82	.82	-.61	.10	Y8	4.00	.67	.00	.08
Y9	.38	.80	1.54	2.39	Y9	.40	1.27	.25	.34
Y10	.37	.88	1.01	1.89	Y10	1.54	1.43	-.23	-1.32
Y11	.34	.86	.61	2.43	Y11	1.50	1.12	.00	-1.76
Y12	4.81	1.12	-1.12	1.30	Y12	3.70	2.11	-.30	-1.63
Y13	4.80	1.18	-1.23	1.48	Y13	4.20	2.30	-.78	-1.47
Y14	4.94	1.16	-1.23	1.29	Y14	4.40	1.96	-.93	-.95
Y15	4.50	1.36	-.95	.31	Y15	2.50	1.58	.00	-2.57
Y16	4.49	1.28	-.83	.37	Y16	2.50	1.84	.80	-.72
Y17	4.47	1.14	-.99	1.25	Y17	3.30	1.89	-.17	-1.64
Y18	4.34	1.21	-.90	.83	Y18	2.60	1.58	.20	-1.84
Y19	4.75	1.09	-1.04	1.51	Y19	3.90	1.73	-.45	-.95
Y20	4.48	1.13	-.61	.15	Y20	3.10	1.79	.10	-1.26
Y21	4.47	1.04	-.87	.99	Y21	3.60	1.58	-.23	-.82
Y22	4.73	1.12	-1.04	1.45	Y22	3.10	1.85	.61	-.89
Y23	4.37	1.15	-.94	1.02	Y23	3.60	1.96	-.41	-1.24

Note. SD = standard deviation.

There has been considerable interest among SEM users regarding the issue of composite versus item level analysis in that the traditional practice of using item level analysis has been challenged (i.e., Catanzaro, 1997; Cui-Cui, 1997; Gregorich, 1997; Muliak, 1997). The primary issue appears to be whether it is appropriate to fit models based on composite variables without first testing the item level measurement model. Those who favor the use of composite variables (i.e., Gregorich, 1997; Russell, 1997) have argued that (a) composite scores are more reliable than single items, (b) in many situations researchers use measures with well known psychometric properties (i.e., IQ tests) so further item level analysis is redundant, and (c) it is very difficult to fit measurement models with established measures because they generally have a large number of items, which results in issues related to sample size and correlated error. Moreover, many researchers feel that much of the psychometric work done on established measures would be lost by reverting to item level analyses (Catanzaro, 1997; Russell, 1997). In the case of the current study, four of the six measures were developed specifically for the study and thereby did not have an extensive psychometric history. Using composites would have inferred that I was certain of the factor structure of each composite. Furthermore, none of the marker variables used in the measurement model violated the assumption of multivariate normality, which can occur with single item indicators. Therefore, it was considered more appropriate to fit the measurement model with marker variables than with composites.

The measurement model for the current study had 23 observed variables and eight latent variables. All observed variables included in the measurement model are described in Appendix E. The overall fit of the initial measurement model was $\chi^2(218, n = 315) =$

345.79, $p < .000$, GFI = .91, AFGI = .89, RMSR = .06. This fit was considered unacceptable given that acceptable fit levels for the chi-square fit index should be $p > .05$. Three paths were added to the initial measurement model as a means of improving the fit. Based on one large modification index, a thorough understanding of the variables, and inspection of residuals, variable Y21 (The accreditation information resulted in implementation of new ideas for how to improve school practices) was allowed to load on Ksi 7 (Instrumental Use); variable Y18 (The accreditation information made me a better advocate of my school) was allowed to load on Ksi 6 (Process Outcomes); and variable Y20 (The accreditation information convinced staff members of need for maintaining existing school practices) was allowed to load on Ksi 1 (Cooperative Goal Interdependence) (see Figure 5). The additional paths resulted in the final measurement model, which had an acceptable fit: χ^2 (199, n = 315) = 230.16, $p = .06$, GFI = .94, AFGI = .92, RMSR = .04, TLI = .99. Figure 5 depicts the measurement model for the current study.

Once the measurement model was determined to have good fit, it was compared to a series of alternate or nested models. Models with different observed variables can not be nested (Hayduk, 1987). The ideal situation is to have the alternate models fit the data as well as the measurement model (not significantly worse). It is therefore important that the measurement model has good fit in that it makes little sense to compare alternate models to a poorly fitted measurement model.

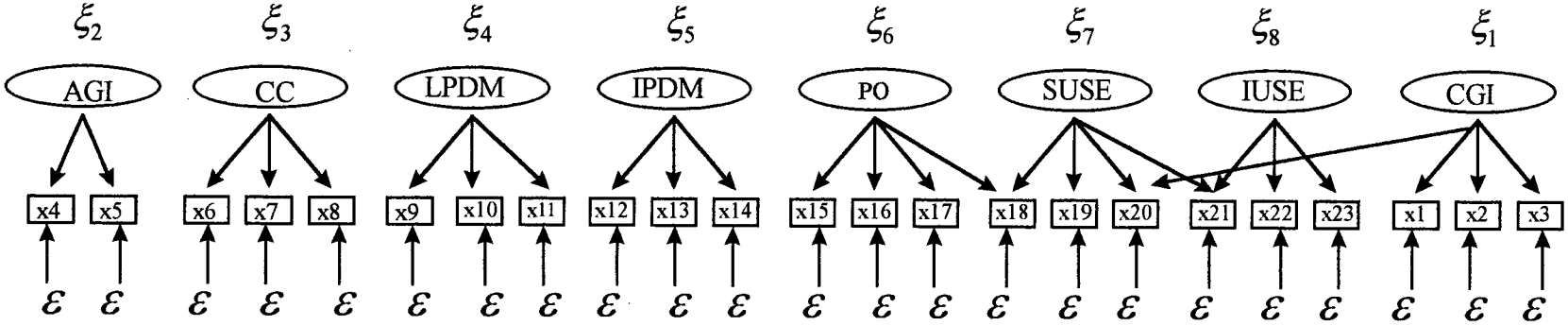


Figure 6. Measurement model.

The typical drawing conventions employed in structural equation diagrams are illustrated in Figure 6. Circles represent latent variables, squares represent observed variables, and relationships are depicted by curved lines[4] (Schumacker & Lomax, 1996). The overall fit of the measurement model in the current study was acceptable: χ^2 (199, n = 315) = 230.16, p = .06, GFI = .94, AFGI = .92, RMSR = .04, TLI = .99. The chi-square was not significant and the GFI, AGFI, and TLI were all at an acceptable level. As well, the RMSR was low which indicated small residuals between the implied Σ and observed S covariance matrices.

Step Two: Comparing the Measurement Model to Alternate Models

Once it was established that the measurement model had good fit, it was compared to several alternate models. The alternate model of most interest was the original theorized target model[5]. The target model for the current study is depicted in Figure 7. The exogenous variable in the target model, which is labeled Participative Climate, is a second-order latent construct. Participative Climate (Ksi 1) has five latent indicators: (1) Level of PDM (Eta 4), a latent variable with three observed indicators, (2) Influence in PDM (Eta 5), a latent variable with three observed indictors, (3) Process Outcomes (Eta 6), a latent variable with four observed indicators, (4) Symbolic Use (Eta 7), a latent variable with four observed indicators, and (5) Instrumental Use (Eta 8), a latent variable with three observed indicators. The target model also illustrates that Level of PDM (Eta 4) predicts Influence in

---

[4] To simplify the measurement model in Figure 6, the curved lines connecting the latent constructs were omitted from the diagram. However, it was still assumed all constructs were intercorrelated.

[5] The original theorized target model was developed from a review of the literature and has not been modified by post-hoc re-specification.

PDM (Eta 5). As well, Level of PDM (Eta 4) and Influence in PDM (Eta 5) predict Process Outcomes (Eta 6). And, Process Outcomes predicts both Instrumental Use (Eta 8) and Symbolic Use (Eta 7).

Table 11 includes maximum likelihood parameter estimates, standard errors, and *t*-values for the target model. The covariance matrix for the target model is shown in Table 12. The squared multiple correlations, theta epsilon and psi estimates, and standard errors for the Y variables are shown in Table 13. All the paths, except Beta (6,4) which is the relationship between Level of PDM and Process Outcomes, were statistically significant at $p < .05$. LISPOWER[6] (Joreskog & Sorbom, 1993) analysis indicated that the relationship between Level of PDM and Process Outcomes had a power value of .21.

There are three types of structural models included in the current study: (a) the original theorized target model (see Figure 6), (b) alternate models which are also theory-based (see Figure 7), and (c) equivalent models which are not necessarily based on theory but have statistical equivalence to the target model (see Figure 8). Statistical equivalence means two or more hypothesized models generate the same covariance matrix which makes it impossible to distinguish between equivalent models based on statistical criteria (Stelzl, 1986). The theorized target model being of most interest has been distinguished from the other alternative models.

---

[6] LISPOWER is a component of LISREL8 (Joreskog & Sorbom, 1993) used to calculate power for testing coefficients.

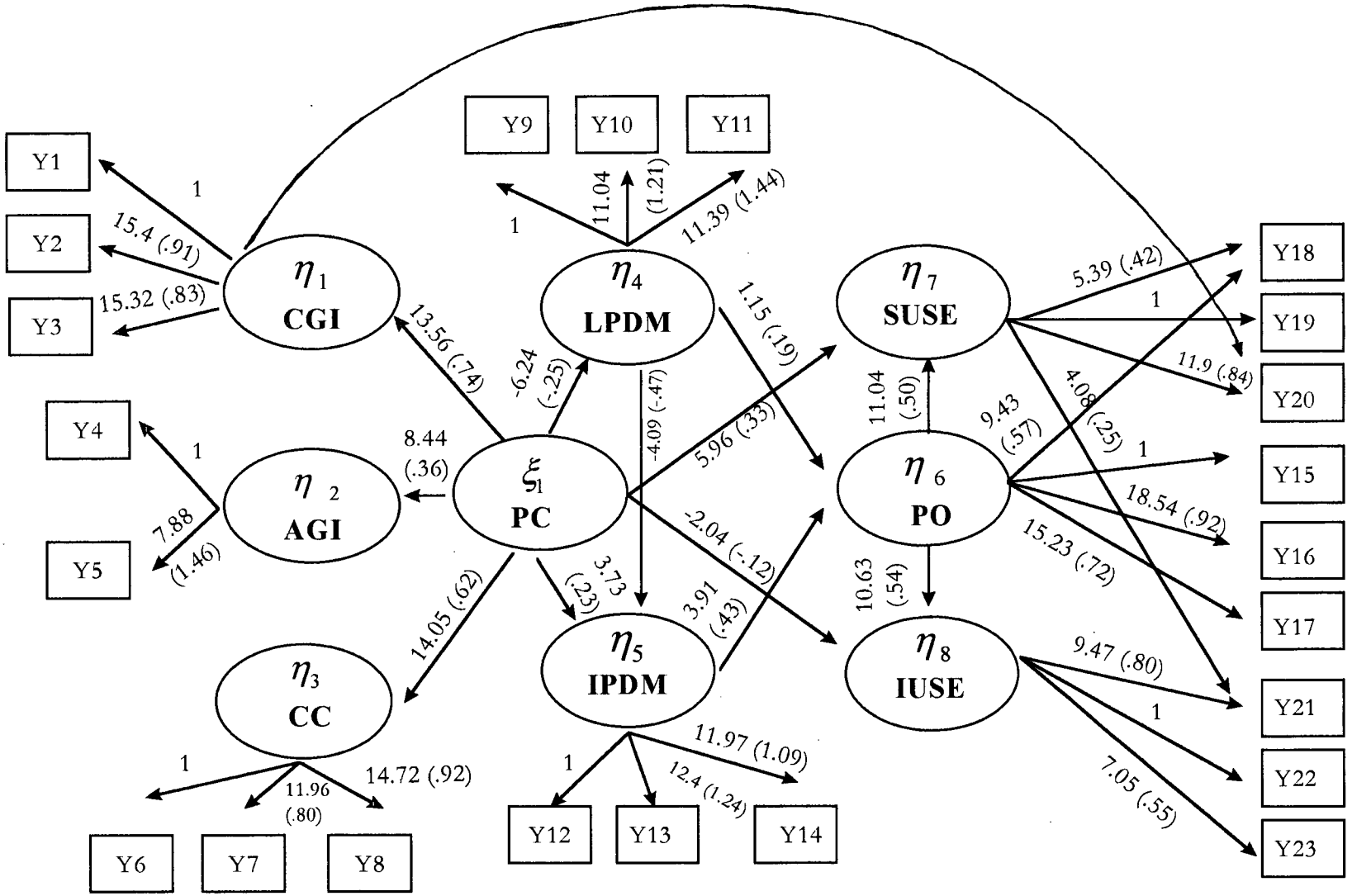


Figure 7. Original theorized target model.

Note. Estimates are standardized (t-values). Maximum-Likelihood Estimates in parentheses. As a means of ensuring that the measurement scale for the latent and observed variables was the same, one factor loading for each latent construct was fixed to 1.

Table 11

T-values for Parameter Estimates of Target Model

Path	Estimate	SE	t-value
Lambda (λ)			
2,1	.91	.06	15.40
3,1	.83	.05	15.32
20,1	.22	.07	3.12
5,2	1.46	.19	7.88
7,3	.80	.07	11.96
8,3	.92	.06	14.72
10,4	1.21	.11	11.04
11,4	1.44	.13	11.39
13,5	1.24	.10	12.40
14,5	1.09	.09	11.97
16,6	.92	.05	18.54
17,6	.72	.05	15.23
18,6	.57	.06	9.43
18,7	.42	.08	5.39
20,7	.84	.07	11.90
21,7	.25	.06	4.08
21,8	.80	.09	9.47
23,8	.55	.08	7.05
Beta (β)			
5,4	-.47	.12	-4.09
6,4	.19	.16	1.15
6,5	.43	.11	3.91
7,6	.50	.05	11.04
8,6	.54	.05	10.63
Gamma (γ)			
1,1	.74	.05	13.56
2,1	.36	.04	8.44
3,1	.62	.04	14.05
4,1	-.25	.04	-6.24
5,1	.23	.06	3.73
7,1	.33	.06	5.96
8,1	-.12	.06	-2.04

Note. t-values > 2.0 are significant at $p < .05$; SE = standard error of estimates.

Table 12. Variance-Covariance Matrix for Target Model

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	1.03																						
2	.641	.899																					
3	.577	.532	.751																				
4	.314	.279	.231	.543																			
5	.410	.384	.314	.334	.970																		
6	.465	.408	.396	.195	.325	.678																	
7	.349	.357	.328	.169	.261	.376	.706																
8	.397	.353	.346	.200	.279	.448	.358	.658															
9	-.219	-.137	-.157	-.143	-.159	-.145	-.134	-.182	.685														
10	-.172	-.175	-.110	-.082	-.138	-.157	-.108	-.176	.372	.799													
11	-.258	-.231	-.190	-.109	-.146	-.265	-.196	-.225	.440	.532	.825												
12	.143	.234	.166	.131	.188	.252	.224	.207	-.161	-.305	-.249	1.38											
13	.296	.287	.283	.172	.154	.323	.240	.295	-.238	-.323	-.343	.881	1.52										
14	.212	.217	.192	.084	.175	.242	.205	.228	-.153	-.316	-.354	.773	.959	1.41									
15	.081	.126	.093	.055	-.004	.143	.127	.100	-.005	-.127	-.084	.268	.330	.196	1.99								
16	.125	.195	.174	.032	.060	.164	.131	.068	.020	-.075	-.006	.343	.335	.260	1.41	1.80							
17	.088	.108	.072	.009	-.060	.101	.095	.079	.055	-.009	.024	.214	.254	.188	1.08	1.00	1.41						
18	.178	.191	.146	.041	.077	.203	.221	.154	-.006	-.119	-.064	.270	.252	.232	1.19	1.08	.862	1.57					
19	.291	.278	.242	.106	.244	.263	.216	.290	-.088	-.126	-.113	.279	.231	.301	.797	.761	.578	.860	1.25				
20	.406	.364	.333	.153	.279	.316	.293	.314	-.182	-.174	-.243	.219	.285	.335	.660	.707	.521	.786	.886	1.38			
21	.057	.037	.037	.048	.064	.066	.063	.046	-.024	-.009	-.041	.219	.149	.124	.815	.704	.558	.746	.593	.486	1.15		
22	-.049	-.024	-.009	.029	-.011	-.021	.097	-.015	.079	.073	.077	.191	.195	.105	.808	.757	.572	.714	.450	.396	.834	1.39	
23	-.093	-.061	-.079	-.095	.047	-.024	-.092	-.060	.078	.021	.035	.248	.081	.029	.418	.372	.338	.321	.256	.133	.510	.418	1.40

Table 13

Squared Multiple Correlations and Standard Error of Estimates

Squared Multiple Correlations	Standard Error of Estimates	
	Theta Epsilon	Psi
Y1 = .68	1,1 = .33(.04)	1,1 = .16(.04)
Y2 = .65	2,2 = .32(.03)	2,2 = .10(.03)
Y3 = .64	3,3 = .27(.03)	3,3 = .10(.03)
Y4 = .42	4,4 = .31(.04)	4,4 = .25(.04)
Y5 = .50	5,5 = .48(.07)	5,5 = .54(.08)
Y6 = .71	6,6 = .20(.03)	6,6 =1.41(.15)
Y7 = .43	7,7 = .40(.04)	7,7 = .44(.07)
Y8 = .62	8,8 = .25(.03)	8,8 = .46(.08)
Y9 = .45	9,9 = .38(.04)	
Y10 = .56	10,10 = .35(.04)	
Y11 = .77	11,11 = .19(.04)	
Y12 = .51	12,12 = .67(.07)	
Y13 = .72	13,13 = .42(.07)	
Y14 = .59	14,14 = .58(.07)	
Y15 = .76	15,15 = .48(.06)	
Y16 = .71	16,16 = .52(.06)	
Y17 = .55	17,17 = .64(.06)	
Y18 = .66	18,18 = .53(.05)	
Y19 = .78	19,19 = .27(.06)	
Y20 = .60	20,20 = .54(.06)	
Y21 = .70	21,21 = .35(.05)	
Y22 = .65	22,22 = .49(.08)	
Y23 = .19	23,23 = .14(.10)	

Note. Standard error of estimates in parentheses for theta epsilon and phi.

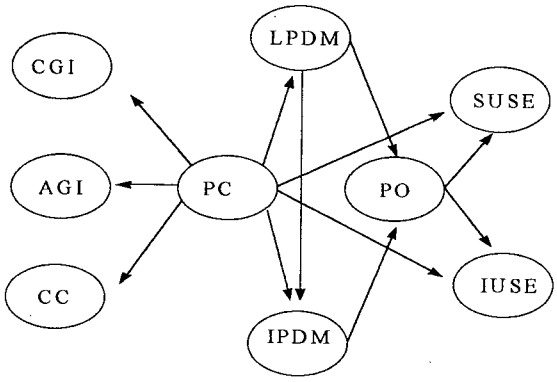
The alternate models (b) and (c) included in Figure 8 are based on the same literature as the target model; however, they are less parsimonious versions of the target model. The paths in the target model indicate that there is (a) a direct relationship between Process Outcomes and Instrumental and Symbolic Use, (b) an indirect relationship between Level of PDM, Process Outcomes, and Symbolic Use, and (c) an indirect relationship

between Influence in PDM, Process Outcomes, and Instrumental Use. The additional paths in the alternate models (b) and (c) create additional paths so that some of the relationships in the target model become direct as well as indirect. In other words, alternate models (b) and (c) have fewer constraints and are thereby less parsimonious than the target model because they hypothesize direct as well as indirect effects between Level of PDM, Influence in PDM, Process Outcomes, Symbolic Use, and Instrumental Use.

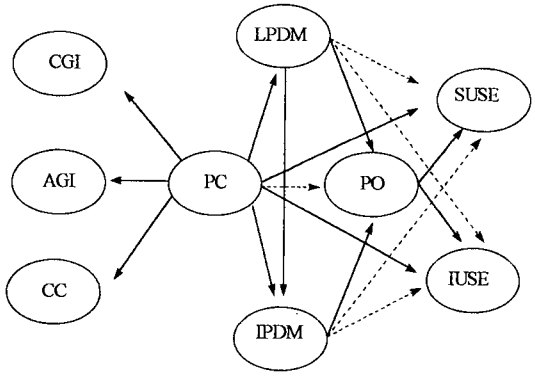
The alternate models were compared using sequential chi-square difference tests. This process determined whether additional constraints imposed by the alternate models significantly worsened the fit compared to the fit of the measurement model. The chi-square difference test is calculated by subtracting the measurement model chi-square from an alternate model chi-square using the function: $\chi_2^2 - \chi_1^2 / \chi^2_{df_2 - df_1}$. The degrees of freedom are also subtracted. The result is a maximum likelihood chi-square difference based on the degrees of freedom difference. A significant chi-square difference indicates that constraints added by the alternate model create a fit that is significantly worse than the fit of the measurement model. Conversely, a nonsignificant chi-square difference indicates that the alternate model fits the data as well as the measurement model, has additional degrees of freedom, and is more parsimonious (cf. Hayduk, 1987).

Using the chi-square difference test in conjunction with a decision tree framework proposed by Anderson and Gerbing (1988), the measurement and alternate models in the current study were compared. The primary purpose of the decision tree is to sequentially test a series of hypothesized models against a baseline model to determine how to proceed if an alternate model *significantly worsens the fit* of the baseline model. Based on the labeling

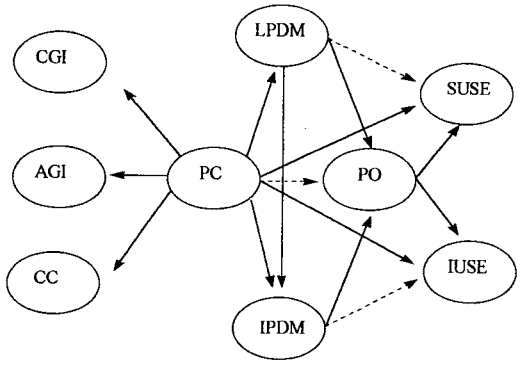
structure for the decision tree used by Anderson and Gerbing, the measurement model is labeled Ms, the target model is labeled Mc, and the alternate models are labeled Mu and Mt (see Figure 8).



(a) Theorized target model (Mc)



(b) Alternate model (Mu)



(c) Alternate model (Mt)

Figure 8. Theorized target and alternate models.
The dotted lines in (a) and (b) indicate differences between target and alternate models.

The target and alternate models were compared based on their degrees of freedom (see Table 14). For example, model Ms had the least number of constraints and the fewest

degrees of freedom. Model Mc had the greatest number of constraints and the largest degrees of freedom.

Table 14 illustrates the results of model comparison. The difference chi-square tests were not significant so the additional constraints on the target model Mc made it more parsimonious without significantly worsening the fit. The decision tree process did not incorporate the alternate model Mu because the chi-square difference between model Mt and Ms was not significant; however, if it had been, model Mu would have been included and tested against model Ms.

Table 14

Goodness-of-Fit of Alternative Models

Model (n =315)	χ^2	*df*	*p*-value	GFI	AGFI	TLI	RMSR
Measurement (Ms)	230.16	199	.06	.94	.92	.99	.04
Alternate (Mu)	248.14	210	.04	.94	.92	.99	.04
Alternate (Mt)	249.79	212	.04	.94	.92	.99	.04
Target (Mc)	251.96	215	.04	.94	.92	.99	.04

Model Comparison Sequential Chi-Square Difference Tests

	$\Delta\chi^2$	Δdf	χ^2_{crit}	*p*-value	Favored Model
Mt - Ms	19.63	13	22.40	$p > .05$	Mt
Mc - Mt	2.17	3	7.18	$p > .05$	Mc
Mc - Ms	21.80	16	26.30	$p > .05$	Mc

Comparing Equivalent Models to the Target Model

As previously stated, equivalent models produce the same covariance matrix, which means they produce the same goodness-of-fit values (Stelzl, 1986). It is, therefore, not possible to choose a "best" model based on statistical criteria alone. The problem of equivalence is especially prevalent when models contain a small number of latent variables.

In models with three or four latent variables, many equivalent models are possible depending on the causal ordering of the variables. However, in more complex models, equivalence is still possible but less likely. The target model in the current study, which had nine latent variables, did not have an overall equivalent model due to the temporal order of the variables. For example, Participative Climate preceded Level of PDM and Influence in PDM in temporal order in that the Participative Climate of the school existed prior to the accreditation program. Also, Level of PDM preceded Influence in PDM because in the case of B.C. School Accreditation it was not possible to have influence in decision making without participating. Furthermore, Process Outcomes could not be addressed without participating and Instrumental and Symbolic Use follow Process Outcomes because they were expected consequences of the Process Outcomes variable.

It must be noted that a grouping of variables within the target model did have several equivalent models. If the relationship among Process Outcomes, Instrumental Use, and Symbolic Use were to be fitted independent of the remainder of the target model, it would have several equivalent models. Figure 9 depicts two possible equivalent models. As depicted in Figure 9a, it was hypothesized that Instrumental Use predicted Process Outcomes, which in turn predicted Symbolic Use. Figure 9b depicts Process Outcomes predicted by Instrumental Use and Symbolic Use. The equivalent models illustrate that there was no sequential ordering for either Instrumental Use or Symbolic Use. It was not possible to make similar claims of equivalence with other parts of the target model. For example, it could not be claimed that Level of PDM predicted Influence in PDM because in

the case of B.C. School Accreditation one can not have influence in participation without participating.

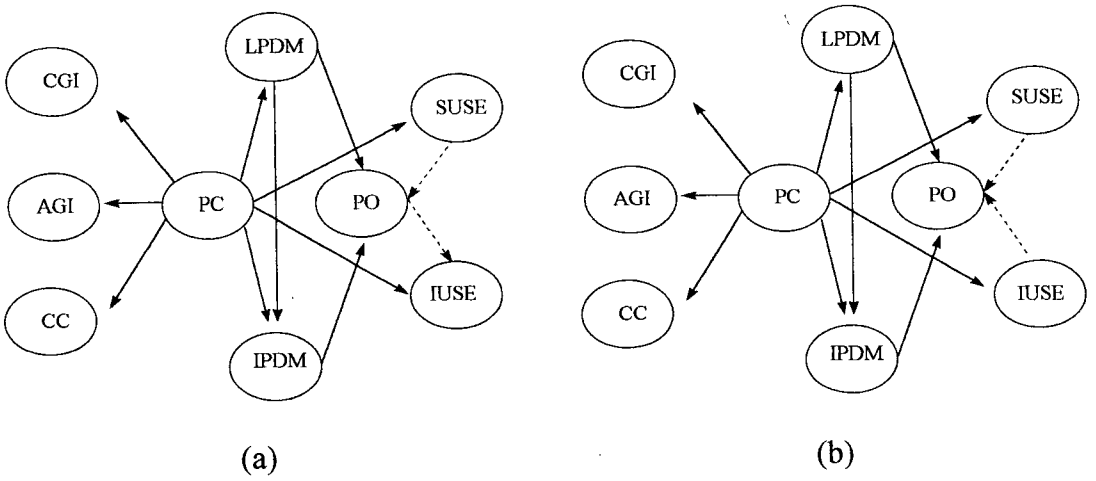


Figure 9. Equivalent models. Dotted lines indicate differences between target (Figure 8a) and equivalent models.

Table 15 includes the goodness-of-fit indices for the target and two equivalent models. Because all three models had the same degrees of freedom, it was possible to make direct comparisons among them without using the decision tree process and chi-square difference tests.

Table 15

Goodness-of-Fit of Equivalent Models

Model (n = 315)	χ^2	*df*	*p*-value	GFI	AGFI	RMSR	TLI
Equivalent Model A	258.96	215	.02	.94	.92	.05	.98
Equivalent Model B	312.94	215	.00	.92	.90	.09	.97
Target Model	251.96	215	.04	.94	.92	.04	.99

Based on the fit indices listed in Table 15, the two equivalent models had poorer fit than the target model so they were discarded as competing models. The target model was accepted as the best fitting model. Although Process Outcomes, Instrumental Use, and Symbolic Use were equivalent when considered independently, when included in the target model, they were not truly equivalent. The models depicted in Figure 9 were compared against the target model, not with the expectation that they would indicate equivalent goodness-of-fit, but to test which one had the best fit.

Expected Cross-Validation

The best way to validate a structural model is to fit it on a second set of data. This was not possible in the current study due to limited resources and time factors, in that the accreditation program occurs only once a year. Furthermore, the accreditation program is currently being revised and changes are likely to be made for the 1996/1997 program. There are currently several strategies which provide estimates of cross-validation using single samples. In the current study, two single-sample cross-validation strategies were used: (1) comparing the expected cross-validation index (ECVI) across alternate models and (2) testing invariance of the target model across grouping variables.

Table 16 shows the ECVI for the target and alternate models. The ECVI is available through LISREL8 (Joreskog & Sorbom, 1993) and is calculated as, $(c / n) + 2(q / n)$, where c = the chi-square for the model, q = the number of parameters estimated, and n = N - 1. LISREL8 also provides confidence intervals for the ECVI. The ECVI is interpreted by comparing the ECVI values across alternate and equivalent models. The model with the smallest ECVI is likely to be the most stable in the population (Schumacker & Lomax,

1996). Table 16 shows that the target model has the smallest ECVI and the narrowest confidence interval, therefore, it is the most likely to replicate.

Table 16

Expected Cross-Validation for Target and Alternate Models

Model (n =315)	χ^2	*df*	*p*-value	ECVI	Confidence Intervals
Target Model Mc	251.96	215	.04	1.19	1.08 - 1.33
Alternate Model Mu	248.14	210	.04	1.21	1.10 - 1.36
Alternate Model Mt	249.79	212	.04	1.20	1.09 - 1.34
Equivalent Model A	258.96	215	.02	1.21	1.10 - 1.36
Equivalent Model B	312.94	215	.00	1.39	**

Note. ** Confidence intervals not calculated because of small *p*-value for chi-square.

The second model validation strategy was to test invariance of the target model across categorical grouping variables. Table 17 includes the results of all invariance tests across groups. First, tests of factor invariance were conducted on two random samples (n = 157); second the total sample (n = 315) was split by elementary teachers (n = 173) and secondary teachers (n = 110); and third, the total sample was split by the questionnaire completion time lags of (a) one-day (n = 103) and (b) one-week (n = 213).

Based on a strategy suggested by Joreskog and Sorbom (1988), invariance was tested sequentially from the least to the most constrained model: (1) factor patterns were held invariant across groups, (2) factor patterns and loadings were held invariant across groups, (3) factor patterns, loadings, and error variances, were held invariant across groups, and (4) factor patterns, loadings, error variances, and variances/covariances were held invariant across groups. For each analysis, covariances matrices for both groups were analyzed

simultaneously, with one factor fixed at 1.0 as a means of establishing a common scale between the observed and latent variables (see Table 17).

Table 17

Invariance of Factor Structures

Group	χ^2	*p*-value	*df*	GFI (a) / (b)	RMSR (a) / (b)
(a) Random 1 (n = 157)					
(b) Random 2 (n = 157)					
• Equal factor pattern	467.96	1.00	430	.89/.89	.06/.06
• Equal factor pattern and loadings	483.36	.12	448	.89/.89	.06/.06
• Equal factor pattern, loadings, and errors	507.00	.12	471	.88/.88	.07/.06
• Equal factor pattern, loadings, errors, and variances/covariances	507.00	.12	471	.88/.88	.07/.06
(a) Elementary (n = 173)					
(b) Secondary (n =110)					
• Equal factor pattern	571.16	.00	430	.90/.82	.07/.09
• Equal factor pattern and loadings	612.10	.00	448	.89/.81	.07/.09
• Equal factor pattern, loadings, and errors	684.97	.00	471	.88/.78	.08/.11
• Equal factor pattern, loadings, errors, and variances/covariances	684.97	.00	471	.88/.78	.08/.11
(a) One-Day (n = 102)					
(b) One-Week (n = 213)					
• Equal factor pattern	579.98	.00	430	.91/.80	.06/.09
• Equal factor pattern and loadings	595.49	.00	448	.91/.80	.06/.11
• Equal factor pattern, loadings, and errors	632.39	.00	471	.90/.78	.06/.10
• Equal factor pattern, loadings, errors, and variances/covariances	632.39	.00	471	.90/.78	.06/.10

LISREL8 provided estimates of the GFI and RMSR for each group and an overall chi-square fit of the model for all groups. The invariance of factor structures, overall chi-square, and GFI and RMSR for each group are included in Table 17. LISREL program

codes for the tests of invariance are included in Appendix F. Results of invariance tests indicated that the chi-square goodness-of-fit indices for the elementary and secondary groups and the one-day and one-week groups were significant. However, the chi-square for the two random samples was not significant. The GFI and RMSR were comparable across all three sets, but somewhat better for the random samples. It is often considered acceptable to use chi-square difference as a test of the hypotheses that factor loadings are equal. This strategy has been used (i.e., Pitts, West, & Tein, 1996; Rahim & Magner, 1995) when chi-square coefficients indicated poor model fit but other indices such as AGFI, GFI, TLI, and RMSR suggested acceptable model fit.

In the current study, tests of factor invariance across the random samples were all non-significant so chi-square difference tests were not necessary. However, for the elementary-secondary and one-day/one-week groups, invariance tests indicated significant chi-square values for all tests and marginal fit from the GFI and RMSR. Therefore, chi-square difference tests were used to determine whether there was invariance across factor loadings (see Table 18). The results indicated that the added constraints worsened the fit of the model in all cases except for the one-day and one-week group. Overall, results of the two strategies suggested the model would likely cross-validate on an independent sample. The expected cross-validation indices (ECVI) were as expected in favor of the target model and the tests of hierarchical invariance on the two random samples indicated that the model was invariant across groups.

Table 18

Chi-Square Difference Tests

Group	$\Delta\chi^2$	Δdf	χ^2_{crit}	p-value
Elementary and Secondary				
EFPL-EFP	40.94	18	28.90	$p < .05$
EFPLE-EFPL[7]	72.87	23	35.20	$p < .05$
One-Day and One-Week				
EFPL-EFP	15.51	18	28.90	$p > .05$
EFPLE-EFPL	36.90	23	35.20	$p < .05$

Note. EFP = Equal Factor Pattern, EFPL = Equal Factor Pattern and Loading, EFPLE = Equal Factor Pattern, Loading, and Error.

Chapter Summary

Based on the standardized coefficients (see Table 11), all the relationships in hypotheses one and two were statistically significant with the exception of the path between Level of PDM and Process Outcomes. The following section includes a summary of the results for paths in hypotheses one and two.

Hypothesis One (Action Theory): Participative Climate, Level of PDM, and Influence in PDM predict Process Outcomes.

- Participative Climate predicted Level of PDM (coefficient -.25, t = - 6.24, $p < .05$)
- Participative Climate predicted Influence in PDM (coefficient .23, t = 3.73, $p < .05$)
- Participative Climate predicted Symbolic Use (coefficient -.12, t = - 2.04, $p < .05$)
- Participative Climate predicted Instrumental Use (coefficient .33, t = 5.96, $p < .05$)

[7] Because the model in the current study only has one KSI construct, there is only one element in the PHI matrix. This resulted in the same outcome when testing invariance of PHI, TE, and LY. Therefore, in this example, it was only necessary to include LY and TE.

- Influence in PDM predicted Process Outcomes (coefficient .43, t = 3.91, $p < .05$)
- Level of PDM predicted Influence in PDM (coefficient -.47, t = - 4.09, $p < .05$)
- Level of PDM did not predict Process Outcomes (coefficient .19, t = 1.15, $p > .05$)

Hypothesis Two (Conceptual Theory): Process Outcomes predicts Instrumental and Symbolic Use.

- Process Outcomes predicted Instrumental Use (coefficient .50, t = 11.04, $p < .05$)
- Process Outcomes predicted Symbolic Use (coefficient .54, t = 10.63, $p < .05$)

The preliminary and main analysis suggested the model is a plausible representation of the data. According to overall fit indices (χ^2 , GFI, AGFI, TLI, and RMSR) the measurement model had good fit. The target model had good fit on the GFI, AGFI, TLI, and RMSR; however, the chi-square index was significant ($p = .04$) which indicated a poor fit. All paths in the target model were significant except for the influence of Level of PDM on Process Outcomes, which had low power. The chi-square difference tests indicated the target model was more parsimonious than the other alternate models. Also, the target model fit the data better than the equivalent models. The cross-validation strategies indicated the target model would likely replicate in other independent samples for the models examined.

The current study tested two principal hypotheses, which represented the action and conceptual theories of the intervening mechanism design (Chen, 1990). An intervening mechanism design is an evaluation design strategy used to test the underlying causal processes of a program. The action theory in the current study tested whether Participative Climate, Level of PDM, and Influence in PDM predicted Process Outcomes. The conceptual theory tested whether Process Outcomes predicted Instrumental and Symbolic

Use. Results of individual path hypotheses indicated both the action and conceptual theories were tenable and that Process Outcomes mediated between the action and conceptual theories. The path between Level of PDM and Process Outcomes was not significant (possibly due to low power). However, the significant path between Influence in PDM and Process Outcomes maintained the important link between the action and the conceptual theories.

CHAPTER SIX:
DISCUSSION AND CONCLUSIONS

This chapter begins with a summary and discussion of the results in relation to prior research. Next, implications for participatory program evaluation in general and B.C. School Accreditation in particular are discussed. Then, directions for future research and limitations of the study are examined. The study concludes with a discussion of contributions and conclusions.

Summary and Discussion of Results in Relation to Prior Research

Action and Conceptual Theories

The primary purpose of the current study was to develop and test a theory-based model of participatory evaluation using an intervening mechanism design that involved testing an action theory and a conceptual theory. Hypothesis one, the action theory, tested the effect of the antecedent program variables (Participative Climate, Level of PDM, and Influence in PDM) on the intervening variable (Process Outcomes). All paths in hypothesis one are tenable with the exception of the path between Level of PDM and Process Outcomes. Hypothesis two, the conceptual theory, tested the effect of the intervening variable (Process Outcomes) on evaluation use (Instrumental Use and Symbolic Use). All paths in the conceptual theory are tenable. Overall, results indicate that both the action and the conceptual theories are tenable, which supports the mediating effect of Process Outcomes and indicates that the model is a plausible and defensible explanation of how participatory evaluation can be expected to work. Furthermore, Chen's (1990) intervening

mechanism design in conjunction with structural equation modeling is considered to be a viable strategy for studying the causal processes of participatory evaluation in general and for assessing the influence of mediating variables in education program evaluation in particular. The intervening mechanism design has been an effective means of conceptualizing the underlying causal processes of the participatory evaluation program in the current study.

Action Theory: Hypothesis One

The relationships among Participative Climate, Level of PDM, Influence in PDM, and Process Outcomes were as expected. Figure 10 illustrates a conceptual interpretation of the relationship between Participative Climate and Level of PDM.

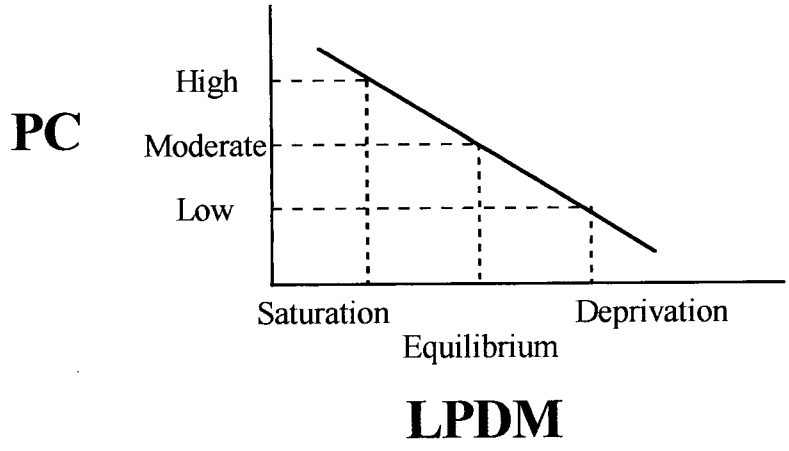


Figure 10. Relationship between Participative Climate and Level of Participation in Decision Making.

Teachers who perceived their schools to have high Participative Climate tended to experience saturation levels of participation (actual participation was higher than preferred participation). Conversely, teachers who perceived their schools to have low Participative

Climate experienced deprivation levels of participation (preferred participation was higher than actual participation). Moreover, teachers who perceived their schools to have moderate Participative Climate experienced equilibrium levels of participation (preferred equals actual participation).

Similarly, the relationship between Participative Climate and Influence in PDM indicated that high influence was related to high Participative Climate and low influence was related to low Participative Climate. Figure 11 illustrates a conceptual interpretation of the relationship between Participative Climate and Influence in PDM. As with Level of PDM, teachers who perceived their schools as highly participative also believed they had high levels of influence in decision making whereas those who perceived their schools as less participative had lower perceived influence in decision making.

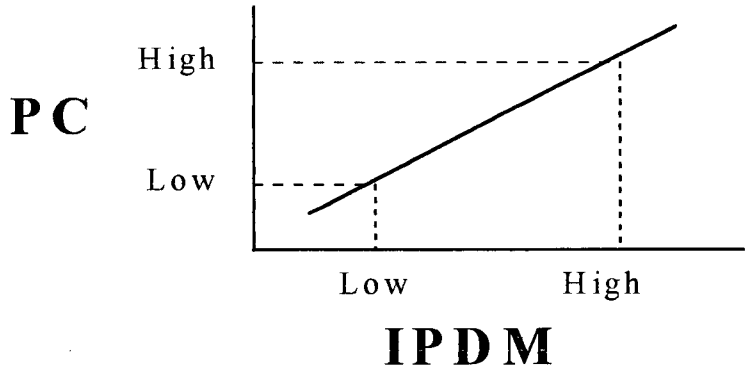


Figure 11. Relationship between Participative Climate and Influence in Participative Decision Making.

The relationships among Participative Climate and Level of PDM and Influence in PDM are consistent with research on goal theory and controversy (i.e., Erez & Arad, 1986; Latham & Saari, 1979; Tjosvold, 1985, 1986, 1987; Tjosvold et al., 1983) in which

workplace climates characterized by supportive atmospheres, group discussion, cooperation, and constructive controversy led to increases in employee satisfaction with participation and increased performance. Therefore, based on this prior research, the positive relationship between Participative Climate, Level of PDM, and Influence in PDM was as expected in that teachers who perceived their school to have cooperative or autonomous goals and constructive controversy believed they had high degrees of participation and influence in decision making.

The relationship between Level of PDM and Influence in PDM was also as expected (see Figure 12).

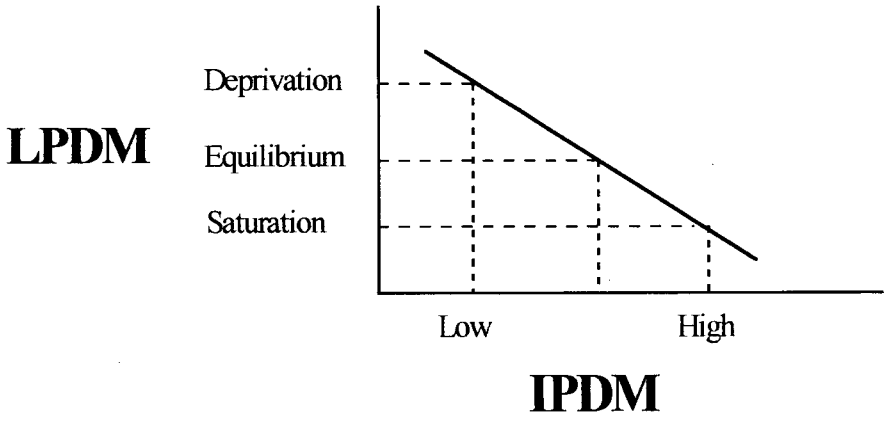


Figure 12. Negative relationship between Level of Participative Decision Making and Influence in Participative Decision Making.

Figure 12 illustrates a conceptual interpretation of the negative relationship between Level of PDM and Influence in PDM in which deprivation levels of participation are related to low perceived influence and saturation levels of participation are related to high perceived influence. Teachers who had equilibrium levels of participation perceived themselves to have moderate influence. These findings correspond with prior research on

participation in school contexts (i.e., Duke et al., 1980; Imber et al., 1990; Taylor & Bogotch, 1994) in which teachers differentiated between participation and influence in school decision making and indicated that participation in decision making did not necessarily result in influence in decision making. The results of the current study confirm this in that teachers with deprivation and equilibrium levels of participation perceived themselves to have lower influence in decision making than teachers with higher levels of participation. This finding indicates that participation does not necessarily predict influence, rather, it is the level of participation that predicts influence.

Results of hypothesis one (action theory) revealed two unexpected relationships. First, the path between Participative Climate and Instrumental Use was expected to be positive rather than negative (see Figure 13).

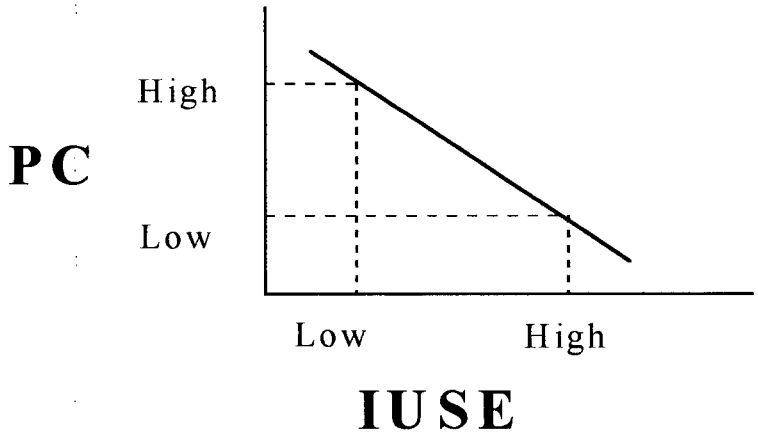


Figure 13. Negative relationship between Participative Climate and Instrumental Use.

The negative relationship between Participative Climate and Instrumental Use suggests that teachers who perceived their school to have high Participative Climate tended to believe the accreditation program *did not* result in changes to school practices.

Conversely, teachers who believed their school had low Participative Climate tended to believe the accreditation program *did* result in changes to school practices. This finding suggests that non-constructive controversy led to integration of opposing ideas and in turn, instrumental use, whereas constructive controversy did not.

The negative relationship between Participative Climate and Instrumental Use contradicts prior research and is somewhat illogical. Based on the research discussed in Chapter One (i.e., Tjosvold, 1986; Tjosvold et al., 1983), workplace climates with cooperative goals and constructive controversy were most effective in facilitating the dialogue necessary to understand conflicting points of view. Based on this it was assumed that, in participative school climates, opposing points of view would be assimilated into decisions which would lead to increased implementation of decisions and thereby increased instrumental use. The results, however, indicate the opposite. Participative Climates with competitive goals and non-constructive controversy had higher levels of perceived Instrumental Use than cooperative climates with constructive controversy.

It is possible that the unexpected negative relationship between high Participative Climate and low Instrumental Use was due to a "group think" mentality (Janis, 1972). The term group think has been used to describe participative decision making situations where poor decisions were made because members did not discuss or integrate conflicting points of view into the decision process . For example, Janis (1972) used group think to describe participatory processes involved in famous political decisions such as the Bay of Pigs in which the decision making capacity of the group was undermined as a result of not considering or questioning alternate ideas. Similarly, Tjosvold and Deemer (1980) reported that cooperative or competitive climates, where conflict had been avoided, resulted in

problematic outcomes such as group think. They posited that participants who avoided discussing conflicting ideas had a, "... false sense of understanding that masked actual ignorance" (p. 590).

In the current study, it is possible that teachers who perceived their schools to have constructive controversy could actually have experienced controversy avoidance. Some teachers may have kept opposing opinions to themselves in order to avoid conflict. If teachers perceived that their input would cause arguments or be used to attack their position on an issue, it is likely they may have chosen to keep their opposing views to themselves. It is also possible that complacency or fatigue during the accreditation process may have resulted in teachers agreeing with the majority or deciding on the easiest alternatives in order to avoid conflict and complete the process more quickly. In such cases, due to the absence of conflict, teachers could have interpreted controversy avoidance to be constructive controversy, which would account for the relationship between high Participative Climate and low Instrumental Use.

An explanation for the negative relationship between low Participative Climate and high Instrumental Use is difficult with the data in the current study. This finding suggests that low Participative Climate is related to high Instrumental Use and further contradicts what was expected based on prior research (i.e., Tjosvold, 1986; Tjosvold et al., 1983). A possible explanation for this finding could be that, in low Participative Climate with high Instrumental Use, the decisions and subsequent changes were not representative of the staff as a whole but represented a vocal minority. In the case of very non-participative climates, it is likely that reaching consensus would be nearly impossible. However, decisions would

have to be made and changes would have to be based on the decisions regardless of whether they represented a vocal minority or the staff as a whole.

Based on the questionnaire structure in the current study, the Instrumental Use questionnaire enabled participants to acknowledge that changes did occur, but there was no way for them to indicate whether they agreed with the changes. Therefore, it is possible that in highly non-participative schools, many decisions and changes were made, which reflected high Instrumental Use, but did not represent the opinions of the staff as a whole. This explanation would account for the negative relationship between Participative Climate and Instrumental Use. However, it raises additional questions regarding the relationship between participation, agreement with decisions, and Instrumental Use. The unexpected negative relationship between Participative Climate and Instrumental Use warrants further study.

The second unexpected result in the action theory was the statistically non-significant relationship between Level of PDM and Process Outcomes. As discussed in Chapter Four, this path had low power, which could be the reason for the non-significant relationship. However, it is also possible that the relationship between Level of PDM and Process Outcomes is an example of "complete" rather than "partial" mediation (James & Brett, 1984). Based on prior research (i.e., Brandon et al., 1994), I predicted the path between Level of PDM and Process Outcomes would be significant and that there would be partial rather than complete mediation between Level of PDM and Process Outcomes. Had I predicted complete mediation, the path between Level of PDM and Process Outcomes would have been omitted and all of the effect of Level of PDM on Process Outcomes would have been mediated by Influence in PDM. In either case, low power or complete versus

partial meditation, further research is needed to clarify this relationship in that it is a key component in the conceptual and action theories of the intervening mechanism design.

The positive relationship between Participative Climate and Symbolic Use was as expected. As discussed in Chapter One, affective models of participation predict links between employee satisfaction and increased performance. For example, through affective models, it is generally assumed that participation will meet certain basic work needs such as respect, independence, and equality, which are believed to increase morale and job satisfaction (Miller & Monge, 1986). I assumed high Participative Climate would be related to these basic work needs and would thereby result in greater school advocacy. The results confirmed this in that teachers who believed their school had high Participative Climate *also* believed the accreditation process (a) made them better advocates of their school, (b) resulted in feelings of overall satisfaction with school achievement, and (c) convinced them to maintain existing school practices. Conversely, teachers who believed their school had low Participative Climate *did not* believe the accreditation process (a) made them better advocates of their school, (b) did not have feelings of overall satisfaction with school achievement, nor (c) were they convinced to maintain existing school practices.

Conceptual Theory: Hypothesis Two

The positive relationship between Process Outcomes, Instrumental Use, and Symbolic Use was as expected. Teachers who had positive attitudes toward the outcomes of the participatory processes also tended to perceive the greatest amount of Instrumental and Symbolic Use. The relationship between Process Outcomes and Instrumental and Symbolic Use is crucial to the action and conceptual theories of the intervening mechanism design. Process Outcomes is the link and mediating component between the action and conceptual

theories. The significance of Process Outcomes as a mediating variable substantiates the intervening mechanism model in the current study and thereby substantiates the importance of Process Outcomes in participatory program evaluation in general and in B.C. School Accreditation in particular.

Implications

Participative Climate is a key factor in understanding the nature and function of participatory evaluation. The model in this study indicates that low Participative Climate results in (a) deprivation participation, (b) low influence, (c) low Process Outcomes, and (d) low Symbolic Use. Therefore, Participatory program evaluation is not a panacea for all evaluations in that it appears to be less effective in non-participative climates. Consequently, it is important that Participative Climate be considered prior to implementing a participative evaluation approach. I make this point fully aware of the realities of conducting participatory evaluations in contexts where the democratic appeal is high and thereby difficult to challenge. From an employee's point of view, disputing a participatory approach could be interpreted to mean employees can't make decisions as well as management. Similarly, from management's point of view, challenging a participatory approach because of concerns over the participative climate of a program could be interpreted to mean management is responsible for a poor work climate.

However, in spite of the popularity of participatory evaluation and the potential problems connected with challenging the use of participatory processes, it is likely that participatory evaluation will be ineffective, in terms of facilitating use, unless Participative Climate is examined closely. In most cases, it is probable that the cost and time

commitment would be better spent on alternate forms of evaluation. Only if there is clear evidence of participants' ability to function effectively with a participatory approach is it likely that participation will be effective in facilitating use.

In the case of B.C. School Accreditation, the rationale for using a participatory approach is not well documented. Lim (1989) reported that revisions for the 1994 Accreditation Manual attempted to take into account research literature on effective schools, which was reviewed and updated each year. Ministry of Education documentation and literature relating to B.C. School Accreditation are not explicit in describing specifically what research on effective schools has been integrated into the accreditation program. Based on my own review of effective schools literature (i.e., Bacharach et al., 1990; Bacharach & Conley, 1986; Cohen, 1982; Conley, 1991; Duke et al., 1980; Imber at al., 1990; Purkey & Smith, 1983; Rosenholtz, 1985) it appears that the 1995/1996 accreditation program has incorporated much of the effective schools ideology in terms of teacher participation.

For example, it is generally agreed in effective schools literature that teachers ultimately control the fate of efforts to alter a school's instructional climate and process, and that attempts at change are more successful when teachers work together. Furthermore, the findings of effective schools literature suggest that teachers should participate in the entire school organization, and that a participatory approach is central to how a school moves toward change. Teachers who work together at reaching consensus are more likely to own their decisions and the resulting attempts at change. These notions are very much a part of the B.C. School Accreditation process, which is staff driven and consensus based.

Purkey and Smith (1983) pointed out, however, that although effective schools research emphasizes teacher participation and consensus building processes, there is little research on how consensus is reached with staff who are unaccustomed to participatory administrative styles. Furthermore, there is little research on what methodology is best suited to studying participatory processes. Therefore, I would agree with Lim (1989) that effective schools research is evident in the current accreditation process, but I would question how the accreditation program addresses the participation issues raised by Purkey and Smith. In other words, how does the current accreditation program address the nature and function of participation? I would argue that it does not. Rather, it appears that participation is used for its democratic, economic, and symbolic benefits.

Ministry literature on the accreditation program does not provide an explanation or rationale for the participatory component of the program. That is, there is no explanation of the purpose, benefits, or expected outcomes of teacher participation. Nor is there an explanation of how teacher participation is linked to improved decision making, teacher performance, or school performance. Furthermore, there is no documentation on how participation is expected to work in schools where staff have had little exposure to shared decision making. Based on the lack of literature on the participatory component of the program, it appears that the Ministry is using a participatory approach because the effective schools literature indicates that participation is related to teacher and school performance. However, as discussed in Chapter One, the research tends not to support this assumption. There is little evidence to suggest that teacher participation in decision making results in increased teacher or school performance. Current changes to the accreditation program have implications for the participatory component of the program.

In a recent report by the Office of the Comptroller General (Ministry of Finance and Corporate Relations, 1997), recommendations were made to have the accreditation program focus as much on measuring product as process. This means that teachers will most probably be required to collect data on objective and independent performance indicators as well as keep track of changes in the indicators, the result of which will be used as evidence of school performance. These changes will affect the nature and function of teacher participation. For example, teachers will likely be required to measure performance indicators at regular intervals, interpret the data, and then make decisions as to what change strategies are needed. This ongoing process of collecting, interpreting, and implementing is going to require a great deal more participation from teachers than is currently required.

If the accreditation program and the school growth plan are to be seen as a continuous improvement process, the participatory structure of the accreditation program should be echoed in the entire improvement process. That is, if teachers are given influence in the accreditation program and school growth plan, then it would be expected that they also have influence in subsequent decisions regarding the implementation of the growth plan. Changes such as these require that more thought be put into the nature and function of teacher participation.

The relationship between teacher participation in decision making, teacher performance, and school performance must be considered. Furthermore, how decisions are made, who has authority, and what strategies are in place to ensure that problems such as group think or dominance of a vocal minority do not occur. Moreover, what training will be available to staff and administrators who have little understanding of shared decision making. These are difficult questions and it is likely that they will never be fully addressed.

Rather, individual schools will be required to deal with these issues in the context of limited time and resources, which is the situation that has occurred with these issues in past accreditation programs. Also, the participatory components of the accreditation program have not been addressed and schools have been required to deal with them under the guidance of their administrator who may or may not be sensitive to the issues of participative decision making.

Based on the current changes to the accreditation program, it is important that consideration be given to the effects of participative climate on expected outcomes. Furthermore, it can not be assumed that mandatory participation will force staff to be participative; requiring teachers to participate will not solve participative climate issues and the accreditation process, as it is now, will not teach or train teachers in effective participative decision making. The Ministry in conjunction with teachers must decide what purpose teacher participation has in the accreditation process, document this purpose, and then build the accreditation program around this purpose. The nature and function of participation must be defined.

Level of participation in pre-evaluation decisions is a key factor in understanding the nature and function of participatory evaluation. The findings indicate the importance of teacher participation in both pre-evaluation and evaluation content decisions in terms of predicting use. As previously explained in Chapter Four, the items for Level of PDM (see Appendix E) were based on pre-evaluation decisions such as deciding how final decisions would be made, how meetings would be conducted, and who sub-committee leaders would be. The statistically significant relationship between Level of PDM and Influence in PDM (see Figure 7) indicates that deprivation participation in pre-

evaluation decisions is related to low influence in evaluation content decisions. I believe that this is a very important finding. It suggests that participation in pre-evaluation decisions is as important as participation in content decisions in terms of predicting Instrumental and Symbolic Use. Furthermore, it suggests that teachers are interested in the process of how decisions are made.

The concern for pre-evaluation decision participation has implications for participatory evaluation. It suggests that participants are interested in how decisions are made and thereby believe that the decision process has an effect on their influence and in turn the eventual use of the evaluation information. For example, it was not surprising that teachers in the current study believed that participation in pre-evaluation decisions was important in that they had the greatest understanding of their program and therefore were most likely to determine which decision processes were best suited to their individual situation.

Participation in the pre-evaluation decisions enables participants to choose decision processes that give them the greatest control. For example, in a context with low Participative Climate, a traditional decision process, such as voting with a show of hands, could create uncomfortable situations for those who vote against the majority, which may result in participants voting against their better judgment. It is likely that anonymous voting strategies would be more appropriate in such contexts. If participants have the opportunity to discuss such issues and concerns, it is likely that they would choose a decision strategy best suited to their particular context. Overall, participation in pre-evaluation decisions allows participants to direct their own participation. It enables them to choose their leader, decide how decisions are to be made, and how meetings are to be conducted, all of which

predict influence in evaluation content decisions and give participants considerable control over the evaluation process.

In the case of B.C. School Accreditation, I believe most schools insufficiently addressed the issue of how decisions would be made. In spite of the fact that there are several examples of different processes for making decisions and reaching consensus included in the Ministry Accreditation Manual, it is likely that most teachers made decisions in the traditional manner (through votes using a show of hands). This assumption is based on a conversation with a Ministry Accreditation official in which I was told that not all the decision making processes were being taught in the accreditation training sessions. It is, therefore, probable that most schools used traditional voting methods and spent little time discussing alternative decision methods.

By choosing to omit some of the decision strategies from accreditation training, the Ministry is sending a message to teachers that the pre-evaluation decision process is not important, which in turn is undermining teachers' pre-evaluation participation. Putting less emphasis on pre-evaluation decisions could be interpreted as a means of limiting teacher influence in the decision process. If the pre-evaluation decisions were to be emphasized as an integral part of the accreditation process, it would force teachers to discuss which decision processes would be best for their school. I believe discussion and consideration regarding how to make decisions would greatly improve teachers' perceptions of the importance of their participation and the eventual use of the accreditation information.

I am aware that there are time issues involved with increased emphasis on pre-evaluation decisions; however, in light of recent changes to the accreditation program, which include reducing the number of criteria statements by one third, more time will be

available to focus on the participative nature of the program. I am also aware that it is often the case that administrators will make many of the pre-evaluation decisions based on their understanding of their staff and as a means of saving time. If the findings of this study are directly applicable, staff should be required to make these decisions, particularly decisions regarding the sub-committee leaders and how decisions are to be made.

Influence in Participative Decision Making and Process Outcomes are key factors in understanding the nature and function of participatory evaluation. The relationship between Influence in Participative Decision Making and Process Outcomes indicates that influence in participatory evaluation is related to positive perceptions of participatory outcomes. Teachers who believed they had influence in the evaluation process also tended to have positive opinions of the outcomes of the participative process. Specifically, teachers who had positive participatory experiences also (a) had positive perceptions of the usefulness and necessity of the information, (b) believed the information enabled them to contribute to the development of the school growth plan, (c) had a better understanding of colleagues' point of view, and (d) had higher levels of school advocacy. Conversely, teachers who believed they had low influence tended to have less favorable opinions regarding the participatory outcomes.

The concept of Process Outcomes is an important factor in the intervening mechanism design. Process Outcomes mediates the relationship between the conceptual and action theories and thereby functions as a link between antecedent program variables and program outcome variables. The significance of Process Outcomes as a mediator variable contradicts prevailing assumptions regarding participatory evaluation. As previously discussed, participatory evaluation is generally accepted as a means of

facilitating conceptual, instrumental, and symbolic use (Garaway, 1995; Greene, 1988a). The significance of Process Outcomes as a mediator in the intervening mechanism design indicates that it is the participants' perception of the participative Process Outcomes that facilitates use. Therefore, it is no longer appropriate to simplify the relationship between participation and use by assuming that participation alone will cause the use of evaluation information. The relationship between participation and use is much more complex than this assumption implies. The findings of the current study indicate that participation must be considered in terms of four components: (a) participative climate, (b) participation in pre-evaluation decisions, (c) participation in content decisions, and (d) perceptions of participative Process Outcomes. All four components affect the use of evaluation information.

Directions for Future Research

Further research is necessary to clarify the unexpected relationship between (a) Participative Climate and Instrumental Use and (b) Level of PDM and Process Outcomes. Both of these relationships contradict what was expected based on prior research. In the case of the non-significant relationship between Level of PDM and Process Outcomes, further study should examine complete versus partial mediation. In the case of the relationship between Participative Climate and Instrumental Use, future research should examine the mediating effects of controversy avoidance and agreement with evaluation decisions.

Another aspect of the model which should be addressed in future research is the relationship between Instrumental and Symbolic Use and Process Outcomes. The current

study did not address possible recursive relationships among these constructs. A more thorough examination of these constructs was problematic in the current study in that there was a considerable amount of correlation among the items for Instrumental Use, Symbolic Use, and Process Outcomes. Attempts at post-hoc model analysis of non-recursive relationships were unsuccessful. Additional work should be done on the measures to find marker variables which load only on a single construct. There are few developed measures for these constructs so continued psychometric work on the measures developed in the current study is recommended.

One last issue that would benefit from further investigation is the distinction between conceptual use and Process Outcomes. As defined in the introduction, conceptual use is increased understanding or learning as a result of the evaluation (Shadish et al., 1991). On the basis of this definition, I assumed conceptual use to be an expected outcome of participative evaluation. It could be argued that almost any credible participatory evaluation would most certainly be expected to result in some degree of learning or increased understanding among the participants. Therefore, I considered Process Outcomes, rather than conceptual use, to be the construct of interest and, therefore, melded conceptual use with Process Outcomes. It is possible, however, that this assumption was wrong and that conceptual use is in fact a separate and thereby important link in the participation use relationship. If the conceptual use construct were to be added to the participatory evaluation model in this study, I would hypothesize that it would precede Process Outcomes and mediate the relationship between Level of PDM, Influence in PDM, and Process Outcomes. It would, however, be important to distinguish the conceptual use construct from Process Outcomes, which may be possible by using a difference score strategy similar to that used

with Level of PDM. The difference score could be used to distinguish actual and expected learning as a result of the participative process. Further research is needed to test this assumption and the relevance of conceptual use in the participatory model.

Limitations of the Study

There are two primary limitations in this study. The first is related to the Level of PDM and Influence in PDM measures. The Level of PDM and Influence in PDM were each measured by three observed variables. Each of the observed variables is related to either a pre-evaluation decision (Level of PDM) or an evaluation content decision (Influence in PDM). Although it was hypothesized that Level of PDM and Influence in PDM were two distinct constructs, which was also supported by the results, developing measures that distinguished between them was problematic.

Participation in evaluation can be voluntary or mandatory. In the case of the B.C. School Accreditation, it was mandatory. Measuring Level of PDM in a mandatory setting was problematic because everyone was required to participate to the same degree in that they had to attend the same meetings and spend the same amount of extra-curricular time in the process. It was, therefore, decided to measure participation in terms of differences between preferred and actual contributions to evaluation decision making. Through the use of a difference score, distinctions were made between deprivation, saturation, and equilibrium levels of participation.

Although the use of a difference score was effective in terms of distinguishing Level of PDM from Influence in PDM, there was considerable correlation between (a) items within the Level of PDM measure, (b) items within the Influence in PDM measure, and (c)

between Level of PDM and Influence in PDM items. Each item on the Level of PDM measure matched an item on the Influence in PDM measure; where the Level of PDM requested participation ratings, the Influence in PDM requested influence ratings. This resulted in high correlations, which made it difficult to fit the measurement model. The final decision was to include only those items which had the highest factor loadings and the lowest correlation. Subsequently, only six of the original 24 items were used to measure these two constructs. Although this was an acceptable number of observed variables by statistical criteria, it emphasizes the need for further study on the measurement of these closely related constructs.

The second limitation of the study is related to the time-frame in which use occurs. The accreditation process unfolds throughout the school year with the development of the school growth plan occurring near the end of the year. The growth plan is intended to guide school improvement for a number of years (approximately six) so not all intended changes are implemented immediately but occur sequentially. The data for the current study was collected in the last two months of the school year when in most cases, the accreditation process was finished or near finished. However, even though most schools were finished with the accreditation process, there would have been considerable variance among schools in terms of how much of their school growth plan was already implemented or was in the process of being implemented. Therefore, when the participants in this study were asked to respond to questions regarding Instrumental and Symbolic Use, it is unclear whether they were responding based on changes they were currently implementing or changes they believed they would be implementing in the future. This issue would benefit from further

study because it is currently not known to what degree school growth plans are actually implemented or to what degree Process Outcomes affects long term implementation.

Finally, it could be considered appropriate to use the school as well as the individual as a unit of analysis. Although viewing participation from the school level would address a somewhat different research question than addressed in the current study, it is an interesting idea and thereby warrants further consideration.

Contributions and Conclusions

The impetus for this dissertation was the need to clarify the underlying causal processes of participatory program evaluation and to learn if and how participatory evaluation influences the use of evaluation information. Specifically, this study was designed to gain a better understanding of participatory evaluation constructs by testing causal relations in a proposed theory-based model of participatory evaluation. As part of the process of identifying the constructs of participatory evaluation, a model was developed based on the action and conceptual theories of an intervening mechanism design (Chen, 1990). A new mediator variable, Process Outcomes, was identified as the causal link between the action and conceptual theories and thereby demonstrated the potential of Chen's (1990) intervening mechanism design as a means of describing how participation influences the use of evaluation information. The model developed for this study can be used as a framework on which new developments are related to previous work and thus, decrease the prevailing assumptions regarding participatory evaluation.

Although participatory evaluation is popular, its application in school settings has outpaced any attempts to analyze its effectiveness. Thus, this study should provide

practitioners with the confirmation that participation is related to use. However, caution should be exercised in that participatory evaluation is not a panacea for all education evaluation. If the findings from the current study are directly applicable, participatory evaluation can foster Process Outcomes and increased use of evaluation information, but it can also foster decreased use, deprivation levels of participation, low influence in decision making, and low Process Outcomes. Overall, this study has provided a theory-based and defensible model of how participatory evaluation can be expected to work to increase the use of evaluation information. This, in itself, is a contribution because there are currently no other models of this kind. The model of participatory evaluation developed and tested in the current study has established a framework on which a theory of participatory evaluation can be further developed.

REFERENCES

Alutto, J. A., & Belasco, J. A. (1972). A typology for participation in organizational decision-making. Administrative Science Quarterly, 9(1), 27-41.

Ambrosie, F., & Haley, P. W. (1988). The changing school climate and teacher professionalism. NASSP Bulletin, January, 82-89.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. Psychological Bulletin, 103, 411-423.

Bacharach, S. B., Bamberger, P., Conley, S. C., & Bauer, S. (1990). The dimensionality of decision participation in educational organizations: The value of a multi-domain evaluative approach. Educational Administration Quarterly, 26, 126-167.

Bacharach, S. B., & Conley, S. (1986). Education reform: A managerial agenda. Phi Delta Kappan, 67(9), 641-645.

Bandura, A. (1983). Self-efficacy determinants of anticipated fears and calamities. Journal of Personality and social Psychology, 45, 464-469.

Bandura, A. (1986). Social foundations of thought and action: A social cognitive view. Englewood Cliff, NJ: Prentice Hall.

Bandura, A., Reese, L., & Adams, N. E. (1982). Microanalysis of action and fear arousal as a function of differential levels of perceived self-efficacy. Journal of Personality and Social Psychology, 43, 5-21.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology, 51, 173-182.

Barrick, C. B., & Cogliano, J. F. (1993). Stakeholder involvement: Mythology or methodology? Evaluation Practice, 14, 33-37.

Bentler, P. M., & Bonett, D. G. (1980). Significance test of goodness-of-fit in the analysis of covariance structures. Psychological Bulletin, 88(3), 588-606.

Bentler, P. M., & Chou, C. (1987). Practical issues in structural modeling. Sociological Methods and Research, 16, 78-117.

Berlyne, D. E. (1963). Motivational problems raised by exploratory and epistemic behavior. In S. Koch (Ed.), Psychology: A study of science (Vol. 5). New York: McGraw-Hill.

Berlyne, D. E. (1965). Curiosity and education. In J. Krumboltz (Ed.), Learning and the educational process. Chicago: Rand McNally.

Bollen, K. (1986). Sample size and Bentler and Bonett's nonnormed fit index. Psychometrika, 51, 375-377.

Bollen, K. (1989). Structural equation models with latent variables. New York: Wiley.

Bollen, K. (1990). Overall fit in covariance structure models: Two types of sample size effects. Psychological Bulletin, 107(2), 256-259.

Boomsma, A. (1982). The robustness of LISREL against small sample sizes in factor analysis models. In K. G. Joreskog & H. Wold (Eds.), Systems under indirect observation: Causality structure, prediction (Part I). Amsterdam: North-Holland.

Boomsma, A. (1983). On the robustness of LISREL against small sample size and nonnormality. Amsterdam: Sociometric Research Foundation.

Brandon, P. R., Wang, Z., & Heck, R. H. (1994). Teacher involvement in school conducted needs assessments: Issues of decision-making process and validity. Evaluation Review, 18(4), 458-471.

Brunner, I., & Guzman, A. (1989). Participatory evaluation: A tool to assess projects and empower people. In R. F. Conner & M. Hendricks (Eds.), International innovations in evaluation methodology (pp. 9-18). San Francisco: Jossey-Bass.

Campbell, D. J., & Gingrich, K. F. (1986). The interactive effects of task complexity and participation on task performance: A field experiment. Organizational Behavior and Human Decision Processes, 38, 162-180.

Campbell, J. P. (1982). Editorial: Some remarks from the outgoing editor. Journal of Applied Psychology, 67, 691-700.

Carnegie Commission on Teaching as a Profession. (1986). A nation prepared: Teachers for the 21st Century. Hyattsville, MD: Carnegie Forum on Education and the Economy.

Carnegie Foundation for the Advancement of Teaching. (1988). Teacher involvement in decision making: A state-by-state profile. Washington, DC: Author.

Catanzaro, S. (1997, June). Composite variables, parcels, a SEMNET FAQ? [On-line]. Available E-mail: Listserv@ualvm.ua.edu: Subscribe: sement@ualvm.ua.edu

Cattell, R. B. (1966). The meaning and strategic use of factor analysis. In R. B. Cattell (Ed.), Handbook of multivariate experimental psychology. Chicago: Rand McNally.

Chen, H. (1990). Theory-driven evaluations. Newbury Park, CA: Sage.

Cohen, M. (1982). Effective schools: Accumulating research and findings. American Education, 18, 13-16.

Conley, S. (1991). Review of research on teacher participation in school decision making. Review of Research in Education, 17, 225-266.

Conway, J. (1976). Test of linearity between teachers' participation in decision making and their perceptions of their school as organizations. Administrative Science Quarterly, 21(1), 130-139.

Conway, J. A. (1984). The myth, mystery, and mastery of participative decision making in education. Educational Administration Quarterly, 20(3), 11-40.

Cousins, J. B., & Earl, L. M. (1992). The case for participatory evaluation. Educational Evaluation and Policy Analysis, 14(4), 397-418.

Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. Toronto: Harcourt, Brace, Jovanovich College Publishers.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16, 297-334.

Cudeck, R., & Henly, S. J. (1991). Model specification in covariance structures analysis and the "problem" of sample size: A clarification. Psychological Bulletin, 109(3), 512-519.

Cui-Cui, C. (1997, June). Composite variables [On-line]. Available E-mail: Listserv@ua1vm.ua.edu: Subscribe: sement@ua1vm.ua.edu

Dembo, M. H., & Gibson, S. (1984). Teachers' sense of efficacy: An important factor in school improvement. The Elementary School Journal, 86(2), 173-184.

Deutsch, M. (1980). Over fifty years of conflict research. In L. Festinger (Ed.), Four decades of social psychology. New York, NY: Oxford University Press.

Ding, L., Velicer, W. F., & Harlow, L. L. (1995). Effects of estimation methods, number of indicators per factor, and improper solutions on structural equation modeling fit indices. Structural Equation Modeling, 2, 119-143.

Duke, D. L., Showers, B. K., & Imber, M. (1980). Teachers and shared decision making: The costs and benefits of involvement. Educational Administration Quarterly, 16(1), 93-106.

Estler, S. E. (1988). Decision making. In N. J. Boyan (Ed.), Handbook of research on educational administration (pp. 305-319). New York: Longman Publishers.

Erez, M., & Arad, R. (1986). Participative goal-setting: Social, motivational, and cognitive factors. Journal of Applied Psychology, 4, 591-597.

Festinger, L. A. (1957). A theory of cognitive dissonance. Evanston, IL: Row, Peterson.

Fetterman, D. M., Kaftarian, S. J., & Wandersman, A. (1996). Empowerment evaluation: Knowledge and tools for self-assessment and accountability. Thousand Oaks, CA: Sage.

Fiske, S. T., Kenny, D. A., & Taylor, S. E. (1982). Structural models of the mediation of salience effects on attribution. Journal of Experimental Social Psychology, 18, 105-127.

Gable, R. K., & Wolf, J. W. (1993). Instrumental development in the affective domain: Measuring attitudes and values in corporate and school settings (2nd ed.). Boston: Kluwer Academic Publishers.

Garaway, G. B. (1995). Participatory evaluation. Studies in Educational Evaluation, 21, 85-102.

Gist, M. E. (1987). Self-efficacy: Implications for organizational behavior and human resource management. Academy of Management Review, 12, 472-485.

Gist, M. E. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. Academy of Management Review, 17, 183-211.

Goodlad, J. I. (1984). A place called school. New York: McGraw-Hill.

Greene, J. G. (1987). Stakeholder participation in evaluation design: Is it worth the effort? Evaluation and Program Planning, 10, 379-394.

Greene, J. G. (1988a). Communication of results in utilization on participatory program evaluation. Evaluation and Program Planning, 11, 341-351.

Greene, J. G. (1988b). Stakeholder participation and utilization in program evaluation. Evaluation Review, 12(2), 91-116.

Gregorich, D. (1997, June). Composite variables, parcels, a SEMNET FAQ? [On-line]. Available E-mail: Listserv@ualvm.ua.edu: Subscribe: sement@ualvm.ua.edu

Hayduk, L. A. (1987). Structural equation modeling with LISREL. Baltimore: John Hopkins University Press.

Hoelter, J. W. (1983). The analysis of covariance structures: Goodness-of-fit indices. Sociological Methods and Research, 11, 325-344.

Hoyle, R. H. (1995). Structural equation modeling: Concepts, issues, and applications. Thousand Oaks, CA: Sage.

Imber, M., Neidt, W. A., & Reyes, P. (1990). Factors contributing to teacher satisfaction with participative decision making. Journal of Research and Development in Education, 23(4), 216-225.

Ivancevich, J. M. (1977). Different goal setting treatments and their effects on performance and job satisfaction. Academy of Management Journal, 20, 406-419.

James, L. R., & Brett, J. M. (1984). Mediators, moderators, and tests for mediation. Journal of Applied Psychology, 69(2), 307-321.

Janis, I. L. (1972). Victims of groupthink. Boston: Houghton Mifflin.

Joint Committee on Standards for Educational Evaluation. (1994). The program evaluation standards: How to assess evaluations of educational programs (2nd ed.). Thousand Oaks, CA: Sage.

Joreskog, K. G., & Sorbom, D. (1988). LISREL 7: A guide to the program and applications (2nd ed.). Chicago: SPSS Inc.

Joreskog, K. G., & Sorbom, D. (1993). LISREL8 user's reference guide. Chicago: Scientific Software International Inc.

Latham, G. P., & Saari, L. M. (1979). Importance of supportive relationships in goal setting. Journal of Applied Psychology, 64, 151-156.

Latham, G. P., Winters, D. C., & Locke, E. A. (1994). Cognitive and motivational effects of participation: A mediator study. Journal of Organizational Behavior, 15, 49-63.

Leithwood, K. (1992). The move toward transformational leadership. Educational Leadership, 49(5), 8-12.

Lim, S. (1989). The revision of elementary school assessment in British Columbia. Research Forum, Winter, 38-40.

Little, R. J. (1987). Statistical analysis with missing data. Toronto: Wiley & sons.

Locke, E. A., & Latham, G. P. (1990). A theory of goal setting and task performance. Englewood Cliff, NJ: Prentice Hall.

Locke, E. A., & Schweiger, D. M. (1979). Participation in decision-making: One last look. Research in Organizational Behavior, 1, 265-339.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. Biometrika, 57, 519-530.

Mark, M. M., & Shortland, R. L. (1985). Stakeholder-based evaluation and value judgments. Evaluation Review, 9, 605-626.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indices in confirmatory factor analysis: The effect of sample size. Psychological Bulletin, 103(3), 391-410.

McDonald, R. P., & Marsh, H. W. (1990). Choosing a multivariate model: Noncentrality and goodness-of-fit. Psychological Bulletin, 107(2), 247-255.

Miller, D. I., Ismail, S., Giesen, J. M., Adams-Price, C., & Topping, J. S. (1993). Ecological dissonance in decision-making participation systems as a predictor of job satisfaction, involvement, alienation, and formalization. Bulletin of the Psychometric Society, 31(2), 146-148.

Miller, K. I., & Monge, P. R. (1986). Participation, satisfaction, and productivity: A meta-analytic review. Academy of Management Journal, 29, 727-753.

Ministry of Education. (1995). British Columbia school accreditation manual. Ministry of Education, Province of British Columbia.

Ministry of Education. (1995). British Columbia school accreditation: Resource guide for external teams. Ministry of Education, Province of British Columbia.

Ministry of Education, Skills and Training. (1996). British Columbia school accreditation: Guide for schools. Ministry of Education, Skills and Training, Province of British Columbia.

Ministry of Finance and Corporate Relations. (1997). Report on accountability in the K to 12 education system. Ministry of Finance and Corporate Relations, Province of British Columbia.

Mulaik, S. A., James, L. R., Van Alstine, J. V., Bennett, N., Lind, S., & Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. Psychological Bulletin, 105(3), 430-445.

Muliak, S. (1997, June). Composite variables, parcels, a SEMNET FAQ? [On-line]. Available E-mail: Listserv@ualvm.ua.edu: Subscribe: sement@ualvm.ua.edu

Newman, D. L., & Cai, M. (1995). Teacher involvement in program evaluation and evaluation use: An empirical examination of the participatory evaluation approach. A paper presented at the Annual Meeting of the American Evaluation Association in conjunction with the International Evaluation Conference, November, 1995, Vancouver, B.C., Canada.

Owen, S. V., & Froman, R. D. (1992). Academic self-efficacy in at-risk elementary students. Journal of Research in Education, 2, 3-7.

Papineau, D., & Kiely, M. C. (1996). Participatory evaluation in a community organization: Fostering stakeholder empowerment and utilization. Evaluation and Program Planning, 19, 79-93.

Patton, M. Q. (1986). Utilization-focused evaluation. Newbury Park, CA: Sage.

Pitts, S. C., West, S. G., & Tein, J. (1996). Longitudinal measurement model in evaluation research: Examining stability and change. Evaluation and Program Planning, 19(4), 333-350.

Purkey, S. C., & Smith, M. S. (1983). Effective schools: A review. The Elementary School Journal, 83(4), 427-452.

Ramirez, N. A. (1985). Development of an evaluation feedback process and an evaluation utilization assessment instrument [CD-ROM]. Abstract from: ProQuest File: Dissertation Abstracts Item: 8522749

Raudenbush, S. W., Rowan, B., & Cheong, Y. K. (1992). Contextual effects on the self-perceived efficacy of high school teachers. Sociology of Education, 65, 150-167.

Rahim, M. A., & Magner, N. R. (1995). Confirmatory factor analysis of the styles of handling interpersonal conflict: First-order model and its invariance across groups. Journal of Applied Psychology, 80(1), 122-132.

Rinne, C. (1993). The impact of anxiety as a mediating variable on health educators' utilization of evaluation results [CD-ROM]. Abstract from: ProQuest File: Dissertation Abstracts Item: 9334464

Rosenholtz, S. J. (1985). Effective schools: Interpreting the evidence. American Journal of Education, 93(3), 352-388.

Russell, D. W. (1997, June). Items parcels [On-line]. Available E-mail: Listserv@ualvm.ua.edu: Subscribe: sement@ualvm.ua.edu

Schumacker, R. E., & Lomax, R. G. (1996). A beginner's guide to structural equation modeling. New Jersey: Lawrence Erlbaum.

Schweiger, D. M., & Leane, C. R. (1986). Participation in decision-making. In E. A. Locke (Ed.), Generalizing from laboratory to field setting. Lexington, MA: Lexington Books.

Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), Perspective of curriculum evaluation (pp. 39-83). Chicago: Rand McNally.

Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). Foundations of program evaluation: Theories of practice. Newbury Park, CA: Sage.

Shalley, C. E., Oldham, G. R., & Porac, J. F. (1987). Effects of goal difficulty, goal setting method, and expected external evaluation on intrinsic motivation. Academy of Management Journal, 30, 553-563.

Smylie, M. A., Lazarus, V., & Brownlee-Conyers, J. (1996). Instructional outcomes of school-based participative decision making. Educational Evaluation and Policy Analysis, 18(3), 181-198.

Stein, M. K., & Wang, M. C. (1988). Teacher development and school improvement: The process of teacher change. Teaching and Teacher Education, 4(2), 171-187.

Stelzl, I. (1986). Changing causal relationships without changing the fit: Some rules for generating equivalent LISREL-models. Multivariate Behavioral Research, 21, 309-331.

Tabachnick, B. G., & Fidell, L. S. (1989). Using multivariate statistics. Northridge, CA: HarperCollins.

Tanaka, J. S. (1987). How big is big enough? Sample size and goodness-of-fit in structural equation models with latent variables. Child Development, 58, 134-146.

Taylor, D. L., & Bogotch, I. E. (1994). School-level effects of teachers' participation in decision making. Educational Evaluation and Policy Analysis, 16(3), 302-319.

Tjosvold, D. (1985). Dynamics of participation: An experimental investigation. Group and Organization Studies, 10(3), 260-277.

Tjosvold, D. (1986). Organizational test of goal linkage theory. Human Relations, 39, 517-528.

Tjosvold, D. (1987). Participation: A close look at its dynamics. Journal of Management, 13, 731-750.

Tjosvold, D., Andrews, & Jones, (1983). Cooperative and competitive relationships between leaders and subordinates. Human Relations, 36(12), 1111-1124.

Tjosvold, D., & Deemer, D. K. (1980). Effects of controversy within a cooperative or competitive context on organizational decision making. Journal of Applied Psychology, 65(5), 590-595.

Tjosvold, D., & Johnson, D. W. (1978). Controversy within a cooperative or competitive context and cognitive perspective taking. Contemporary Educational Psychology, 3(4), 376-386.

Tjosvold, D., & McNeely, L. T. (1988). Innovation through communication in an educational bureaucracy. Communication Research, 15(5), 586-581.

Tjosvold, D., Wedley, W. C., & Field, R. H. G. (1986). Constructive controversy, the Vroom-Yetton model, and managerial decision-making. Journal of Occupational Behaviour, 7, 125-138.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. Psychometrika, 38, 1-10.

Vanderslice, V. J., Rice, R. W., & Julian, J. W. (1987). The effects of participation in decision-making on worker satisfaction and productivity: An organizational simulation. Journal of Applied Social Psychology, 17, 158-170.

Wagner, J. A. & Gooding, R. Z. (1987a). Effects of societal trends on participation research. Administrative Science Quarterly, 32, 241-262.

Wagner, J. A., & Gooding, R. Z. (1987b). Shared influence and organizational behavior: A meta-analysis of situational variables expected to moderate participation-outcome relationships. Academy of Management Journal, 30, 524-541.

Weiss, C. H. (1983). The stakeholder approach to evaluation: Origins and promise. In A. S. Bryk (Ed.), Stakeholder-based evaluation: New directions for program evaluation (Vol. 17, pp. 3-14). San Francisco, CA: Jossey-Bass.

Weiss, C. H., Cambone, J., & Wyeth, A. (1992). Trouble in paradise: Teacher conflicts in shared decision making. Educational Administration Quarterly, 28(3), 350-367.

Wholey, J. S. (1981). Using evaluation to improve program performance. In R. A. Levine, M. A. Solomon, G. M. Hellstern, & H. Wollman (Eds.), Evaluation research and practice: Comparative and international perspectives (pp. 92-106). Beverly Hills, CA: Sage.

Appendix A

Letters

Letter to principals sent prior to receiving questionnaire package

I NEED YOUR FEEDBACK! I'm a former teacher in school district #42 who has gone back to school to complete a Ph.D. in Educational Psychology at UBC. My study involves the B.C. School Accreditation Program. I'm interested in determining whether teachers perceive the accreditation program to be a learning experience, and whether they use what they learn to make changes to their school. Participation in the study would require that your teachers' complete a series of questionnaires. The total time commitment for each of the three sets of questionnaires would be less than 10 minutes.
I'm sending this letter to you after having contacted your district office. I understand that it is ultimately your decision whether your school participates in any type of survey study involving teachers. There is a package of questionnaires being sent to you, which I hope you will distribute to your staff. I believe that they will find the questions relevant and thought provoking.

Your schools' participation is very important! All the accreditation schools across the province are being asked to participate. The identity of teachers who choose to participate will be strictly confidential. You should receive the questionnaires in a few days. There will be additional information about the study included with the questionnaires, but please feel free to fax or call if you have any questions.

Letter sent with principal questionnaire package

Your questionnaires have arrived. As I mentioned in my previous letter, additional information pertaining to the project is attached to the teacher questionnaire booklets. For analysis reasons, it is better if there is a time lag between the completion of the questionnaires. Therefore, each questionnaire booklet has been divided into three sections. The instructions in the booklets ask that the teachers leave about a week between completion of each section.

Teacher identity will be kept strictly confidential. The codes will be cut off the questionnaires as soon as the data is entered.

It would be greatly appreciated if you could collect the questionnaires in three weeks and return them in the postage-paid envelope included in this package.

I appreciate your time and consideration. I hope that you distribute the questionnaires and give your staff the opportunity to participate.

If you have any further questions, I would be happy to talk with you. Or, if you're in the lower mainland, I could visit your school. If you would like a copy of the study, just let me know and I'll gladly send it.

Appendix B

Questionnaires

Level of Participation in Decision Making

Steps involved in the accreditation process are listed below. Beside each step is a scale to indicate the amount you **actually contributed to** the decision making processes and the degree to which you would have **preferred** to have contributed. The following numbers represent varying amounts of contribution:

1	2	3	4
very infrequent contributions	occasional contributions	regular contributions	very frequent contributions

Circle the number that best represents your **actual** contribution to the decision making activities and the degree to which you would have **preferred** to contribute for each of the statements.

	Actual Contribution	**Preferred Contribution**
1. deciding who would be the self-assessment chair	1 2 3 4	1 2 3 4
2. deciding who would be the sub-committee leaders	1 2 3 4	1 2 3 4
3. deciding how final decisions would be made (i.e., voting, Delphi method etc.)	1 2 3 4	1 2 3 4
4. deciding when meetings would be held	1 2 3 4	1 2 3 4
5. deciding how meetings would be conducted	1 2 3 4	1 2 3 4
6. deciding how long meetings would be	1 2 3 4	1 2 3 4
7. clarifying criteria statements	1 2 3 4	1 2 3 4
8. gathering evidence	1 2 3 4	1 2 3 4
9. determining satisfaction levels	1 2 3 4	1 2 3 4
10. adding internal comments	1 2 3 4	1 2 3 4
11. categorizing school priorities	1 2 3 4	1 2 3 4
12. building school growth plan	1 2 3 4	1 2 3 4

Influence in Participative Decision Making

Each statement below refers to a decision making step in the accreditation process. Please indicate the degree to which you agree or disagree with each statement below by circling the appropriate number to the right of each statement.

	Strongly disagree	Moderately disagree	Disagree slightly	Agree slightly	Moderately agree	Strongly agree
1. I had influence on deciding who was chosen to be the self-assessment chair	1	2	3	4	5	6
2. I had influence on deciding who the subcommittee leaders would be	1	2	3	4	5	6
3. I had influence on deciding how final decisions would be made (i.e., voting, Delphi method etc.)	1	2	3	4	5	6
4. I had influence on deciding when meetings would be held	1	2	3	4	5	6
5. I had influence on deciding how meetings would be conducted	1	2	3	4	5	6
6. I had influence on deciding how long meetings would be	1	2	3	4	5	6
7. I had influence on decisions about clarifying the criteria statements	1	2	3	4	5	6
8. I had influence on decisions about gathering evidence	1	2	3	4	5	6
9. I had influence on decisions regarding satisfaction levels	1	2	3	4	5	6

10. I had influence on decisions regarding the internal comments	1	2	3	4	5	6
11. I had influence on decisions about categorizing school priorities	1	2	3	4	5	6
12. I had influence on decisions about building our school growth plan	1	2	3	4	5	6

Goal Interdependence

(Tjosvold, Andrews, & Jones, 1983)

Please indicate the degree to which you agree or disagree with each statement below by circling the appropriate numeral to the right of each statement.

	Never	Seldom	Sometimes	Usually	Always
My colleagues:					
seem pleased when I succeed.	1	2	3	4	5
like to show that they know more than I do.	1	2	3	4	5
do not know what I want to accomplish.	1	2	3	4	5
work best when they work alone rather than with me.	1	2	3	4	5
show as much concern for what I want to accomplish as to what they want.	1	2	3	4	5
seem to be threatened when I learn new skills and knowledge.	1	2	3	4	5
and I work separately.	1	2	3	4	5
help me find ways to achieve my objectives.	1	2	3	4	5
seem to get in the way of my growth and development.	1	2	3	4	5
prefer to work alone rather than with me.	1	2	3	4	5

give high priority to the things I want to accomplish.	1	2	3	4	5
are unconcerned about whether I get ahead in the school organization.	1	2	3	4	5
like to demonstrate their superiority.	1	2	3	4	5
My colleagues:					
like to get their rewards through their own individual work.	1	2	3	4	5
take pride in my accomplishments.	1	2	3	4	5
show much more concern for what they want to accomplish than for what I want to accomplish.	1	2	3	4	5
structure things so that their goals and my goals can be achieved.	1	2	3	4	5
are disturbed by my accomplishments.	1	2	3	4	5
are uninterested in the things I want to accomplish.	1	2	3	4	5
are interested in things that I want to accomplish.	1	2	3	4	5
structure things that favor their goals rather than mine.	1	2	3	4	5
are most concerned about what they accomplish when working by themselves.	1	2	3	4	5
give high priority to the things they want to accomplish and low priority to the things I want to accomplish.	1	2	3	4	5

Constructive Controversy

(Tjosvold, Wedley, & Field, 1986)

Please indicate the degree to which you agree or disagree with each statement below by circling the appropriate numeral to the right of each statement. Please base your responses on how your school staff generally interacts when it solves problems and makes decisions.

	Never	Seldom	Sometimes	Usually	Always
1. Our staff has a "we are in it together" attitude	1	2	3	4	5
2. Staff members express their own views fully	1	2	3	4	5
3. Staff members feel understood and accepted by each other	1	2	3	4	5
4. We all influence each other	1	2	3	4	5
5. We first try to understand the problem fully	1	2	3	4	5
6. Staff members try to win by pushing and keeping their original views	1	2	3	4	5
7. All staff members' views are listened to, even if they are in the minority	1	2	3	4	5
8. Disagreeing with another staff member's idea is not a rejection of that staff member	1	2	3	4	5
9. Staff members try to control each other	1	2	3	4	5
10. The staff understands the problem before we seek a solution	1	2	3	4	5
11. We seek a solution favorable and acceptable to all staff members	1	2	3	4	5

12. Opposing views aid in the full consideration of the issues	1	2	3	4	5
13. We try to blame each other for problems	1	2	3	4	5
14. There is a lot of give and take discussion	1	2	3	4	5
15. All ideas are expressed before we begin to evaluate them	1	2	3	4	5

Process Outcomes

Please indicate the degree to which you agree or disagree with each statement about the **accreditation process** below by circling the appropriate numeral to the right.

	Strongly disagree	Moderately disagree	Disagree slightly	Agree slightly	Moderately agree	Strongly agree
1. Through the accreditation process I've learned about problems in my school that I wasn't aware of.	1	2	3	4	5	6
2. I learned **new** things about my school because of the accreditation program.	1	2	3	4	5	6
3. The accreditation process has made me think about what's important for my school.	1	2	3	4	5	6
4. Because of the accreditation program, I learned **useful** things about my school.	1	2	3	4	5	6
5. The accreditation program did **not** teach me anything **new** about my school.	1	2	3	4	5	6
6. The accreditation process gave me a better understanding of the strengths of our school.	1	2	3	4	5	6
7. The accreditation activities provided me with the necessary **knowledge** to contribute to the development of our school growth plan.	1	2	3	4	5	6

8. The accreditation provided me with a better overall understanding of our school practices.	1	2	3	4	5	6
9. The steps in the accreditation were **not useful** in giving me ideas for the development of our school growth plan.	1	2	3	4	5	6
10. The accreditation increased my understanding of my school's role in the community.	1	2	3	4	5	6
11. The accreditation gave me a better understanding of the weaknesses of our school.	1	2	3	4	5	6
12. I did not learn anything **useful** about my school as a result of going through the accreditation program.	1	2	3	4	5	6
13. The accreditation gave me a better understanding of other staff members' needs.	1	2	3	4	5	6
14. The accreditation provided no new details about our school.	1	2	3	4	5	6
15. The accreditation gave me a better understanding of the points of view of other staff members.	1	2	3	4	5	6
16. The accreditation activities gave me **ideas** for developing the various components of the school growth plan.	1	2	3	4	5	6

Use of Evaluation Information

During the **accreditation process**, your staff gathered various types of information about your school. This questionnaire addresses your perceptions of this information.

Please indicate the degree to which you agree or disagree with each statement below by circling the appropriate numeral to the right of each statement.

The accreditation information:	Strongly disagree	Moderately disagree	Disagree slightly	Agree slightly	Moderately agree	Strongly agree
1. raised discussion about overall school achievement.	1	2	3	4	5	6
2. is useful for persuading parents to support what we are doing in our school.	1	2	3	4	5	6
3. has made me a better advocate of my school.	1	2	3	4	5	6
4. resulted in a feeling of satisfaction with overall school achievement.	1	2	3	4	5	6
5. convinced staff members of need for maintaining existing school practices.	1	2	3	4	5	6
6. is useful in gaining community support for what we are doing in our school.	1	2	3	4	5	6
7. resulted in no action to make changes in our school practices.	1	2	3	4	5	6
8. resulted in some existing school practices being terminated.	1	2	3	4	5	6

9. confirmed the effectiveness of the existing school practices.	1	2	3	4	5	6
10. resulted in modifications of existing school practices.	1	2	3	4	5	6
11. resulted in suggestions for how to improve the school.	1	2	3	4	5	6
12. was used as a starting point for further evaluation of specific school practices.	1	2	3	4	5	6
13. resulted in implementation of new ideas for how to improve school practices.	1	2	3	4	5	6
14. resulted in new school policy being implemented.	1	2	3	4	5	6

Appendix C

Percentage of Missing Data

Percentage of Missing Data

Observed Variable	Number of Observations Missing	Percent of Observations Missing
Y1	7	.02
Y2	5	.01
Y3	7	.02
Y4	6	.01
Y5	10	.03
Y6	2	.006
Y7	3	.03
Y8	2	.006
Y9	25	.08
Y10	13	.04
Y11	8	.03
Y12	3	.01
Y13	6	.01
Y14	3	.01
Y15	8	.03
Y16	7	.02
Y17	7	.02
Y18	11	.03
Y19	9	.03
Y20	10	.03
Y21	12	.03
Y22	12	.03
Y23	15	.04

N = 315

Appendix D

Tests for Non-Random Pattern of Missing Data: Tests of Mean Differences Across Groups

Tests for Non-Random Pattern of Missing Data: Tests of Mean Differences Across Groups

T-tests for Equality of Means

Variable	n	Mean	SD	t-value	df	2-tail Sig
Y13				-1.32	313	.19
Grp 1	290	4.75	1.25			
Grp 2	25	5.03	.97			
Y16				-1.50	313	.14
Grp 1	290	4.39	1.36			
Grp 2	25	4.75	1.12			
Y20				-.82	313	.41
Grp 1	290	4.42	1.18			
Grp 2	25	4.58	.99			
Y23				-.06	313	.95
Grp 1	290	4.33	1.14			
Grp 2	25	4.34	1.19			

Appendix E

Description of Marker Variables

Description of Marker Variables

Variable	Questionnaire Title	Item #	Question/Statement
Y9	Level of PDM	2	Actual and preferred participation in deciding who would be the sub-committee leaders.
Y11	Level of PDM	3	Actual and preferred participation in deciding how final decisions would be made (i.e., voting, Delphi method etc.).
Y10	Level of PDM	5	Actual and preferred participation in deciding how meetings would be conducted.
Y12	Influence in PDM	7	I had influence on decision about clarifying criteria statements.
Y13	Influence in PDM	9	I had influence on decisions regarding satisfaction levels.
Y14	Influence in PDM	11	I had influence on decisions about categorizing school priorities.
Y1	Cooperative Goal Interdependence	5	My colleagues show as much concern for what I want to accomplish as to what they want.
Y2	Cooperative Goal Interdependence	8	My colleagues help me find ways to achieve my objectives.
Y3	Cooperative Goal Interdependence	11	My colleagues give high priority to the things I want to accomplish.
Y4	Autonomous Goal Interdependence	4	My colleagues work best when they work alone rather than with me.
Y5	Autonomous Goal Interdependence	22	My colleagues are most concerned about what they accomplish when working by themselves.
Y6	Constructive Controversy	3	Staff members feel understood and accepted by each other.
Y7	Constructive Controversy	4	We all influence each other.
Y8	Constructive Controversy	11	We seek a solution favorable and acceptable to all staff members.
Y15	Process Outcomes	4	Because of the accreditation, I learned useful things about my school.
Y16	Process Outcomes	8	The accreditation activities provided me with the necessary knowledge to contribute to the development of our school growth plan.

Variable	Questionnaire Title	Item #	Question/Statement
Y17	Process Outcomes	16	The accreditation gave me a better understanding of the points of view of other staff members.
Y18	Symbolic Use of Evaluation Information and Process Outcomes	3	The accreditation information has made me a better advocate of my school.
Y19	Symbolic Use of Evaluation Information	4	The accreditation information resulted in a feeling of satisfaction with overall school achievement.
Y20	Symbolic Use of Evaluation Information	5	The accreditation information convinced staff members of need for maintaining existing school practices.
Y23	Instrumental Use of Evaluation Information	10	The accreditation information resulted in modifications of existing school practices.
Y22	Instrumental Use of Evaluation Information	12	The accreditation information was used as a starting point for further evaluation of specific school practices.
Y21	Symbolic and Instrumental Use of Evaluation Information	13	The accreditation information resulted in implementation of new ideas for how to improve school practices.

Appendix F

LISREL Program Code for Invariance Tests

LISREL Program Code for Invariance Tests

The following command file was adjusted for each test of invariance. The "MO" line for the second group was adjusted as follows:

```
LISREL
  /"RANDOM SPLIT"
  /DA NG=2 NI=23 NO=157
  /MO NY=23  NE=8  NK=1  C
    LY=FU,FI  BE=FU,FI C
    GA=FU,FI  PH=ST  PS=DI,FR  TE=DI,FR
  /FR LY(2,1) LY(3,1) LY(20,1)
  /FR LY(5,2)
  /FR LY(7,3) LY(8,3)
  /FR LY(10,4) LY(11,4)
  /FR LY(13,5) LY(14,5)
  /FR LY(16,6) LY(17,6) LY(18,6)
  /FR LY(18,7) LY(20,7) LY(21,7)
  /FR LY(21,8) LY(23,8)
  /FR BE(5,4) BE(6,4) BE(6,5) BE(7,6) BE(8,6)
  /FR GA(1,1) GA(2,1) GA(3,1)
  /FR GA(4,1) GA(5,1) GA(7,1) GA(8,1)
  /VA 1.0 LY(1,1) LY(4,2) LY(6,3) LY(9,4) LY(12,5)  C
      LY(15,6) LY(19,7) LY(22,8)
  /ST 1 ALL
  /OU  AD=OFF NS
  /"TWO"
  /DA NO=157
  /MO
  /ST 1 ALL
  /OU AD=OFF NS.
```

Invariance Test	Adjustment to MO in second group
1. Equal factor pattern	MO LY = PS BE = PS GA = PS
2. Equal factor pattern and loadings	MO LY = IN BE = PS GA = PS
3. Equal factor pattern, loadings, and errors	MO LY = IN TE = IN BE = PS GA = PS
4. Equal factor pattern, loadings, errors, and variances/covariances	MO LY = IN TE = IN PH = IN BE = IN GA = IN

The model tested in the current study was a second-order model with only one KSI construct. Subsequently, because there was only one value in the PHI matrix, there was no difference between invariance test numbers three and four. Also, it was necessary to constrain patterns for Beta and Gamma matrices in order to identify the model. Constraining the patterns for the Beta and Gamma matrices created additional constraints to all of the tests, which resulted in different results than had they been omitted. The following table shows the difference between the tests of invariance with and without Beta and Gamma constrained to BE=PS and GA=PS.

Group	χ^2 without BE=PS, GA=PS	χ^2 with BE=PS, GA=PS	*p*-value
(a) Elementary (n = 173) (b) Secondary (n =110)			
Equal factor pattern	774.73(*df*=442)	571.16(*df*=430)	.000
Equal factor pattern and loadings	813.45(*df*=460)	612.10(*df*=448)	.000
Equal factor pattern, loadings, and errors	882.72(*df*=483)	684.97(*df*=471)	.000
Equal factor pattern, loadings, errors, and variances/covariances	882.72(*df*=483)	684.97(*df*=471)	.000
(a) One-Day (n = 102) (b) One-Week (n = 213)			
Equal factor pattern	825.22(*df*=442)	579.98(*df*=442)	.000
Equal factor pattern and loadings	852.66(*df*=460)	595.49(*df*=460)	.000
Equal factor pattern, loadings, and errors	888.04(*df*=483)	632.39(*df*=483)	.000
Equal factor pattern, loadings, errors, and variances/covariances	888.04(*df*=483)	632.39(*df*=483)	.000