UCLA
UCLA Electronic Theses and Dissertations

Title

Tracing the Pathways of Evaluation Influence: From Theory to Practice

Permalink

https://escholarship.org/uc/item/101864tw

Author

Banuelos, Noelle Vargas

Publication Date

2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Tracing the Pathways of Evaluation Influence:

From Theory to Practice

A dissertation submitted in partial satisfaction of the

requirements for the degree of Doctor of Philosophy

in Education

by

Noelle Vargas Banuelos

2012

ABSTRACT OF THE DISSERTATION

Tracing the Pathways of Evaluation Influence:

From Theory to Practice

by

Noelle Vargas Banuelos

Doctor of Philosophy in Education

University of California, Los Angeles, 2012

Professor Christina Christie, Chair

This study sought to explore the ways that evaluation influenced changes in attitudes and actions at the individual and interpersonal levels and the pathways through which this influence occurred, all in a real-world evaluation context. Specifically, Mark & Henry's (2004) proposed model of evaluation influence was used as the framework onto which various components of a completed evaluation were mapped. This was accomplished through mixed-method analyses of secondary data from a federally-funded school district evaluation of Small Learning Communities (SLCs). Survey data yielded potential evaluation outcomes which included indicators of behavioral and cognitive change as a result of evaluation findings and hypothesized pathways leading from one type of outcome to another. A series of logistic regressions were used to predict the likelihood of outcome occurrence. Predictors included characteristics such as school assignment and affiliation, measures of organizational context, and awareness. Qualitative analyses

of interview and focus group data supported the quantitative findings and provided more nuanced insight into the emergent outcomes and pathways.

Ultimately, the findings showed that the greater an individuals' awareness of the evaluation's findings and the greater their understanding of the SLC program's vision, purpose, and goals as well as their own role and comfort in how their specific part of the program had developed, the greater the odds of both behavioral and cognitive change as a result of those evaluation findings. These outcomes, or indicators of evaluation influence, took place largely at individual and interpersonal levels. Furthermore, the presence of some evaluation outcomes increased the odds of other outcomes occurring, providing supporting evidence for Mark & Henry's hypothesized pathways of evaluation influence.

The findings generated from this study contribute to research on evaluation in the areas of use and the broader conception of evaluation influence by bridging the gap between theory and practice though systematic study. This expanded conception of evaluation consequences and influence will likely inform practitioners, guide evaluation practice, and promote a more focused awareness from the user perspective.

The dissertation of Noelle Vargas Banuelos is approved.

Marvin C. Alkin

Todd Franke

Robert Cooper

Christina Christie, Chair

University of California, Los Angeles

2012

DEDICATION

I dedicate this work to my daughter who never fails to inspire me through her own persistence and determination, to my husband who beams with love and pride at my accomplishments, and to my mother who, in spirit, smiles down at me from above and brightens my world.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

It is not possible to officially acknowledge all of the many people who supported me academically, emotionally, and at times, financially in completing this work. I could not have accomplished all that I have without them. I would like to formally recognize Professor Marvin C. Alkin who has gone above and beyond his role as my advisor and mentor, often with paternal inclinations, and to whom I am eternally indebted. I am proud to be one of his students, now and in the future, as I look forward to our continuing relationship. I would also like to acknowledge Professor Christina Christie for providing both the guidance and the encouragement I needed during this process. I am privileged to call her colleague and friend.

VITA

1987	B.A., Psychology New York University New York, New York
1989-2001	*Classroom Teacher* New York City Public Schools Bronx, New York
	ATID Academy West Covina, California
	Northwest College Pomona, California
	El Rancho Unified School District Pico Rivera, California
2001	M.A., Educational Psychology California State University, Long Beach Long Beach, California
2001-2006	*Educational Research Analyst* Los Angeles Unified School District Los Angeles, California
	Action Learning Systems Monrovia, California
2006-2011	*Research/Evaluation Associate* National Center for Research, Evaluation, Standards, & Student Testing University of California, Los Angeles Los Angeles, California
	SRM Evaluation Group University of California, Los Angeles Los Angeles, California
2011-2012	Program Results Consultant Los Angeles County Office of Education Santa Fe Springs, California

PUBLICATIONS AND PRESENTATIONS

Rivera, N. V., Vo, A. T., & Lee, J. (2009, November). *Organizational structures and attitudes toward evaluation: An examination of contextual factors that shape evaluation*. Presented at the 24th Annual Conference of the American Evaluation Association, Orlando, FL.

Rivera, N. V. (2007, November). *Challenges in evaluating adult education programs: How theory can help fill in the gaps and connect the dots.* Presented at the 22nd Annual Conference of the American Evaluation Association, Baltimore, MD.

Sass, J. S. & Rivera, N. V. (2005, October). *Using qualitative methods to study schools as organizations: Techniques from evaluations of charter schools and school-based professional development.* Paper presented at the Annual Conference of the American Evaluation Association, Toronto, CAN.

Ai, X., & Rivera, N. V. (2005). Linking ideas to practice: Effectiveness of coaching upon teacher practice. *Research on alternative and non-traditional education: Teacher education yearbook XIII*. (Dangel, J. R. & Guyton, E. M, Eds.) Maryland: Association of Teacher Educators.

Rivera, N. V. (2004, November). *Studying the impact of organizational context on school-based professional development.* Presented at the 19th Annual Conference of the American Evaluation Association, Atlanta, GA.

Rivera, N. V. (2003, November). *Broadening the scope of outcomes in the study of Adult ESL programs.* Paper presented at the 82nd Annual Conference of the California Educational Research Association, San Francisco, CA.

Ai, X., & Rivera, N. V. (2003, April). *Linking ideas to practice: Effectiveness of coaching upon teacher practice.* Paper presented at the 84th Annual Meeting of the American Educational Research Association, Chicago, IL.

Rivera, N. V. (2003, April). *Exploring the effectiveness of a teacher training program through the lens of adult learning principles.* Paper presented at the 84th Annual Meeting of the American Educational Research Association, Chicago, IL.

Sass, J. S. & Rivera, N. V. (2002, November). *Coaching as school-based professional development: A first-year evaluation report*. Paper presented at the 17th Annual Conference of the American Evaluation Association, Washington, DC.

Rivera, N. V. (2002, November). *The role of adult learning principles in professional development evaluation.* Poster session presented at the 17th Annual Conference of the American Evaluation Association, Washington,

CHAPTER 1

Introduction

There has been a longstanding interest and increasing number of calls for research on evaluation (e.g., Mark, 2008; Miller, 2010). These calls have prompted inquiries in two prominent areas – bridging the gap between theory and practice and the impact of evaluation (and research) on policy and practice. There are significant gaps in the literature despite noted benefits to an increased evidence base about evaluation. Furthermore, the evaluation research literature offers little in the way of empirical studies examining how evaluation is used, the influence evaluation may have on potential users, or the consequences of evaluation. Mark (2008) pointed out that even though research on evaluation and its value is debated among evaluators, a larger evidence base may provide guidance for evaluation practitioners, allow for documentation and increased understanding of evaluation's contributions, motivate efforts to improve practice, and facilitate appropriate claims about what evaluation can actually do.

Various efforts at promoting research on evaluation have included the presentation of frameworks and typologies to facilitate study, the formation of topical interest groups in evaluation organizations, and RFP's. For example, the William T. Grant Foundation in 2010 sought to fund studies that would contribute to the understanding of what affected policymakers' and practitioners' acquisition, interpretation, and use of research evidence. Their premise that little attention was devoted to understanding the "user side" added a facet to this type of research that previously may not have had much attention in the research literature. Miller (2010) suggested that the lack of "well-developed frameworks" and guidance for

studying critical aspects of theory to practice may have contributed to the limited research base. Miller developed a preliminary framework that included five criteria for embarking on research in this area: operational specificity, range of application, feasibility in practice, discernable impact, and reproducibility. Mark (2008) offered a taxonomy to organize and focus subjects of inquiry in research on evaluation.

Evaluation use has been one of the most studied aspects of evaluation practice and prevalent in the evaluation research literature. To put it simply, use may be described as behavioral and/or cognitive actions that result from an evaluation, be it findings or process. Much of the research has focused on what facilitates use (i.e., human and context factors (Alkin & Taut, 2003), stakeholder participation (Cousins, 2003), and focus on primary intended users (Patton, 1997)). Research on explicit evaluation use such as the degree of use and the specific changes that may occur are much less prevalent in the literature. Building on Kirkhart's argument (2000) that evaluation use is too narrow a conception and a broader notion of evaluation "influence" should be considered, Mark & Henry (2004) posited that evaluation influence may travel through pathways leading to change, or the consequences of evaluation. Their model expands on traditional definitions of use (i.e., conceptual, instrumental) by drawing on social and psychological theories and research. Studying the pathways of the influence an evaluation may contribute to evaluation research in that it has the potential to identify and understand how and why evaluation works to bring about changes (behavioral, cognitive, and motivational) at various levels (individual, interpersonal, and collective).

Purpose of This Study

This study sought to explore the ways that evaluation influenced changes in attitudes and actions at the individual and interpersonal levels and the pathways through which this influence occurred, all in a real-world evaluation context. Specifically, Mark & Henry's (2004) proposed model of evaluation influence was used as the framework onto which various components of a school district evaluation were mapped. The goal was to "test" this model of evaluation influence using data from a five-year federally-funded evaluation on Small Learning Communities (SLCs) in a large urban school district. This evaluation was selected because of its user-oriented approach to the design and the evaluators' secondary interest in stakeholder awareness and use of findings. It was my informed belief that evaluation designs which are predominantly user-oriented, utilization-focused, collaborative, and/or other inclusive of various forms of participatory evaluation would serve as good contexts for studying the pathways of evaluation influence. Stakeholders in such approaches and/or designs are more likely to be aware of, have knowledge of, and even participate to varying degrees in the evaluation process. The participatory involvement of stakeholders has also been associated with increased evaluation use (Cousins, 2003). In this particular SLC evaluation, the emphasis on use and the collaborative efforts employed by the evaluation team provided data on stakeholder awareness, knowledge, and use of findings that correspond to the proposed model of evaluation influence and the various forms that influence may take. Had the SLC evaluation and its team members not taken this approach or had this approach failed, it is likely that the pool of stakeholders

who were knowledgeable and aware of the evaluation would be much smaller, adversely affecting any collectible evidence of evaluation influence.

The findings generated from this study contribute to research on evaluation in the areas of use and the broader conception of evaluation influence by bridging the gap between theory and practice though systematic study. This expanded conception of evaluation consequences and influence will likely inform practitioners, guide evaluation practice, and promote a more focused awareness from the user perspective.

Research Questions

This study addressed the concept of evaluation influence and endeavored to empirically study influence processes and the pathways through which they were manifested in a real-world evaluation context. The research questions that guided this study are as follows:

1. How was this federally-funded, large scale SLC evaluation designed and implemented? How do the components of the evaluation align to Mark & Henry's program theory of evaluation influence?

2. What are the "inputs" of this evaluation that may lead to and may impact potential underlying processes, outcomes, and pathways to evaluation influence?

3. What are the underlying processes and outcomes that emerged throughout the course of this evaluation?

4. To what extent do these processes/outcomes link together to form "pathways of influence"?

5. To what extent does the evaluation help explain the potential underlying processes, outcomes, and pathways to evaluation influence? In what ways does this study offer preliminary evidence to support Mark & Henry's theory of evaluation influence?

The design of this study incorporated secondary data from a large-scale school district evaluation and utilizes a mixed-method analytic approach. The school district evaluation was completed prior to this study and generated data through interviews, focus groups, and surveys. The evaluation proposal, design, and implementation data were also available and analyzed to inform the broader research questions. The combination of qualitative and quantitative methods used in this study promotes credibility and validity to the findings and conclusions and overcomes the limitations in working with secondary data sources.

CHAPTER 2

Review of the Literature

In the first section of this chapter, a review of key evaluation literature as it pertains to evaluation use is presented. This includes a discussion of traditional forms of use, the factors affecting use, and the broader notions of evaluation's influence. Research on the impact of organizational context on evaluation process and use is also considered and discussed as it pertains to evaluation within institutions such as schools and districts. Finally, key current empirical research studies on evaluation use are reviewed. The second section of this chapter discusses Mark & Henry's (2004) program theory of evaluation influence in more specific detail as the framework utilized in this study.

Evaluation Use

It is important to begin by noting that although evaluation theorists and practitioners may differ in their definitions of evaluation's purpose, many would agree that an important goal of evaluation if not the main goal, is social betterment. Without evaluation use, however, this goal of social betterment is not likely to take place. Use, therefore, may be considered "a means by which evaluation achieves social betterment" (Henry, 2000). Cousins & Shulha (2006, p. 271) noted that for evaluation, "of central interest is the use of evaluation to help ameliorate social injustice and inequity." In the evaluation literature, there are several, more practical ways in which the notion of evaluation use has been defined. Among them, one of the simplest yet practical definitions was offered by

Alkin & Taut (2003) as "the way in which an evaluation and information from the evaluation impacts the program that is being evaluated" (p.2).

As evaluations are often commissioned for the purpose of systematically gathering information that will guide decision makers (Alkin, 1975, Patton, 1997), there must be an 'it' to utilize and the presence of users (Alkin, Daillak, & White 1979). The use of the information generated from an evaluation is referred to as findings use and can also vary in nature. Evaluation findings can serve three purposes: to render overall judgment of merit, worth, or value; to facilitate improvement; and/or to generate or increase knowledge (Patton, 1997). Evaluation findings used for judgment are generated primarily through summative evaluations while findings that are used for improvement and knowledge generation are primarily generated through formative evaluations. The three categories of use most often referred to in the evaluation literature are instrumental use, conceptual use, and symbolic use. Instrumental use pertains to occurrences whereby knowledge generated by an evaluation is used to impact a direct action (Alkin & Taut, 2003). A government agency's decision to fund implementation of an educational program based on an evaluation of effectiveness is an example of instrumental use (Leviton & Hughes, 1981). The use of findings to render judgment or to facilitate improvement both fall under the umbrella of instrumental use. Conceptual use (sometimes described as enlightenment) occurs when conceptual understandings about the program, organization, etc. have been made that changes the way users think about the program or aspects of the program. With conceptual use, no direct decisions are made or expected (Alkin & Taut, 2003; Patton, 1997). Symbolic use occurs when an evaluation and its findings are used to

enhance reputations or as act as a status symbol. Evaluation findings that are used

to justify decisions that have already been made have been described as

legitimative use (Alkin & Taut, 2003). Leviton & Hughes (1981) describes an

analogous category of persuasive use, where evaluative evidence is used to

persuade others to support or defend a particular position.

More recently, it has been noted that evaluation use can also come from

participating in the evaluation process, also known as process use (Patton, 1997;

Hofstetter & Alkin, 2003, Alkin & Taut, 2003). Patton (1997) defined process use

as the "individual changes in thinking and behavior, and program or organizational

changes in procedures and culture that occur among those involved in evaluation as

a result of the learning that occurs during the evaluation process." As with findings

use, process use can be instrumental or conceptual and may occur during the

course of an evaluation or at its conclusion. Instrumental process use may occur

when a decision is made as a result of the evaluation process. Conceptual process

use may occur, for example, when participation in the evaluation changes the

attitudes of users about evaluation.

Kirkhart (2000) argued that the term "use" was limiting and did not allow for

an expanded understanding of the influence that an evaluation may exert.

Furthermore, the term "use" focused only on evaluation findings and was an

"imprecise fit with non-results-based applications, the production of unintended

effects, and the gradual emergence of impact over time (p. 6). Kirkhart proposed a

shift in terminology to "influence" as a broader and more inclusive term. Henry &

Mark (2003) supported this shift and proposed a focus on evaluation influence,

arguing that the evaluation literature did not adequately address the change

processes through which the evaluation process and findings may lead to the ultimate goal of evaluation – social betterment. Furthermore, Mark (2011) stated that the term "use" has implications of immediacy, intentionality, and awareness. He added that evaluations may have consequences that are removed from the evaluation location or from the original parties involved. Finally, he proposed that "use" and "influence" served as complementary concepts, without diminishing the importance of the classic forms of use or of intended use. In Cousins & Shulha's (2006) comparative analysis of evaluation use literature, the authors noted that a "unique emergent theme concerns a call to move beyond use to a theory of influence" (p. 286). They concluded that this move was likely to stimulate debate and further research in evaluation.

Factors Affecting Evaluation Use

Several researchers have provided insight into factors associated with evaluation use as well as produced factor lists and categories based on findings from empirical evaluation research. Research syntheses of evaluation use factors conducted by Cousins & Leithwood (1986), Leviton & Hughes (1981), and Alkin (1985, as cited in Hofstetter & Alkin, 2003) continue to be among the more prominent in the evaluation literature.

In their review of 65 empirical studies on the use of evaluation findings, Cousins & Leithwood (1986) developed a conceptual framework within which 12 specific factors that influence use were identified. Six of these factors were associated with characteristics of evaluation implementation: evaluation quality, credibility, relevance, communication quality, findings, and timeliness. The

remaining six factors were associated with characteristics of decision or policy settings: information needs, decision characteristics, political climate, competing information, personal characteristics, and commitment and/or receptiveness to evaluation. Their analyses revealed that the relative influence of these factors was dependent largely on the type of use. Leviton & Hughes (1981) identified five major clusters of variables consistently related to utilization: relevance of evaluation to the needs of potential users, communication between evaluators and users, information processing, credibility or trust placed in the evaluation, and user involvement, commitment, and advocacy. Alkin (1985, as cited in Alkin & Taut, 2003) identified three main categories of factors: evaluation factors, context factors, and human factors. Evaluation factors relate to the actual conduct of the evaluation, including the procedures used, and the information collected and reported. Context factors reflect on the context of the program being evaluated such as pre-existing evaluation bounds and organizational characteristics as well as individual program characteristics. Human factors include both user and evaluator characteristics (although evaluator characteristics can also be considered as part of the evaluation factor), and such things as organizational responsibility, personal and professional style.

One of the most important determinants associated with use is considered to be the "personal factor" (Hofstetter & Alkin, 2003). The personal factor, as described by Patton (1997), is the "presence of an identifiable individual or group of people who personally care about the findings it generates" (p. 44). Through his research on use, it was found that information generated from evaluations was

used when the personal factor was present and when it was absent, there was a notable absence of impact stemming from the evaluation.

Shulha & Cousins (1997) described the effort by evaluation researchers and theorists to clarify the role of context and the effects on evaluation use (utilization). Various perspectives emphasized the role of context in creation of knowledge, the inextricable link between context and political activity, the relationship between use and organizational context and processes, and the interaction between the evaluator and the program context. Furthermore, they discussed emerging evidence supporting the contribution of collaborative and participatory evaluation models to evaluation use (Cousins, 2003).

Impact of Organizational Context on Evaluation Process and Use

Context may be broadly defined as the situations, conditions, or circumstances that directly influence an organization. Searches for relevant research show that these constructs have also been described as "school organization," "organizational culture," "school climate," and "school culture," to name a few. Regardless of title or keyword descriptor, the significance of these constructs in studying and understanding the evaluation process, outcomes, and influence, is evident. In this study, context refers to internal factors such as the organization's sense of purpose and mission, attitudes and beliefs, perceptions of administrator and teacher support, and level of collaboration among staff (Brady, 2008; Kowalski & Hermann, 2008). These internal factors are of greater importance as they may help explain why organizations that appear similar in

structure may function differently and yield varying levels of effectiveness (as cited in Kowalski & Hermann, 2008).

Studies have shown that the organizational climate of a school has an impact on student and staff performance. Furthermore, a district's climate plays a role in affecting change or reform (DiPaola & Smith, 2008). Noted aspects of positive climate such as trust, collaboration, and shared understanding were emergent across the research literature as promoting improved performance. Therefore, the relevance and importance of studying organizational context is paramount as it impacts both program outcomes as well as evaluation outcomes. It is these latter outcomes which include the "evidence" provided to stakeholders for decision-making and change. Although there is an expanding body of research literature establishing data-driven decision-making at schools as an important part of creating effective schools (Lachat & Smith, 2005), there is less available on evaluation use of findings-based decision-making in schools and/or districts. Coburn, Touré & Yamashita (2009) found that research evidence played an important, but indirect role in school district decision-making. Rather decision-making was about interpretation, argumentation, and persuasion. Decisions were also influenced by organizational structure, resource constraints, and leadership turnover.

In public education, organizational hierarchies may also have a direct influence on how evaluations are used. In most cases, district administrators solicit and coordinate program evaluations but the outcomes of the evaluation may lie at the school and teacher levels. Evaluations are typically commissioned by those entities responsible for the program and not by those who receive its benefits (Weiss, 1993) or are affected by the findings. Awareness of the organizational

structure may help to illuminate the political dynamics that affect if and how evidence-based decisions, including those involving evaluation findings, will be made (Honig & Coburn, 2008). Furthermore, it can inform the pathways to potential use and evaluation influence.

In their study of district and school-level evaluation, Pechman & King (1986) concluded that use was more of a developing process as opposed to a "static ideal." The evaluation process itself may be unfamiliar and sometimes intimidating to school staff and, akin to any introduced innovation, there was rarely any immediate change as a result. This study raises two points, still relevant in current circumstances. First, use of evaluation findings, whether instrumental or conceptual is not likely to be instantaneous or fixed. Second, schools are continually inundated with new programs, reform efforts, or other innovations. Whereas evaluators and higher level stakeholders may view use as making changes based on findings and recommendations, those stakeholders at the ground level (i.e., school teachers) may view it as another burden imposed on them by the district. Therefore, it may be reasonable to assume that in a context that does not promote improvement or is over-burdened with innovations, use may not be overtly evident. Limited understanding of evaluation in general, lack of trust, and threat of negative results have also been found to be a barrier to the evaluation process (Taut & Alkin, 2003) and, by logical extension, a barrier to use.

Evaluation use: Credible evidence?

Evaluation use has received substantial attention in the evaluation literature and is the most researched aspect of evaluation (Christie, 2007). But the question

of whether this evidence base is entirely credible has also stimulated dialogue among evaluation researchers and theorists as well as examinations and reviews of such research. A primary deficiency in the evidence base lies in the lack of or insufficient reporting of methodological information in published articles or reports on studies of evaluation use. Without such descriptions, the quality of the study and its design is difficult, if not impossible, to determine. Furthermore, the ability of these studies to be replicated is limited. Brandon & Singh (2009) argued that if the methodologies employed in studies of evaluation use are not sound then the strength of their conclusions cannot be known. These authors reviewed of 52 empirical studies of evaluation use (from 1971 to 1998) looking for evidence of methodological soundness and validity. In addition to the abovementioned lack of solid methodological reporting in many of the studies, the authors found that narrative reflections were used very often as results or to support conclusions. These reflections are essentially anecdotal descriptions not based on systematic methods of data collection or research and are written from the point of view of the evaluator without explanations of the qualitative development for the purposes of replication or verification. Consequences have included overstatements of such reflections as support for study conclusions. The use of narrative reflections and the implications have also been acknowledged in other reviews. Cousins & Shulha (2006), for example, have suggested using such reflections in conjunction with more conventional approaches with explicit methodological processes to enhance credibility.

Other issues in methodological reporting emerged related to validity. Brandon & Singh's review noted that attention to the validity related to instrument

development, self-report bias (particularly due to recall error), and measurement error resulting from questions of past duration and frequency have been largely ignored in the evaluation use literature. A general lack of attention has been paid to the quality of studies by both researchers and reviewers. Leviton (2003) has suggested that the evidence base suffered from a "standard of evidence that many of us would never dream of applying to the conduct of evaluations, too often predominates in the study of evaluation use" (p. 526).

Another area of deficiency concerns the fact that a majority of the published research on evaluation use takes place within education contexts (Brandon & Singh, 2009). This only presents a problem when the results are combined with results from studies in other contexts or overgeneralized without regard to contextual issues and factors. However, because of the greater number of evaluation use studies in educational settings and the balance of methods across these studies, Brandon & Singh (2009) concluded that there was more support leading to a credible evidence base for research on evaluation use in education.

A review of evaluation use research by Johnson, Greenseid, Toal, King, Lawrenz, & King (2009) presents a positive shift from the past by only including studies that met a predetermined set of criteria for methodological quality. These criteria included, but were not limited to, soundness of design, measurement of variables, methodological appropriateness, and extent to which bias was addressed. After this in-depth screening by independent reviewers, 41 studies from 1986 to 2005 met the minimum quality rating and were subsequently reviewed. This study, which used Cousins & Leithwood (1986) framework for categorizing empirical studies of evaluation use, provided a significant contribution to the credible

evidence base for evaluation use. However, Johnson et al. did not publish which studies received what rating or sufficient information on the specific levels which constitutes each rating.

As evaluation researchers and theorists continue to study, debate, and re-conceptualize the effects and consequences of evaluation, the need for research and empirical inquiry in these areas is of paramount importance. Understanding broader notions of evaluation use and evaluation influence – the levels at which they can occur (i.e., individual, collective), the forms they take (i.e., behavioral, affective), and the potential pathways influence may travel, may reveal previously unknown conceptions of the impact of evaluation or, at the very least, promote inquiry and research among scholars and practitioners alike.

Program Theory of Evaluation Influence

Henry & Mark (2003) proposed a conceptual theory of evaluation influence that focuses on evaluation consequences ("evaluation outcomes") which could plausibly lead toward or away from social betterment, the ultimate goal of evaluation. The theory has been conceptualized in a traditional logic model format with inputs, activities, outputs, and outcomes. This format was inspired by Cousins (2003), who offered a similarly structured model of utilization based on a participatory evaluation approach. The format enables the reader to visualize evaluation practice as activities fueled by inputs leading to outcomes. Mark & Henry adapted and expanded this model to be used with a general evaluation approach and incorporated their proposed outcomes, processes, and pathways. These outcomes (not program outcomes but evaluation outcomes) are the main

thrust in this theory of evaluation influence as they encapsulate the processes and mechanisms that mediate change as a result of an evaluation. It is within these outcomes that the traditional forms of use (i.e., conceptual, instrumental) fit. The authors asserted that current conceptions of use do not adequately establish the link between an evaluation and its effects or outcomes. This program theory addresses the broader notion of influence and the underlying mechanisms through which evaluation has its effects. Figure 1 (on page 18) shows the graphic conceptualization of this program theory in logic model format, adapted from the graphic originally presented by Mark & Henry (2004). Some descriptive content from the within the outcome boxes was removed from the figure due to space limitations.

The first column of this model, the inputs, refers to the evaluation and organization context and settings. The arrows leading from the inputs indicate that these inputs lead to (and contribute to) the evaluation activities. It is important to note that the there are arrows leading from these inputs directly to the outputs/short-term outcomes, acknowledging that these inputs may have an impact on some outcomes regardless of the evaluation activities. The second column of evaluation activities includes all aspects of the evaluation process, including the generation and dissemination of evaluation findings. The third column contains two boxes, the top representing evaluation outputs. The first box, knowledge production attributes, was also adapted by the authors from Cousins (2003), and represents the judgments made by stakeholders about the evaluation process and/or findings. In Cousins' model, "evaluation production knowledge" is a consequence of participatory evaluation that includes interactive processes such as

stakeholder participation and control of technical decision-making (Cousins, 2003). In this model, however, these attributes are not specific to participatory evaluation approaches and are more directly a stakeholder response.

Figure 1. Program Theory of Evaluation Influence, Adapted from Mark & Henry (2004).
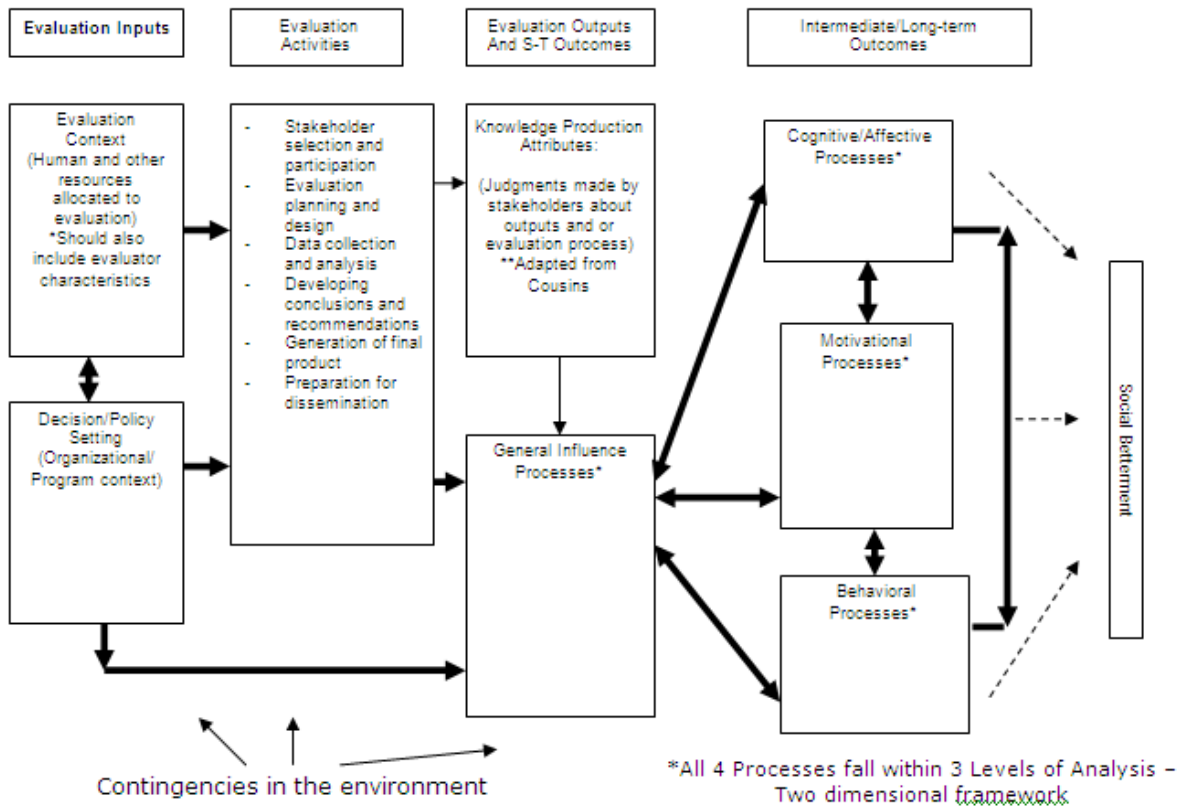
The second box in this column, General Influence Processes, may be considered as more of a short-term outcome than an output and is described more comprehensively with the other outcomes in the next section. The box in the lower left hand portion of the model titled "contingencies in the environment" represent the competing processes, facilitating factors, and inhibiting conditions that may impact an evaluation at any point.

The authors' proclaimed that current models of use were "generally silent on

the range of underlying mechanisms through which evaluation may have its effects"

(p.37, Mark & Henry, 2004). They argued for the need to go beyond describing

types of use and factors promoting use to a focus on these underlying processes

and the pathways that form the connections.

The arrows in this model, particularly those leading to the various outcomes,

correspond to the hypothesized pathways. The bi-directionality of some arrows

were not intended to imply simultaneous causation but rather iteration, as one may

lead to another and back again. The outcome boxes correspond to the various

processes that indicate change as a consequence of an evaluation. Together, it is

this portion of the model that illustrates the authors' main objective in

conceptualizing the underlying mechanisms of evaluation influence.

Table 1 presents a detailed description of the contents of the outcome boxes

in the model or, as labeled, the four types of evaluation influence processes. The

authors used the term "processes" to also refer to outcomes because of their

proposed duality – potentially serving as evaluation outcomes as well as the

underlying processes leading to other outcomes. The terms "mechanisms" and

"processes" have also been used interchangeably in their descriptions of the model.

Table 1 also illustrates the division of the processes into three levels of analyses –

individual, interpersonal, and collective. The authors referred to each level as the

"locus" of analysis as it denotes the location of evaluation influence. The majority

of the empirical literature on evaluation use focuses on the individual level. The

interpersonal level is where changes occur as a result of interactions between

individuals. The collective level depicts the influence of evaluation on organizations

or groups. Each of the three levels is subsumed within each process, creating a two-dimensional taxonomy of evaluation outcomes.

The specific indicators listed within each cell are drawn from the social and behavioral science research literature as indicators of a particular process/outcome. The authors suggested that these within-cell processes were not the only possible outcomes but represented a sample of what could occur in a given evaluation. Furthermore, moving across the levels of analysis does not suggest greater magnitude from the individual level to the collective level. Rather, the authors intended the processes to be analogous across the levels.

Table 1.

Underlying Mechanisms That Potentially Mediate Evaluation Influence (Mark & Henry, 2004).

Type of Process/Outcome	Level of Analysis		
	Individual	Interpersonal	Collective
General Influence Processes (Likely to set into motion some change/outcome)	Elaboration Heuristics Priming Skill Acquisition	Justification Persuasion Change Agent Minority-opinion influence	Ritualism Legislative hearings Standard Setting Policy consideration
Cognitive/ Affective	Salience Opinion/attitude valence	Local descriptive norms	Agenda setting Policy-oriented learning
Motivational	Personal goals and aspirations	Injunctive norms Social reward Exchange	Structural incentives Market forces
Behavioral	New skill performance Individual change in practice	Collaborative change in practice	Program continuation, cessation or change Policy change Diffusion

The first set of processes in the column, General Influence, were described by the authors as the "fundamental architecture of change" as they may be thought

of as short-term outcomes or stepping stones to the other outcomes of interest. General Influence processes may occur at all three levels of analyses and may be cognitive, motivational, or behavioral in nature. Their importance is judged less by the form it takes and more by whether or not it leads to other intermediate or long-term outcomes or processes.

The last three sets of processes/outcomes appear in the program model as intermediate and/or long-term outcomes. Cognitive/Affective, refer to changes in thoughts and feelings as a result of the evaluation. The traditional form of conceptual use would correspond to this particular outcome. Motivational processes refer to goals, aspirations, and responses to perceived positive and negative consequences. The authors have noted that although this particular outcome is rarely addressed in literature on use, it may serve as an important intermediate outcome (or mechanism) leading to behavior change. The last set of processes, Behavioral, corresponds to the traditional form of instrumental use in that it refers to changes in actions. Process use, changes that occur as a result of the evaluation process, is not explicitly addressed in this program theory of evaluation influence. The authors noted that since the occurrence of process use is triggered by evaluation activities (and not specifically evaluation findings), it could not be translated into the processes/mechanisms as proposed in the model.

Mark & Henry acknowledged that this conceptual model was general and should be tailored to specific contexts. The specific indicators of influence (as described in Table 1) were suggested as starting points for future research and development. As the authors promoted this model for several purposes (e.g., as a tool for understanding the pathways of evaluation influence, to guide evaluation

practice), they also invited empirical and systematic study of this model and advocated for its potential contribution to the growing evidence base for evaluation practice.

This review of evaluation literature brought together the concepts of use and influence. Evaluation use, the more theoretically and empirically studied concept, has been shown to be impacted by many factors ranging from contextual to personal characteristics. The study of evaluation influence, as a concept and as an empirical focus, is relatively new and ripe with possibilities. If one views evaluation influence as a bigger umbrella under which traditional definitions of evaluation use are housed, then much of the literature of use can be extended and applied to the broader notion of influence. Alternatively, if the viewpoint is of influence as unintentional and outside the boundaries of primary users, then the literature reviewed here may be considered a springboard from which to guide future study. Mark & Henry's program theory of evaluation influence appears conceptually related to the first viewpoint – influence as a broader conception of use. This is most clearly depicted in the similarities between the various processes (e.g., Behavioral processes) and traditional terms of use (e.g., instrumental use). They have unpacked evaluation outcomes and hypothesized pathways proposing that influence is more than an unintended effect of evaluation practice and is not bound by the constraints of the traditional definitions. As evaluation scholars become more open to the concept of influence, empirical research and theoretical debate shall ensue. As a final point, what can be learned from the empirical evidence thus far is that greater attention needs to be paid to methodology and issues of validity in order to

add to the credible evidence base of evaluation research, regardless of whether the

focus is on use or influence.

CHAPTER 3

Methodology

The setting for this study is the federally-funded evaluation of a Small

Learning Communities (SLC) program in a large urban school district. This chapter

begins with a general description of the SLC program to provide the necessary

background and reference for this study. A comprehensive description and

discussion of the design of the SLC evaluation and its implementation are presented

in Chapter 4 as it pertains directly to the results for Research Question #1 – How

was this federally-funded, large scale SLC evaluation designed and implemented?

The remainder of the chapter describes the study design, the data sources used to

address the research questions, definitions of key terms, and an overview of the

analyses performed.

Description of the SLC Program

In 2005, a large, urban school district received a federal grant to develop and

implement Small Learning Communities (SLC) at three high schools. Small learning

communities, a school reform effort, are developed and implemented for the

purposes of impacting student achievement by personalizing education. This

personalization is typically achieved through the development of structures and

processes that facilitate positive relationships among teachers, among students,

and between teachers and students. Federal grants have been issued in support of

SLC implementation in large public high schools with enrollments of 1,000 or more

students (U. S. Dept. of Education, 2009). The district's proposal for this SLC grant

included goals that were also aligned to the high school reform initiative already in

place. In addition to improving student achievement, closing the achievement gap, building capacity for leadership among teachers, and improving the schools' culture and climate were included among these proposed goals.

The three high schools served by this grant were comprised of diverse student populations with significant proportions of students of color as well as those classified as "socioeconomically disadvantaged." Each school was also challenged by varying levels of low achievement. School Accountability Report Cards showed that for the 2007-08 school year, on average for the three schools, about one-third of the students met or exceeded the state standards in English while about one-sixth of the students met or exceeded the state standards in mathematics, as measured by the California Standards Tests (CSTs). At the start of this five-year grant, each of the three high schools already had some form of learning communities in place on their campuses. By Year 5, all three schools had implemented wall-to-wall SLCs, which meant that all staff and students were affiliated with an SLC. The high schools had seven to nine functional SLCs each by the end of the grant period.

Organizationally, at there was an SLC Director at the district level who managed the overall implementation of the SLC program across the sites. The SLC director organized monthly meetings throughout the academic years to support the implementation of the activities specified in the initiative and to monitor progress. At the school level, SLC Coordinators managed SLC implementation (often in collaboration with school administrators) and generally served as a liaison between the school and the district. Each SLC also had a lead teacher. Although this position was not supervisory or particularly administrative, these lead teachers

worked primarily within their SLCs to develop, strengthen, and sustain but also worked across SLCs to develop greater cohesion as a school. In addition to the monthly meetings, the SLC coordinators and SLC lead teachers participated in quarterly workshops, or Lead Teacher Institutes, where they received support, guidance, and resources to enhance SLCs and encourage progress in meeting the grant goals.

There were many challenges to SLC implementation over the five years, some of which were beyond the scope of the grant such as budget cuts and teacher layoffs. Other challenges included teacher buy-in, scheduling, and adequate time for planning and collaboration. Successes of the program at the student level included increased academic motivation and aspiration toward college and career goals and greater connectedness toward peers and teachers. At the teacher level, there was more opportunity for cross-discipline communication and collaboration.

As part of the accountability portion of the grant, the district was required to have an external evaluation component related to grant implementation. The SRM Evaluation Group, housed within UCLA's Graduate School of Education and Information Sciences and supervised by Dr. Marvin Alkin, was contracted by the district as the external evaluation team for the five year grant period. The program administrators, at both the district and school levels, had access to the evaluation team and evaluation findings throughout the course of the grant and relied on these resources for continued program development and implementation.

Study Design

This study addressed the concept of evaluation influence and empirically studied the levels of influence and the pathways through which they were manifested, all in a real-world evaluation context. The design of this study included a mixed-method analytic approach using the interview, focus group, and survey data generated from the SLC evaluation. The qualitative portions of this study drew from evaluation data sources to map out the SLC evaluation and its relevant components as they corresponded to Mark & Henry's program theory of evaluation influence. Interview, focus group, and open-ended survey responses were the data sources for the qualitative analyses of associations, links, and pathways between context, activities, and underlying mechanisms/outcomes. Where appropriate, I hypothesized alternative explanations. The quantitative portions of this study drew upon qualitative findings as well as quantitative survey data to establish variables for statistical analyses, specifically for testing specific pathways of influence. The determinants or predictors of influence were either taken directly from the data or were derived through factor analysis. Relationships were tested through multiple logistic regression procedures.

Data Sources

This study used extant data that had been collected as a function of the SLC evaluation. Since the evaluation team collected more purposeful data on evaluation use during the final year of the evaluation, this study primarily focused on and utilized this secondary data for Year 5. The available data from Years 1-4 of the evaluation varied both in quality and in purpose and were used primarily to

supplement and support the analyses, findings, and interpretations of this study as necessary. All of the data was provided by the evaluation team in a de-identified format. In other words, all data that were potentially identifiable (i.e., school names, administrator names, etc.) were removed from the dataset prior to distribution. Codes such as "School A" were assigned in order to make important distinctions for the analyses and interpretations. This de-identification process was also applied to qualitative data and any evaluator field notes used in this study. Evaluation reports were available and downloadable from public websites.

Table A-1 in the Appendix provides an overview of all of the data sources and instruments utilized through all five years of the evaluation. Table 2, on the following page, contains only the data sources and instruments used for this study. These sources pertained only to staff and administrator respondents – no parent or student data were used.

For Year 5 of the evaluation, a staff survey was administered online in the spring to all staff at the three grant high schools. The resulting respondent size, encompassing staff from all three schools, was 272. The response rate for this survey was unknown. In addition to survey items and questions specific to the program and evaluation purposes, the evaluation team added items/questions about evaluation awareness, use, and outcomes. The findings from these data were included in the summative report provided to the district at the end of the evaluation period.

Table 2.

Data Sources Used for this Study

Year of Evaluation	Data Sources	Description
Year 5 (2009-2010) PRIMARY SOURCES	School Staff Survey	Online survey administered to staff in Winter 2010. Across the 3 schools, N=272. Approximately 83% were teachers, 5% school administrators, 3% counselors, and 9% in other school-site positions, both certificated and classified.
	Teacher Focus Groups	A total of 11 focus groups (N=38). 4 of 11 groups comprised of SLC Lead Teachers.
	Principals, SLC Coordinators, Administrator Interviews	Principal from each school (N=3), SLC Coordinators (N=4), and SLC Project Director (former and current)
	Evaluator Field Notes	Field notes from SLC professional development meetings, notes on process and/or findings from Year 5 data and analyses
	Evaluation Reports	Publicly available annual evaluation report and summative report
Years 1-4 (2005-2009) SUPPLEMENTAL SOURCES	School Staff Surveys Interviews Evaluator Field Notes Evaluation Reports	Similar descriptions as above.

Also in Year 5, as in other years, school principals and SLC Coordinators were interviewed in person and audiotaped. Teachers and SLC Lead Teachers participated in focus groups separately. The participants were representative of the site's SLCs; however, they were not selected at random or in any systematic way. The SLC Coordinators "invited" teachers to participate in the groups and the criteria for selection was not known or recorded by the evaluation team in the annual reports. For both interviews and focus groups, there was an added set of questions for Year 5 regarding evaluation use and awareness, in addition to questions on SLC development, implementation, challenges, and effects.

Definitions of Key Terms

　　The following key terms are defined as they have been utilized and addressed in this study.

　　Demographics. At the school level, demographic characteristics included school affiliation, school assignment, years of experience, and past leadership roles. These data come largely from the survey population and was hypothesized to have an impact on evaluation influence.

　　Context. This construct was broadly defined as the situations or circumstances in which a district and the school sites functioned. Factors which indicated a school's context included the culture and climate of the school, administrator/teacher support, and level of collaboration among staff. Staff buy-in of the program and its implementation was another aspect of the school's culture and/or context. Factors related to the district context included the ways in which information about the evaluation (both process and findings) were relayed to the schools and the ways in which the district facilitates or inhibits the evaluation. Searches for relevant research showed that this construct has also been described as "organizational culture," "organizational context," "organizational climate," and "school culture," to name a few. Regardless of title or keyword descriptor, the significance of this construct in studying and understanding evaluation influence was evident. Context was hypothesized to have an impact on evaluation influence.

　　In Mark & Henry's program theory of evaluation influence, context falls under the general heading of *Inputs* (see Figure 1 on page 18). Also included under this heading is the evaluation context, or the conditions in which the evaluation has taken place. This context may include the purpose of the evaluation, the resources

allocated to the evaluation, and characteristics of the evaluation team members. Data on some aspects of the evaluation context in this SLC evaluation was not directly collected but was available through other data sources (e.g., field notes, evaluation reports).

Awareness. This term signifies whether or not the stakeholders in the evaluation were aware that an evaluation was taking place, if they had any knowledge of any findings generated from the evaluation, and if they were aware of the use of such findings. Awareness was considered an external "contingency on the environment" as opposed to an input into the process, largely because lack of or limited awareness was hypothesized to have an impact on evaluation influence not evaluation activities.

Activities. Evaluation activities comprised any action, practice, and/or process that the evaluation team engaged in, with or without stakeholder participation, to conduct the SLC evaluation. The following list of activities included, but was not limited to:

- Meetings – for such purposes as planning, communication, information gathering
- Methodological practices (e.g., designing data collection instruments, data collection activities, analysis)
- Generation of findings

Processes/Outcomes. As described earlier, the terms "processes," "mechanisms," and "outcomes" have been used sometimes interchangeably by Mark & Henry to refer to the elements in their program theory of influence through which the effects of evaluation may be mediated. Each element, as further

proposed, may play a dual role – it may serve as an outcome (or consequence of

the evaluation) or it may serve as an underlying mechanism (which leads, in turn,

to another outcome). For the purpose of this study and for clarity, I largely used

the terms "outcome" and "processes" to be more consistent with the model as

presented in Figure 1.

Pathways. The term "pathway of influence" is used by Henry & Mark

(2003) to denote how, in practice, specific elements (underlying mechanisms,

outcomes) may link together resulting in a pathway. For example, if an evaluation

report stimulates a stakeholder to think more about the program's effectiveness

and subsequently this individual changes his/her attitude or opinion about the

program, that would be considered a pathway from a General Process to a

Cognitive/Affective Process.

Overview of Analyses

This study employed a mixed-method approach to the analyses and in

response to the guiding research questions. The overarching goal of this study was

to "test" this model of evaluation influence. In order to accomplish this task,

quantitative analyses were required of the various model components and the

relationships to and among the outcomes. Due to the nature of the available data,

multiple logistic regression techniques were the most appropriate statistical

approach. Furthermore, in preparation for the inclusion of some variables into

these analyses, factor analysis procedures were utilized in the formation of

constructs. Descriptive analyses were also conducted for the variables of interest.

Due to small sample sizes and the structural limitations of the data, qualitative methods were used to analyze the open-ended survey data and interview/focus group data. These data were iteratively reviewed, coded, and then categorized, depending on the focus of the particular analysis. Qualitative reviews of existing documents also involved coding and categorizing. Both the quantitative and qualitative results were then brought together, synthesized, and interpreted as a meaningful whole in response to the goal of this study.

Because each analytic approach and specific technique was specific to a particular research question, more detailed descriptions of these analyses is presented in Chapter 4 alongside their results and interpretation.

CHAPTER 4

Analyses, Results, and Conclusions

In this chapter, each research question is followed by descriptions of the analyses, the results, and conclusions appropriate to that research question. This format was adopted as the most feasible way to meaningfully present all of the information necessary to respond to each research question, given that the questions are cumulative.

Research Question #1

How was this federally-funded, large scale SLC evaluation designed and implemented? How do the components of the evaluation align to Mark and Henry's program theory of evaluation influence?

The two questions that comprise Research Question #1 set the stage for this study and lay the foundation for the ensuing research questions, analyses, and conclusions. To respond to these questions, the analyses involved systematic qualitative review of several reports, documents, and data collection instruments. Although the primary focus was on Year 5 of the evaluation, some attention was paid to the first four years of implementation, particularly changes and/or challenges, as they led to the final year of the evaluation.

The first major section responded to the foremost question of how the evaluation was initially designed and then actually implemented. The evaluation design, as proposed in 2005 is presented. Following this section, the implementation of the evaluation is presented and discussed as it corresponded to the main components of a logic model: inputs, activities, outputs, and outcomes.

Formatting in this manner facilitated the interpretation of subsequent findings and reference to Mark & Henry's model. The second major section answered the question of how the evaluation components mapped onto Mark & Henry's model.

Proposed SLC Evaluation Design

The original grant proposal submitted by the school district in 2005 was the primary data source used for this analysis. During the grant proposal phase, members of the evaluation team worked together with district staff to propose an evaluation design to be submitted to the federal grantors. A primary objective of the evaluation team was to be responsive to the district as well as the individual schools' needs and concerns about implementation of the grant and of the evaluation. The evaluation's theoretical framework incorporated aspects of Alkin's (2004) "user-oriented" and participatory evaluation approaches. The goal in using these approaches was to work closely with key stakeholders in designing and conducting the evaluation that was "collaborative and user-focused" and would include "timely and regular feedback to schools."[1]

According to the grant proposal, the purpose of the evaluation was to assess each school's progress on the main goals from the High School Reform Initiative of which SLCs were the vehicle for change: improving student achievement, closing the achievement gap, creating culture/climate that supports improvement, and building leadership capacity to sustain the reform efforts. The evaluation team proposed to work with SLC Design Team members and program staff in developing theories of action, or logic models, to understand the underlying mechanisms that

---

[1] These quotes were taken from the original grant proposal submitted by the school district in 2005.

would lead to progress on the four main program goals. It was anticipated that these theories of action would also serve to refine site-level goals and objectives. Since the district's grant proposal did little to address how proposed activities (or various "strategies") were expected to lead to change and how outcomes were to be defined and measured, the collaborative development of theories or logic models was determined to be beneficial to all.

Data was to be collected to address the major indicators of progress toward the goals of improved achievement, creation of culture supporting achievement, and building leadership capacity to sustain efforts. The proposed data sources included annual surveys for parents, students, and staff; annual interviews/focus groups for all three groups; and attendance, graduation, enrollment, and achievement databases provided by the district research office.

Proposed evaluation products included annual reports for each year of the evaluation and a cumulative report incorporating results across the grant period. Other proposed activities included assistance with the district's annual performance reports (APR) to their federal grantors and discussions/presentations of findings. Allowing for potential adjustments to evaluation activities, instruments, and reporting to fit the needs of the stakeholders was also implied throughout the proposal.

Evaluation Inputs

Inputs in an evaluation are generally comprised of the resources allocated for the evaluation, not for the program being evaluated. However, the concept is similar in that both financial and human resources are needed to support the

process. Human resources consist of those available to conduct the evaluation as well as program staff support. The evaluation team, on their end, was well staffed and had additional internal support when needed. Staff support provided by the district came largely through the district's research unit and school site SLC Coordinators. The research unit generally assisted with hard copy surveys (e.g., scantron formatting, scanning, managing databases, and cleaning data) and online surveys through posting links on school websites. The SLC Coordinators served as the primary facilitators of data collection activities at each school site. The district-level SLC Director provided general support throughout the evaluation process, particularly in Year 5, often ensuring that the evaluation team was able to implement the activities with minimal problems.

The setting or the context of the evaluation is also an input into the process. The broader setting was the district in which the schools were housed and definitely played a role in the implementation of evaluation activities. At the district level, administrators had the ability to both hinder and facilitate the progress of the evaluation. The change in SLC Director from Year 3 to Year 4 and the resulting positive changes in support and collaboration highlighted the role of the district context as an input. The school sites, however, were the settings at which both the program and the evaluation took place. Whereas it was found that the district context had more of an impact than the schools on evaluation activities (Rivera, Vo & Lee, 2008), the school-level context played more of a role in the evaluation outcomes. A comprehensive analysis and discussion of school-level context is presented under Research Question #2. It is important to note here that context – represented by school culture, climate, teacher buy-in, etc. – was included

throughout the evaluation by the team as an indicator of progress and was addressed on both the survey and interview protocols across all five years of the evaluation. This resulted in multiple data sources available for this study from which to analyze context and understand its impact on evaluation outcomes.

Demographic information on the school as a whole and of the staff and students were also inputs in that demographic variations had the potential to impact how activities were implemented and how outcomes were achieved. This information was collected by the evaluation team annually through state websites and survey items. The evaluation reports typically included tables of demographic information of survey respondents and there did not appear to be significant variations across the years of the evaluation. Furthermore, the reports did not include analytic results of variations in responses disaggregated by demographic variables.

Awareness of the evaluation process and findings was an external factor, or contingency in the environment, that had the potential to impact the evaluation process and/or outcomes at various points. For this reason, the evaluation team included measures of awareness in Year 5 on the survey and focus group protocols. For this study, its presence or absence was considered a predictor of evaluation influence as an outcome. Although more than an input, the analyses and discussion of this predictor is presented under Research Question #2.

Evaluation Activities

The evaluation activities comprised all of the actions and/or practices that the evaluation team members engaged in to conduct this program evaluation. The

Year 5 evaluation activities covered three major areas: data collection, data

analyses, and reporting. Table 3 provides a list of the specific evaluation activities

conducted in Year 5 divided by major area.

Table 3.
Evaluation Activities for Year 5 of the SLC Evaluation.

Data Collection	Review/Revise data collection instruments
	Attend various district level meetings
	Coordinate with district staff on preparing surveys for online dissemination
	Coordinate with SLC Coordinators for time/participants for focus groups
	Online Surveys of Parents, Students, and Staff at each of 3 school sites
	Student Focus Groups at each of 3 school sites
	Teacher and Lead Teacher Focus Groups at each of 3 school sites
	Principal and SLC Coordinator Interviews at each of 3 school sites
Data Analyses	Clean survey data and translate into appropriate software
	Quantitative descriptive analyses of survey data
	Qualitative analyses of open-ended survey and interview/focus group data
	Quantitative/Qualitative analyses of changes across 5 years*
Reporting	Generation and dissemination of Year 5 Annual Report
	Generation and dissemination of Summative Report*
	Revisions of reports based on feedback from district
	Presentation of findings at district level meeting(s)

*These activities were specific to the Summative Report

For the most part, the evaluation activities presented in the table were

typical of the activities that took place during Year 4 as well as Year 5 of the SLC

evaluation. The only changes or additions to the evaluation activities during that

period were the addition of the summative analyses and report and the move from

paper surveys to online administration[2]. Most of the changes that did take place

occurred in the first three years of the evaluation.

---

[2] There was an earlier attempt to administer surveys online but due to limited district support, the process was not successful until later in the evaluation period.

The first major change took place in Year 1 with the development of the "theory of action" plans. Limited understanding of key stakeholders of the definition, development, and implementation of SLCs hindered the progress of this evaluation activity. Despite the district's attempt to invite content experts to assist stakeholders in this foundational process, challenges and confusion arose which led to the termination of the logic modeling process by the SLC Project Director. As the theories of action never fully developed, the evaluation design (as originally proposed) was modified slightly to achieve a broader set of goals, as opposed to school-specific goals.

The data collection instruments, their development and administration, also went through changes early in the evaluation process. In Year 1, district administrators declared that staff and student surveys had to accommodate the needs of the SLC evaluation as well as two other evaluations taking place at these sites. This resulted in the loss of many survey items designed to establish a baseline measure of various relevant SLC-related constructs. The surveys were revised in Year 2 to better meet the needs of the SLC evaluation. Furthermore, due to perceived challenges in attaining parent participation, parent surveys were not administered until Year 3 of the evaluation and only at two sites. This was not intentional but apparently due to a miscommunication between site and district.

In Years 2 and 3, both a fall and a spring version of the staff survey were administered. The fall survey was intended to focus on SLC implementation while the spring survey focused on SLC effects. However, logistical challenges in survey administration as well as greatly fluctuating response rates led to a single, more comprehensive spring staff survey.

Interview and focus group activities also saw some changes in the earlier years of the evaluation. In Year 2, teacher interviews were changed to teacher focus groups and remained that way thereafter. Student focus groups did not begin until Year 2 and had to be conducted in conjunction with another evaluation which was given district priority. Therefore, the evaluation team had little input into the question design and only reported summaries of findings. For the same reasons, the evaluation team was not able to conduct site administrator interviews during this year. Finally, given the challenges faced by the administration of the parent survey, the evaluation team determined that parent interviews/focus groups would not be a feasible data collection effort.

The most notable evaluation activity that never really got off the ground was the collection of district student data. As proposed, these data were to be collected to address numerous progress indicators related to student achievement, course enrollment, and graduation. Initially, there appeared to be some concern at the district level of sharing student information. In Year 3, the district agreed to release achievement data with the condition that it was linked to student surveys at the district's research office and presented to the evaluation team after identifying information had been removed. However, the resulting database was found to contain many errors and after several attempts to repair, it was deemed unusable by the evaluation team. After this year, no requests for student data were made.

With regard to evaluation reporting, an annual report was generated and disseminated for all five years of the evaluation. Supplemental reports were also intermittently prepared by the evaluation team in the form of interim and/or mini-reports, often initiated by the team to stimulate buy-in and use of findings. In

Years 2 and 3, site-specific mini-reports and SLC-specific mini-reports (Year 3) were generated and disseminated. In Year 4, the final report format was re-structured to facilitate interpretability, discussion, and use by individual school sites. Invitations to meetings to present and discuss findings appeared to begin as early as Year 2. However, an increase in collaboration and inclusion of the evaluation team in formal and informal discussions of findings began to take shape in Year 4 – which all coincided with the placement of a new district-level SLC administrator.

Outputs and Outcomes

In its simplest form, implementation of an evaluation involves the utilizing of inputs, taking into account external contingencies, and performing the evaluation activities. The outputs and outcomes of an evaluation are the results of this implementation. Therefore, in this SLC evaluation, although the outputs and outcomes were not "implemented," they were planned components of the evaluation with the intention that they would occur as a result of the implementation.

Outputs are generally the products or services provided to the clients, program participants, or other relevant stakeholders. In the SLC evaluation, the actual reports would be considered an output even though the generation of the report was an evaluation activity. Similarly, presentation of findings may be considered both an output, in terms of the service provided, and an evaluation activity. In this study, these outputs were viewed as part of the evaluation

activities all leading to potential outcomes. The rationale for this decision was largely to maintain consistency with Mark & Henry's model.

The outcomes described in a program logic model are generally the changes or benefits resulting from the program's activities and outputs. Outcomes may be sequential or chronological – occurring in the short-term, long-term, or somewhere in between (i.e., intermediate). The intended outcomes of an evaluation may be tied to the program and would likely be connected to the purposes of the evaluation. In this SLC evaluation, there were two overarching purposes of the evaluation as documented. The first purpose was *to understand* how SLCs were developed and implemented during the grant period. The second purpose was *to understand* the effects of SLCs on student achievement, attitudes towards achievement, culture, and climate at the school sites, and the structure and process of developing leadership capacity. Based on the findings generated in both the annual reports and the summative evaluation report, the evaluation team produced appropriate information for the stakeholders to "understand." However, the outcome was not that the information was available – that, in fact, was more of an output or a very short-term outcome. Rather, the outcomes were *evidence* of stakeholder understanding, specifically, actions and/or changes that took place as a result of the information provided. In the case of the SLC evaluation and this study, this evidence of understanding was indicated by some behavioral or cognitive/affective action or change. These outcomes (or processes), therefore, were indicators of the evaluation's influence.

In Year 5, the evaluation team built in measures of use and influence into the evaluation itself in order to determine whether or not these evaluation outcomes

had been achieved. Specifically, they added a small subset of items and questions

to the data collection instruments and presented these results in the summative

report. Comprehensive analyses and discussion of these outcomes, both behavioral

and cognitive, are presented under Research Question #3.

Mapping SLC Evaluation onto Mark & Henry's Model

The most meaningful way to discuss how the components of the SLC

evaluation mapped onto Mark & Henry's program theory of evaluation influence was

to first construct a new model incorporating and integrating the relevant (and

accessible) components from each (see Figure 2 on the following page). This figure

also serves as a graphic reference for this study's subsequent analyses and

discussion of results.

As shown, the inputs that emerged through the available evaluation data

correspond to Mark & Henry's general input categories. As previously discussed,

awareness of the evaluation is a "contingency in the environment" but since its

occurrence (or lack of) was not independent of the evaluation process, it was re-

categorized as an input. Analytically speaking, these three inputs serve as the

predictors of the evaluation outcomes. The SLC evaluation activities, as described

in the previous section, also correspond to Mark & Henry's model. However,

because this study focused on Year 5, the activities did not vary. In other words,

they occurred as described and there was no variation in implementation within the

year. Therefore, the activities did not serve as a variable or a predictor. They were

simply included in this model to serve as a reminder of where in the logic they fit.

Figure 2. Reconstructed Model of Influence after Mapping SLC Evaluation Components
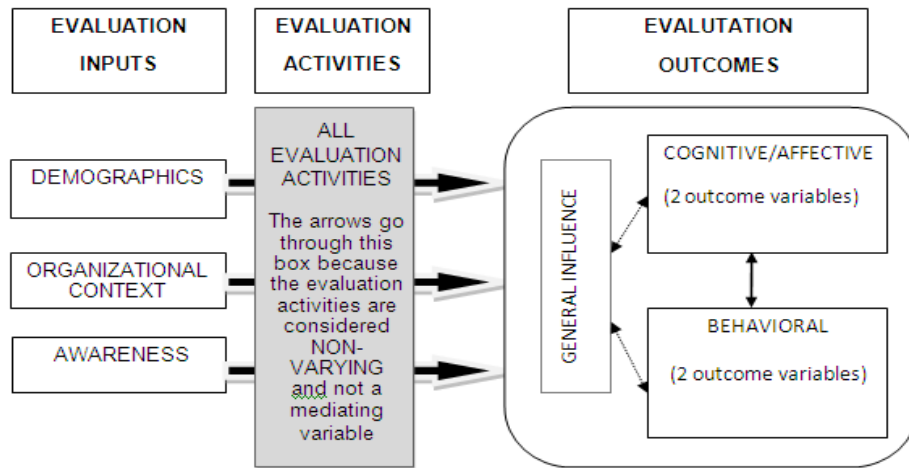
Conceptually, the inputs of an evaluation feed into the activities and variations in the inputs have a potential impact on the execution of the activities. For Year 5, all activities that were planned were also performed. Therefore, it was assumed that for Year 5, the activities were like a "constant" and did not mediate the potential effects of the inputs on the outcomes in any significant way. The thick solid arrows passing through from the inputs to the outcomes further illustrate this assumption. It is important to note that it is the evaluation findings (not the participation in the full range of evaluation activities) that, according to Mark & Henry's program theory, trigger influence. This distinction needs to be explicit given that the generation and dissemination of evaluation findings is subsumed under the heading of Evaluation Activities in both models.

Grouped under the heading of Evaluation Outcomes are three of Mark & Henry's proposed processes. The available SLC evaluation data provided this study with two quantitative outcome variables each pertaining to the Cognitive/Affective

and Behavioral Processes. General Influence outcomes were not directly measured but rather emerged through qualitative analyses of open-ended survey items and interview/focus group data.[3] Motivational Processes do not appear in this reconstructed model because there was no available data on this outcome. The dark bi-directional arrow between Cognitive/Affective and Behavioral outcomes indicates that the available data allowed for studying each as an outcome with the other as a predictor. This corresponded with Mark & Henry's notion that each indicator of influence had a potential for serving as an outcome and/or as a process leading to another outcome. The lighter bi-directional arrows between these two outcomes and General Influence processes indicate a similar iterative notion.

Mapping out the available data from the SLC evaluation also resulted in removing the sequential labels attached to the outcomes (e.g., short-term, intermediate, and long-term). The determination of long-term outcomes was beyond the scope of this study. The implication would be that the outcomes of the SLC evaluation were short term or intermediate. However, if the program theory of influence proposed that the achievement of evaluation outcomes is iterated through processes and underlying mechanisms, placing a pre-determined temporal label on these outcomes does not make sense. Therefore, in this reconstructed model, the temporal label was less important than their occurrence and the pathways to other outcomes.

Overall, the components of the SLC evaluation mapped onto the Mark & Henry's theory in a manner that allowed for this study of the proposed outcomes and pathways of evaluation influence. Knowledge Production Attributes, included in

---

[3] A comprehensive analysis and discussion of the outcomes is presented under Research Question #3.

Mark & Henry's model as an evaluation output (see Figure 1), was removed from the reconstructed model because there was no data collected or available from the SLC evaluation to consider this component. This and other components of Mark & Henry's model that were removed or adapted did not alter the goal of this study in any significant way. The available data from the evaluation allowed for both qualitative and quantitative analyses, albeit on a limited scale.

<div align="center">

Research Question #2

</div>

What are the "inputs" of this evaluation that may lead to and may impact potential underlying mechanisms, outcomes, and pathways to evaluation influence?

As previously described, the inputs in Mark & Henry's model refer to the evaluation context and the organizational/program context. Consistent with the traditional linear structure of program logic models, the arrows in the model leading from the inputs column indicate that these inputs contribute to the evaluation activities. The authors expanded beyond the linear structure by placing arrows that lead from the inputs directly to the outputs/short-term outcomes, hypothesizing that these inputs may have an impact on outcomes regardless of the evaluation activities. Furthermore, the authors posited that there were contingencies in the environment that may have an impact on the pathway from inputs to activities and outcomes. This is consistent with commonly used logic model templates that cite contingencies, or constraints on the program, as inputs (United Way, 1996). Building on this portion of Mark & Henry's model, the inputs have been re-conceptualized into three main categories – Demographics, Context, and

Awareness. These categories emerged from the available data and were determined to be satisfactorily representative of the setting, context, and environmental contingencies offered by the authors.

In response to this second research question, the formation of each of the three inputs is addressed in the following sections. Each section presents a discussion of definition and relevance of the input as well as the qualitative and quantitative analyses and results to be used in response to subsequent research questions. Specifically, this includes the methods and results for analyzing and restructuring the inputs to serve as quantitative predictors.

Demographics

Other than the evaluators themselves, the participants in the program, in this case, the school-level individuals who implemented and participated in the program, served as the primary human resource into the evaluation. Without this resource, the evaluation activities could not have taken place, as there would have been no program to evaluate. These participants also played an important role in the evaluation in that they participated in data collection activities, had access to evaluation findings, and had opportunities to provide feedback on the evaluation. Some of these participants also took part in evaluation planning, dissemination of evaluation products (e.g., reports, memos), discussions, and presentations of evaluation findings.

The demographic information was available at the school-site participant level and based solely on the Year 5 staff survey respondents. There were 272 respondents to the survey from the three study schools. Unlike the surveys from

previous years, there were no questions requesting gender or ethnicity responses.

Rather, demographic questions asked for position or current assignment, teaching

experience in years, and current grade level taught, if applicable. Table 4 presents

the distribution of the two demographic variables used as input predictors. The

other variables (i.e., teaching experience and grade level) were specific only to

teachers and not representative of the entire survey respondent group.

Table 4.
Demographic Information for Y5 Survey Participants (N=272)

	n	%
School		
School C	81	30%
School P	67	25%
School J	124	46%
Position		
Teacher	218	80%
Non-teaching or classified	14	5%
School Administrator - any type[1]	14	5%
Counselor	9	3%
Other - not described	9	3%
Missing (Unknown)	8	3%

[1]This includes principals, assistant principals, and coordinators

Context

Context was analyzed both qualitatively and quantitatively to provide a well-

rounded depiction of culture and climate, shared understanding and purpose,

administrator and teacher support, and level of collaboration among staff. The data

for the qualitative analyses of context came largely from administrator interviews,

teacher focus groups, and field notes.

In Year 5 of the evaluation, the discussion of context with school

administrators and teachers centered on the notion of buy-in, particularly staff buy-

in of the program and its sustainability beyond the grant. Therefore, the contextual

themes that emerged in these data relate largely to buy-in, whether improved or not, and the conditions or circumstances that were attached to it. Internal to each school, buy-in was thought to be related to the inner workings of the program and other logistical issues. Some administrators saw buy-in as related to building better relationships and promoting collaboration and therefore associated their efforts in these areas as working toward a more positive context. One principal, in particular, noted transparency in decision-making, shared discussions, and encouraged input as efforts made to increase buy-in.

Another important theme connected to buy-in was the common sentiment that the schools and staff had, as one teacher put it, "been through these hoops before." The belief was that the district imposed initiatives and programs continuously which tended to disappear when the funding was no longer available. This led to decreased buy-in and even resistance on the part of teachers, particularly veteran teachers. A few teachers spoke of buy-in being at its highest levels at the start of the initiative when the promise of things to come was new and exciting but dwindled off as structures were not put into place (or fell out of place) and larger problems emerged. The majority did not have buy-in because they were waiting for the district to "discard" this program for the next one, a pattern many had experienced before. It was shared that although staff had the opportunity to attend a workshop which emphasized the purpose behind the SLC program and that they finally "understood" the purpose and why the district promoted it, this shared understanding did not seem to be enough to improve the context – particularly the buy-in for the program.

Two of the three schools (School C and School J) were relatively new to this initiative whereas School P has had more experience with implementing this type of program. This important distinction is reflected in the responses offered by both administrators and teachers. Although all schools noted issues with buy-in, which may also be considered as impacting their context, the data yielded from School P showed a more flattened effect. In other words, the administrators and teachers noted some lack of buy-in but didn't seem as concerned as the other schools. Some teachers believed that the administrators could have been more assertive or direct in their efforts but lacked the fervency in their comments that some respondents from the other schools demonstrated. This is not to suggest that the context of School P was not positive, in general or in relation to the implementation of this initiative. On the contrary, as is shown in the subsequent section, School P scored highest on two quantitative measures of context. It was perhaps because of this positive context that the respondents went along with the flow and were not as reactive as some respondents from other schools.

__Context-related indicators__. Survey items related to context were used as the basis for creating composites or indices to serve as both predictors and control variables for quantitative analyses. A principal component analysis (PCA) was used to condense the information contained in these context-related survey items into one or two factors (indicators) with a minimal loss of information.

There were eight survey items in total. These items were selected based on their association with shared understanding, climate, and perceived support in relation to the SLC program. All if the items were scored on a 4-point scale ranging

from 1 ("Strongly Agree") to 4 ("Strongly Disagree"). Table 5 presents the survey

items, their mean scores, and standard deviations.

Table 5.
Means and Standard Deviations of Context-related Survey Item Scores Included in Principal
Component Analysis (N=200)

Survey Items	Mean	SD
C1. There is a climate of trust here among students, teachers, and administrators.	2.28	0.85
C2. The *faculty and staff* understand the purpose and goals for SLCs *at this school*.	2.03	0.67
C3. I understand the purpose and goals for small learning communities *at this school*.	1.57	0.67
C4. The faculty and staff at this school have opportunities to suggest modifications to the SLCs.	2.15	0.84
C5. The school provides time on a regular basis for SLC teams to meet to share information, discuss students' academic progress, curriculum needs, etc.	2.04	0.90
C6. I understand the vision and goals for *my particular SLC*.	1.78	0.78
C7. I understand what my role is in *my SLC*.	1.82	0.83
C8. I am comfortable with how *my SLC* has developed.	2.15	0.92

Reviewing the general assumptions for factor analysis, there were no major

normality or linearity violations. Table 6 presents the correlations for each of the

eight survey items listed in Table 5. As shown, all items were moderately

correlated with one another ($p<.001$).

Table 6.
Correlation Matrix of Context-Related Survey Items (N=200).

	C1.	C2.	C3.	C4.	C5.	C6.	C7.	C8.
C1.	1.000							
C2.	.486	1.000						
C3.	.265	.492	1.000					
C4.	.475	.451	.364	1.000				
C5.	.261	.348	.277	.450	1.000			
C6.	.240	.359	.675	.392	.252	1.000		
C7.	.265	.317	.635	.383	.152	.790	1.000	
C8.	.378	.412	.542	.458	.233	.651	.544	1.000

p<.001 for all of the above correlations

All eight items were entered into a principal component analysis using a Varimax method with no pre-set factor solution. The factors were rotated orthogonally using the Varimax procedure so that the relationships among the items were maintained. The Varimax method maximized the variance of loading within the factors across the items, increased the larger loadings and decreased the smaller loadings to facilitate interpretability, and spread the variance from the first factor to the smaller factors. The results indicated the presence of two factors with acceptable eigenvalues valued at 1.0 or greater (3.99 and 1.28). The first component accounted for the largest amount of the variance (50%) and the second component accounted for 16% of the variance. After rotation, the eigenvalues were 3.04 and 2.23 with variances of 38% and 28%, respectively.[4] Table 7 presents the rotated factor loadings for the eight survey items. Factor loadings equal to or greater than .630 were the pre-set values for inclusion of an item into the factor. This value is considered "very good" (Tabachnick & Fidell, 2007).

---

[4] Table A-2 in the Appendix contains the initial eigenvalues for all eight items.

Table 7.
Factor Loadings for Context-Related Survey Items (N=200).

Survey Items	Factor 1	Factor 2
C1. There is a climate of trust here among students, teachers, and administrators.	.159	**.736**
C2. The *faculty and staff* understand the purpose and goals for SLCs *at this school*.	.344	**.672**
C3. I understand the purpose and goals for small learning communities *at this school*.	**.774**	.271
C4. The faculty and staff at this school have opportunities to suggest modifications to the SLCs.	.329	**.722**
C5. The school provides time on a regular basis for SLC teams to meet to share information, discuss students' academic progress, curriculum needs, etc.	.028	**.712**
C6. I understand the vision and goals for *my particular SLC*.	**.893**	.163
C7. I understand what my role is in *my SLC*.	**.903**	.109
C8. I am comfortable with how *my SLC* has developed.	**.760**	.313

The two emergent factors represented two distinct, but related dimensions of context. The items that loaded onto the first factor represented the individual's understanding of the SLC program's vision, purpose, and goals, as well as their own role and comfort in how their specific part of the program had developed. The second factor was more global in that it represented how the faculty and school were supportive, demonstrated a trusting climate, and shared a common understanding of the program. The internal consistency coefficients (alpha) for each factor were .88 for factor 1 and .72 for factor 2.

To create the two new context-related indicators (CONTEXT1 and CONTEXT2) from these items, the scores were summed across the four items in each factor. The resulting variable scores ranged from 4 to 16. The basic descriptive

information for the two variables was CONTEXT1 (M=7.43, SD=2.81) and

CONTEXT2 (M=8.38, SD=2.40). Inspections of the histograms as well as measures

of skewness and kurtosis suggested that these variables were normally distributed.

Based on the original coding scheme for the individual items, the lower the variable

scores the more positive the context, as measured.

   Differences across schools. Looking at the two context indicators across

the three schools, significant differences were found. These differences provided

the rationale for including school as a control variable in the subsequent

quantitative analyses. Table 8 displays these findings.

Table 8.
Differences Among Schools on Context Indicators (N=200).

	M (SD)	F	p
CONTEXT1 - Individual		3.50	.032
School P	6.94 (2.81)*		
School C	8.08 (3.08)		
School J	7.31 (2.84)		
CONTEXT2 - School		21.27	.000
School P	7.39 (1.83)*		
School C	8.79 (2.32)		
School J	9.69 (2.68)		

   Post hoc analyses[5] revealed that for both CONTEXT indicators, School P

demonstrated the more positive context. This finding was counter to what would

be expected given the somewhat flattened responses and seemingly accepted lack

of buy-in from some of the teachers in this school. It was perhaps indicative of the

---

[5] Tukey's Honestly Significant Difference (HSD) Test

longer experience that this school has had with this type of program and the implied awareness that they were likely to sustain it beyond the grant. It also raised the question of whether their climate of trust, shared understanding of the purpose and their roles, and provided support will have an impact on evaluation influence, as they seemed to have a more complacent attitude toward buy-in of the program itself.

Awareness

Awareness signifies whether or not the stakeholders in the evaluation were aware that an evaluation was taking place and if they had knowledge of any findings generated from the evaluation. Although this contingency was not formally or explicitly hypothesized in the research questions, lack of or limited awareness of the evaluation, in its various forms, was considered an input which may have an impact on influence.

There were two questions on the staff survey that addressed awareness of the evaluation, both as an activity and evaluation findings. These questions asked respondents "When did you first learn that the SLC initiative was being evaluated?" and "What years were you aware of the SLC initiative evaluation findings?" For both questions, there were six multiple response options – the first four options listed specific years (e.g., Y1 (2005-06)), the fifth option acknowledged awareness but could not recall year, and the last option indicated that the respondent "never knew." The responses were recoded to Aware = 1, Not Aware = 0. Missing data was left blank. This recoding was in preparation for inclusion of these two variables in the subsequent quantitative analyses.

Table 9 presents the response percentages for each of these items as well as the new variable name. The first, AWARENESS1, corresponds to awareness of the evaluation as a general activity that took place for any one or more years of the evaluation period and AWARENESS2 corresponds to awareness of the evaluation findings for any or all years.

Table 9.
Response Frequencies of Evaluation Awareness Survey Items.

Survey Question (Variable name)	N	Aware*	Not Aware
When did you first learn that the SLC initiative was being evaluated? (AWARENESS1)	245	81%	19%
What years were you aware of the SLC initiative evaluation findings? (AWARENESS2)	243	69%	31%

*Note: Aware is an aggregated category for all responses to these survey questions that indicate awareness except the "Never Knew" response.

Teachers comprised most of the respondents who reported either not being aware of the evaluation in general or of the evaluation findings. Although all of the school administrators that completed the survey indicated that they were aware of the evaluation in general, 29% (4 of 14) reported that they were not aware of the evaluation findings at all. The same pattern was seen among the counselors – 29% (2 of 7) were unaware of the findings despite all reporting awareness of the evaluation.

In order to gain a more informative understanding of how awareness was generated, survey respondents were presented with two open-ended questions asking them to explain *how* they find out about the evaluation in general and the evaluation findings. Only 33% of the 272 survey respondents provided answers as

to how they found out about the SLC evaluation in general. Responses generally described either one or more individuals as the source of the information (42 responses) or a setting in which the information was transmitted (71 responses). The majority of the responses that indicated one or more individuals as the source of awareness listed those closest to the SLC program at the school site: SLC Lead Teachers (24% of 42) and SLC Coordinators (36%). School or district administrators were the second most common sources. The three most commonly described settings in which awareness occurred were some type of meeting (e.g., staff meeting, SLC meeting) (37% of 71), actual participation in one or more of the evaluation activities (e.g., survey, focus group) (24%), or via email (17%). Additionally, four respondents indicated that they took part in some portion of the program's initial development and therefore learned of the evaluation component.

In response to how participants became aware of the evaluation findings, only 19% (of 272) offered a description. Similar to the above pattern of responses, the most common responses were of an individual closest to the SLC program as a source (25% of 56 responses) and/or some type of meeting as the setting in which they learned of the evaluation findings (38%). Other notable sources included the evaluation report and email communication.

Data from teacher focus groups also followed similar response patterns. All 38 focus group participants expressed an overall awareness that an evaluation had taken place. As for awareness of any evaluation findings, those teachers closest to the program and its implementation (i.e., SLC Lead Teachers) were the most aware and most able to describe how they became aware. The most common sources were sharing information at meetings, reading reports (in their entirety or particular

sections), and through discussing with others (i.e., SLC Coordinators or other lead teachers). Other teachers who said they were aware of some evaluation findings varied in their descriptions of how they became aware. For example, several teachers indicated that they "may have seen" information in an email, in a meeting, or in a report. At least two teachers mentioned being invited by an administrator to "come over" and look at the findings if they were interested.

Overall, this examination of awareness as an input that may impact influence revealed that the further away the participants (or those involved in the evaluation) were from the program's implementation, the less aware they were of the evaluation's findings. Despite many teachers' acknowledgements of their awareness of the evaluation or its findings (as shown by the survey responses), their ability to describe how they became aware or in what setting seemed vague.

Research Question #3

What are the underlying processes and outcomes that emerged throughout the course of this evaluation?

In response to this research question, the focus is on the "outcomes" portion of the model(s) and all of the data subsumed under this heading. These outcomes, processes, or indicators of influence, are based on action and/or change, whether behavioral or cognitive. Both the staff survey and interview protocols contained items and questions designed to gather information on these outcomes. The subsequent sections describe the methods for identifying these outcomes in the data, through both analyses and interpretation. The sections are divided by the format of the data and the analyses – quantitative and qualitative.

Quantitative Indicators of Influence

The quantitative indicators of influence were derived from staff survey responses and only available in dichotomous format. There were four questions on the survey directly related to changes made based on the influence of the evaluation. The original response options as offered on the survey differentiated these changes by year. The responses were recoded into a "Yes" or "No" response to provide an overall indication of change. Blank responses were left blank, as it could not be assumed that an absence of response meant no change. The dichotomous nature of the responses indicated whether changes occurred or not, according to the respondent, and did not describe or explain what the changes may have been.

The first two questions represented changes that fall into Mark and Henry's category of Behavioral processes (see Table 10). These processes refer to actions taken or changes made as influenced by the evaluation. The first question asked respondents if the evaluation findings were used to make changes to their SLCs.

Table 10.
Response Frequencies of Outcome Items Related to Behavioral Processes.

Survey Item	N	Yes	No
Did you use any of the evaluation findings to make changes to your SLCs?	233	36%	64%
Did you use any of the evaluation findings to make changes to your own practice?	228	53%	47%

As Table 10 shows, 36% of those who responded indicated that they made a change. The second question asked if changes were made to their practice based

60

on the evaluation findings. More than half of the respondents (53%) reported that they had. Although both questions represented a general behavioral change, it was interesting to note the difference in the number of positive responses. Making "changes" to an SLC is considerably more difficult because its structure may only be changed at a collective and/or administrative level. These respondents may have worked as part of a collective group or have made small, indirect changes to aspects of their SLCs. Alternatively, making changes to one's own practice falls within the realm of an individual's control. If this type of behavioral change was to be influenced by the evaluation, it would likely be at this individual level, hence the greater percentage of positive responses.

The second group of questions represented changes that fall into the category of Cognitive/Affective Processes. These processes refer to changes in thoughts, opinions, perceptions, and/or feelings as influenced by the evaluation. The questions asked respondents if the evaluation findings led to changes in the way they thought about SLCs and school reform in general. For both questions, over half of the responses indicated that changes in thinking had occurred. The disparity in responses seen in the above two questions did not appear here, as shown by the response frequencies in Table 11.

Table 11.
Response Frequencies of Outcome Items Related to Cognitive/Affective Processes.

Survey Item	N	Yes	No
Did the evaluation findings change the way you think about SLCs?	220	58%	42%
Did the evaluation findings change the way you think about school reform in general?	221	53%	48%

Similar numbers of respondents indicated that they had made some type of cognitive change based on the evaluation findings. This similarity was probably due to the fact that changes at the cognitive/affective level are not bound by what can actually take place. It was therefore possible that the evaluation had a greater influence on cognitive change.

Qualitative Indicators of Influence

There were two primary sources that contained more in-depth qualitative data on the nature of these influence outcomes. It was through analyses of these data that the specificity of the outcomes and their level of focus emerged. These sources included four open-ended survey questions and a subset of focus group questions.

Two of the four survey questions asked respondents to describe how they used the evaluation findings to make changes – both to their SLCs and to their practice. Both of these questions attempted to capture behavioral actions or changes that were influenced by the evaluation's findings. The last two survey questions asked respondents to describe how the evaluation findings changed the way they "think" about SLCs and about school reform in general. These two questions attempted to capture cognitive and/or affective processes or changes that were influenced by the evaluation findings. Only 25% (or 68) of the survey population (N=272) provided meaningful responses to one or more of the above questions. It was for this reason that the responses could not be subsequently recoded for inclusion in the quantitative analyses. Of the 25% who responded, the

vast majority were teachers (85%). This is in proportion with the larger distribution of survey respondents.

Multiple rounds of qualitative analyses were conducted on these data. The coding process began with first flagging all meaningful responses that addressed one of the four questions asked. Although Mark & Henry provided potential, specific indicators of evaluation influence as part of their program theory,[6] these indicators were not utilized as codes during the first few rounds of analyses. The primary rationale for this was because it was anticipated that these survey data might not correspond directly to the proposed indicators yet still yield emergent outcomes related to influence and use that may be otherwise missed, particularly since the number of responses was small and the questions were very broad. Therefore, the only categories employed in the initial analyses corresponded to the overall process referred to by the question – Behavioral or Cognitive/Affective. Many responses had to be re-located into the appropriate category as some responses provided an "action" when asked for a description of cognitive outcome and vice versa.

For the next round, responses were coded as "individual" or "interpersonal." These two codes corresponded to Mark & Henry's levels of analysis denoting the location of influence. If the respondent provided a description of something that "I" did, thought, or changed, it was coded as "individual." If the respondent provided a description of something they did as part of a group (i.e., "we" did, thought, or changed), it was coded as "interpersonal." The interpersonal level was described by Mark & Henry as changes that occurred as a result of interactions between

---

[6] Refer to Table 1, p.20.

individuals. The authors intended these levels of analysis (and the corresponding processes) to be analogous and not suggestive of greater magnitude.

Also in this second round, the responses were reviewed to distinguish those in the Behavioral or Cognitive categories from what may be considered General. As the authors proposed, General Influence processes may be thought of as a short term outcome or as a stepping stone to other outcomes of interest. Its importance can be judged by whether or not it leads to other outcomes. However, one could argue that any outcome has the potential to lead to other outcomes. Therefore, this recoding was largely a judgment call I made based on the broadness or limited specificity of the response. Responses that indicated a "focus" on something was coded as a General process. Also, such actions that were expressed in terms of planning, discussing, meeting, looking at, or updating were also coded as General Influence processes.

The third round of analyses involved a closer look at the Cognitive/Affective responses. These responses were coded to reflect such cognitive changes as reinforcing/validating opinion of program, increased awareness of program effects, greater understanding of what was needed for program to be successful, and so forth. The final round of analyses sought to find indications of processes that may have led or influenced others, or pathways. The results of this round are discussed under Research Question #4.

Regarding how the evaluation influenced behavioral action or change to the SLCs, early rounds of analyses showed that 26 respondents provided a total of 32 responses with descriptions of these actions or changes made to the SLC program. The majority occurred at the interpersonal level of influence. The final analyses

revealed that more than half of the responses (58%) fit better under the category of General influence processes than under Behavioral processes as they were either very broadly described and/or judged to be short-term in leading to a more intermediate or long-term outcome. The 42% of the responses that remained as Behavioral process outcomes were indicative of actions or changes that took place at the interpersonal level as they involved the respondent interacting with others. Examples of reported change/action included changing the name and direction of the SLC, modifying curriculum to "match SLC's direction and mission," and "set smart goals" to improve the SLC. There was one response that indicated change to the SLC at the individual level – "suggested service learning projects with SLC's vision and mission."

The majority of responses indicating General influence also took place at the interpersonal level. These actions mostly addressed "focus" (e.g., "helped to focus our academy"), modifying or updating (e.g., "we used to modify our program"), and planning or discussing (e.g., "discussed in our meetings). There were two broadly described responses that appeared to take place by others but it was unclear as to whether the respondent interacted or simply noted others' as the performers of change. Examples included "meetings and efforts were planned carefully by the administration" and "[findings] were used by administrators."

There were 39 respondents who reported that, at the individual level, the evaluation findings influenced an action or change to their practice. These reported actions or changes ranged in their descriptive specificity. Examples of broad responses included "connected my teaching more to my SLC objectives" and "conferenced with more of my students." Examples of more specific responses

included "used information to assist with intervention counseling for students," "talked more [with students] about going to college and the A-G requirements," and "made changes to my curriculum to incorporate more stratification of learning for struggling students." There were four respondents who reported influence at the interpersonal level. For example, one respondent reported forming a "grade-level team with my other core content teachers and began to pilot shared and complementary lessons." This influence occurred at the interpersonal level because interacting with others led to this new behavioral action. This was the most specific response at the interpersonal level – the other three were broad (e.g., "we [teachers] are sharing input on shared students).

Overall, the influence of the evaluation on behavioral action and/or change occurred primarily at the interpersonal level when it concerned making changes to the larger program and occurred at the individual level when it concerned making changes to practice. Furthermore, these descriptions of behavioral action/change corresponded to two of Mark & Henry's proposed indicators of influence (i.e., individual change in practice and collaborative change in practice). Some of the indicators of influence that fell out of the Behavioral category and into the General category also corresponded to Mark & Henry's proposed indicators. For example, at the interpersonal level, the proposed "change agent" outcome is when the findings influenced work toward organizational change. In the data, several of the reported actions were of staff working together to change and/or improve the program (e.g., planning, meeting, and discussing). In fact, one could argue that this "change agent" outcome should not be limited to the General category as working toward organizational change may also be considered an intermediate outcome. For

66

example, there were many reported actions in the Behavioral category that depicted staff working together on specific actions targeted toward organizational change (e.g., "[We] lead teachers analyzing data and implementing changes on a school-wide basis").

There were two survey questions regarding the influence of the evaluation on cognitive/affective change. The first question asked how the evaluation influenced thinking about SLCs. Early rounds of analyses showed that 43 respondents provided a total of 51 responses indicating the nature of these thoughts or changes in thinking about the SLC program and almost all took place at the individual level. Similar to the analyses of behavioral processes, the final analyses revealed that more than half of the responses (55%) fit better under the category of General Influence processes than under Cognitive/Affective processes as they were either very broadly described and/or judged to have the potential to lead to other outcomes.

Analyses of the 45% of the responses that remained under the category of Cognitive/Affective processes revealed that the reported cognitive change emerged as an expanded understanding and/or awareness of the SLC program – including the effects of the program, what was needed for the program to be successful, their roles or the roles of others in the program, others' experiences in the program, the limitations of the program, and the function/purpose of the program. I judged these reports of expanded knowledge/awareness were judged to be evidence of cognitive change even though it may be argued that awareness was a General Influence process (or short-term outcome) that may lead to knowledge (an intermediate outcome). However, since the interpretation was bounded by the

67

specificity of the response, it was very difficult to distinguish between the two cognitive functions. Mark & Henry proposed two indicators that corresponded to this emergent outcome of expanded knowledge/awareness – elaboration (a General Influence process) and attitude change (a Cognitive/Affective process). Elaboration implies that the evaluation influences individuals to think more about the program and their expectations for its outcomes. Essentially, elaboration could lead to an attitude change about the program. As stated, the data did not allow for these distinctions to be made and therefore the judgment was made to keep these cognitive indicators of influence together.

Those responses of cognitive change related to SLCs that fell into the General Influence category also took place at the individual level. One prominent indicator of influence that emerged from the analyses was the reinforcement or validation of a previously held opinion or perception. In their model, Mark & Henry proposed an indicator termed "justification" which implies a similar process, although they suggested that this process takes place at the interpersonal level. These findings suggest that a similar process occurs at the individual level as well. Other themes that emerged within this General category included the formation of negative opinions about the program and its effects and negative opinions of the findings' relevance to the individual. In this latter group, responses suggested that because of their roles (e.g., elective teacher, RSP teacher), the findings, and sometimes the program itself, was not particularly relevant to them. Across the remainder of the responses, some addressed a "focus" on certain aspects of the program, some offered positive opinions of the program, and some suggested a specific cognitive activity (i.e., "reflected on my teaching this past year"). The variety of content in

these responses made it difficult to categorize them any further. There were two responses that indicated change at the interpersonal level. One was cognitive and expressed a positive opinion as occurring among a group of teachers (e.g., "We believe we are making progress."). The second was affective in that the evaluation findings helped to influence "our SLC feel more like a team."

For the last outcome question, how the evaluation influenced thinking about school reform in general, early rounds of analyses showed that 14 respondents provided responses indicating the nature of these thoughts or changes in thinking about school reform and almost all took place at the individual level. The emergent categories and response patterns were similar to the above findings of cognitive change related to SLCs. Half of the responses fit better under the category of General influence processes than under Cognitive/Affective processes. Within this latter category, responses that described an expanded awareness or understanding of what was needed for reform to be successful or the benefits of reform were considered Cognitive/Affective indicators of influence (as opposed to General Influence). Within the General influence category, some responses were "justification" (i.e., reinforcing or validating previously held opinion), while the remainder was of negative opinions of how the school implemented reform or doubt that the school could successfully implement any true reform.

Given the smaller number of participants in the focus groups and interviews and the non-systematic method in which questions were asked, it was not feasible to apply the same analytic procedures with these data. Rather, the transcripts were coded for any incidence of action as influenced by the evaluation and whether or not it supported the survey findings. The analyses revealed that those in

positions closest to the SLC program and its implementation (i.e., SLC Coordinators, SLC Lead Teachers) had the most to report about use of the evaluation findings and any changes influenced by the findings, whether at the individual or interpersonal levels. School principals provided less detailed descriptions of this influence. Overall, the analyses yielded individual and interpersonal actions corresponding to the General Influence process – discussions, review, questioning of implementation, sharing, and reflection. There were a few reported actions emerged as Behavioral processes at the interpersonal level. For instance, at one school the evaluation findings influenced those closest to the program to create and set three goals for the year and drive their implementation monitoring. Additionally, it was reported that the evaluation findings influenced SLC Lead Teachers to make changes as a group (e.g., collaborate more and function more uniformly). There were no responses which suggested cognitive change other than what was placed in the General Influence category (e.g., reflection).

Non-lead teachers participating in the focus groups varied in their responses ranging from those who reported not being aware of the findings to those who reported use by others. The questions were not asked systematically across these teacher focus groups or across the schools which made it difficult to present any findings related to use or influence for this group.

Research Question #4

To what extent do these processes/outcomes link together to form

"pathways of influence"?

It is in response to this research question that the model is "tested."

Quantitatively, the relationships and pathways to and across the evaluation

outcomes are examined. Additional qualitative analyses and results supplement

these findings to provide a more meaningful depiction of these pathways.

Quantitative Results

Logistic regression equations were used for two purposes: 1) to determine

which of the three sets of inputs (Demographics, Awareness, and Context) were

significant predictors of the outcome indicators of influence (both Behavioral and

Cognitive/ Affective); and 2) controlling for these inputs, to determine if one set of

outcomes (i.e., Behavioral) could predict the likelihood of another set of outcomes

(i.e., Cognitive/Affective). Because General Influence outcomes were only

emergent through the qualitative analyses of a small subset of data, these

outcomes were not included in any quantitative analyses.

The first round of logistic regression analyses placed all six predictor

variables (two for each set of inputs) into an equation for each of the four

outcomes. Table 12 lists the predictor variables. The predictors were entered into

the regression model in three separate blocks so that each block contained the two

variables corresponding to the particular input.

Table 12.

Name and Description of Input Predictors in Logistic Regression.

Demographics	
School	School Affiliation
Position	School Assignment
Awareness	
AWARENESS 1	Awareness of evaluation as an activity
AWARENESS2	Awareness of evaluation findings
Context	
CONTEXT1	Individual-level understanding of vision, purpose, role in the program
CONTEXT2	School-level support, climate, shared understanding

Since logistic regression analyses deleted cases with missing data list wise, several cases were dropped from the model. For all of the models, the resultant samples sizes were shrunk to 61% to 65% of their original sizes (N=272). A discussion of this dropped data is presented at the end of this section.

The models were all significant and across all four, the same two variables emerged as statistically significant predictors of the odds of the particular outcome occurring. The first significant predictor, AWARENESS2, represented the individual's awareness of the evaluation's findings. The second significant predictor, CONTEXT1, represented the individual's understanding of the program's vision, purpose, and goals, as well as their role and comfort in how their particular part of the program had developed. The results of the four regression models are presented in Tables 13 through 16 on the following pages.

Table 13.
Logistic Regression Results for Inputs into the Evaluation Influencing the Odds of Behavioral
Changes to the SLC Program (N=177).

	Findings influenced changes to SLCs	Findings did not influence changes to SLCs			
	N	N	OR[1]	SE	p Value
Demographics					
School					
School C	25	41	1.44	0.45	0.42
School J	15	26	1.98	0.60	0.26
School P	29	41	◊		
Position					
Teacher	58	94	3.30E8	20370.74	0.99
Classified	3	3	5.79E8	20370.74	0.99
School Admin	4	3	4.77E8	20370.74	0.99
Counselor	3	4	2.62E8	20370.74	0.99
Other	1	4	◊		
Awareness					
AWARENESS1	69	108	5.17E8	7484.07	0.99
AWARENESS2	69	108	**15.39**	0.79	**0.001**
Context					
CONTEXT1	69	108	**0.78**	0.09	**0.01**
CONTEXT2	69	108	0.87	0.11	0.21

Note: ◊ indicates referent group
[1]Odds ratios computed from one regression model

Table 14.
Logistic Regression Results for Inputs into the Evaluation Influencing the Odds of Behavioral
Changes to Practice (N=175).

	Findings influenced changes to practice	Findings did not influence changes to practice			
	N	N	OR[1]	SE	p Value
Demographics					
School					
School C	33	34	0.46	0.49	0.12
School J	22	20	0.56	0.61	0.35
School P	46	20	◊		
Position					
Teacher	89	61	8.89E8	21561.38	0.99
Classified	3	3	6.44E8	21561.38	0.99
School Admin	4	3	3.74E8	21561.38	0.99
Counselor	5	2	1.15E9	21561.38	0.99
Other	0	5	◊		
Awareness					
AWARENESS1	74	101	2.62	0.80	0.23
AWARENESS2	74	101	**13.10**	0.53	**0.000**
Context					
CONTEXT1	74	101	**0.76**	0.10	**0.003**
CONTEXT2	74	101	0.97	0.11	0.81

Note: ◊ indicates referent group
[1]Odds ratios computed from one regression model

Table 15.
Logistic Regression Results for Inputs into the Evaluation Influencing the Odds of Cognitive Changes About SLC Program (N=170).

	Findings influenced thinking about SLCs	Findings did not influenced thinking about SLCs			
	N	N	OR[1]	SE	p Value
Demographics					
School					
School C	37	28	0.60	0.46	0.26
School J	19	21	0.51	0.57	0.24
School P	46	19	◊		
Position					
Teacher	90	56	1.28	1.51	0.87
Classified	3	3	0.81	1.74	0.90
School Admin	5	2	1.09	1.74	0.96
Counselor	2	5	0.15	1.77	0.29
Other	2	2	◊		
Awareness					
AWARENESS1	102	68	1.34	0.65	0.65
AWARENESS2	102	68	**5.40**	0.49	**0.001**
Context					
CONTEXT1	102	68	**0.82**	0.08	**0.02**
CONTEXT2	102	68	0.92	0.10	0.41

Note: ◊ indicates referent group
[1]Odds ratios computed from one regression model

Table 16.
Logistic Regression Results for Inputs into the Evaluation Influencing the Odds of Cognitive Changes About School Reform (N=167).

	Findings influenced thinking about reform	Findings did not influence thinking about reform			
	N	N	OR[1]	SE	p Value
Demographics					
School					
School C	34	30	0.76	0.43	0.53
School J	18	23	0.62	0.53	0.37
School P	41	21	◊		
Position					
Teacher	82	62	1.90	1.35	0.63
Classified	3	3	1.33	1.59	0.86
School Admin	4	2	1.79	1.62	0.72
Counselor	2	4	0.51	1.64	0.68
Other	2	3	◊		
Awareness					
Awareness1	93	74	1.02	0.62	0.97
Awareness2	93	74	**3.45**	0.46	**0.01**
Context					
CONTEXT1	93	74	**0.84**	0.08	**0.03**
CONTEXT2	93	74	0.89	0.10	0.24

Note: ◊ indicates referent group
[1]Odds ratios computed from one regression model

As Table 13 shows, the odds of the findings influencing behavioral changes to SLCs were over 15 times more likely if there was an awareness of the evaluation findings (OR=15.39). Similarly, in Table 14, a similar pattern was found for the odds of the findings influencing behavioral changes to the individual's practice (OR=13.10). For the outcomes related to cognitive influence, the odds of the findings influencing changes in thinking about SLCs were over five times as likely for those who were aware of the findings (OR=5.40) while the odds of the findings influencing changes in thinking about school reform in general were three times as likely for those aware of the findings (OR=3.45 – see Tables 15 and 16). Tables 13 through 16 also show that the odds of any influenced change were less likely if the context was less than positive. In fact, the odds ratios were relatively consistent in magnitude across all four regression models, ranging from 0.76 to 0.84.

The second round of logistic regression analyses sought to test if one set of outcomes (i.e., Behavioral) could predict the likelihood of another set of outcomes (i.e., Cognitive/ Affective). As before, four models were run for each of the four outcomes. In addition to the two significant "inputs" from the first round of analyses, the three outcomes were also included as potential predictors of the fourth outcome. Table 17 shows the results of the two regression models that included the behavioral outcomes.

Table 17.
Logistic Regression Results (Second Round) for Both Behavioral Outcome Models (N=171)

Predictors	Findings influenced changes to SLCs			Findings influenced changes to practice		
	OR[1]	SE	p Value	OR[1]	SE	p Value
AWARENESS2	**7.98**	0.68	**0.02**	**5.52**	0.54	**0.001**
CONTEXT1	**0.82**	0.09	**0.002**	0.88	0.09	0.16
Behavior						
Findings influenced changes to SLCS	---	---	---	**3.49**	0.52	**0.02**
Findings influenced changes to practice	**3.48**	0.52	**0.02**	---	---	---
Cognitive/Affective						
Findings changed thinking about SLCs	2.13	0.61	0.22	**6.49**	0.58	**0.001**
Findings changed thinking on reform	1.34	0.56	0.60	1.94	.058	0.25

[1]Odds ratios computed from one regression model

As may be seen in Table 17, controlling for the significant inputs of context and awareness, the odds of the findings influencing behavioral changes to SLCs were over three times more likely if the findings also influenced changes to one's practice (OR=3.48). Table 17 also shows the reciprocal relationship within the Behavioral process – that the odds of the findings influencing behavioral changes to one's practice were over three times more likely if the findings also influenced changes to SLCs (OR=3.49). Furthermore, controlling for the inputs, the odds of the findings influencing behavioral changes to one's practice were over six times more likely if the findings also influenced changes to thinking about SLCs (OR=6.49). This particular result provided evidence to support a "pathway" of influence from a process to an outcome.

Table 18 shows the results of the two models that included the cognitive outcomes. As shown in the table below, controlling for the two inputs, the odds of

the findings influencing thinking about SLCs were over 28 times more likely if the

findings also influenced changes in thinking about school reform (OR=28.82). The

odds of the findings influencing cognitive change about SLCS was over six times as

likely if the findings also influenced changes to one's practice (OR=6.65). This

finding also supports the same pathway as described above but in the opposite

direction, providing evidence of the bi-directionality of the pathway as proposed.

Table 18 also shows that within the Cognitive process, the odds of the findings

influencing thinking about school reform were over 27 times more likely if the

findings influenced thinking about SLCs.

Table 18.
Logistic Regression Results (Second Round) for Both Cognitive Outcome Models (N=171)

| | Findings changed thinking about SLCs | | | Findings changed thinking about reform | | |
	OR[1]	SE	p Value	OR[1]	SE	p Value
AWARENESS2	1.58	0.61	0.46	0.84	0.61	0.78
CONTEXT1	0.98	0.10	0.85	0.90	0.09	0.24
Behavior						
Findings influenced changes to SLCS	2.52	0.63	0.14	1.35	0.55	0.59
Findings influenced changes to practice	**6.65**	0.56	**0.001**	1.83	0.58	0.30
Cognitive/Affective						
Findings changed thinking about SLCs	**---**	---	**---**	**27.60**	0.53	**0.00**
Findings changed thinking on reform	**28.82**	0.54	**0.00**	---	---	---

[1]Odds ratios computed from one regression model

Figure 3 presents a graphic depiction of these analytic results. As posed by

Mark & Henry, the bi-directionality of the arrows implied iteration as one outcome

may lead to another and back again. As shown, within each major category

(Behavioral and Cognitive processes) and across categories, the indicators of

influence formed bi-directional pathways in which they served as both outcomes

and processes leading to another outcome. The within category pathways, although not explicitly accounted for in the author's model, offer further supporting evidence for the underlying mechanisms of evaluation influence. As tested in this study, all of the arrows show the relationships as the odds of the outcomes' occurrence.

Figure 3. Graphic presentation of the results from Round 2 logistic regression analyses.
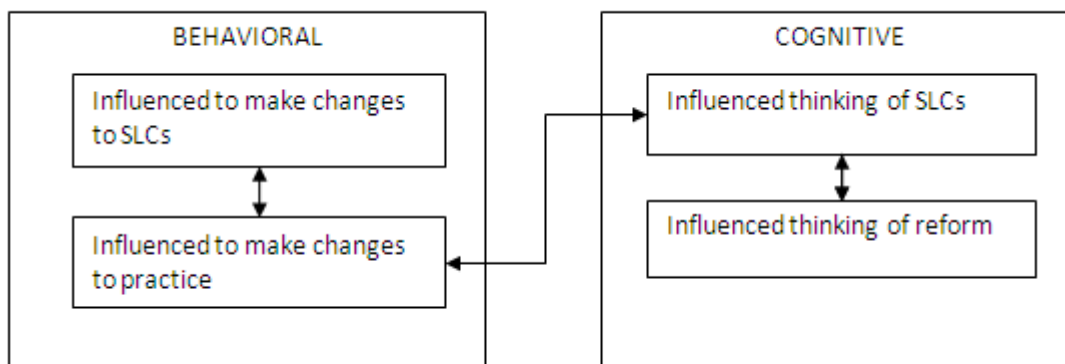
Missing data. The logistic regression analyses deleted cases with missing data listwise, resulting in several cases being dropped from the models. For all of the analytic models, the resultant samples sizes were shrunk to 61% to 65% of their original sizes (N=272). Looking closely at the data revealed that 37% (n=101) of the respondents were missing at least one of six predictor variables and/or at least one of four outcome variables. Further review of cases with missing predictor data did not show any systematic pattern and although these data may not be necessarily missing completely at random, the review yielded no reason to believe that the "missing predictor" sample was really different from the "complete predictor" sample of respondents.

In order to examine whether or not there were meaningful differences between the respondent sample with "missing outcome data" and the sample with "complete outcome data," the cases were first flagged if an outcome variable was missing. Cross-tabulations and chi-square analyses were run to see if the two groups were proportionately different on the demographic variables available from the survey. The information on years teaching and grade level taught were not included as they only applied to staff with current teaching assignments.

The only significant difference that emerged between the two groups of respondents was in the distributions of staff positions (χ^2=25.15, p=.000). Teachers were the primary respondent category in both groups; however, the differences seemed to be the most evident in the proportions of the classified and other (non-described) positions. In other words, in these job positions, there were twice as many respondents with missing outcome data than without. Although it cannot be stated definitively, it may be possible that some respondents in these positions left these items blank because of their perceived distance from the implementation of the program and/or the evaluation. While it may be hypothesized that had these respondents provided responses, they would have been negative or indicate reflecting the absence of influenced change, it was difficult to tell. Furthermore, the small sizes of these particular respondent groups would not have impacted the analytic results significantly.

Qualitative Results

The outcomes derived from the open-ended survey questions were also analyzed to determine if there was any evidence of pathways between outcomes.

Of the 68 respondents that provided at least one response to these open-ended questions, five cases suggested the emergence of a pathway. In two of the cases, an indicator of General influence led to Cognitive/Affective change. In the first, the respondent was at first "bothered" by the focus of the program, as it did not seem to match the finding's recommendations (General). This developed into an expanded understanding of what was needed for the program to be successful (Cognitive/Affective). In the second case, the respondent "focused" on certain aspects of the program (General) and subsequently expanded awareness of the potential effects of the program (Cognitive/Affective). The other three cases revealed pathways within the General influence process. For example, one respondent reported that the evaluation influenced communication with other SLC teachers which helped them "feel more like a team." Another respondent reported that the evaluation findings showed that SLCs were "here to stay" which led the teacher to increased involvement in his/her SLC.

Research Question #5

To what extent does the evaluation help explain the potential underlying processes, outcomes, and pathways to evaluation influence? In what ways does this study offer preliminary evidence to support Mark & Henry's theory of evaluation influence?

This research question, with both its parts, is where all of the findings come together. The SLC evaluation served as the milieu with which to "test" or explain Mark & Henry's program theory of evaluation influence. The logic model format of the program theory facilitated the mapping of the SLC evaluation and its

80

components. Referring back to the reconstructed model from Research Question #1,[7] I have established that the available data from the evaluation adequately corresponded to the model components in so much that testing the emergence of outcomes and potential pathways was possible.

The two inputs into this evaluation that had a direct impact on the evaluation outcomes were context and awareness. Specifically, the findings showed that the greater the individuals' understanding of the SLC program's vision, purpose, and goals as well as their own role and comfort in how their specific part of the program had developed, the greater the odds of action and/or change as a result of the evaluation findings. Interestingly, the more global context (shared understanding with others, perceived support from administrators and staff, and demonstration of a school-wide trusting climate) did not have a direct impact on the evaluation outcomes. Significant school differences in context did not have an impact on the evaluation outcomes, as school affiliation was not a significant demographic input. In general, this link between measures of the organizational context and the evaluation outcomes add to the current literature on the effects of context. Although it has been concluded by some (e.g., Pechman & King, 1986) that due to contextual factors in school district evaluations use may not be immediately evident, these findings show that targeted study of various indicators of influence may capture more immediate effects.

Awareness of the SLC evaluation findings also led to the odds of these findings influencing evaluation outcomes. Though this seems a rather obvious statement, it is distinguished from the findings that awareness of the SLC

---

[7] Refer to Figure 2 on page 45.

evaluation in general did not have any impact. In other words, awareness that an evaluation of the SLC program was taking or had taken place was not enough to impact the odds of the findings influencing change. One interpretation was that the SLC evaluation did not result in process use; in other words, the awareness of or engagement in the evaluation process was not sufficient to lead to the evaluation outcomes. However, Mark & Henry's position was that process use could not be translated into the model because the proposed influence processes/outcomes were triggered by evaluation findings not process. These findings correspond to that position. Ultimately, stakeholder knowledge (or at least awareness) of the evaluation findings promoted the odds of action and/or change.

The explanatory power of these findings is relevant in offering evidence to support Mark & Henry's model but remains at a surface level. What the quantitative findings could not explain was the impact on 1) the outcomes characterized as General influence, 2) the location of influence (i.e., individual or interpersonal), and 3) specific emergent outcomes (i.e., elaboration, change agent). This was largely due to the fact these outcome distinctions were based on the qualitative analyses of a small sample of data. Because the sample sizes were small, quantifying these data and including them in the regression analyses was not possible. Furthermore, through focus group and interview data, it was established that teacher buy-in (as an indicator of context) was a challenge to moving SLC program implementation forward. What could not be shown through these analyses was whether buy-in impacted the odds of action/change as a result of the evaluation findings.

The outcomes of the evaluation (as measured in this study) were determined both through the evaluation's pre-existing data and through analyses of qualitative data. The findings gave forth not only evidence to support the emergence of outcomes influenced by the evaluation findings but also evidence to support different types of outcomes, different locations of influence, the proposed duality of outcomes also serving as processes to other outcomes, and the existence of connecting pathways.

Corresponding to Mark & Henry's model, the outcomes included both behavioral and cognitive processes. On the Year 5 staff survey, the items that addressed these outcomes asked respondents about change in specific areas (i.e., SLC program, individual practice). Because of the structure of these items, it was possible to test the proposed dual nature of these outcomes/processes both within the same category as well as across categories. Behavioral outcomes, or outcomes that indicate action or change, occurred in two areas – changes to SLCs and changes to individual practice. Furthermore, the occurrence of one led to the likelihood of the other occurring all as a result of the evaluation findings and controlling for any variations in context and awareness. Within the larger category of Behavioral processes, there was evidence of pathways from one specific type of behavioral outcome to another. Although not explicit in Mark & Henry's model, these within-category pathways provide evidence to support the underlying mechanisms of evaluation influence. The same occurred within the category of Cognitive/Affective processes. In other words, the occurrence of cognitive change or new thought about SLCs led to the likelihood of cognitive change about school reform in general. Given that SLCs are a type of school reform showed that the

evaluation findings influenced cognitive outcomes that were not just program-specific but extended outward.

There was evidence of one pathway that extended across categories. Behavioral change in individual practice led to the odds of cognitive change about SLCs occurring, and vice versa. Both of these outcome/processes took place primarily at the individual level. This type of pathway, from one major influence category to another and back again, corresponds directly to Mark & Henry's model and offers empirical evidence to support their program theory. In addition, the fact that the emergent pathways were all bi-directional further supports the authors' proposed duality of the outcomes. In each of the significant relationships both within and across categories, the specific indicators served as both outcomes and processes leading to other outcomes.

The presence of the General Influence process category and the specific outcome indicators really only became evident through the analyses and not prior to it. By definition, their importance lies in their ability to lead to other outcomes. The results have shown, however, that their importance is also dictated by the purpose of the evaluation itself, and possibly even the program. Furthermore, outcomes that may be considered general in one evaluation may not be general in another. Although these notions were not implicit in the authors' model, the findings showed that more was going on in this category than originally suggested. The potential for pathways both within this General Influence category and from the General category across to the Behavioral or Cognitive categories was also made evident through the small sample of qualitative data.

The findings also showed that the outcome indicators listed under specific levels (or locations) of influence were not always exclusive to that location. Justification as an evaluation outcome, for instance, was seen at both the individual and the interpersonal levels.

Building on the reconstructed model of evaluation influence that emerged after mapping on the SLC evaluation components,[8] this final model (see Figure 4) was created which incorporates and summarizes the key study findings and illustrates the influence processes, their predictors, and the pathways between them.

Figure 4. Post-Analyses Model of Evaluation Influence

---
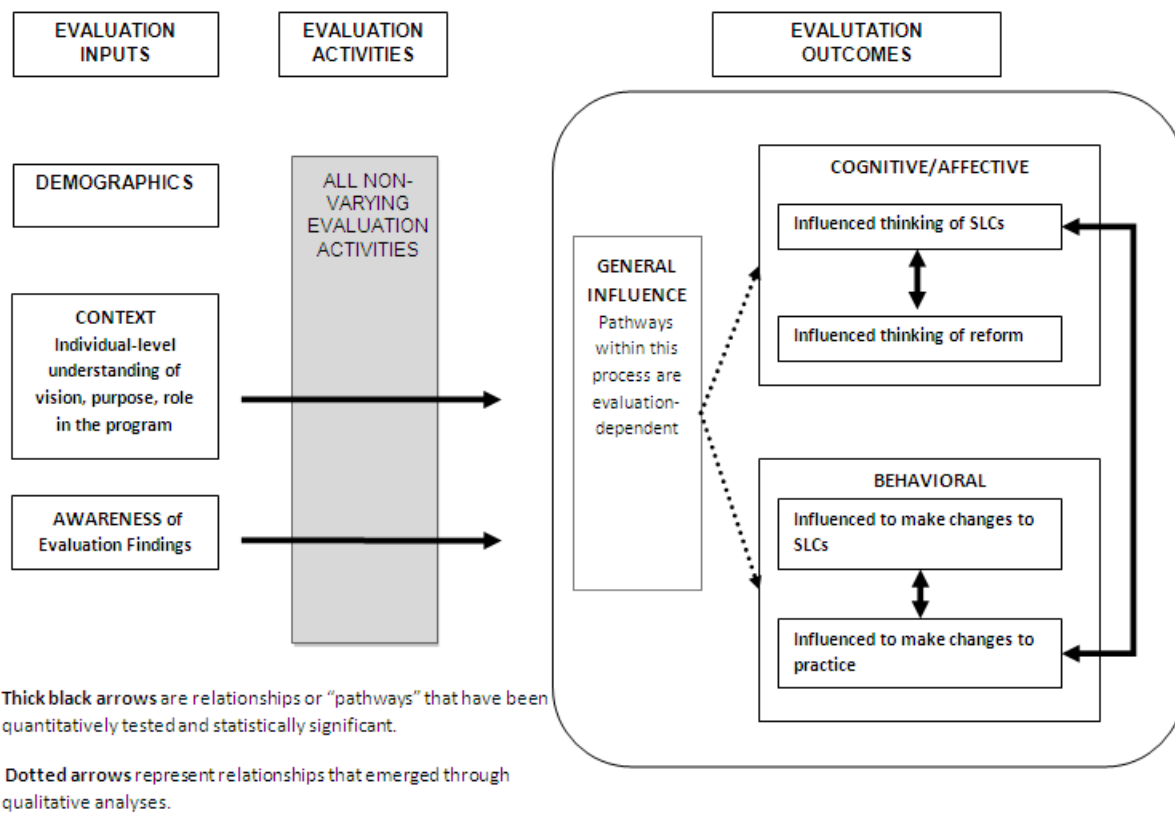
[8] See Figure 2 on page 45.

In general, the response to this final research question described how this study provided preliminary evidence to support Mark & Henry's program theory of evaluation influence. Although much of the empirical evidence was at the surface level, the qualitative findings yielded substantial information to support the presence of deeper, underlying mechanisms and processes of evaluation influence. Despite the fact that this study was limited to secondary data sources, the findings also shone light on the possibility that more was going on than was originally proposed by the authors.

CHAPTER 5

Discussion

There are noted gaps in the literature related to empirical research on evaluation theory and practice. These gaps have stimulated theorists, researchers, and practitioners to call for such research and to build an evidence base for evaluation. Broadly speaking, this study responds to this call and contributes to this growing evidence base through the systematic examination of theory and practice. More specifically, this study sheds brighter light onto the forms and processes of changes that occur as a consequence of evaluation, employs mixed methods to address some of the validity issues that have limited past research in their claims as "credible evidence," and adds food for thought in the influence vs. use conversation.

Evaluation use is predominant in the research literature as the most widely studied aspect of evaluation practice. The traditional definitions and various types of use have remained consistent over time, with theorists and authors adding to the list but with little challenge to the standard versions. Kirkhart's introduction of "influence" as a more encompassing concept than "use" was, in a way, a challenge to the traditional definitions. Simply, the argument wasn't that these definitions were inappropriate but that they were too narrow. Conceptually, influence was more inclusive and took into account impact over time and space. Mark & Henry's program theory of evaluation influence built upon these notions and focused on the consequences of evaluation and the underlying mechanisms that mediated change.

The empirical study of this program theory of evaluation influence was guided by a series of cumulative research questions. The questions were designed

to map a real-world evaluation onto the model, focus on each major component of the model, test the existence of hypothesized pathways, and determine the extent to which this study offered preliminary evidence to support this theory. Discussions of each question cover key findings as well as the broader implications that were generated through the process. Prior to this, however, I would like to comment on the methodology. Various scholars and authors have written of the lack of "credible evidence" related to studies of evaluation use and have attributed much of this to poor methodological reporting, soundness, and validity issues. Every effort has been made in this study to describe the data, report and describe every level of analysis, and employ the most appropriate analytic techniques. Details of the evaluation and the data collection process were also provided to enhance credibility. The mixed-method approach to the analyses supports the credibility of the findings, despite the limitations of secondary data and small sample sizes. I am confident that the findings fulfill this study's purpose: to make a significant contribution to evidence-based research on evaluation in the areas of use and the broader conception of influence by bridging the gap between theory and practice.

The first research question laid the foundation with a qualitative study of the design and implementation of the SLC program evaluation and how the components aligned to the model of evaluation influence. The initial objective was to lay out the components of the evaluation to determine how it mapped onto the various components of the model. This was effectively done through systematic review and discussion. The "findings" generated were specific to both the SLC evaluation and Mark & Henry's model. In the end, it was determined that the components of the evaluation aligned to the model appropriately which allowed for subsequent

analysis. The "lessons learned" from this first question, however, were broader in scope. In general, this exercise highlighted the benefits of using real-world evaluations in considering various evaluation theories and approaches. Deconstructing the evaluation design implementation was a worthwhile method of bringing this program theory of evaluation to life. This could be particularly advantageous for evaluation practitioners who subscribe to a particular evaluation theory or approach, as it would help inform their working knowledge of how practical and realistic their espoused theory really is. Mark & Henry (2004) suggested that the model may help illuminate differences among evaluation theories. This model of influence has the added benefit of tracing the pathways to evaluation outcomes and emphasizing the various forms these outcomes can take. The relevance of these exercises is evident as evaluation scholars and practitioners continue to offer frameworks, guidance, and taxonomies (Miller, 2010; Mark, 2008) to promote research on evaluation and bridging the gap between theory and practice.

What also emerged through this exercise was an understanding of the feasibility of the logic model format in preparing an evaluation design. Simply, as evaluators, we may construct logic models of a program to understand program theory, program implementation, and program outcomes. We often use these models to design an evaluation, whether it is an outcomes-based evaluation, an implementation evaluation, or some other variation. Adopting a logic model format as the framework of the evaluation design can expand the evaluator's thinking from the start – beyond the evaluation questions, resources, and methodology. This format allows the evaluator to better articulate the purpose of the evaluation, the

89

inputs necessary to conduct the evaluation activities, the activities themselves, and most importantly, what they expect to occur as a result of the evaluation. Mark & Henry's model served well as a guide to understanding the processes, mechanisms, and pathways of evaluation influence and as a framework to map out an evaluation design which will stimulate consideration of and efforts toward desired evaluation outcomes.

The second research question dealt specifically with the inputs into the evaluation process. Inputs, as noted in previous discussions, are inclusive of resources, contextual factors, and other contingencies that feed into the evaluation process. In fact, the inputs of a program evaluation share some similarity to those inputs that impact a program itself. It was for this reason that the evaluation team included measures of demographics and context in the SLC evaluation.

Demographic information of the participants and organizational context are not mutually exclusive as characteristics such as experience, position, and gender can play a role in developing the context in which the participants all function. Various evaluation scholars have written of the impact of contextual and human factors (Alkin & Taut, 2003) and personal characteristics and climate (Cousins & Leithwood, 1986) on evaluation use. Furthermore, the study of change or reform in a school or district, in this case the development of Small Learning Communities, necessitate understanding individual level context (teachers and administrators) and the collective level context (sociopolitical climate) (Fullan, 2007). Awareness of the evaluation process and of the evaluation findings was also included as a contingency/input.

The inputs in this study were examined and developed into predictors for subsequent testing of their impact on evaluation influence. The notion of context emerged in three forms. Qualitatively, context was indicated by the level of staff buy-in for the SLC program. This was no surprise given that school districts are often in the midst of initiatives, reforms, and new programs at various stages of implementation and effectiveness which may result in pessimistic opinions from those at the ground level. Quantitatively, two types of indicators emerged. The first indicator reflected individual understanding of the SLC program's purpose, goals, and role, as well as comfort in how the program had developed. The second indicator was more global, encompassing perceived support, trusting climate, and shared understanding.

Overall, the examination and development of inputs served as a stepping stone to the greater analyses and did not yield any new information regarding evaluation. Rather, the value of the inputs emerged through testing their impact on the evaluation outcomes. Corresponding to Cousins & Shulha's (2006) discussion of contextual factors, understanding the fit of evaluation outcomes with the "users" construction of reality and the extent to which uncertainty exists or collective assumptions may be in error is likely to have potent explanatory value for evaluation use" (p.273), or in this case, influence.

The examination of evaluation outcomes, specifically indicators of evaluation influence, was the focus of the third research question. Mark & Henry's model proposed several process categories (i.e., General, Behavioral) as mechanisms of evaluation influence that could take place at the individual, interpersonal, and collective levels. Housed within each process are the indicators or evidence of that

influence.   Together, these are the potential outcomes of this model. The

secondary data sources limited this study to a narrower scope of potential

processes but nevertheless yielded support for the existence of outcomes and their

correspondence to the program model of influence.

Similar to the examination of inputs, the outcomes were both studied "as is"

and derived through systematic qualitative study. The "as is" outcomes consisted

of data that specifically addressed behavioral and cognitive change as a result of

the evaluation. For the most part, half of the survey respondents reported change.

These outcomes were used in the quantitative testing of the model. The qualitative

analyses of the nature of these outcomes yielded more nuanced explanations and

findings.

Generally speaking, indicators of evaluation influence emerged as behavioral

and cognitive processes at the individual and interpersonal levels. Specifically,

there was qualitative evidence to support that many participants made behavioral

changes to their practice at the individual level and changes to their SLCs at an

interpersonal level, all as a result of the evaluation findings. This was not

surprising since making change (or taking action) to an SLC would be something

that is likely to require interactions with multiple individuals. In terms of cognitive

change or new thinking, the locus was more individual than interpersonal. Cognitive

change is not bound by interpersonal or collective interactions. The more specific

indicators of evaluation influenced posed by Mark & Henry were not easily

distinguishable in the data and those that were (e.g., change agent), emerged as a

behavioral outcome whereas the authors had posited as a more general outcome.

Nevertheless, the deeper examination of cognitive and behavioral processes has

illuminated several underlying mechanisms that transpire within the traditional forms of conceptual and instrumental use.

One of the more important findings that emerged from this examination of outcomes was that the processes and indicators of influence were more likely than not to be program-specific. The examination of outcomes in this study showed that the specific descriptions of change were bounded by the scope of the program, the roles of the participants, and the limitations imposed by the organization. This was most notable in the study of the General Influence processes. In the program model, these outcomes are considered more short-term or as stepping stones to other outcomes of interest. Many of the emergent indicators of influence fell into this category. These included actions of "focusing" on a particular area or planning to take action and corresponded to the model's proposed outcomes. Alternatively, there were other described behavioral or cognitive changes that also corresponded to the General Influence processes that were up for debate. For example, the SLC program as a district reform may not be open to major change at the school level. Moreover, changes at the school level may appear small in scope but may be significant to the participants involved. Therefore, what may appear as a General process or short-term outcome (according to the model), may actually reflect the extent to which that outcome could occur. For example, "change agent" is described as a general process in which the evaluation findings influence individuals to work together toward organizational change. Evidence of this outcome emerged through the data in various forms (i.e., teachers working collaboratively for to implement an SLC-wide or school wide change). For this evaluation and program, this outcome could be interpreted as a behavioral change and more substantive

93

than a short-term outcome. This brings the discussion back to the initial point –

that evaluation outcomes, their significance, and their chronological importance are

more program-specific than generic. Mark & Henry (2004) acknowledged that the

indicators proposed were not set in stone but rather potentialities drawn from social

science literature. The structure and elements were intended as starting and could

be tailored to specific contexts. Overall, the analyses supported these intended uses

and provided an adaptable taxonomy for studying the various possibilities of

influence.

The analysis of the model determined the impact of demographics, context,

and evaluation awareness on the likelihood of behavioral and/or cognitive change

as a result of the evaluation findings. Reciprocal pathways within and between the

processes were also tested, to determine if the occurrence of one could impact the

likelihood of another occurring. This was the primary function of the fourth

research question. Two major findings can be pulled from the quantitative analyses

performed. The first was that awareness of the evaluation findings and personal

understanding of the program's goals and purpose as well as role and buy-in, were

predictive of behavioral and cognitive action. This is not a novel finding as the

impact of context on evaluation use and practice exists in the research literature.

Nevertheless, the finding emphasizes the need for contextual factors to be included

in research on evaluation outcomes.

The second was that the occurrence of evaluation outcomes, or indicators of

influence, promoted the likelihood of other evaluation outcomes. These pathways

emerged across processes (e.g., from Behavioral to Cognitive) and within processes

(e.g., cognitive change about SLCs to cognitive change about school reform in

general). Moreover, there was some qualitative evidence for the existence of other pathways, such as from General Influence processes to Cognitive. This is the essence of Mark & Henry's model – the existence of processes and pathways through which evaluation may achieve influence and the outcomes which would indicate that influence had occurred (Mark & Henry, 2004).   In a recent review of the 41 empirical studies of evaluation use (Johnson, Lija, Toal, King, Lawrenz & Voikov, 2009), it was very evident that much of the research focused on factors that challenged or promoted use, whether or not findings were useful, or simple indicators of whether use had occurred. There was little evidence, in this review, that these studies sought to unpack use as an outcome. Mark & Henry's model does the unpacking through the processes and pathways, albeit under the umbrella of influence as opposed to use.

        The final research question tied all of the previous questions and findings together. On the whole, Mark & Henry's program theory expanded our thinking about the forms that change could take and the various levels in which they could occur. The study showed that not only were there behavioral and cognitive changes as a result of the evaluation findings but that these changes occurred at both individual and interpersonal levels and that some changes promoted the likelihood of other changes. The results provided supporting evidence to the proposed pathways between these processes, suggesting that forms of use/change are not always isolated incidents but may inspire or stimulate. Finally, the study shows that other aspects of the evaluation setting, primarily the organizational context of the evaluand, can have a significant impact on the evaluation

95

consequences.   These are not only the significant findings of this last question but the entire study.

Through engaging in this study and the literature on influence (and use), additional findings emerged that were unintended but equally insightful. Although the proposed outcomes and processes in their model correspond to traditional forms and definitions of use, Mark & Henry argued that these definitions were too limiting and did not promote further understanding of the processes involved and the potential pathways among them. Some opponents to the notion of influence over use have argued that influence is unintentional (Alkin & Taut, 2003) and can occur long after the evaluation has ended reaching a broader range of individuals. Influence, then, appears contrary to those who hold use as intended, time sensitive, and focused on a group of primary users. Alternatively, Johnson et al. (2009) noted that evaluation scholars were increasingly viewing evaluations as having influence and that focusing solely on direct use may not have adequately captured broader levels of influence.

I began this study with no firm opinions on which concept or definitions were more appropriate; however, some conclusions about the influence vs. use distinction have emerged throughout the course of my research. At the big picture level, the concept of influence is not inherently different from use but encompasses the traditional forms and definitions. Furthermore, Mark & Henry's model and proposed processes and pathways do not differentiate from use or oppose it but rather expands it, opening the box to enable us to understand the underlying processes that mediate change. In this study, the "indicators of influence" could also have been described as indicators of use – they were the same actions and

changes that occurred as a result of the evaluation and giving them different labels would not change that. It is my conclusion that this is just as much, possibly even more in fact, a study of use as it is influence and that the two concepts are related in practice, in ways that make the distinctions academic and abstract.

So why is this important? Because this program theory and the results of this study contribute to the greater understanding of the complexity of evaluation consequences, adds to the toolkit of how to measure its various forms, and promotes future research that will add to the evidence base. The definitive value of this model is in its emphasis on the consequences of evaluation – whether we call it use or influence – and the focus should be on the processes and mechanisms promoting evaluation outcomes regardless of traditional labels, intentionality vs. un-intentionality, or if it immediately follows the evaluation or long after. Cousins & Shulha (2006) noted the prospect of moving toward a theory of influence as "intriguing" and "likely to stimulate ongoing debate and inquiry in the field" (pp. 283-284).

Research on evaluation, specifically the empirical testing of evaluation theories and models, can be challenging to conduct. A major challenge may lie in translating the theory into measurable terms and/or testable questions. Mark & Henry's use of the logic model format to frame their program theory of evaluation influence offered an advantage in this respect in that potential predictors, relationships, pathways, and outcomes were identifiable and more amenable to quantitative testing. A notable advantage of testing logic models is the potential to move beyond the study of individual components.

Other challenges in using models to test theory can result from matching real world data to the theoretical model, particularly when the data are from secondary sources. In this study, for instance, not all of the program model components were included due to lack of supporting data. Consequently, the format of available data may limit the sophistication of the analytic approaches that would be most suitable for testing a logic model, or program theory. Again, the inability to perform a more preferred statistical method such as structural equation modeling in this study was due to the nature of the available data. Designing a study with the goal of testing a logic model would appear to be the best way to overcome the aforementioned challenges. Nevertheless, there is an advantage to using logic models to test evaluation theory and subsequent studies of Mark & Henry's program theory should include the most appropriate data to fit the model.

Limitations

There were several limitations in working with secondary data sources. The first involved the lack of control over the data collected, the instruments used, or the procedures followed. The data used for this study came from a federally-funded large scale school district evaluation and was limited to the scope of the evaluation's purpose and the wants and needs of the stakeholders. It was because evaluation use and various factors associated with use were of interest to the evaluation team in Year 5 of the evaluation that data was available for this study. However, this limited much of the analyses to this final year. The study would have certainly been strengthened by including time factor as a component of the analyses. Missing data and response rates were also products of relying on

secondary data. Procedures to increase participation or built-in procedures to minimize missing data did not appear to be systematically followed by the evaluation team.

Other limitations were specific to the quantitative analyses. The study's focus on Year 5 and the large amounts of missing data limited the sample sizes available for analyses. The possibility of recoding open-ended survey data into a format appropriate for quantitative analyses was impeded by the very small sample sizes. This prevented the testing of more specific processes and evaluation outcomes. Furthermore, the format of the quantitative outcome data was dichotomous and did not readily lend itself to more sophisticated analytic procedures, particularly those for model testing (i.e., path analysis, structural equation modeling).

Finally, there were components (full or partial) of Mark & Henry's program theory that were not addressed in this study. For example, the Motivational processes outcome component was excluded from the quantitative analyses because there was no available data corresponding to this particular outcome. For similar reasons, the inclusion and examination of knowledge production attributes as an evaluation output were also excluded. Also, the examination of outcomes that would indicate influence in the longer term was impossible since the official evaluation period had ended.

Implications for Future Research

There are several avenues emergent in this study that can guide future research. Challenges for practicing evaluators who want to study or examine the

outcomes of their evaluation (e.g., use, influence) include the constraints of time, resources, and budget. These constraints apply to the planning of this additional study, data collection, and fitting it into the existing evaluation with minimal burden to the stakeholders and/or participants. The logic model structure of Mark & Henry's program theory lends itself to practical application and provides a suitable framework for planning inquiry or examination. Furthermore, it can be mapped onto an existing evaluation and provide guidance for where, when, and how to collect data. The model is also customizable to promote desired areas of focus. In essence, this model (and this study) has shown that it may not require much in additional resources to build components into the evaluation to systematically study the evaluation's outcomes.

Given the limitations of using secondary data, there were components of Mark & Henry's model that were not included in this study but are deserving of attention. In particular, study of change at the collective level may provide greater insight into the processes of change at the organization or group level. Such research would be greatly informative in evaluations where promoting "instrumental use" or behavioral change is not feasible at the individual level – where one person may not have the ability or power to make program or policy decisions as a result of an evaluation. Another area of potential research may lie in the study of the motivational processes included in the model. As proposed by the authors, these processes may serve as important processes leading to behavior change. Furthermore, time and intention may be built into a research study to further understand the scope of influence.

APPENDIX

Table A-1.
Data Sources of SLC Evaluation (2005-2010)

Instrument		Total *N*	Description/Purpose
Staff Survey			
	Y1	278	To measure teacher attitudes towards involvement in SLC
	Y2 – F2006	273	implementation and development, personalization as it relates to
	Y2 – S2007	30	both student/teacher (staff) and peer/ peer relations, and teacher
	Y3 – F2007	317	perception of student achievement expressed in terms of their
	Y3 – S2008	193	students' academic behavior and college preparation/knowledge.
	Y4	190	
	Y5	272	
Staff focus groups/ interviews			
	Y1	4	To gain an understanding of the teacher experience and perception
	Y2	5 FG*	of SLC implementation and development, needed resources,
	Y3	28	challenges/ dilemmas, and solutions.
	Y4	55	
	Y5	38	
Principal Interview			
	Y1	5	To gain an understanding of the principal's experience and
	Y2	0	perception of SLC implementation and development, needed
	Y3	3	resources, challenges/ dilemmas, and solutions.
	Y4	4	
	Y5	3	
SLC Coordinator Interview			
	Y1	3	To gain an understanding of the SLC coordinator's experience and
	Y2	0	perception of SLC implementation and development, needed
	Y3	3	resources, challenges/dilemmas, and solutions.
	Y4	3	
	Y5	4	
Student Survey			
	Y1	6,533	To measure students' attitudes towards involvement in SLC
	Y2	3,682	implementation and development, personalization as it relates to
	Y3	5,505	both student/teacher (staff) and peer/ peer relations, and student
	Y4	6,713	achievement expressed in terms of academic behavior and college
	Y5	2,335	preparation/knowledge.
Student focus groups			
	Y2	4 FG*	To gain an understanding of the student experience with SLC
	Y3	8 FG*	implementation and development, personalization as it relates to the
	Y4	59	state of teacher/student and peer/peer relations, and academic
	Y5	54	achievement as expressed through academic behavior and college knowledge.
Parent Survey			
	Y3	1,715	To measure parents' attitudes towards involvement in SLC
	Y4	2,584	implementation and development, personalization as it relates to
	Y5	514	parent perception of their children's interactions with teachers (staff), and parent perception of student achievement expressed in terms of their child's academic behavior and college preparation/knowledge.

*Exact number of focus group participants was not noted.
Note: This table was adapted from the 2005-2010 Summative Report

Table A-2.
Initial Eigenvalues for Eight Context-Related Survey Items

Survey Item	Eigenvalue	% of Variance	Cumulative Variance
C1. There is a climate of trust here among students, teachers, and administrators.	3.995	49.941	49.941
C2. The *faculty and staff* understand the purpose and goals for SLCs *at this school*.	1.278	15.969	65.911
C3. I understand the purpose and goals for small learning communities *at this school*.	0.762	9.522	75.432
C4. The faculty and staff at this school have opportunities to suggest modifications to the SLCs.	0.617	7.710	83.143
C5. The school provides time on a regular basis for SLC teams to meet to share information, discuss students' academic progress, curriculum needs, etc.	0.462	5.771	88.914
C6. I understand the vision and goals for *my particular SLC*.	0.395	4.941	93.855
C7. I understand what my role is in *my SLC*.	0.290	3.630	97.485
C8. I am comfortable with how *my SLC* has developed.	0.201	2.515	100.000

REFERENCES

Alkin, M. (2004) *Evaluation roots: Tracing theorists' views and influences.* Thousand Oaks, CA: Sage.

Alkin, M., Daillak, R., & White, P. (1979). *Using Evaluations: Does Evaluation Make a Difference?* Beverly Hills, CA: Sage. Chapter 1.

Alkin, M. & Taut, S. (2003). Unbundling Evaluation Use. *Studies in Educational Evaluation, 29*, 1-12.

Brandon, P. R. & Singh, J. M. (2009). The strength of the methodological warrants for the findings of research on program evaluation use. *American Journal of Evaluation, 30* (2), 123-157.

Christie, C. A. (2007). Reported influence of evaluation data on decision makers' actions: An empirical examination. *American Journal of Evaluation, 28* (1), 8-25.

Coburn, C. E., Toure, J., & Yamashita, M. (2009). Evidence, interpretation, and persuasion: Institutional decision making in the central office. Teachers College Record, 111, 1115-1161.

Cousins, B. J. (2003). Utilization effects of participatory evaluation. In D. L. Stufflebeam, & T. Kellaghan (Eds.), *International Handbook of Educational Evaluation*. Boston, MA: Kluwer Academic Press.

Cousins, J. B. & Earl, L. M. (1995). Participatory evaluation: Enhancing evaluation use and organizational learning capacity. *The Evaluation Exchange, 1*(3 & 4). Retrieved 10/27/10 from http://www.hfrrp.org/evaluation/the-evaluation-exchange/issue-archive.

Cousins, B. & Leithwood, K. (1986). Current Empirical Research on Evaluation
Utilization. *Review of Educational Research*, *56 (3)*, 331-364.

Cousins, B. J. & Shulha, L. M. (2006). A comparative analysis of evaluation
utilization and its cognate fields of inquiry: Current issues and trends. In I.
F. Shaw, J. C. Greene, & M. M. Mark (Eds.), *The Sage Handbook of
Evaluation*. Thousand Oaks, CA: Sage Publications, Inc.

DiPaola, M. & Smith, K. (2008). Measuring district climate. In W.K. Hoy & M.
DiPaola [Eds.], *Improving schools: Studies in leadership and culture (pp.
117-135)*. Charlotte, NC: Information Age Publishing, Inc.

Fullan, M. (2007). *The New Meaning of Educational Change. Fourth Edition*. New
York, NY: Teachers College.

Henry, G. T. (2003). Influential evaluations. *American Journal of Evaluation*,
24(4), 515-524.

Henry, G. T & Mark, M. M. (2003). Beyond use: Understanding evaluation's
influence on attitudes and actions. *American Journal of Evaluation*, *24* (3),
293-314.

Hofstetter, C.H., & Alkin, M.C. (2003). Evaluation use revisited. In D.L.
Stufflebeam & T. Kellaghan (Eds.), International handbook of educational
evaluation. Boston, MA: Kluwer.

Honig, M. I. & Coburn, C. E. (2008). Evidence-based decision-making in school
district central offices: toward a policy and research agenda. *Education
Policy, 22*, 578-608.

Kirkhart, K.E. (2000). Reconceptualizing evaluation use: An integrated theory of
influence. *New Directions for Evaluation, 88*, 5-23.

Kowalski, T. J. & Hermann, K. (2008). Exploring the cultures of high schools in
    relation to their demographic characteristics and performance. In W.K. Hoy
    & M. DiPaola [Eds.], *Improving schools: Studies in leadership and culture*.
    Charlotte, NC: Information Age Publishing, Inc.

Johnson, K., Greenseid, L. O., Toal, S. A, King, J. A., Lawrenz, F. & Volkov, B.
    (2009). Research on evaluation use. A review of the empirical literature
    from 1986 to 2005. *American Journal of Evaluation, 30*(3), 377-410.

Lachat, M. A. & Smith, S. S. (2005). Practices that support data use in urban high
    schools. *Journal of Education for Students Placed At Risk, 10*(3), 333-349.

Leviton, L. C. (2003). Evaluation use: Advances, challenges, and applications.
    American Journal of Evaluation, *24* (4), 525-535.

Leviton, L. C., & Hughes, E. F. X. (1981). Research on the Utilization of
    Evaluations: A Review and Synthesis*. Evaluation Review, 5(4),* 525-548.

Mark, M. M. (2011). Toward better research on – and thinking about – evaluation
    influence, especially in multisite evaluations. In J. A. King & F. Lawrenz
    (Eds.) Multisite evaluation practice: Lessons and reflections from four cases.
    New Directions for Evaluation, 129, 107-119.

Mark, M. M. (2008). Building a better evidence base for evaluation theory.
    In N.L. Smith & P. R. Brandon (Eds.), *Fundamental Issues in Evaluation* (pp.
    111-134). New York, NY: Guilford Press.

Mark, M. M. & Henry, G. T. (2004). The mechanisms and outcomes of evaluation
    influence. *Evaluation*, *10* (1), 35-57.

Miller, R. L. (2010). Developing standards for empirical examinations of evaluation
    theory. *American Journal of Evaluation*, *31(3*), 390-399.

Patton, M. Q. (1997). *Utilization-focused evaluation: A new century text (3rd ed.).* Thousand Oaks, CA: Sage.

Pechman, E. M. & King, J. A. (April, 1986). *Analyzing the school evaluation use process to make evaluation worth the effort.* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA. (ERIC Document Reproduction Service No. ED 294 299).

Rivera, N. V., Vo, A. T. & Lee, J. A. (November, 2009). *Organizational structures and attitudes toward evaluation An examination of contextual factors that shape evaluation.* Paper presented at the Annual Meeting of the American Educational Research Association, Orlando, FL.

Shulha, L.M., & Cousins, J.B. (1997). Evaluation use: Theory, research and practice since 1986. *Evaluation Practice*, 18 (3), 195-208.

Taut, S. M. & Alkin, M. C. (2003). Program staff perceptions of barriers to evaluation implementation. *American Journal of Evaluation*, *24* (2), 213-226.

Weiss, C. H. (1993). Where politics and evaluation research meet. *American Journal of Evaluation*, *14(1*), 93-106.