TRANSFORMATIONAL LEADERSHIP AND COLLECTIVE EFFICACY:  
A MODEL OF SCHOOL ACHIEVEMENT

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy
in the Graduate School of The Ohio State University

By

Michael Raymond Nicholson, B.S., M.A., C.A.S

The Ohio State University
2003

Dissertation Committee:
Dr. Wayne K. Hoy, Adviser
Dr. Anita Woolfolk Hoy
Dr. Franklin Walter
Dr. Scott Sweetland

Approved by

Adviser
College of Education
ABSTRACT

Recent research has demonstrated that collective efficacy is an important factor that positively correlates with school achievement. Antecedents to collective efficacy, though, have been rarely tested, prompting this study of transformational school leadership as a viable influence on teachers’ competency beliefs. Hypotheses were formed, based on previous scholarship and theory, concerning various relationships among school context, prior achievement, teacher norms, principal leadership, and school achievement. A model of school achievement based on these variables was developed, using socioeconomic status and prior school achievement as control variables. Achievement data, federal lunch program participation rates, and teacher perceptions were collected from 146 Ohio elementary schools, which constituted the sample in this research. Factor and correlational analyses along with structural equation modeling were used in the statistical analyses.

The results of this research failed to support the hypothesis that transformational leadership influenced collective efficacy; however, collective efficacy was positively related to student achievement even controlling for SES and prior achievement. Socioeconomic status was positively related to collective efficacy and was positively related to individualized consideration, but it was not related to intellectual stimulation. Only intellectual stimulation of transformational leadership was directly related to student
achievement, but it was not indirectly related to student achievement through collective efficacy. Further, socioeconomic status was both directly related to student achievement and indirectly related to student achievement through collective efficacy and prior school achievement. Finally, theoretical, practical, and research implications were discussed
This is dedicated to my wife Leigh, my Mom, and my Dad
ACKNOWLEDGMENTS

There is no way that I could have undertaken this doctoral program without the help I received from a number of people. Among the many I leaned on, Leigh, my wife, took the brunt of my educational indulgence. She, in essence, took care of our three children, Danny, Will, and Caroline, for the last three years as a single parent might. The degree I have earned is as much hers as it is mine. Her sacrifices were clearly observed, daily. Thank you, Leigh. I hope the benefits of our investment are at least in proportion to the sacrifices you made for their attainment.

My two broader families have encouraged me in many ways since I shared with them my wishes to attain my doctorate. Their good wishes were constant reinforcers, forming an important familial push. Thank you Dad, Sylvia, Mom and Dad Dolen, Ann & Jon, Kevin and Sandy, Sue and Steve, Steve and Wendy, and Mike and Carrie. I hope you all didn’t grow tired of the sweatshirts.

The meaningful experience I have had can be attributed, at least in part, to many other good people, not the least of whom is Dr. Hoy. It seems appropriate that one of Dr. Hoy’s research interests includes the concept of trust. He clearly embodies the underlying dimensions of trust: benevolence, reliability, openness, competence, and honesty. My trust and admiration of Dr. Hoy was found early and grows deeper with every shared experience. I can’t thank him enough for calling me “out of the blue” two years ago.
Dr. Woolfolk Hoy’s, Dr. Walter’s, and Dr. Sweetland’s vast intellects are matched by their genuine interest in facilitating student success. Although they accommodated my request to serve on the committees, my friction-free experience suggests providence had a hand in their decision. I can’t imagine all committees function so congenially and efficiently, and with such care and concern for student development.

Tim Cybulski’s friendship had a most positive effect on my studies and personal growth. Few other people, if any, could I share with so many of my educational interests. Much to my delight, our conversations quickly allowed for many other topics, even if time rarely could. May the Juntos continue far into the future.

Though many others doubtlessly aided me, the glow these three people cast on OSU deserve special mention. Deb Zabloudil flawlessly guided me through the necessary bureaucratic channels at every juncture of the program. Patient, dependable, and knowledgeable, Deb serves a crucial role in making the program as user-friendly as possible. Diane Baugher provided not only signatures and smart advice, but also her good nature, sense of humor, and quick wit balanced the program with appreciated levity. I am glad she is back. Dr. Montemayor’s teaching and shared discourse lifted my curiosities in psychology far beyond my initial interests. His approvals were among the earliest validations that I might belong here.

Dr. Anne Hyland and Dr. Martha Alcock showed me immediate and important applications of my studies. Their enthusiasm for my interests, patience with my stumblings, and advocacy for my development have accelerated my contributions, as modest as those offerings might be. Their collaboration represents a rare strong link.
between higher education and k-12 schools. Our shared work has only started and my learning has just begun.
VITA

October 27, 1965............................................................... Born – Buffalo, New York

1987................................................................................... B.S., Computer Science
Hartwick College, Oneonta, NY

1990................................................................................... M.A., Elementary Education
SUNYAB, Buffalo, NY

1992-1998 ................................................................. Middle School Teacher,
Williamsville, NY

1999................................................................................... C.A.S., Educational Admin.
SUNYAB, Buffalo, NY

1998-2000 ................................................................. Assistant Principal,
Niagara-Wheatfield CSD
Niagara Falls, NY

2000-2002 ................................................................. Graduate Research Assistant,
The Ohio State University

2002-present................................................................. District Administrator, Research
Bexley CS, Bexley, OH

FIELDS OF STUDY

Major Field: College of Education
Educational Policy and Leadership
Educational Administration
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>v</td>
</tr>
<tr>
<td>Vita</td>
<td>viii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xiii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xiv</td>
</tr>
<tr>
<td>Chapters:</td>
<td></td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2. Literature review</td>
<td>12</td>
</tr>
<tr>
<td>3. Teacher efficacy definition</td>
<td>12</td>
</tr>
<tr>
<td>4. Teacher efficacy effects</td>
<td>13</td>
</tr>
<tr>
<td>5. Teacher efficacy construct development</td>
<td>17</td>
</tr>
<tr>
<td>6. Teacher locus of control</td>
<td>21</td>
</tr>
<tr>
<td>7. Responsibility for student Achievement</td>
<td>23</td>
</tr>
<tr>
<td>8. Second conceptual strand of teacher efficacy</td>
<td>25</td>
</tr>
<tr>
<td>9. Social cognitive theory</td>
<td>26</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Assumptions</td>
<td>26</td>
</tr>
<tr>
<td>The original research</td>
<td>29</td>
</tr>
<tr>
<td>The effects of self-efficacy beliefs</td>
<td>29</td>
</tr>
<tr>
<td>Sources of self-efficacy beliefs</td>
<td>30</td>
</tr>
<tr>
<td>The motivating processes involving efficacy beliefs</td>
<td>34</td>
</tr>
<tr>
<td>Webb efficacy scales</td>
<td>37</td>
</tr>
<tr>
<td>Teacher efficacy scale</td>
<td>39</td>
</tr>
<tr>
<td>Dimension reconceptualization</td>
<td>42</td>
</tr>
<tr>
<td>Locus of control revisited</td>
<td>44</td>
</tr>
<tr>
<td>Outcome efficacy</td>
<td>45</td>
</tr>
<tr>
<td>Studies confirming teacher efficacy’s confusion</td>
<td>47</td>
</tr>
<tr>
<td>An integrated model of teacher efficacy</td>
<td>49</td>
</tr>
<tr>
<td>Appraisal of the integrated teacher efficacy scale</td>
<td>52</td>
</tr>
<tr>
<td>An integrated Teacher Efficacy Model’s validation study</td>
<td>54</td>
</tr>
<tr>
<td>Collective Efficacy</td>
<td>56</td>
</tr>
<tr>
<td>Assumptions</td>
<td>57</td>
</tr>
<tr>
<td>Collective efficacy in schools</td>
<td>58</td>
</tr>
<tr>
<td>Past school studies approximating collective efficacy</td>
<td>59</td>
</tr>
<tr>
<td>Cohesion and ideologies</td>
<td>59</td>
</tr>
<tr>
<td>Teacher efficacy and parent involvement</td>
<td>60</td>
</tr>
<tr>
<td>Efficacy and alienation</td>
<td>61</td>
</tr>
<tr>
<td>Efficacy and demography</td>
<td>62</td>
</tr>
<tr>
<td>More relevant studies of collective efficacy</td>
<td>62</td>
</tr>
<tr>
<td>Pioneering research in collective efficacy</td>
<td>62</td>
</tr>
<tr>
<td>Supporting research of collective efficacy</td>
<td>65</td>
</tr>
<tr>
<td>Mastery experiences and collective efficacy</td>
<td>66</td>
</tr>
<tr>
<td>Teacher and collective efficacy</td>
<td>68</td>
</tr>
<tr>
<td>Implications for present study</td>
<td>70</td>
</tr>
<tr>
<td>Leadership theory</td>
<td>71</td>
</tr>
<tr>
<td>The origins and pervasiveness of leadership</td>
<td>72</td>
</tr>
<tr>
<td>Leadership trait theory</td>
<td>74</td>
</tr>
<tr>
<td>Early leadership trait theory</td>
<td>74</td>
</tr>
<tr>
<td>Later leadership trait theory</td>
<td>75</td>
</tr>
<tr>
<td>Leadership behavior theory</td>
<td>77</td>
</tr>
<tr>
<td>Leadership behavior inquiry via the LBDQ</td>
<td>77</td>
</tr>
<tr>
<td>Derivatives of the LBDQ</td>
<td>80</td>
</tr>
<tr>
<td>The Managerial Grid</td>
<td>82</td>
</tr>
<tr>
<td>Contingency theory</td>
<td>83</td>
</tr>
<tr>
<td>Least Preferred Coworker</td>
<td>84</td>
</tr>
<tr>
<td>Path-Goal theory</td>
<td>86</td>
</tr>
<tr>
<td>New leadership paradigm</td>
<td>89</td>
</tr>
<tr>
<td>Charismatic authority</td>
<td>90</td>
</tr>
<tr>
<td>Charismatic leadership authority</td>
<td>91</td>
</tr>
<tr>
<td>Transformational leadership</td>
<td>95</td>
</tr>
<tr>
<td>Theoretical beginnings</td>
<td>95</td>
</tr>
</tbody>
</table>
5. Discussion of results .............................................................. 176
   Summary of findings .................................................................. 178
   Discussion ................................................................................ 179

   Collective efficacy and student achievement .......................... 179
   Socioeconomic status and collective efficacy ......................... 180
   Intellectual stimulation and student achievement .................. 181
   Socioeconomic status and student achievement ..................... 184
   Transformational leadership and collective efficacy ............... 185
   Socioeconomic status and individualized consideration ........ 186

   Practical, research, and theoretical implications ..................... 187
   Practical implications .............................................................. 187
   Theoretical implications ......................................................... 189
   Research implications ........................................................... 191

   Conclusions ............................................................................ 193

Appendices .................................................................................. 195
   Appendix A Factor analysis of transformational leadership ....... 195
   Appendix B 4-factor structure of transformational leadership ...... 196
   Appendix C 1-factor structure of transformational leadership ...... 197
   Appendix D Exploratory factor analysis of Trans. Leadership ...... 198
   Appendix E EFA of TL after variable elimination (orthogonal) .... 199
   Appendix F EFA of TL after variable elimination (oblique) ......... 200
   Appendix G Conceptual design of 2-factor model of TL .......... 202
   Appendix H Standardized path coefficients of 2-factor model .... 203
   Appendix I Review of CFA’s for the TL structure models ........... 204
   Appendix J A Review of both school achievement models ......... 206

References .................................................................................. 207
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Factor loadings of items representing the five dimensions</td>
<td>129</td>
</tr>
<tr>
<td>3.2 Transformational leadership item description</td>
<td>131</td>
</tr>
<tr>
<td>3.3 Respondent Group Comparison</td>
<td>133</td>
</tr>
<tr>
<td>4.1 Comparison of sample’s and state’s elementary schools</td>
<td>141</td>
</tr>
<tr>
<td>4.2 Grade configurations</td>
<td>142</td>
</tr>
<tr>
<td>4.3 Items constituting the first factor of the exploratory factor analysis</td>
<td>152</td>
</tr>
<tr>
<td>4.4 Items constituting the second factor of the exploratory factor analysis...</td>
<td>155</td>
</tr>
<tr>
<td>4.5 Items comprising the two factors of transformational leadership</td>
<td>158</td>
</tr>
<tr>
<td>4.6 Research variables’ descriptive statistics</td>
<td>165</td>
</tr>
<tr>
<td>4.7 Correlations among research variables</td>
<td>167</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Generic social cognitive model of human agency</td>
<td>27</td>
</tr>
<tr>
<td>2.2</td>
<td>Correspondence between Transformational Leadership Behaviors and Sources of Efficacy Information</td>
<td>115</td>
</tr>
<tr>
<td>2.3</td>
<td>A theoretical model of school achievement</td>
<td>122</td>
</tr>
<tr>
<td>4.1</td>
<td>Model of math achievement</td>
<td>169</td>
</tr>
<tr>
<td>4.2</td>
<td>Model of reading achievement</td>
<td>170</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

Educators today face challenges from a multitude of sources. Many schools need to meet ambitious state-mandated achievement standards in a context unlike any existing during previous times. The growing diversity of our K-12 student body pose new problems for teachers faced with growing constituent expectations. Kaestle (1995) indicates that the familiar techniques of past educational practices fail to meet all the needs of today's students. Twenty-five years ago Resnick and Resnick (1977) noted that contemporary teaching methodologies were not meant to prepare students for the literacy standard sought then, nor were they meant to bring all students to the same level of competence. Alarmingly, family poverty and divorce continue to climb, taking from children the developmental assets they need for academic success (Sadker & Sadker, 1997). Identifying school factors that strengthen teacher resolve for answering these challenges is crucial.

Yet, another feature of our society holds back our potential for educating youth. Our individualistic culture, manifesting in our behaviors, attitudes, and politics (Moskos, 1988; Putnum, 2000; Wilson, 1997), and the balkanized orientation of our schools’
classrooms (Huberman, 1993; Miller, 1999) exacerbates school problems needing collaborative resolutions. School-level factors that promote cohesion among staff need to be identified.

Individual teachers need not, and should not, deal with the aforementioned problems alone, as many of their struggles stem from common challenges facing their colleagues, too. Indeed, the collective efforts of teachers seem the best chance of overcoming their shared difficulties. The quality of the collective efforts partly stem from teachers’ sense of collective efficacy, their belief they can overcome such problems in a unified endeavor (Bandura, 1997a, p. 35). Collective efficacy is the degree “to which perceptions of efficacy, either high or low, are shared across teachers in a school building” (Tshannen-Moran, Woolfolk Hoy, & Hoy, 1998, p. 221). Collective efficacy is one of the few school-level characteristics as influential as socioeconomic status in affecting school achievement (Bandura, 1993; Goddard, Hoy, & Woolfolk Hoy, 2000). Determinants of collective efficacy hold much promise for improving education for our students.

Problem Statement

Collective efficacy and principal leadership are the major independent variables of this research. Collective efficacy is derived from Bandura’s (1977) theory of self-efficacy – the cornerstone of his model of human agency. The premise of his work lies in its conception of personal competence beliefs, or efficacy perceptions, and the factors that influence them. Bandura’s (1986) sociocognitive theory provides an explanation for how principal behavior, teacher behavior, and teacher cognitive and personal factors reciprocally determine teachers’ individual and collective efficacy. Self-efficacy theory
extends the study of human agency to a broader realm, beyond just personal and psychological factors, to include one’s environmental and behavioral influences. Bandura’s framework of efficacy invites research into contextual factors that affect teachers’ beliefs in personal and collective competence.

Principal leadership is conceptualized in terms of transformational leadership, which is also known as "charismatic" or "visionary" leadership. Transformational leadership affects followers in ways that are measurably greater and categorically different than the effects claimed in previous theories of leadership (Shamir et al., 1993, p.577). Such leaders heighten the needs, values, preferences, and aspirations of followers from self-interests to group interests. Moreover, they inspire followers to commit strongly to the leader's mission, to make significant personal sacrifices in the interest of the group, and to perform above and beyond the call of duty (House and Shamir, 1993).

Subordinate motivation stirred by transformational leaders’ behavior triggers better employee performance than subordinate actions influenced by non-transformational leaders (Avolio, Waldman, & Einstein, 1988; Howell & Avolio, 1993).

Transformational leadership, as conceived by Bass (1985), not only emphasizes leader behaviors that galvanize a unified spirit among subordinates, but also stresses individual attention to each follower. Thus, collective and personal efficacy beliefs should be nurtured under such direction. Research and theory suggest a positive relationship between transformational leadership behaviors and efficacy (Fuller & Is, 1986; House & Shamir, 1993; Shamir, House, & Arthur, 1993; Shear & Howell, 1999).
This study reviews the research on collective efficacy and transformational leadership behavior and ties the two conceptual strands together using Bandura’s (1986) sociocognitive theory to explain teacher agency.

Bandura (1997a) asserts that “perceived personal and collective efficacy differ in the unit of agency, but in both forms efficacy beliefs have similar sources, serve similar functions, and operate through similar processes” (p. 478). Shear and Howell (1999) found that self-efficacy mediated the relationship between the interaction effect of leadership behavior and task feedback on performance quality over time. Consequently, applying Bandura’s (1997a) observation regarding the parallel concepts of self- and collective efficacies, with regard to their sources, functions, and processes, collective efficacy may mediate the relationship between organizational leadership and group performance. The potential for leaders to not only develop teachers' instructional behaviors, but to indirectly improve students' achievement, compels further study of the relationship between transformational leadership and collective efficacy.

Sociocognitive theory suggests contextual factors that influence sources of self-efficacy beliefs stand the best chance of stimulating individuals’ capacity and desire for human agency (Bandura, 1997a). This study posits that transformational leader behavior is instrumental in developing sources of self-efficacy: mastery experiences, verbal persuasion, vicarious learning, and physiological states. The sources’ value in promoting efficacy depends on their ability to trigger reflective thought (p.79) more than their ability to impart technical information and resources, though both effects can sometimes coexist.
Transformation leadership has four components (Bass, 1985), known as the four “I’s” -- intellectual stimulation, individualized consideration, idealized influence, and inspirational motivation. These four elements describe the conduct of transformational leaders to elevate the aspirations, needs, motivations, and desires of their followers. Hence, it seems likely, that the elements of transformational leadership will influence efficacy beliefs because they motivate individuals to “expend substantial effort in pursuit of goals, persist in the face of adversity, rebound from temporary setbacks, and exercise some control over events in their lives” (Tschannen-Moran et al., 1998, p.207). If the four “I’s” stimulate followers’ reflective thought processes, then they should reinforce sources of self and collective efficacy (House & Shamir, 1993). By focusing followers’ sights on an appealing future vision and supporting their sources of competency beliefs, transformational deeds arguably inspire greater frequency of subordinate self-reflective thought than those activities associated with instructional tasks (i.e., providing technical feedback, making available task-enhancing resources, etc).

Leithwood and his colleagues have extensively studied transformational leadership in schools (e.g., Leithwood, 1994; Leithwood & Duke, 1999; Leithwood, Jantzi, & Fernandez, 1994; Leithwood, Jantzi, & Steinbach, 1999; Leithwood & Jantzi, 1999; Leithwood & Jantzi, 2000; Leithwood, Menzies, Jantzi, & Leithwood, 1999;). Their models of such leadership stem from a conceptualization of school leadership containing eight dimensions, some of which overlap Bass’s (1985) transformational leadership factors: creating vision, developing group goals, maintaining high performance expectations, modeling, providing individual support, providing intellectual stimulation, building a productive school culture, and building structures for
collaboration (Leithwood et al., 1994). Leithwood et al. have studied various effects of school transformational leadership on school, teacher, and student outcomes, for example teacher burnout (Leithwood et al., 1999), student engagement with school (Leithwood & Jantzi, 1999), teacher commitment to change (Leithwood et al., 1994), and school improvement (Leithwood et al., 1993). Although Leithwood et al. have studied transformational leadership’s effect on school and teacher characteristics more than most, if not all, other researchers, to date they have not studied its effect on collective efficacy, a potentially important school-level predictor of school-level performance.

Since Coleman et al.’s (1966) study of educational equality in the United States, SES has largely stood as the most consistent, and often strongest, correlate of school-level achievement. Indeed, Hoy and Sabo (1998) report that SES "typically overwhelms other variables in predicting achievement" (p. 115). Nonetheless, Edmonds (1979) demonstrated more than a decade after the Coleman Report that school-level measures, other than SES, related to school performance, such as frequency of student evaluation, basic skill emphasis, and school orderliness, among others. Concurrently, Brookover, Beady, Flood, Schweitzer, and Wisenbaker (1979) argued that school climate more strongly predicted school achievement than did SES. Since then, Bandura (1993) demonstrated collective efficacy related to school achievement more than SES did. A few other studies (Hoy & Sabo, 1998; Hoy, Tarter, Kottkamp, 1991) have also found that certain aspects of school climate as well as collective efficacy (Goddard, Hoy, & Woolfolk Hoy, 2000) influence student achievement. Because SES is so difficult for schools to alter, researchers are searching for other properties of schools that are more
amenable to change (Bandura, 1993; Hoy, Smith, Sweetland, 2001). Transformational leadership of school principals and the collective efficacy of schools are two such candidates that may influence student achievement.

This study is concerned with a number of questions. First, is there a relationship between principal leadership behavior and collective efficacy of the school? What dimensions of transformational leadership are the best predictors of collective efficacy? Are collective efficacy and transformational leadership predictors of student achievement in mathematics and reading? Do transformational leadership and collective efficacy go beyond socioeconomic status in predicting student achievement in schools? Does socioeconomic status relate to transformational leadership and collective efficacy? Do transformational leadership and collective efficacy both have direct, independent effects on student achievement or does one work through the other? In order to answer these latter questions, a causal model using socioeconomic status, transformational leadership, and collective efficacy will be proposed.

**Research Hypotheses**

The preceding research questions give rise to the following series of hypotheses that will guide the empirical phase of this research:

1) All dimensions of transformational leadership behavior are positively related to teachers’ collective efficacy; that is, idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration are all positively related to collective efficacy.

2) Collective efficacy is positively related to student achievement, in particular reading and math achievement.

3) Socioeconomic status is positively related to collective efficacy.

4) Socioeconomic status is positively related to transformational leadership.
5) Transformational leadership behavior is directly related to student achievement and indirectly related to student achievement through collective efficacy.

6) Socioeconomic status is directly related to student achievement and indirectly related to student achievement through collective efficacy.

These hypotheses suggest the basis for a causal model that will be developed and refined after a thorough review of the literature for each of the major variables of this study. Hence, the purpose of the study is not only to examine bivariate relationships among the variables but also to develop a theoretical explanation of how the properties of schools work to increase the achievement level of students, measured at the school level. The study is concerned with explanation as well as prediction; indeed a casual model will be formulated and tested.

Definition of Terms

Various terms will be expressed throughout this study. For the sake of clarity, their meanings are intended as follows:

**Collective efficacy** – “a group’s shared belief in its conjoint capabilities to organize and execute the courses of action required to produce given levels of attainment” (Bandura, 1997a, p. 477)

**Human Agency** - a person’s intentional execution of action resulting from his or her control over his/her thoughts, behavior, and environment.

**Self-efficacy** - “Self-efficacy is a future oriented belief about the level of competence a person expects he or she will display in a given situation.” (Tschannen-Moran et al., 1998)
Social Cognitive Theory - A model for explaining psychosocial functioning in terms of reciprocal causation between behavioral expressions, cognitive and other personal factors, and environmental qualities that interactively influence each other. (adapted from Wood & Bandura, 1989)

Transformational Leadership - a supervisory model advocating the elevation of followers’ needs, values, preferences, and aspirations from self-interests to collective interests, causing them to become highly committed to the leader’s mission, to make significant personal sacrifices in the interest of the mission, and to perform above and beyond all expectations. (adapted from House & Shamir, 1993). The four dimensions of Transformational Leadership are often referred to as the four "I's" of transformational leadership. A description of each dimension follows:

Idealized Influence

The first "I" describes behaviors that develop followers' trust. Idealized influence demonstrates high standards of ethical and moral conduct, shares risks with followers in setting and attaining goals, considers the needs of others over their own desires, uses power to move individuals or groups toward accomplishing their mission, vision and cause, but never for personal gain (Hoy and Miskel, 2001, p. 415).

Inspirational Motivation

The second “I” describes spirited appeals and images to increase participants’ focus on a vision. Bass and Steidlmeier (1999) state these appeals and images encourage commitment to mutual goals and responsibilities.

Intellectual Stimulation

The third “I” prompts creative thinking about school problems, especially related to teaching and learning. The evaluation, formulation, and implementation processes of problem solving are all open for participant input. Leaders encourage the abandonment of traditional thinking. Stimulation of new ideas is promoted by questioning current strategies, posing problems from different perspectives, and supporting new procedures for work (Hoy and Miskel, 2001).
Individualized Consideration

The fourth “I” gives organizational members personal attention. The leader focuses attention on each teacher and staff person’s need for achievement and growth. Bass and Steidlmeier view individualized consideration as service to others, underscoring "… the necessity of altruism if leadership is to be anything more than authoritarian control (1999, p. 189)." Transformational leaders take time to know members individually and help them develop their capacity to grow and develop as educators.

Assumptions and Limitations

An overarching assumption throughout this research is that teacher perceptions accurately portray their own feelings as well as their principals' actions. Bandura’s research (1997a) with self-efficacy supports the notion that individuals' notions of self-efficacy are quite useful in making predictions about their behavior. Self-efficacy and collective efficacy are teacher beliefs about their own and their faculty's capabilities, respectively, which are important determinants of motivation and persistence in behavior.

The leadership behavior of the principal of a school is also determined by the perceptions of teachers, following Bandura's (1997a) sociocognitive theory. It is assumed that having multiple observations recorded about the same principal improves the accuracy of principal behavior perceptions; in fact, Halpin’s (1966) research on leader behavior supports such an assumption.

The current research was limited to elementary schools in Ohio. Fourth grade math and reading proficiency exams were used to measure student achievement. These are the elementary grade levels that the state uses to judge student proficiency in elementary schools. A diverse sample of elementary schools representing suburban,
rural, and urban schools was selected, but participation was voluntary. Hence, care should be taken when generalizing the findings to a larger population of Ohio elementary schools.

Relying on principal approval for school participation in our study invites questions about school self-selection. Perhaps factors worth further study drive a principal's decision to partake in such research, but that investigation is beyond the scope of this study. Perhaps such a distinction between participating and nonparticipating schools manifests in differing levels of teacher collective efficacy or transformational principal behavior. Although the sample was not random, an attempt was made to get a representative sample of Ohio elementary schools in terms of size, socioeconomic status, and population density.
CHAPTER 2
LITERATURE REVIEW

This chapter provides the conceptual and historical development of teacher efficacy, transformational leadership, and student achievement. The evolution of teacher and collective efficacy will be addressed before discussing transformational leadership and student achievement. Within the teacher and collective efficacy section of this chapter, a brief review of teacher efficacy definitions and some notable effects of teacher efficacy precede a survey of the concept's development since inception. Afterward, an integrated model of teacher efficacy is presented before collective efficacy's conceptual and construct frameworks are established. Separate prefaces will introduce the transformational leadership and student achievement sections later in this chapter.

Teacher Efficacy Definition

Teacher efficacy definitions have varied for most of its theoretical life. One of the first uses of the term "teacher efficacy" surfaced in research conducted by Barfield and Burlingame (1974), the earliest reference to teacher efficacy contained the Educational Resources Information Center system (Woolfolk & Hoy, 1990). Barfield and Burlingame's definition expressed teacher efficacy as "a personality trait that enables one to deal effectively with the world" (p. 10). Two years later, Armor et al. (1976) define teacher efficacy as "the extent to which the teacher believes he or she has the capacity to produce an effect on the learning of students" (p. 23).
Ten years hence, Ashton and Webb (1986) define teacher efficacy as the "teachers' situation-specific expectation that they can help students learn" (p. 3). Guskey and Passaro (1994) suggest teacher efficacy is a "teacher's belief or conviction that they can influence how well students learn, even those considered difficult or unmotivated" (p. 2). This definition serves as the working definition of teacher efficacy for this study.

Teacher Efficacy Effects

Teacher efficacy has been found to positively correlate with many teacher behaviors, teacher attitudes, and school contexts that are believed to facilitate student achievement. Indeed, teacher efficacy itself predicts student achievement (Ashton & Webb, 1986; Bandura, 1997a; Midgley, Feldhausser, & Eccles, 1989). The link between teacher efficacy and student achievement appears to be that "teachers' beliefs of personal efficacy affect their instructional activities and their orientation toward the educational process" (Pajares, 1997, p. 31). Some of these teacher behaviors and dispositions merit review for demonstrating the importance of this psychological construct.

Teacher Attitudes

Teachers with a low sense of efficacy tend to have very different views of student classroom control then teachers with a high sense of efficacy. Woolfolk, Rosoff, and Hoy (1990) studied the relationship between teacher efficacy and pupil control ideology with 55 religious schoolteachers. They found that the stronger the belief teachers had in their teaching efficacy, the more benevolent their pupil control orientation was. Conversely, the lower the view the teachers had of their teaching efficacy, the more rigid, untrusting, and controlling their classroom policies were toward pupil management.
Preservice teachers' beliefs about their capabilities appear related to their views of classroom management and their beliefs about bureaucratic control. Woolfolk and Hoy (1990) surveyed 182 aspiring teachers about their sense of instructional efficacy, pupil control, and bureaucratic orientation. Their findings suggest prospective teachers with a high sense of efficacy held more humanistic beliefs about student control and a greater bureaucratic orientation. Perhaps counterintuitive at first glance, prospective teachers low in their efficacy beliefs displayed a more rigid pupil control ideology and were less bureaucratic in their beliefs. The apparent disjoint between pupil control orientation and bureaucratic orientation might be explained by the composition of Woolfolk and Hoy's (1990) sample. New teachers confident in their abilities may be optimistic about a lock-step role of a school because the loyalty such feelings foster might provide a sense of security in a new position, while their beliefs in a humanistic orientation toward pupil control may be part of their effective teacher notions. "Loyalty is more appropriate in a context where one anticipates success" (p. 90).

Teachers' beliefs in their self-efficacy predict the likelihood they refer students for special education consideration. The relationship between teacher efficacy and their likelihood of making special education referrals emerged in an interesting study conducted by Meijer and Foster (1988). Two hundred thirty Dutch elementary teachers were given vignettes of supposed pupils illustrating their behaviors, work abilities, and school achievement. They were asked to indicate, on a scale from 0 to 100, the degree to which the students could realize a satisfactory educational experience in a regular classroom setting. Additionally, the teachers were prompted to rate the likelihood they would refer the student for special education consideration. Their results showed that
teachers with a higher sense of teacher efficacy were more willing to work with difficult students than were teachers with a lower sense of teacher efficacy. Soodak and Podell (1993) directed a very similar study with 192 regular and special education teachers and found very similar results.

Teacher Instructional Behavior

Teacher efficacy predicts not only teacher attitudes but also it predicts teacher behavior. Ashton and Webb's (1986) study of 48 basic skill teachers in a southeastern university community found a positive correlation between teacher efficacy and several "classroom process variables [teacher behaviors]" (p. 142). Paying close attention to students, giving students full attention, providing students individual attention, and monitoring students' understanding before starting new subject matter all positively related to teachers' belief in their ability. Ashton and Webb's (1986) finding that high efficacy teachers give students more individual attention complements Gibson and Dembo's (1984) observation that high efficacy teachers were more effective and persistent in leading individual students to correct responses during class discussions than low efficacy teachers. By contrast, low efficacy teachers would go on to other students, other questions, or just give the answer rather than continue working with the confused student.

The instructional practices of teachers correlates with teacher efficacy. Allinder (1994) surveyed 800 special education teachers from four Midwestern states to investigate the relationship between teacher efficacy and instructional practices. Statistically significant positive relationships between teacher efficacy and "instructionally relevant" (p. 92) behaviors were found. Those teachers who held
stronger beliefs in their teaching abilities were more likely to explore various teaching methods, approach their profession in an organized and "planful" manner, and manifest teaching confidence and enthusiasm.

In addition to special education, other teaching areas have provided contexts for similar teaching belief-teaching behavior dynamics. Enochs, Scharmann, and Riggs (1995) studied the relationship between science teaching efficacy of 73 preservice teachers and science teaching methods. Their findings indicate that aspiring teachers with greater science teaching efficacy preferred activity-based teaching methodologies to the lesser "hands-on" instructional modes.

Although more effects will be discussed during the review of the teacher efficacy construct development, their role in influencing student achievement merits attention now. The effects of teacher efficacy may play an important role in shaping the attitudes and behavior of teachers as they prepare for and engage in their craft, teaching students. Ross (1994) offers a brief review of teacher efficacy correlates as they relate to student achievement. His meta-analysis of 88 teacher efficacy articles summarizes some indirect means by which teacher efficacy beliefs may influence student achievement. In particular, Ross states that the greater a teacher's belief in his or her teaching competence, the greater the probability he or she may (a) explore and implement powerful teaching strategies, (b) employ classroom management methods that promote the development of self-directed learners, (c) tend to the deficiencies of lesser-able students, (d) encourage the self-efficacy of students to initiate contacts with the him or her, (e) work toward more
ambitious teaching goals, and (f) demonstrate persistence while working with academically difficult students. Thus, teacher efficacy may facilitate teaching behaviors that help students succeed.

Teacher Efficacy Construct Development

Although concepts related to efficacy have existed in psychology for more than four decades (White, 1959), teacher efficacy notions have existed in educational research only since the late 1970's (Armor et al., 1976; Berman et al., 1977). The reasons for psychological constructs inquiry behind teaching and education are found, partly, in the political fabric of our country's civil rights movement. A review of the political backdrop spurring school research into factors predictive of school achievement precedes a survey of teacher efficacy construct development.

The Origin of the Teacher Efficacy Research

The pursuit of school characteristics predictive of student achievement originated, primarily, from research commissioned by the 1964 Civil Rights Act. The historic law charged the Secretary of Education to authorize a study researching the state of public education in the country. Coleman et al. (1966) conducted the landmark studies to investigate the equity of educational opportunity in the United States. Their major findings illustrated a bleak picture for the effectiveness of schools and teachers:
It is clear that no strong [school] stimulus is making its impact felt in such a way as to interfere with the general relation of [student] background to achievement; that is, it is clear that schools are not acting as a strong stimulus independent of the child's background, or the level of the student body … This is not to say, of course, that schools have no effect, but rather that what effects they do have are highly correlated with the individual student's background, and with the educational background of the student body in the school; that is, the effects appear to arise not principally from factors that the school system controls, but from factors outside the school proper. The stimulus arising from variables independent of the student background factors appears to be a relatively weak one. (p. 311)

Indeed, educators began to question the effectiveness of schools and their influence on the underprivileged (Ashton & Webb, 1986). Subsequent to the civil rights era, educational research remained interested on minority students' progress in schools. Limiting research methodologies may have aided in the bleak findings of school research. Berman, McLaughlin, Bass, Pauly and Zellman (1977) propose that teacher attributes eluded the attention of research because those traits were felt to be out of the influence of school policy. Interestingly, although Berman et al. credit the Coleman Report as one of the few studies to consider teacher-level characteristics, the report's researchers admit that not many teacher variables were measured and no conclusions could be drawn regarding important teacher characteristics. Eventually, though, with the help of Rotter (1966), research surfaced that began to challenge the dire conclusions made about schools' effectiveness in the Coleman Report.

The First Conceptual Strand Of Teacher Efficacy

Rotter (1966) devised a social learning theory predicated strongly on behaviorism. He theorized that much of human behavior is learned because of the reinforcement it produces. What characterizes Rotter's social learning theory from others is the origin of those reinforcements, not necessarily what the reinforcements were. His theory posits
that over time, individuals build a sense of generalized expectancies about the origin of reinforcements, or outcomes. Based on their previous experiences, people tend to favor an internal locus of control, believing they primarily influence the outcomes of their behavior, or an external locus of control, believing luck, or external forces outside of their control, determines the outcomes of their behavior. Consequently, people greatly differ in their beliefs about locus of control because people's experiences so greatly differ (Hock, 1995). Rotter's (1966) Internal-External Locus of Control scale, an instrument used in a wide variety of settings, gauged people's general expectancy orientation.

Interestingly, one is hard pressed to find the word "efficacy" printed anywhere in Rotter's (1966) work, prompting questions about how it influenced the initial studies of teacher efficacy. Bandura (1982) offers a possible explanation for the loose application of the term "efficacy" to such work:

Recent years have witnessed a growing convergence of theory and research on the influential role of self-referent thought in psychological functioning (De Charms, 1968; Garber & Seligman, 1980; Lefcourt, 1976; Perlmuter & Monty, 1979; Rotter, Chance, & Phares, 1972; White, 1959). Although the research is conducted from a number of different perspectives under a variety of names, the basic phenomenon being addressed centers on people's sense of personal efficacy to produce and to regulate events in their lives. (p. 122)

Guskey (1982) and Guskey and Passaro (1994) identify White (1959), as did Bandura above, as being one of the first to conceptualize an efficacy-type theory. White's (1959) concept of competence and effectance motivation describe mankind's desire to "explore and master" their environment. While White's theoretical basis concerning these concepts has evoked some criticisms (Bandura, 1997a), his notion of people desiring control of their surroundings emanates through most notions of efficacy today. Rotter
(1966), in crediting White (1959) as having coined a concept similar to his locus of control theory, lends support to Bandura's (1982) integration of efficacy-related research. 

The Teacher Efficacy Construct's Debut in Research

Los Angeles Schools retained the RAND Corporation in 1975 to study the reading achievement gains made by minority students participating in the district’s Preferred Reading Program. With this endeavor, Armor et al. (1976) became the first researchers to employ teacher efficacy as a predictor variable in school research. Teacher efficacy in this study was based on the conceptual work of Rotter (1966), the first of two theoretical bases supporting teacher efficacy research.

Armor et al. (1976) gauged teachers' efficacy beliefs two ways. The first RAND item, "When it comes right down to it, a teacher really can’t do much because most of a student’s motivation and performance depends on his or her home environment," (p. 73) gauged teachers' feelings that the influence of the school community overwhelms a teacher's ability to positively affect student learning, an external locus of control orientation. The second item, “If I try really hard I can get through to even the most difficult or unmotivated students,” (p. 73) measured teachers' beliefs in an internal locus of control. The product of these two prompts indicated a level of teacher efficacy. The degree to which teachers believed they could control the reinforcements (student performance and motivation) of their behavior (teaching) comprised the RAND notion of teacher efficacy. Although family background variables explained the largest share of reading gain variance, the fact Armor et al. found teacher efficacy significantly and positively related to reading achievement among minority students was a significant discovery in light of Coleman et al.'s (1966) conclusion about the futility of schools.
Berman et al. (1977), another RAND study, investigated the factors associated with sustained federally initiated school programs. Teacher efficacy was gauged using the two prompts employed on the previous RAND survey. The findings of this research demonstrated teacher efficacy positively relating to the continuation of federally funded projects, percentage of project goals accomplished, teacher change, and the continuation of project techniques and resources. The teacher efficacy coefficients indicated some of the strongest relationships in their analysis. As exciting as these findings were, much work lied ahead for refining teacher efficacy as a psychological construct utilized in educational research.

Multiple Item Teacher Efficacy Scales

One serious concern researchers had with the RAND construct of teacher efficacy was its brevity. Single item concept measures manifest problems of reliability, limited expressions of concept representations, and constrained variation of values that weaken relationships between variables (Bandura, 1997a). Perhaps both the interest in and psychometric problems of the RAND teacher efficacy construct inspired subsequent development of related scales. Notable among those successor instruments were Rose and Medway's (1981) Teacher Locus of Control and Guskey's (1981) Responsibility for Student Achievement. A review of these instruments reveals possible conceptual weaknesses of the RAND efficacy measure.

Teacher Locus of Control

Rose and Medway (1981) felt Rotter's (1966) Internal-External Locus of Control scale was carelessly used in school settings, believing it was never intended for such specific applications as classroom instruction. Thus, they created a locus of control scale
tailored to the explicit experiences of teachers. A 28-item forced choice instrument evolving from their work included prompts describing student successes and failures.

Eighty-nine fourth grade teachers responded to their Teacher Locus of Control instrument (TLC) for which data was reported. Factor analysis utilizing principal components extraction and varimax rotation fitting revealed two strong dimensions - internal responsibility for student successes (I+) and internal responsibility for student failures (I-E). Data from teachers responding to both Rotter's (1966) Locus of Control index (I-E) and the TLC scales yielded positive and significant correlations between I-E and I+, yet showed statistically insignificant correlations between I-E and I-, suggesting Rotter's scale may not have considered situations constituting negative outcomes in the manner Rose and Medway (1981) have. Also, the discrepancy in subscale correlations with I-E added credence to the TLC's additional dimensions.

The authors of the TLC scale considered the circumstances surrounding Berman et al.'s (1977) teacher efficacy estimations to be a sound context in which to gauge the validity of TLC instrumentality, reasoning that the RAND variable "teacher efficacy" was conceptually similar to the TLC. Accordingly, Rose and Medway (1981) disseminated their survey to 45 teachers recently completing in-service trainings about innovative instructional techniques. Their findings showed that their I- scale more consistently predicted criterion associated with outcomes of the in-service (e.g., number of reading groups, use of learning centers, number of math groups, etc.) than did Rotter's I-E scale, especially in the middle elementary grades. Rose and Medway's (1981) legacy might be that constructs attempting to measure teacher expectancies about student performance
should distinguish between affirming and negative achievement. Their findings regarding the differentiation of teacher expectancies for student achievement are corroborated by similar locus of control research.

Responsibility for Student Achievement

Guskey (1981) also pursued teacher locus of control research initiated by the RAND studies. Venturing to measure teachers' beliefs regarding responsibility for the classroom successes and failures, Guskey constructed the Responsibility for Student Achievement scale (RSA). Akin to Rose and Medway (1981), Guskey included an equal number of positive and negative student outcomes in the RSA across internal and external control factors, believing that teachers' dispositions toward student achievement accountability form differentially depending on the quality of student performance in question. Guskey cited Crandall, Katkovsky, & Preston (1965) as having constructed a prior instrument for assessing students' attitudes toward internal control factors influencing their school learning outcomes (i.e., the Intellectual Achievement Responsibility Questionnaire). The earlier research showed evidence of two internal responsibility factors, one for positive achievement expectancies and one for negative achievement expectancies.

Deviating from research already reviewed, the RSA scale employed alternative-weighting items to uphold Guskey's belief that few student achievement events manifest because of a singular factor. Thus, teachers allocated one hundred percentage points between two possible explanations for a student performance scenario. Two hundred fifteen elementary school teachers answered Guskey's instrument during a pilot study of his scale. Factor analysis performed on the pilot study results indeed supported the
existence of two scales, each one describing qualitatively different student outcome
cenarios, positive and negative student achievement. The weak correlation between the
two scales provides this support.

Consequently, the responses to the RSA items yielded three indices of teacher
responsibility for student performance: a global score, a subscale for positive student
outcome responsibility, and a subscale for negative student outcome responsibility.
Teacher responses to positive student achievement were averaged to arrive at a mean
percentage allocation of teacher responsibility for positive student outcomes. The same
was done for teacher responses to negative student achievement explanations. Finally, a
global measure of teacher overall responsibility for student achievement emerged from
the averaging of the two subscale indices. The meaning of the global score may pale to
the information gained by interpretation of the subscale score, in light of the differing
dispositions teachers may hold for positive and negative student achievement
responsibility. A global index possibly obscures the psychology underlying teacher's
feelings for student achievement responsibility.

The significance of Guskey's (1981) and Rose and Medway's (1981) discoveries
of multi-dimensional teacher efficacy constructs resonated with contemporary research
concerning the more general psychological concept of self-efficacy. Albert Bandura,
psychology professor from Stanford, developed a unified theory of human behavior, its
origin being an explanation for the mediating effects of self-efficacy on various
psychological treatments. Attention now will be focused on Bandura's (1977) conception
of self-efficacy, its explanation for human motivation and behavior, and its broader theoretical context, social cognitive theory, for its repercussions on later teacher efficacy research endure to this day.

Second Conceptual Strand of Teacher Efficacy

Bandura's self-efficacy theory provided a second, arguably more comprehensive, model for developing the teacher efficacy construct. Whereas Rotter's (1966) locus of control theory suggests people tend to be motivated when they perceive the outcomes of their behavior to be internally determined (outcome expectancies), Bandura (1977) argues that behavior motivation stems from people's self-efficacy beliefs, expectations that they are able to perform a behavior at a certain level of attainment (efficacy expectancies). Although their functions interact in determining human behavior, the two psychological constructs are wholly different, conceptually and empirically (Bandura, 1997a). Differences emerge from a study of self-efficacy theory.

In 1977, Bandura introduced his unified theory of behavioral change, based primarily on his concept of self-efficacy. One of the theory's distinguishing features is that it explains the effects of various interventions on behavior. The power of Bandura's theory rests not only in its ability to account for behavior change based on four different sources of stimuli, but it prescribes the means for altering behavior. Bandura (1971) states that the value of any theory lies in its ability to both explain past phenomena and predict future behavior based of the construct in question. The numerous and central applications in which self-efficacy theory performs explanatory and predictive functions demonstrate its importance in psychology (Bandura, 1997a).
Bandura's (1977) theory of self-efficacy stretches beyond the field of psychology to include social and environmental elements. Foreshadowing Sarason's (1981) suggestion for revitalizing the narrow field of psychology, Bandura formulated his self-efficacy theory within a broader social cognitive theory, a model explaining the interactive influences of the environment, behavior, and personal factors on each other. A review of the major concepts and assumptions of social cognitive theory precedes an examination of the key points of self-efficacy theory, for this will lay the foundation for the second strand of teacher efficacy theory development.

**Social Cognitive Theory**

Social cognitive theory implies several assumptions about human behavior and its relationship to the environment and human cognition. These assumptions are important, for they have not always been shared among prominent psychologists. The research spurring the discovery of the self-efficacy construct actually brings together two divergent strands of the social sciences - psychology and sociology. The results of bridging these two disciplines afford a powerful construct predictive of many valued attitudes and behaviors. The assumptions, development, effects, determinants, and motivational processes of self-efficacy are shared below.

**Assumptions.** Social cognitive theory, an explanation of behavior change, rests on the acceptance of several assumptions. Primary among the assumptions lies the belief that humans are agentive, or can exert influence. A rejection of strict behaviorism, social cognitive theory assumes human behavior to be both constructive and reactive. "The human mind is generative, creative, aspiring, and proactive, not just reactive. Hence, people are both producers and products of their life conditions" (Bandura, 1999, p. 214).
People do not sit passively, waiting to be prompted by the environment before enacting their next behavior. Whereas individuals’ levels of initiative spans the continuum from low to high, Bandura (2001) maintains that every human contains the capacity to influence their environment and their own thinking, despite the most radical views of Skinner (1974).

It is the very notion of influence between human and environment that speaks to the next assumption of social cognitive theory. Bandura (1997a) posits triadic reciprocal causation as the theoretical design linking behavior, socio-contextual features, and personal factors. All three dimensions mutually determine each other in an interactive manner. Envision a triangle with each of the factors at a different vertex. Each side constitutes bi-directional paths of causation between vertices. For example, environment influences behavior and behavior influences the environment. Consider the following diagram for understanding triadic reciprocal causation.

![Diagram of triadic reciprocal causation]

**Figure 2.1: Generic Social Cognitive Model of Human Agency (Bandura, 1999)**
The joining of personal and social factors for explaining human behavior integrates uncommon theoretical bedfellows. Social cognitive theory breaks away from traditional models describing human behavior by making such linkages between the two behavioral sciences. Indeed, concentrating on only one branch of science would provide an incomplete theoretical rationale of human functioning. "A full understanding [of behavior] requires an integrated causal perspective in which social influences operate through self-processes that produce the actions” (Bandura, 1997a, p. 6).

Maddux (1995) summarizes the remaining assumptions relevant to social cognitive theory. People possess the ability to internally symbolize past experiences, hypothetical scenarios, and the conveyance of sophisticated thoughts to others. Due to our potential to mentally encode thoughts, self-reflection enables people to analyze and assess thoughts and experiences - a metacognition. Exploiting our metacognitive capabilities, humans exercise self-regulating functions to change the circumstances of their existence. Such corrective actions manifest in the selection or alteration of environmental conditions and by exerting influence over personal cognitions. Finally, people learn not only by enduring the lessons afforded by trial-and-error experiences, but acquiring knowledge and skills vicariously, or through their observation of other people's behavior. Without such an alternative to trial-and-error, the process of learning would be slow, to the point that obtaining sophisticated information or skills would be impossible (p.5). Together with social cognitive theory's beliefs in human agency and the tripartite structure of reciprocal influence, these assumptions set the stage for a focused review of Bandura's (1977) self-efficacy theory.
*The original research.* Bandura (1977) studied the relationship between behavioral changes in phobic patients and their respective modes of clinical treatment. Noting the lack of an integrated theory for explaining the effects of various treatments on patients, Bandura endeavored to create a unified theory of behavior change. He traced the evolution of clinical therapy in psychological settings from behavior-oriented treatments to cognitively centered interventions. Noting the indiscriminant results of both treatments, he ventured to explain a common psychological process mediating all interventions. With this work, he founded the theory of self-efficacy. The thoroughness of his model manifests in its explanatory framework and its empirically tested suggestions for improving self-efficacy beliefs.

Self-efficacy beliefs are the expectations one holds for garnering the necessary competencies, motivation, and courses of action for attaining a specified performance standard (Bandura, 1997a). The novelty of Bandura's definition of efficacy lies in its focus on performance attainment (Pajares, 1997). Many expectancy theories prior to Bandura's landmark study focused on "people's hopes for favorable outcomes rather than with their sense of personal mastery" (Bandura, 1997a, p. 194). Beliefs in attainment derive from a combination of psychosocial and cognitive processes. The benefits of positive self-efficacy beliefs are many. A survey of self-efficacy effects leads a discussion of self-efficacy sources of information and self-efficacy cognitive processes.

*The effects of self-efficacy beliefs.* As noted earlier, one indication of a psychosocial theory's value rests with its ability to predict behavior based on the central construct in question. The predictive quality of self-efficacy theory, by way of its related effects, more than holds its value across many contexts of human behavior. Bandura's
self-efficacy model has been successfully applied to a variety of settings, including mental health, cardiovascular rehabilitation, smoking behavior, business decision making, school achievement, sports performance, and job selection (Pintrich & Schunk, 1996). The reason self-efficacy theory fits so many domains of experience lies in its central role for explaining general psychosocial functioning. Serving as the chief personal factor, self-efficacy is the main psychological construct of social cognitive theory's account of behavior. Indeed, "Bandura provided a view of human behavior in which the beliefs people have about themselves are key elements in the exercise of control and personal agency" (Pajares, 1997, p. 3).

Observing some of the general effects of self-efficacy beliefs helps in understanding its diverse applications across various domains of behavior. People who hold stronger perceptions of their efficacy seek more challenging tasks, exert more effort during difficult endeavors, persist longer during such efforts, maintain stronger resiliency to adversity, entertain enabling rather than hindering thought patterns during demanding tasks, endure less stress during taxing activities, and realize greater performance attainments than those who hold weaker self-efficacy beliefs (Bandura, 1997a). The common struggle against hardship stands as the common theme throughout these effects, indicating self-competency perceptions may be pivotal for the development of people's coping skills. The means for enhancing competency perceptions comprise another facet of his theory, highlighting the comprehensive quality of Bandura's (1977) self-efficacy model.

Sources of self-efficacy beliefs. Recall that Bandura (1977) set out to explain why various forms of clinical treatment had altered the behavior of patients' psychosocial
functioning in his seminal work on self-efficacy. The results of his research yielded "a common mechanism of behavioral change" through which the four modes of treatment were cognitively processed (Bandura, 1982, p. 127). The four interventions eventually became known as the sources of self-efficacy beliefs: mastery experiences, vicarious learning, social persuasion, and physiological states. In judging one’s self-efficacy beliefs, individuals evaluate information from these four primary sources (Bandura, 1986). A brief description concerning each mode of influence follows.

Mastery experiences entail the favorable recall of behaviors relating to the performance standards in question. Information about personal experiences relevant to the psychosocial functioning provides the strongest indications of self-efficacy beliefs about that behavior (Bandura, 1977). When recall determines past occurrences participating in such activity yielded success, efficacy perceptions will likely be strong. However, past similar events resulting in failure strongly suggest efficacy beliefs to be low. Mastery experiences enhance self-efficacy perceptions by building effective strategies for accomplishing various instructional tasks. Repeated successes with certain strategies strengthen one’s efficacy in performing such tasks. Enactive sources of information are the most influential sources of efficacy information (Bandura, 1977, 1982).

Vicarious learning, or modeling, bolsters self-efficacy by positively comparing another's capabilities with one's own. Put another way, self-efficacy develops from relating similarly skilled individuals’ task performances to one’s own capabilities. This form of efficacy information carries less potency for competency enhancement or decline than mastery experiences, but delivers needed information when past experiences or
clarity of such incidents is lacking (Pajares, 1997). Additionally, the perceived likeness of
the person modeling the behavior implicates the validity with which one attaches to such vicarious information. The more the observer and the onlooker are the same (i.e., ability, age, peer standing, etc), the more weight such efficacy information will hold (Bandura, 1997a). Indeed, "… modeled successes by similar others raise observers' beliefs in the efficacy, and modeled failures lower them" (p. 96).

Social persuasion influences teachers’ self-efficacies through realistic encouragement concerning capabilities from knowledgeable others. A major factor in discerning the value of social persuasion is the "expertness, trustworthiness, and attractiveness" of the source of persuasion (Maddux, 1995, p.11). Those persuaders enjoying more credible statuses enhance the efficacies of their targets more effectively. Shallow, empty, or nonsubstantive feedback adds little to one's competency perceptions. Pajares (1997) reinforces the necessity of substantial content in social persuasion by recalling Erickson's (1980) advice that egos reject feeble outside attempts at inflation. Bandura (1996) states that it is easier to weaken one's self-efficacy beliefs via social persuasion than it is to bolster one's competency expectations, due to the relative ease with which unsubstantiated praise evaporates, as does one's efficacy perceptions, upon unrealized performance attainments.

Finally, physiological states influence self-efficacy perceptions by informing an individual of their physical and emotional readiness for undertaking a given task. The body's condition during performance tasks conveys somatic signals about the relative comfort of pursuing such endeavors. These signals, collectively classified as physiological and emotional states, indicate a visceral sense about the body's
"confidence" in engaging in identified activities. Sweaty palms, shortness of breath, and a dry mouth suggest an activity may be beyond the capabilities of the actor. However, the diffuse and indistinct nature of physiological and emotional states can compromise their value in determining self-efficacy perceptions (Bandura, 1997a). Interpretation of this information alone may prove to be unreliable. Does a racing heart indicate fear, excitement, or exertion? Many times, context clues need to be combined with somatic cues to more accurately interpret the results of a given performance. For example, a high school track star, through his friend's dare, competes against local college runners. During the race, his heart, skin, and breathing rate will likely behave similarly to his opponents. Physiological cues alone will not indicate to the high school runner how his performance is faring. Only through contextual signals can he ascertain the outcome of his efforts. Still, because physiological cues, such as the ones mentioned above, pervade such a wide realm of performance experiences, good physical health and positive mental attitudes positively influence self-efficacy beliefs (Bandura, 1997a).

In sum, each source of self-efficacy beliefs contributes to one's reading of self-competence in its own way. However, one does not perceive self-efficacy based on the reading of only one source of information. Undeniably, it is through an integrated analysis of the four self-efficacy sources that accurate determination can be made. The weight given to each source varies with personal interpretation, selection, and recollection of the information (Pajares, 1997). Moreover, the process of determining one's efficacy for a given performance is just a component of the larger process of motivation-generation, indeed the inducement for much behavior. A review of cognized motivation occurs next.

33
The motivating processes of self-efficacy beliefs. Bandura (1997a) states proactive behavior and intrinsic motivation largely derive from cognitive activity. The future vision compelling one to act is not the motivation for action, though. Bandura argues that desired future attainments generate forethoughts represented as personal goals, efficacy beliefs, and outcome expectancies. This internal motive-generating process is termed cognitive motivation and encompasses three schools of theory: (a) causal attributions [attribution theory], (b) outcome expectancies [expectancy-value theory], and (c) cognized goals [goal theory].

While stressing the strengths of cognized goal theory, Bandura (1997a) shares the subpar effects and indulgent assumptions of causal attribution theory and expectancy-value theory, respectively. The meager relationships causal attributions have with performance motivation compromises its use as an intervention for behavior modification. Expectancy-value models of cognitive motivation, in all its related forms, make assumptions that stretch one's idea about how people make choices to form performance motivation. Bandura disagrees with the systematic process most expectancy-value theoreticians apply to peoples' actual outcome selection behavior. Because comprehensive outcome information for a given performance is rarely available and people often make outcome predictions based on certain biases, the outcome optimization system they assume to be rational is not (p. 125). Coupled with its theoretical shortcomings, expectancy-value theory's empirical ability to explain performance motivation often falls short of hoped for yields. Expectancy-value theory endures, though, as "recent efforts to increase the predictiveness of expectancy-value
models have added an efficacy-like factor to the usual set of predictors" (p.127). The pairing of efficacy and outcome expectancies implicates the development of teacher efficacy constructs, as we will review shortly.

Goal theory, however, shows strong support for predicting performance motivation in both lab and field studies (Bandura, 1997a). The basis for goal theory lies in its reliance on explicit, short-term goals for behavior inducement. Proximal goals, or more immediate goals, are much more practical for personal growth than goals inferred from distant future visions. Bandura (1982) demonstrated the strength of proximal goals over distant ones by his research on children with deficient skills and low-interest in math. The students maintaining short-term goals made greater gains in self-directed learning, developed higher competencies in mathematical operations, and facilitated stronger efficacy beliefs in mathematics than did members of a control group or children utilizing more distant goals or no goals. "Whereas proximal subgoals provide immediate incentives and guides for action, distal goals are too far removed in time to effectively mobilize effort or to direct what one does in the here and now" (Bandura, 1982, p. 134). Goals operate with two other personal factors within the social cognitive model to comprise the self-reactive influences forming motivation: self-efficacy and affective self-reflection.

The interaction of personal goal setting, affective self-reflection, and self-efficacy beliefs produce varying degrees of cognitive motivation, depending on the product of the respective magnitudes and directions. The emotional response to a given performance, when compared against a predetermined goal, describes affective self-reflection. If goals have been met through the execution of a task, the response can be positive and
encouraging. The prospect of experiencing satisfaction from goal attainment is certainly motivating for pursuing similar future tasks. Goals not met after task completion can be negative or motivating, depending on the strength of one's efficacy beliefs for the task at hand.

If efficacy beliefs are strong, then unsatisfied goal attainment, or a newly created proximal goal, creates discrepancy production, a gap between behavior and cognized goals (Wood & Bandura, 1989). The pursuit of such standards through enacted behavior causes discrepancy reduction, the desire to bring behavior in line with mental standards. If self-efficacy beliefs are such that confidence in attaining unmet performance standards, or newly entertained goals, falters then equilibrium will be achieved through lowering the standard. Since people are unlikely to stay in an inert state for long, "The goals people set for themselves at the outset of an endeavor are likely to change, depending on how they construe the pattern and level of progress they are making" (Bandura, 1997a, p. 129). Accordingly, the goals change to fit the progress and efficacy perceived.

Although outcome-expectancy cognitions bowed to cognized goals for predicting performance motivation, outcome expectations paired with efficacy beliefs predict human behavior (Bandura, 1997a, p.20). The two factors of psychosocial functioning suggested by Bandura, outcome expectancies and efficacy beliefs, led to later conceptions of teacher efficacy and collective efficacy (Bandura, 1993; Gibson & Dembo, 1984; Goddard, Hoy, & Woolfolk Hoy, 2000; Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998; Tschannen-Moran & Woolfolk Hoy, 2001). Although not entirely adopting Bandura's factors of human behavior, Ashton, Olejnik, Crocker, and McAuliffe (1982) offers an interesting, if not confusing, mix of RAND and Bandura ideas in their teacher efficacy

*Webb Efficacy Scale*

Ashton et al. (1982) interpreted the RAND teacher efficacy items much differently than the Rose & Medway (1981) and Guskey (1981). Instead of reading the two efficacy items as polar examples of locus of control, they viewed the items as prompts measuring two dimensions of teacher efficacy. The fact the two RAND items insignificantly correlated across several data sets of teacher responses suggested there was no relationship between the underlying factors of the items. Had the RAND items been polar points on a locus of control continuum, one would have expected a significantly negative correlation value between the two prompts.

Ashton et al. (1982) used the RAND items to build another model of teacher efficacy. Utilizing Bandura's (1977) theory to explain teacher efficacy, they built upon the RAND 1 prompt by adding more items querying teachers about their beliefs in external factors affecting student achievement (outcome expectancies). Teachers' beliefs about their ability to influence student learning, despite difficult circumstances, comprised a seven-item scale, called the Webb Efficacy Scale (WES). The format of the scale mirrored that of RAND 1 -- two beliefs regarding the same aspect of teaching preceded a two-option, forced-choice response section. For instance, "Heterogeneously grouped classes provide the best environment for learning" and "Homogeneously grouped classes provide the best environment for learning" (Ashton et al., 1982, p. 21)
embody one set of beliefs to which teachers would respond by indicating their preference. The results of their WES pilot study did not bode well for its future use as an instrument measuring general teaching efficacy. The instrument contained low internal correlations among its items, and its correlations with RAND 1 were not significant. The preceding WES example might indicate why the psychometric properties of the scale yielded subpar qualities. Class composition issues implicate teaching methodology options (i.e., cooperative group learning, peer tutoring, lecture, etc.) as much as they may indicate beliefs about outcome expectancies.

The Ashton Vignettes, Ashton et al.'s (1982) second instrument, were constructed to broaden RAND's conception of personal teaching efficacy, or teachers' beliefs in their own teaching efficacy. RAND II measured teachers' beliefs to reach even the most difficult students. Ashton et al. expanded the personal teaching concept to include explicit scenarios associated with teaching, in the hopes "situational vignettes would elicit more teacher variability" (p. 11). Following each of the 15 scenarios, teachers circled their level of agreement on a sliding 7-point response scale. The results of the vignette studies appreciably bettered those of the WES. There were no significant correlations between RAND 1 and the 15 vignettes, but seven significant correlations emerged between RAND 2, the personal teaching efficacy standard, and the vignettes.

In all, Ashton et al. (1982) chose not to base a longer teacher efficacy scale on the concept of locus of control due to empirical support and theory. Rather they conceptualized teacher efficacy based on the notion of pseudo-outcome and efficacy perceptions. Although they moved the study of teacher efficacy in a new direction, they failed to develop a reliable teacher efficacy scale themselves. Ashton et al's (1982)
conceptual and empirical difficulties may be that they strongly suspected the existence of
two dimensions for a teacher efficacy, but they failed to fashion their scale items to
reflect what those dimensions were. The basis for this conjecture lies in Ashton and
Webb's (1986) acknowledgement admitting their state of thinking in 1982:

Although the Webb Efficacy Scale Questionnaire was developed before we were
aware of the need to distinguish between two dimensions of teachers' sense of
efficacy, five of the seven items on that instrument appear to measure teachers' sense of teaching efficacy. (p. 140)

Moreover, in research published four years later by the same lead author, Ashton, the
RAND items measuring teacher efficacy continued to be the main variables against
which teacher behaviors were correlated (Ashton & Webb, 1986). Indeed, Ashton et al
may have been caught between two theoretical orientations for conceptualizing teacher
efficacy. Research more uniform in its theoretical grounding comes from Gibson and
Dembo (1984). Bandura (1997a), himself, credits the following study with improving the
measurement of teacher efficacy by measuring the two facets of teacher efficacy with
more than one-item prompts (p. 243).

Teacher Efficacy Scale

Concurring with Ashton et al.'s (1982) contention that teacher efficacy manifests
two factors, rather than the one factor evident in Armor et al. (1976) and Berman et al.
(1977), Gibson and Dembo (1984) endeavored to improve upon previous efforts at
constructing a teacher efficacy scale. They embraced Bandura's (1977) self-efficacy
theory for constructing a new teacher efficacy instrument. Their work in establishing a
two-dimension construct met more favorable empirical and succeeding literature support
than did their conception of their second teacher efficacy dimension.
Gibson and Dembo (1984) used as the basis for their teacher efficacy instrument Bandura's (1997a) notion that behavior is best predicted by utilizing one's self-efficacy beliefs and outcome expectancies for a given performance (p. 20). As such, one of their teacher efficacy factors, termed personal teaching efficacy, reflected items gauging a teacher's beliefs in his own ability to "bring about student learning," a self-efficacy perception (Gibson & Dembo, 1984, p. 573). The researchers relate this scale directly back to Berman et al.'s (1977) RAND item 2. The other factor, termed general teaching efficacy, allegedly assessed a teacher's outcome expectancies for his teaching. This factor directly linked to Berman et al.'s (1977) RAND item 1. A 6-point likert scale ranging from strongly agree (1) to strongly disagree (6) provided teachers the means for response. Results from the administration of their Teacher Efficacy Scale to 208 elementary teachers confirmed the presence of these two factors. Factor analysis not only confirmed the existence of the two dimensions, but it also reduced the usable variables from 30 to 16.

Consequently, nine items, exemplified by such statements as "When a student is having difficulty with an assignment, I am usually able to adjust to his or her level," made up the personal teaching efficacy dimension (Gibson & Dembo, 1984, p. 581). Seven items, illustrated by prompts like "Even a teacher with good teaching abilities may not reach many students," (p. 582) comprised the general teaching efficacy dimension. The significant, but low correlation (r = -.19) between both factors supported the idea that both dimensions represented related, but independent components of teacher efficacy.

40
With personal teaching efficacy explaining 18.2% of the variance and general teaching efficacy accounting for another 10.6%, almost 30% of the total variance was predicted by these factors.

Although the personal teaching efficacy component met less resistance in subsequent literature, the general teaching efficacy dimension came under fire. Tschannen-Moran et al. (1998) criticized the TES second factor, general teaching efficacy, for not actually measuring teachers' outcome expectancies of their instruction. In their 1998 respected piece "Teacher Efficacy: It's Meaning and Measure," the authors remind readers that an individual's outcome expectancies for a given behavior, as Bandura conceived of them, can only be measured from that individual's perception of his own outcome beliefs, not from gauging the individual's perception of others' outcome expectancies for the same behavior, as how Gibson and Dembo measured their general teaching efficacy construct. Indeed, Bandura (1997a) states:

> People regulate their performances by beliefs of what they can do rather than by expectations of what effort can do for others. Danehower [1988] shows this to be the case. Efficacy beliefs contribute to performance, whereas generalized effort expectancy does not. (p. 126)

Regardless, though, certain teacher efficacy effects emerged from Gibson and Dembo's (1984) study. Classroom observations of both high and low efficacy teachers, of whom only eight were studied, revealed differentiated classroom teaching behaviors. High efficacy teachers tended to provide less criticism to students, persist longer with students having difficulty, give more informative feedback, and spend less time in small group activities than do low efficacy teachers. Although it may be counterintuitive at first, higher efficacy teachers may be more confident to engage the whole class, whereas,
perhaps, less efficacious teachers feel more secure leading smaller groups of students. Gibson and Dembo comment that their student grouping findings mesh with the teacher effectiveness literature.

The debate revolving around a teacher efficacy construct shifted from whether it contained two dimensions to what the factors represented. As we will see, the dimensions of teacher efficacy carried the burden of scholarly challenges for several years to come. The following studies provide insight into the controversy surrounding the factors contributing to teacher efficacy beliefs.

Dimension Reconceptualization

In a turn of interpretation, Woolfolk and Hoy (1990) reconceptualized the general teaching efficacy factor as it had been previously considered (Ashton & Webb, 1986; Gibson & Dembo, 1984). Woolfolk and Hoy, in their study of efficacy beliefs and student control perceptions of prospective teachers, reconsidered the outcome expectancy interpretation of the general teaching efficacy factor. Noting that outcome expectancy connotes a belief about consequences predicted from enacted behaviors, Woolfolk and Hoy felt previous applications of this concept to teacher efficacy constructs did not fit the items comprising its scale. An item typically found on outcome expectancy scales, or general teaching efficacy constructs, had been "When it comes right down to it, a teacher really can't do much because most of a student's motivation and performance depends on his or her home environment" (p. 82). Such an item implies beliefs about the ability of teachers in general, not beliefs about performance outcomes. Although Woolfolk and
Hoy (1990) still used a modified Teacher Efficacy Scale, they clarified the meaning of its second dimension, for "drawing conclusions about results of studies of teacher efficacy, we must look first at how efficacy is defined in each study" (p. 90).

Woolfolk and Hoy's (1990) discovered that the personal teaching efficacy subscale may be further differentiated into responsibility for positive student achievement and responsibility for negative student achievement. The researchers ultimately decided to stay with the two-factor solution of teacher efficacy due to parsimony. The teacher efficacy measures maintained the same relationships with the pupil control ideology and bureaucratic orientation variables, regardless of how the researchers factored the personal teaching efficacy construct. This study highlights one of the first steps in integrating different models of teacher efficacy conceptualizations and it implicated future thinking about the factors constituting teacher efficacy beliefs.

This study's emphasis on relating the dimensions of teacher efficacy to other variables speaks to the importance thorough self-efficacy study. Case in point, a reverse relationship emerges for teaching efficacy's relationship with pupil control ideology. High teaching efficacy teachers, when dichotomized between low and high personal efficacy beliefs, show divergent trends in pupil control ideologies. High personal efficacy teachers favor more humanistic means of interacting with students, whereas low personal efficacy teachers bias a more rigid management of students. Although divergent trends manifest after dichotomizing the subjects with low teaching efficacy beliefs, the
relationships pupil personal efficacy perceptions and pupil control orientations are reversed. Woolfolk and Hoy (1990) warn that researchers studying relationships only between whole teacher efficacy measures and other concepts risk missing important relationships between subscale concepts and other variables.

Woolfolk and Hoy (1990) has its detractors. Two studies, in particular, take issue with how Woolfolk and Hoy conceptualized the dimensions of teacher efficacy. Guskey and Passaro (1994) disagreed with the researchers interpretation of the negative and positive factors observed in their factor analysis while Soodak and Podell (1996) reconceptualized Woolfolk and Hoy's three dimensions. A brief look at the dissenting articles precedes a review of scholars' calls for a reconsideration of the whole teacher efficacy construct.

_Locus of Control Revisited_

In an attempt to clarify the opposing views of Woolfolk and Hoy (1990) and Gibson and Dembo (1984) regarding the teacher efficacy factor structure, Guskey and Passaro (1994) administered a modified Teacher Efficacy Scale (Gibson and Dembo, 1984) to 342 teachers, both aspirant and veteran, to help ascertain the latent concepts determining teacher efficacy. The factor analysis performed on the subjects' responses forced a two-dimensional factor structure yielding 32% of their score variance. Reminiscent of Rotter's (1966) locus of control theory, Guskey and Passaro suggest the negative and positive items loading on separate factors, characterized by an external and internal teaching orientation, respectively, corroborate their supposition that teachers' beliefs about the instrumentality of their actions drive their efficacy perceptions.
A teacher efficacy construct built on a new locus of control model derived from Guskey and Passaro's findings. During the time of the RAND studies, teacher efficacy was based on locus of control theory conceived as a continuum, with internal factors at one end and external factors at the other. Guskey and Passaro explain this notion well, in that, when locus of control is conceived in this manner, attributions of causality follow zero-sum rules. If more attributions favored internal sources of control, then fewer attributions were available for external explanations. However, their conception of locus of control followed a two-factor dimension, with internal factors of causality loading on one factor and external factors falling on the other. Following this model, teachers could still hold strong feelings of internal causality, regardless of their beliefs about external influences.

As provocative as their discovery was, the construction of their scales limits the reach of their conclusions. Despite the authors' claims to clarify the confusion surrounding the teacher efficacy construct, their linking of internally-oriented items with positive student outcomes and externally-oriented attributions with negative student outcomes makes their assertions about locus of control premature. At best, Guskey's (1982) earlier notion about teacher responsibility for student achievement resonates with their survey results, further confounding the conceptual state of teacher efficacy. Their study was not the only one to fall short of discerning a clear picture of teacher efficacy influences.

**Outcome Efficacy**

Guskey and Passaro's (1994) work had several logical gaps according to Soodak and Podell (1996). First, their contention that locus of control beliefs fall across two
dimensions buck long-held notions about locus of control being a continuum (see Rotter, 1973). Secondly, the internal and external connotations in Guskey and Passaro's (1994) scale items defy intuition and lack explanation. The fact Guskey and Passaro (1994) constrain their factor analyses to extracting two dimensions severely limits their ability to detect any other sources of teacher efficacy beliefs.

Soodak and Podell endeavored to find the dimensions of teacher efficacy by administering a modified Gibson and Dembo (1984) Teacher Efficacy Scale to 310 teachers working in and around New York City. They added 18 items to the original 16, supplementing the scale with items relating to heredity, eating patterns, and TV watching practices, hypothesizing that student behavioral and emotional influences could contribute to teacher efficacy beliefs. Prior research into teacher efficacy sources hadn't considered external influences other than the student's home life. As such, the aforementioned student characteristic variables were included on Soodak and Podell's teacher efficacy instrument.

Principal components factor analysis revealed a three-factor solution explaining 30% of the response score variation. The items comprising the first factor described personal efficacy, a dimension pertaining to a teacher's beliefs in his own teaching competency (e.g., "I can effectively manage my students' behavior"). Similar to past teacher efficacy analyses, items pertaining to educational challenges stemming from adverse home circumstances loaded onto the teaching efficacy dimension, factor two. Also loading onto the teaching efficacy factor were some of the new variables Soodak and Podell introduced, personal student characteristics pertaining to television violence and heredity. Finally, the researchers' third factor, characterized by items relating to
students' attainment of certain outcomes, like "doing better," "getting a better grade," and "mastering a new concept," was termed **outcome efficacy**, a name fraught with problems for other researchers.

Goddard (1998) points out that the fault with Soodak and Podell's interpretation of their second factor starts right with its name, "outcome efficacy." According to Bandura's (1997a) social cognitive theory, self-efficacy and outcome expectancies are two completely different cognitions. Inherently, one scale cannot contain two separate concepts. Self-efficacy perceptions stems from one's belief he or she can complete a task at a certain performance standard, while outcome expectancies develop from one's assessment of the likely outcomes stemming from the successful execution of the same task. For this reason, Soodak and Podell's "outcome efficacy" description of their factor two is a confused label.

**Studies Confirming Teacher Efficacy's Confusion**

The confused state of the teacher efficacy construct became the subject of research published in the mid-1990's. One study, a meta-analysis of teacher efficacy literature, surveyed the teacher efficacy landscape to note the varied treatments of the construct, its antecedents, effects, and other related issues. Another study compared the cross-correlations of several teacher efficacy measures. The results of the quantitative comparisons across instruments revealed moderately surprising results. A review of these two studies follows.

Ross (1994) reviews 88 studies concerning teacher efficacy, each containing empirical measures for efficacy beliefs and identified predictors or effects of teacher efficacy perceptions. From his analysis, Ross observed that instruments based on Gibson
and Dembo's Teacher Efficacy Scale often have flaws in their design. For example, they cross with locus of control ideas, confound efficacy measures with outcome expectancy perceptions, risk response bias due to the prevalence of positively worded items, and lack the specificity called for by Bandura for measuring self-efficacy perceptions (p. 5). Ross, in his vast review, ties together the findings, shortcomings, and future research needs of teacher efficacy study. Questions remain though about the diverse teacher efficacy instruments. Although his article analyzes the findings of teacher efficacy measures across many fronts, one does not get a sense about the degree to which the various scales measure the same thing. Subsequent research aided in this regard.

Coladarci and Fink (1995) study the convergent validity of major teacher efficacy scales. Their parsimonious paper contained credible empirical support for the amorphous nature of teacher efficacy. Simply, they administered many of the teacher efficacy scales already reviewed here (i.e., RAND items, Webb Efficacy Scale, Efficacy Vignettes, Teacher Efficacy Scale, Teacher Locus of Control scale, and the Responsibility for Student Achievement Questionnaire) and two non-efficacy instruments (i.e., Affect for Teaching and Teaching Self-Concept) to 333 K-12 public school teachers in Maine. The resultant correlational analysis revealed relatively weak relationships not expected among commonly conceived measures and relatively strong relationships between theoretically or conceptually different constructs.

The overall moderate correlations indicate the six teacher efficacy constructs were, indeed, measuring something in common. However, a few constructs covary with non-efficacy scales as strongly, or stronger, as they relate to other teacher efficacy instruments. For example, the Teacher Efficacy Scale and the RAND Items shared
highest correlations with each other. The uncommon theoretical grounding of these measures suggests that each might find stronger correlations with other, more similarly conceived instruments. Moreover, TES correlated with non-efficacy instruments, affect for teaching and teaching self-concept, at better than .50 values, larger correlations than it had with other teacher efficacy instruments. Three of its four greatest correlational values were with either non-Bandurian oriented efficacy items or with non-teacher efficacy instruments. Coladarci and Fink's (1995) work raises more questions not only about an instrument that's considered one of the most widely used teacher efficacy scales (Ross, 1994), but about the field of teacher efficacy research.

Reichers and Schneider's (1990) model for scientific construct development calls for three stages of evolution: (a) concept introduction and elaboration, (b) concept evaluation and augmentation, and (c) concept consolidation and accommodation. With the work of Ross (1994) and Coladarci and Fink (1995), teacher efficacy, as a scientific construct, rests in stage two. Hunt (1991) characterizes stage two as "the evaluation and augmentation stage" (p. 131). Furthermore, she states, "[Within this stage], there are critical reviews of the concept and the literature supporting it. Cries of faulty conceptualization, inadequate ways of operationalizing the concept, and equivocal empirical results are heard" (p. 131). With the work of Tschannen-Moran et al. (1998), teacher efficacy underwent more augmentations, though some of the modifications serve to integrate the varying concepts employed thus far.

An Integrated Model of Teacher Efficacy

Responding to the uncertainty surrounding teacher efficacy constructs, Tschannen-Moran et al. (1998) sought to construct a new theoretical model of teacher
efficacy, integrating many of the concepts and theoretical underpinnings previously used. They also added some elements previously not emphasized in such models. In large part, they continued with the two-dimension factor structure, but they reconceptualized its meaning. Their overall model considers the four sources of efficacy beliefs as previously outlined, an analysis of the teaching task, an assessment of personal teaching competence, and an even distribution of positive and negative, internal and external, factors affecting student achievement. Though their model lacks dedicated empirical support, Tschannen-Moran et al. (1998) base many of their suggestions on past research and logic systems.

Adhering to the idea efficacy beliefs are context-contingent, Tschannen-Moran et al. (1998) renamed the dimension commonly known as teaching efficacy to analysis of teaching task. Their acknowledgement of the contextual nature of competency perceptions agrees with Ross, Cousins, and Gadalla's (1996), who found that teachers’ efficacy beliefs rest, in part, on the subject areas for which the beliefs are assessed. A science teacher may manifest different efficacy beliefs for teaching physical science than he may exhibit for teaching life science. The analysis of the teaching task involves two major considerations: (a) weighing the contextual factors that impede teaching and (b) gauging the weight of contextual factors that facilitate teaching. These considerations differ from one situation to the next, as evidenced by the interplay of student characteristics, material resources, time constraints, and professional collaboration opportunities for any given teaching assignment.

Tschannen-Moran et al. (1998) also theorized an even distribution of positive and negative sources of contextual influences and notions of student achievement. For the
reasons given by Woolfolk and Hoy (1990), both qualities of student achievement need to be considered when making self-efficacy judgments. Furthermore, there have not been any items that measured the extent to which positive situational factors affected student outcomes on the general teaching efficacy scale. Gibson and Dembo's (1984) teaching efficacy dimension, and its derivatives, have tapped external attributions for student failure (Guskey and Passaro, 1994). However, student achievement benefits related to a healthy school and community environment are documented (Brookover, Beady, Flood, Schweitzer, and Wisenbaker, 1979; Hoy & Sabo, 1998). The lack of scale items reflecting such possibilities severely limits an instrument's ability to gauge teacher efficacy beliefs accurately.

Another noteworthy aspect of Tschannen-Moran et al.'s (1998) model lies in its temporal verbiage of its personal teaching competence items. Citing inconsistent wordings of past personal teaching efficacy scales, the authors stress competency beliefs should be measured by perceptions of current functioning, like those items comprising the Science Teaching Efficacy Beliefs instrument (Riggs & Enochs, 1990). The comparisons individuals need to make between perceptions of present abilities and perceptions of the needed abilities in order to help determine efficacy beliefs justify their choice of present tense. Assessments of past abilities, aspirations for future proficiencies, and determinations made from hypothetical scenarios will not help form useful efficacy beliefs.

Another contribution Tschannen-Moran et al.'s model made to the teacher efficacy literature centers on the notion of item specificity. Bandura (1997a) maintains
that, while most teacher efficacy scales have made vast improvements over the original RAND items, they still prompt teachers with vague notions about the instructional behavior in question.

Multi-item measures are an improvement over single-item ones, but teacher efficacy scales are, for the most part, still cast in a general form rather than being tailored to domains of instructional functioning. Teachers' sense of instructional efficacy is not necessarily uniform across different subjects. (p. 243)

Tschannen-Moran et al. (1998) raised the theoretical struggle of finding the optimal level of specificity, for items that maximize precision risk minimizing the generalizability of findings and items that maximize a global orientation risk transforming the efficacy instrument into a "personality trait" measure (Pajares, 1996, p.547).

The acceptance of the integrated model into the scholarly community was not obvious for a few years. Few articles outright reviewed its merits. However, a comprehensive assessment of the model's qualities appeared in a paper presented to the Educational Research Exchange in 2001 (Henson, 2001). Despite the new conceptualization of teacher efficacy proposed by Tschannen-Moran et al. (1998), Gibson and Dembo's (1984) Teacher Efficacy Scale continued to be the mainstay of teacher efficacy research, as an instrument reflecting the new model theory did not emerge for another three years (Tschannen-Moran & Woolfolk Hoy, 2001). Before reviewing the new instrument based on the integrated model, an inspection of Henson's (2001) critique will be offered.

**An Appraisal of the Integrated Teacher Efficacy Model**

The balanced review cites the strengths of the integrated model's new conceptualizations and the untested assertions made by the same. On the whole, Henson distinguishes three features of the model and one major conceptual challenge.

Despite the numerous teacher efficacy instruments based on Bandura's (1977) self-efficacy theory (e.g., Ashton et al. 1982; Gibson & Dembo, 1984, Riggs & Enochs, 1990; Soodak & Poddel, 1996), very few, if any, acknowledge the role of self-efficacy sources of information in their frameworks. Henson (2001) credits Tschannen-Moran et al. (1998) with being one of the first models to make self-efficacy sources of information an explicit component of the construct's paradigm. Henson supported the task analysis dimension of the teacher efficacy framework for its greater degree of task specificity and fairer representation of external assets and constraints. Previous GTE factors, to which the task analysis dimension is related, held broader notions of contextual perceptions and overwhelmingly conveyed environmental constraints affecting student achievement. Henson further agreed with the model's portrayal of personal competency belief items, in that their assessments are made in light of current functioning, not past, future, or hypothetical scenarios.

Henson (2001) illustrated the inherent theoretical struggle in attaining optimal teacher task descriptions on teacher efficacy scales. Several theoreticians have voiced this challenge, with regard to the measurement of efficacy beliefs in general (Bandura, 1997a; Maddux, 1995; Pajares, 1997; Pintrich & Schunk, 1996). The dilemma lies in that narrow descriptions of teaching tasks minimize the generalizability of the findings, while broad references to teaching duties muddled the concept being measured. Tschannen-Moran et al.'s (1998) brought this difficult point to the fore with their call for efficacy
beliefs to correspond with specific tasks. Henson (2001) not only acknowledges the integrated model for keeping this point relevant, but offers assurance that efficacy beliefs are generalizable by recalling Bandura's (1997a) conditions for applying efficacy perceptions to behavior outside the measured domain. "The level of generality of the efficacy items within a given domain of functioning varies depending on the degree of situational resemblance and foreseeability of task demands" (p. 13).

The untested basis upon which Tschannen-Moran et al.'s (1998) model evolved invited many questions from Henson (2001). For example, the sources of teacher efficacy information, because of their absence form prior models of teacher efficacy, have no empirical precedent validating their role in teacher competency beliefs, so how do they quantifiably relate to teachers' perceptions of their competence? Moreover, in studies which may have drawn observations about such relationships, their conceptualizations of teacher efficacy lack consistency with the integrated framework of Tschannen-Moran et al. (1998). Thus, research is needed regarding the relationships of mastery experiences, social persuasion, vicarious learning, and physiological states with teacher efficacy beliefs. Henson's questions largely called for empirical inquiries justifying the theoretical relationships asserted in the model. Two of the model's original authors responded accordingly three years later.

An Integrated Teacher Efficacy Model's Validation Study

Tschannen-Moran and Woolfolk Hoy (2001) answered Henson's (2001) call with a teacher efficacy instrument validation study based largely on the assumptions grounding their integrated model. The instrument designers witnessed three phases of the survey's development before determining a final scale of teacher efficacy, namely the
Ohio State Teacher Efficacy Scale (OSTES). Factor strength and parsimony intermittently influenced the addition and deletion of scale items throughout each stage. In the end, though, three factors emerged to describe teacher efficacy: efficacy for instructional strategies, efficacy for classroom management, and efficacy for student engagement. Although other statisticians confirmed much of the factor structure underlying the OSTES (Roberts & Henson, 2001), it manifests differences from the theorized model of Tschannen-Moran et al. (1998).

External validation steps confirm that teacher efficacy beliefs, as measured by the OSTES, compare predictably to other related measures. Some of the concepts correlated with the three subfactors of the OSTES were RAND 1 (GTE proxy), RAND 2 (PTE proxy), TES's GTE, and TES's PTE. Consistently, OSTES and its three dimensions covaried more strongly with RAND 2 and PTE than with RAND 1 and GTE. Not only do the correlations suggest the OSTES and its underlying dimensions focus on personal perceptions of teaching competencies, but they lend further support to the growing body of literature decrying the clouded state of GTE (Coladarci & Fink, 1995; Goddard, 1998; Guskey & Passaro, 1994; Ross, 1994; Tschannen-Moran et al., 1998).

GTE's low correlation with the OSTES and its dimensions has implications for the task analysis factor of Tschannen-Moran et al.'s (1998) model of teacher efficacy. These low correlations with GTE-type instruments make sense in light of the absent OSTES items probing task analysis perceptions. To be sure, task analysis and GTE factors differ in their accommodation of items measuring teachers' perceptions of contextual resources aiding the teaching process. However, one wonders if the addition of such items led to the demise of a GTE-type factor in the OSTES. Discussions with
one of the authors indicate this to be the case (Tschannen-Moran, personal communication, February 2, 2002). Items in which personal competence considerations were paired with contextual reflections had low factor loadings, hence leading to their deletion. "Once again, the lower correlation between GTE and other measures of efficacy suggest this scale is the least successful in capturing the essence of efficacy" (Tschannen-Moran and Woolfolk Hoy, 2001, p. 801). By extension, one wonders if Tschannen-Moran and Woolfolk Hoy's comments concerning GTE's utility bear on the value of a task analysis factor.

Although the OSTES may be vulnerable to response bias, it yields both a total teacher efficacy score and three subscale scores, based on the results of principal-axis and second-order factor analyses. The external validity checks, as noted earlier, indicate equally useful results for the OSTES and its subscales. The merits of the OSTES go beyond the qualities of traditional teacher efficacy instruments. The OSTES gauges more perceptions concerning the varied roles of teachers than the discipline- and resistance-oriented items of the RAND and TES surveys. The face validity of the newest teacher efficacy instrument presents a description that is reflective of sound teaching; for efficacy for instructional strategies, efficacy classroom management, and efficacy for student engagement speaks to the reasonable competencies of an effective teacher.

Collective Efficacy

Social cognitive theory not only helps to explain individual motivation and psychosocial functioning, but it provides a framework for predicting group behavior, too. Indeed, social cognitive theory has sociological implications. The concept of efficacy relates to group-level dynamics, as demonstrated theoretically and empirically (Bandura, 56
Reviewing how the assumptions of social cognitive theory apply to group behavior, such as school performance, facilitates understanding the notion of collective efficacy.

**Assumptions**

The agentive assumption of social cognitive theory applies to schools, just as it applies to individuals. By way of example, the purposeful initiative of schools emerges in their planned instructional programs, designed curriculums, and their fundraising pursuits, to name a few likely endeavors. The very fact that many social institutions, including schools (i.e., local educational agencies), are known as "agencies" supports the assumption that collectives are agentic. To be sure, the purposeful action of schools results from the initiative of their individual members.

Perhaps less obvious, though no less logical, is the link between triadic reciprocal causation and school functioning. School performance (i.e., internal actions facilitating student learning) serves as the behavioral factor of the tripartite. The communication between school members at the organizational level mirrors the adaptive cognitions of the individual at the personal level. The environmental influences acting on school staff behaviors and school staff cognitions include governmental agencies, local businesses, and community members. All three of these factors interact to mutually determine each other's functioning.

The remaining assumptions associated with social cognitive theory emerge in standard school functioning. Vicarious learning, self-reflection, and self-regulation are especially salient behaviors of schools striving to improve their performance. Schools that adopt the effective strategies practiced by other schools demonstrate vicarious
learning behaviors, as they had to discover, by some observable means, the techniques employed by the "effective" school. The desire to consider other approaches constitutes self-reflection and self-regulation, as the presence of such processes seldom leaves one inert and unmotivated to change. The natural fit of social cognitive theory assumptions to both individual and collective functioning suggests the central mechanism facilitating behavior operates for both units of agency. Indeed, for Bandura (1997a) states, "Perceived personal and collective efficacy differ in the unit of agency, but in both forms efficacy beliefs have similar sources, serve similar functions, and operate through similar processes" (p.478).

Collective Efficacy in Schools

Members of social settings share beliefs about their collective utility, explicitly and implicitly, such that an emergent characteristic develops suggestive of the group's ability to perform. Bandura (1997a) states this shared perception of group members concerning “the performance capability of a social system as a whole” (p. 469) constitutes collective efficacy. For schools, the collective efficacy construct gauges teachers' beliefs about the staffs' capability of to influence student achievement (Goddard et al., 2000). Several notions of collective efficacy, based on divergent operationalizations, have been studied, but it has only been fairly recent that linkages with school performance have been found. Noting the ties teacher efficacy has with student achievement, collective efficacy's relationship with school performance warrants investigation.

Immediately, the link between teacher efficacy beliefs and student achievement (Anderson, Greene, & Loewan, 1988; Armor et al., 1976; Ashton & Webb, 1986;
Midgley et al., 1989; Moore & Esselman, 1992; Moore & Esselman, 1994; Ross, 1992) indicates a parallel relationship might exist between schools and school performance, given Bandura's (1997a) linkage between the two levels of agentic beliefs. Moreover, the behaviors and attitudes to which high teacher efficacy beliefs relate (Ross, 1998) imply similar collective behaviors and attitudes would manifest for likewise collective efficacy perceptions. The research exploring such relationships, based on Bandura's (1993) collective efficacy conceptualization, is relatively new, though. One possible reason for the lacking collective efficacy research may be problems associated with schools being the unit of analysis (Pajares, 1996). Nonetheless, studies aggregating efficacy-like perceptions of individual teachers to the school level stand as the precursors to the recent and few collective efficacy inquiries. A review of these early pseudo-collective efficacy studies precedes a survey of two important studies with direct bearing on the research at hand.

Past Studies Approximating Collective Efficacy

Several studies have attempted to use various school-level measures to capture climate-like indices. These studies helped to build interest in studying the "personality" of the school and its effects on various teacher and school outcomes. Four studies are of interest and are reviewed below.

Cohesion and ideologies. Fuller & Is (1986) studied the degree to which internal and external school factors shape the variation of teacher beliefs. Ideologies measured included the import of school wide planning, the trend of teaching of literacy skills, the congruence between personal and school educational philosophies. Testing two causal notions for school belief convergence, the researchers explored both school contextual...
influences, such as leadership, teacher efficacy, and teacher participation and broader socio-economic factors, such as student body diversity, teacher-student ratio, and community social class composition. Fuller and Is utilized the RAND measures of teacher efficacy to gauge teacher competency beliefs. Nine teachers from each of the 145 participating schools completed surveys. Results indicate "internal organizational factors - strong principal leadership, involvement in a participatory reform effort, and a sense of efficacy among teachers - are associated with greater agreement among staff on several different beliefs" (p. 527). The significance of this study lies in that "collective" efficacy beliefs, measured as the mean of teachers' individual competency perceptions, relates to the cohesion of ideologies among staff members. Thus, collective efficacy, as it was measured then, implicates the social norms within a school building.

Teacher efficacy and parent involvement. Hoover-Demsey, Bassler, and Brissie (1987) hypothesized that a school's collective sense of teacher efficacy relates to differing forms of parental involvement, believing that efficacy beliefs among teachers affects the climate of the school. The mean teacher efficacy score gauged by a TES-like instrument constituted the collective efficacy measure employed in this study. School-level parent involvement indices came from the average perceptions of teachers responding to the researchers' parental involvement surveys (average teacher survey return rate was 69% across 66 schools.) Their findings show that a school's average teacher efficacy level positively related to parent attendance rates at parent-teacher conferences, incidence of parent volunteers in the school, frequency of parent tutoring at home, and quality of teacher support by parents. The parallel between the Hoover-Dempsey et al.'s (1987) study and the study at hand lies in their conception that the teacher efficacy beliefs
implicate the school climate. As such, collective efficacy beliefs figured by the mean efficacy perception are justified, Bandura's (1997a) organizational coupling heuristic notwithstanding.

*Efficacy and alienation.* Newmann, Rutter, and Smith (1989) based their study of teacher efficacy, community, and expectations on the notion that the three aforementioned variables greatly contribute to the school's climate, to the extent such they reduce feelings of alienation among teachers. The authors studied the relationship between school variables, like orderly environment, administrative responsiveness, and innovation, and the three climate variables previously mentioned. Up to 30 teachers in 353 public high schools participated in the study. An interesting methodological strategy was employed in this study of high school teacher alienation. In addition to operationalizing collective efficacy as the mean of teachers' individual perceptions of their own competencies, these researchers indexed the variance of teachers' efficacy scores as a group efficacy consensus. When the group efficacy consensus variable was entered as the last variable in a three-step regression analysis, the explained variance of the efficacy perceptions increased by 10%. The authors hypothesized that when teachers vary greatly in their efficacy perceptions, the divisiveness that may be associated with such divergent beliefs causes dissension and a loss of community. Since community and efficacy beliefs moderately correlated (r = .46), the diminished community feeling may erode the efficacy beliefs of teachers. Conversely, when efficacy beliefs converge, feelings of community may be enhanced resulting in enhanced efficacy beliefs. While the innovative use of variance in collective efficacy research deserves merit, the inclusion of two satisfaction variables on the efficacy scale needs noting, as well.
Efficacy and demography. Esselman and Moore's (1992) research studied the relationships between general teaching efficacy and personal teaching efficacy to various demographic and organizational measures, such as suspensions, failure rates, attendance patterns, mobility and dropout rates. Participants in the study included 1,802 teachers from elementary, middle, and high schools, located in an urban midwestern school district. The number of participating schools is not revealed in the study. The efficacy scores, based on scales advocated by Ashton (1985), constituted factor measures aggregated to the school level. Correlational analysis revealed statistically significant negative relationships between collective teaching efficacy and rates of suspension, and between collective personal teaching efficacy and suspension rates as well as drop out rates. The significance of this study reveals task analysis factors' influences on the collective efficacy perceptions of the staff.

More Relevant Studies of Collective Efficacy

The collective efficacy studies noted above contribute to the growing interest in studying climate-oriented efficacy notions effects on important school variables, like parental support, convergence of ideology, as well as other school climate factors. However, none made a link between collective efficacy perceptions and school-level performance, or the aggregated performance of its students. Bandura (1993) pioneered this research using his model of self-efficacy as his conceptual underpinnings. A few other studies followed suit, as will be explicated below.

Pioneering research in collective efficacy. Bandura (1993) pursued the idea of linking school-level efficacy beliefs with school-level achievement. The basis of his research stemmed from the notion that teachers work more in a collective sense than in
an isolated manner. Their relations generate belief systems that invigorate or discourage the collective performance of their social system (Brookover et al., 1979; Good & Brophy, 1986; Purkey & Smith, 1983; Rutter, Maughan, Mortimore, Ouston, & Smith, 1979). Hypothetically, then, the extent to which teachers' beliefs in the collective ability of their staffs is positive, so, too, is the aggregated performance of schools' students. Although the specifications of his sample and explication of his efficacy scale items aren't revealed, the resulting path model depicting the relationships among his theorized variables is explicit.

Three important relationships emerged in his study. Collective efficacy, measured by aggregated teacher perceptions of the staff's collective ability to teach students, more strongly correlates with aggregated performance of the school's students than does the indexed socioeconomic background of the student body. Moreover, the background socioeconomic status of the student body influences collective efficacy beliefs more strongly than it influences students' collective achievement. Further still, the collective efficacy perceptions of the staff predict student body performance as strongly as prior academic achievement. Bandura (1997a) observes the interplay between collective efficacy and student socioeconomic variables this way:

Although student body characteristics have some direct effect on school achievement, they influence school achievement largely by altering the staff members' beliefs about their collective efficacy to motivate and educate their students. The stronger the staff members shared beliefs in their instructional efficacy, the better their schools performed academically. When variations in staff members' beliefs in their collective efficacy are controlled, the relationship between students' characteristics and school attainment is considerably reduced. (p. 481)
Subsequent studies may deliberate about how to operationalize collective efficacy measures due to Bandura's application of organizational coupling to this strand of efficacy research. Previous studies of collective efficacy gauged school-level efficacy beliefs by aggregating teachers' perceptions of their own competencies to the school level. Bandura (1993) operationalized collective efficacy perceptions as the mean score of teachers' "beliefs in their schools' capability as a whole" (p. 141). Judgments about how to operationalize collective efficacy should derive from the extent to which organizational members rely on one another for the successful functioning of the social system. Tightly coupled organizations reflect high interdependence; suggesting collective efficacy beliefs need to account for the cooperative efforts of the staff. Loosely coupled organizations function mainly through the successful endeavors of its individual members; indicating collective efficacy beliefs may emerge from the aggregated measures of members' individual efficacy beliefs. Believing that school staffs interact at a moderate level to successfully execute the schools missions, Bandura chooses to aggregate teachers' perceptions of the staff's competencies. Debate, though, lingers about the appropriate operationalization of collective efficacy (Goddard, 2002, personal communication, January 17, 2002).

Few studies challenge Coleman et al.'s (1966) assertions about the systemic failure of schools to overcome the biases and inequities attributed to SES as directly as Bandura (1993). As he demonstrated, collective efficacy perceptions could be one of the most promising and alterable school variables for improving school achievement, the difficult process of changing teacher beliefs notwithstanding (Vosniadou, 1994). More studies verified the strength of collective efficacy beliefs and their reviews follow.
Supporting research of collective efficacy. One such study to largely replicate Bandura's (1993) research born out of Goddard's (1998) dissertation at the Ohio State University. His dissertation significantly contributed to Goddard, Hoy, and Woolfolk Hoy’s (2000) study. A few differences separated Goddard et al.’s (2000) work from Bandura's (1993). Goddard et al. employed different statistical means for data analysis and utilized a collective efficacy instrument derived from Tschannen-Moran et al.'s (1998) teacher efficacy model. Because of the sophisticated statistical technology utilized by Goddard et al.’s Hierarchical Linear Modeling, student achievement and SES variables remained at the student level, despite this being largely a study of school-level characteristics. By following this strategy, the researchers avoided aggregation bias, misestimated standard errors, and regression heterogeneity associated with student-level variables measured at the school level.

The collective efficacy instrument evolved from the task analysis and personal teaching competence factors discussed earlier (Tshannen-Moran et al., 1998). Using the Gibson and Dembo (1984) TES instrument as a starting point, Goddard et al. modified the orientation of the questions so that it reflected perceptions of the staff's ability to positive influence student achievement. Furthermore, in the spirit of the Tschannen-Moran et al. (1998) framework, the authors balanced the group competence and task analysis items to reflect both positive and negative aspects of student learning and external realities, before submitting the scale to a pilot study involving 70 teachers from 70 different elementary schools. Factor analyses initially revealed the existence of the two scales theorized to form collective efficacy beliefs: group competence beliefs and task analysis perceptions. Second-order factor analysis showed that the subscales
correlated at better than .70. Moreover, four items loaded onto both factors at values
greater than .40. Elementary teachers apparently had trouble distinguishing between
competency beliefs and task analysis perceptions, a phenomenon actually predicted in the
integrated model (Tschannen-Moran et al., 1998). Teachers make efficacy judgments by
conducting competency determinations in light of performing analyses of the task
demands. At the collective level, the line between the two separate processes may
become blurred, suggesting that the collective efficacy scale measures one underlying
factor. That factor, collective efficacy, is defined as "perceptions of teachers in a school
that the faculty as a whole can organize and execute the courses of action required to
have a positive effect on students" (Goddard & Goddard, 2001, p. 809).

The results of the Goddard et al.’s (2000) study corroborated some of the
important findings of Bandura (1993). Collective efficacy beliefs predicted student
achievement in both mathematics and reading more strongly than any of the demographic
variables, including SES. The respectable amount of between school achievement score
variance explained by collective efficacy perceptions attests to the predictive strength of
this construct. Collective efficacy explained 53.27% of math score variance and 69.64%
of science score variance. In summarizing the results of their data, the authors may have
noted the import of the collective efficacy's relationship with school achievement. "[The]
negative association between SES and achievement is more than offset by the positive
association between collective efficacy and student achievement" (Goddard et al., 2000,
p. 500).

*Mastery experiences and collective efficacy.* Using the same data set from his
previous study (438 teachers in 47 schools), Goddard endeavored to test Bandura’s (1977,
1997) contention that mastery experiences serve as strong determinants of efficacy beliefs. Mastery experiences were operationalized as mean student-level achievement scores from a prior statewide standardized reading assessment. Together with SES and race demographic data, mastery experiences were utilized as predictor variables for school-level collective efficacy measures. When the prior reading achievement, SES, and minority population rates were modeled to predict between-school variation in collective efficacy perceptions, unambiguous results occurred. Past reading achievement, the variable for mastery experiences, explained more between-school variance in collective efficacy perceptions (about 65%) than either of the other two demographic variables. Moreover, when all three independent variables are considered together, SES and minority representation become insignificant in their explanation of the dependent measure, collective efficacy.

Goddard (2000) followed up his earlier study (Goddard et al., 2000) by including prior achievement in their HLM model for predicting between-school variation of collective efficacy beliefs. The prior model utilized collective efficacy, SES, minority representation, and gender variables as the independent measures. After including prior achievement, the model revealed that collective efficacy still relates positively and significantly to school performance. Further yet, the amount of predictive power retained by collective efficacy, after including prior achievement, was 50%, the explained variance of school score variation. This is only a ten percent drop from when prior achievement was not included in the model. In this model, too, collective efficacy was more powerful than SES for explaining school performance.
A third research question endeavored to explore the meaning of collective, as some researchers have operationalized collective efficacy by utilizing means (Bandura, 1993; Esselman & Moore, 1992; Goddard et al., 2000) while others tested collective efficacy by employing variances (Newmann et al., 1989). Goddard measured consensus by measuring the extent to which schools' faculties' beliefs converged. The results showed that the collective efficacy consensus variable failed to reach statistical significance in correlational analyses of the variables. Noting collective efficacy's strong predictive power when it is measured by the mean of teachers' perceptions, its two operationalizations strongly suggest two different concepts are being studied.

**Teacher and collective efficacy.** Goddard & Goddard (2001) research the relationship between teacher efficacy and collective efficacy. Noting the theoretical framing of both constructs, namely social cognitive theory, the researchers reasoned that teachers' individual perceptions of competence are affected by the social milieu in which they work, for the environment is one component of the tripartite. At the same time, the social interactions of the teachers help determine the climate of the school, and by implication, the quality of collective efficacy beliefs.

Using the same data set as was used in Goddard et al. (2000), Goddard and Goddard (2001) employed similar variables in the study. Minority representation, SES, prior achievement, and collective efficacy constituted much of their variable set. School size, as indicated by number of students, served as another predictor measure in their tested models. Hierarchical Linear Modeling served as the statistical method for analyzing the data. The results of the research indicate that mean teacher efficacy systematically fluctuates across schools. The pattern of fluctuation appears to be
predicted more by collective efficacy measures than by SES, minority representation, or school size. Indeed, when the explanatory model includes all the predictor variables, collective efficacy remains the only statistically significant measure. Moreover, when comparing different predictor models, the equations containing SES and prior achievement each explained roughly 25% of the variance in between-school teacher efficacy variance and each left a statistically non-zero balance unexplained. The latter indication of these models implies there are other variables not considered by the model that helps predict school-level teacher efficacy. By comparison, the model containing collective efficacy explained almost 75% of the variance and left an unexplained amount not statistically different than zero. Statistically speaking, no other variables, beyond collective efficacy, need be considered for predicting school-level teacher efficacy.

Goddard and Goddard (2001) conclude their study by suggesting some implications of their study for further research. One might find it important to find school-level variables that enhance collective efficacy. While Goddard (2000) demonstrated mastery experiences strongly influence collective efficacy beliefs, prior school achievement presents logistical problems as an alterable variable, as it stands as the ultimate reason for enhancing teachers' collective efficacy. The authors note Bandura's (1997a) call for strong school leadership as a possible variable for bettering teachers' collective competency notions. "Strong leadership may indeed be aspects of organizational life that can build collective efficacy" (Goddard & Goddard, 2001, p. 817). It is appropriate that the most recent and available study of collective efficacy calls for research linking leadership qualities and collective competencies of the staff, for this is the basis of the present study.
Implications for Present Study

The effects of teacher efficacy beliefs underscore their importance in schools. Teacher attitudes that facilitate student learning (Meijer and Foster, 1988; Soodak & Podell, 1993; Woolfolk and Hoy, 1990) and instructionally effective teacher behaviors (Allinder, 1994; Ashton & Webb, 1986; Gibson and Dembo, 1984) broadly address the type of teacher outcomes associated with teacher efficacy beliefs. Student outcomes related to teacher competency beliefs include increased student interest in school, more favorable student evaluations of their teachers, and increased student efficacy beliefs (Midgley et al., 1989; Woolfolk, Rosoff, & Hoy, 1990). Perhaps the most important teacher efficacy effect is its relationship with student achievement (Anderson et al., 1988; Armor et al., 1976; Ashton & Webb, 1986; Midgley et al., 1989; Moore & Esselman, 1992; Moore & Esselman, 1994; Ross, 1992). Little doubt remains about the importance of teacher efficacy beliefs for teaching and learning.

Several aspects emerged from the review of teacher efficacy literature that suggest certain school characteristics may influence the competency beliefs of teachers. Notable among them is leadership. Although several conceptualizations of teacher efficacy have transpired over the years, strong leadership has consistently related to teacher efficacy beliefs, regardless, for the most part, of how teacher efficacy was measured. A brief, but by no means exhaustive, summary of literature relating leadership behaviors to teacher efficacy perceptions follows.

Lee, Dedrick, and Smith (1991) found that teachers who perceive their principals to be strong (e.g., "Knows what kind of school he wants," "Gets resources for the school," "The principal sets priorities, makes plans," etc.) tend to have higher efficacy
beliefs. Coladarci and Breton (1991) observe higher efficacy scores for teachers who perceive their principal supervision to be useful. A positive school climate focused on instruction enhances teachers' sense of efficacy (Moore & Esselman, 1992). Newmann et al. (1989) and Hoy and Woolfolk (1993) note that teachers' efficacy beliefs benefit from principals' who respond to their needs, help them solve instructional and classroom issues, and who maintain an orderly school environment. Hipp and Bredeson (1995) found that principals who model professional behavior and inspire collective purpose facilitate teachers' senses of efficacy.

Bandura (1997a) provides the conceptual framework for not only linking school leadership to teacher efficacy, by way of the triadic reciprocal causation model, but he theorizes the link between self-efficacy and collective efficacy beliefs, too. The only difference between self-efficacy beliefs and collective efficacy beliefs is the unit of agency, as noted above. They both share analogous sources, processes, and functions. The research noting statistical relationships between principal behaviors and teacher efficacy beliefs leads one to theorize that principal behaviors should relate to collective efficacy, as well. Before exploring this relationship, leadership theory and research warrants a review.

Leadership Theory

This section of the chapter provides the conceptual and historical development of transformational leadership. Generally, the evolution of leadership theory will be addressed before discussing transformational leadership. A discussion of the historical and cultural pervasiveness of leadership precedes an illustration of the early theories of leadership. Subsequently, the newer leadership models are explained, during which...
transformational leadership is discussed in detail, including its relevant instruments, factors, validation, and effects. The application of transformational leadership to schools follows, after which its limitations and criticisms are discussed. Finally, the theories linking transformational leadership to collective efficacy, thereby serving as rationale for this study's research hypotheses, conclude the chapter.

The Origins and Pervasiveness of Leadership

Societies have long been interested in the study of leadership. Stories of leadership, dating back to days of the Old Testament and earlier, provide explanations for "the dominance of ... leaders and the submission of ... subordinates" (Bass, 1990, p. 3). These legends not only cross time, but they span geography, too. Evidence of leadership myths exist from ancient China, Greek Mythology, and Egyptian hieroglyphics, among other early writings (Hunt, 1999). The superior-subordinate relationship manifests in animals as well as humans. Bass (1990) cites studies of primates, birds, and other animals that describe animals yielding to their own for feeding, mating, and territory advancement. Thus, leadership phenomena pervade temporal, cultural, and animal boundaries.

Despite its pervasiveness, a convincing definition of leadership is elusive. Bass (1990) contends there are as many definitions as there have been theorists' attempts to write one. Although many people think the idea of leadership is important (Bolman & Deal, 1991), leadership definitions lack consensus and precision. Bennis and Nanus (1985) liken the concept of leadership to the notion of love in that everyone knows it exists but nobody can define it, despite more than 350 attempts. Yukl (1998, p. 2) provides eight representative definitions of leadership from the past fifty years that
display minimal likeness to one another. The little shared among the various definitions expresses "a process whereby intention is exerted by one person over other people to guide, structure, and facilitate activities and relationships in a group or organization" (p. 3). Noting the prevalence and universal quality of leadership, its influence appears to be a necessary function in society. Bass (1990) stresses there are no known cultures that exist without leadership in some form.

Scientific inquiry of leadership started roughly a hundred years ago, though some (Bass, 1990; Hunt, 1999) credit Thomas Carlyle's (1847) essays, collectively titled "Heroes and Hero Worship," as one of the initial attempts to study leaders for commonalities. His thesis, initiating the "Great Man" theory, suggested "heroes" are born with unique qualities. Under such a framework, a nation's fortune rested on the chance that birth bequeathed countrymen the needed abilities to overcome any societal hardships. As with several subsequent theories, this model inadequately explained the successes of several important people—for the "Great Man" theory didn't account for women!

Several other leadership theories emerged from the literature during the course of the next century. Trait, behavior, and situational frameworks received much scholarly attention for much of the twentieth century. Transformational and related theories came into being during the last twenty years. Although the "Great Man" paradigm fell out of favor, some of its contentions remains in subsequent leadership models. The traits distinguishing "Great Men" must be identifiable and, as such, form the next major theory explaining leadership.
Leadership Trait Theory

The idea that effective leaders had various traits in common held appeal for leadership researchers during the early to mid 1990's. Much scholarship was invested into this line of inquiry. A description of early leadership trait research and its findings precede the review of a second wave of such scholarly pursuits.

Early leadership trait theory. Studies of traits among leaders, as opposed to non-leaders, proliferated during the first half of the 1900's. Researchers wished to find a set of traits that consistently portrayed successful leaders. Just about every trait imaginable was considered. For example, age, height, weight, talkativeness, appearance, fluency, athletic prowess, tone of voice, intelligence, and self-sufficiency, among many others, were the focus of such studies. The results of these studies met little fanfare as no pattern of traits distinctly described leaders. Whereas one trait may characterize a leader in one situation, it just as likely did not describe a leader in another context. Many of these studies investigated one trait at a time without much theory to guide the researchers' efforts. Moreover, the conception of leader came into question, as the static notion of leadership gave way to a more dynamic theme (Bass, 1990).

Nonetheless, Stogdill (1948) categorized the wide-ranging characteristics found in the leadership trait research between 1904 and 1947. Five categories initially emerged, including capacity, achievement, responsibility, participation, and status. Soon afterward, Stogdill (1948) added a sixth classification - situation. This last quality was added as a result of Stogdill's observation that a leader's characteristics must align, at least in part, with the goals and composition of his followers. Stogdill's (1974) comment on this necessity provides further clarification:
Strong evidence indicates that different leadership skills and traits are required in different situations. The behaviors and traits enabling a mobster to gain and maintain control over a criminal gang are not the same as those enabling a religious leader to gain and maintain a large following. (p. 72)

The disparity and divergence of the early trait studies surface in the findings of Bird (1940). His analysis of 20 trait studies, covering 79 qualities, revealed that 65% of the traits were never mentioned in 19 of the 20 studies. Extroversion, humor, intelligence, and initiative were the only attributes researched in more than one study. The lack of convergence noted by Bird (1940) and Stogdill (1948) is expressed well by Yukl (1998), as he remarked that none of the traits found during this time of study were sufficient or necessary for leadership success. The apparent futility in trait research did not, however, stop industrial psychologists from endeavoring to profile the prototypical management supervisor (1998).

Later leadership trait theory. The "selection research" era referred to the time following Stogdill's (1948) meta-analysis and to the recruiting process industrial psychologists designed for pursuing exemplary managers. Changes occurred in research orientation and design for investigating traits of significance. Emphasis shifted from focusing on leadership status to managerial effectiveness. New methodologies included factor analyses, multivariate regression, and theory consideration as newer technology allowed more variables to be studied at one time. The influence and importance of context swayed researchers to study more specific administrative duties, technical skills, and managerial motivations pertaining to the requirements of a particular job. The confluence of these changes netted results strikingly similar to the first forty years of study in leadership traits.
Stogdill (1974), again, performed a meta-analysis of the trait literature, this time spanning the years 1949 to 1970. Several of the same traits that were reported during Stogdill's (1948) earlier literature review appeared again. The profile of leadership emerging after this analysis of 163 trait studies is as follows:

The leader is characterized by a strong drive for responsibility and task completion, vigor and persistence in pursuit of goals, venturesomeness and originality in problem solving, drive to exercise initiative in social situations, self-confidence and sense of personal identity, willingness to accept consequences of decision and action, readiness to absorb interpersonal stress, willingness to tolerate frustration and delay, ability to influence other persons' behavior, and capacity to structure social interaction systems to the purpose at hand. (Stogdill, 1974, p. 81)

Bass (1990) notes that the similarity in results between Stogdill's literature reviews affirms that personality traits distinguish leaders from non-leaders, leaders from followers, and higher-grade leaders from lower-level leaders. This is not to say, a uniform trait approach to leadership study is preferred. However, as Bass notes, the inclusion of context into the second round of trait studies affirms the notion that trait research needs to recognize the importance of situational factors. A middle ground, of sorts, was realized between trait and situational schools of thought. Yukl (1998) remarks that the results of the second literature review by Stogdill showed that some scholars probably abandoned trait study too hastily after the 1940’s assessments of trait research.

Perhaps, the most salient application of trait research came in the growth of assessment centers, facilities in which candidates for leadership positions were observed to glean a sense of their traits under various contrived circumstances. Over 1,000 centers surfaced by the early 1970’s. Bass (1990) reminds us, though, that despite the convergence of almost 70 years of studies on some general trait patterns, “there is no
overall comprehensive theory of the personality of leaders” (p. 87). As such, the leadership theory field eventually focused research efforts in new directions. If a description of a leader proves elusive, then a description of leadership behavior and work may be more promising. In the 1950’s, many researchers began to focus on what managers actually do while working, and with this line of inquiry, the behavior approach to leadership theory developed.

*Leadership Behavior Theory*

Perhaps the major contribution of leader behavior research stems from its study of managerial behavior and the resultant generalizations made about leaders’ dominant activities. This stream of leadership behavior research explores the actions of managers via instruments used to measure subordinate perceptions of their superiors’ endeavors, namely questionnaires. Factor and correlational analysis stand as the primary analytic techniques for generalizing the findings of subordinate observations. Although attention eventually shifted away from leader behavior inquiry, important theories and instruments emerged during this period.

*Leadership behavior inquiry via the LBDQ.* Much of the early research into leadership behavior began at the Ohio State University in the early 1950’s. Hemphill (1949) brought his early investigations of leadership behavior from the University of Maryland to OSU when he joined its staff. He and his colleagues created a list of 1,800 statements describing various dimensions of leader behavior. The list was consolidated to 150 statements, based on their exclusiveness to only one subscale. These descriptors
developed into the first version of the Leader Behavior Description Questionnaire (LBDQ) (Hemphill, 1950; Hemphill & Coons, 1957), the core instrument dominating the early research of leadership behavior (Yukl, 1998).

After preliminary administrations of the LBDQ to military and civilian subordinates, staff responses reflecting perceptions of supervisor behavior revealed two broad category of behaviors: initiating structure and consideration (Fleischman, 1957; Halpin & Winer, 1957). These two factors, as measured by the LBDQ, describe much of the behavior of leaders and are related to the task-oriented and people-oriented notions of leadership functioning. For much of the past fifty years, leadership behavior has been conceived as comprising these two dimensions, for they appear in or highly correlate with a wide variety of leadership constructs (e.g., Bass, 1985; Blake & Mouton, 1964; Fiedler, 1967; Likert, 1961). A description of each dimension follows.

Initiating structure describes the leader’s actions that define his or her functions and the tasks of the subordinates for the realization of organizational goals. Yukl (1998) suggests the following behaviors exemplify initiating structure: criticizing inadequate performance, stressing the value of meeting deadlines, coordinating staff work responsibilities, communicating expected standards of performance, and maximizing the work capacity of subordinates. “He lets group members know what is expected of them,” serves as an example item from the LBDQ measuring the production–orientation of the leader (Bass, 1990, p. 512).

Consideration implies the leader’s personal concern for the follower and can manifest in numerous ways. A highly considerate leader may be one who listens
attentively to subordinate matters, supports employees’ positions on various issues, performs personal favors, consults staff regarding impending decisions, and takes time to learn about subordinate interests, in addition to other acts of consideration. “He is friendly and approachable,” demonstrates an LBDQ item measuring personal concern by the leader (Bass, 1990, p. 512).

Yukl (1998) cites Fleishman and Harris (1962) as one of the best examples of correlational research on the two dimensions of the LBDQ. The researchers investigated leadership behavior at a truck manufacturing facility. Although no production quality or output measures were mentioned, the number of grievances and voluntary employee turnover served as indicators of leadership effectiveness. Perhaps intuitively congruent, the results of the study showed that leaders with higher ratings of consideration had fewer grievances and incidences of turnover. Correspondingly, supervisors with lower ratings in initiating structure had higher rates of grievances and employee departures. The value of the study may lie in the researchers’ observation that there appear to be an upper bound for consideration and lower limit for initiating structure for which employee grievance and turnover yields. Skinner’s (1969) work supports the curvilinear findings of Fleishman and Harris (1962). Yukl (1998) credits Fleishman and Harris as being some of the only leadership theorists to study initiating structure and consideration for interactions or curvilinear effects.

The LBDQ and modified versions of the same have been used in hundreds of inquiries relating initiating structure and consideration to various dependent variables
(Yukl, 1998). However, not satisfied with the limited scope of LBDQ’s two behavioral constructs, several researchers developed instruments to measure a broader range of leadership actions. Some of these instruments follow.

Derivatives of the LBDQ. Seltzer and Bass (1987) supplemented the LBDQ dimensions with three transformational leadership behaviors, charisma, intellectual stimulation, and individualized consideration, to capture more of the behaviors likely to account for the second-order changes taking place in organization. Bass (1985), a leadership scholar and industrial psychologist who conceived of transformational leadership, felt higher motivations in subordinates needed to be tapped by leaders in order for organizations to successfully navigate the rapidly changing economic, political, and technological world in which businesses operated. Initiating structure and consideration conceptually lacked the logic systems to account for the numerous examples of organizations transforming their work routines to accommodate vastly changing environmental demands. Empirically, these dimensions failed to explain much of the leadership behavior in many companies. As will be seen later, the genesis of the Bass’s (1985) Multifactor Leadership Questionnaire is found with the LBDQ.

Halpin and Croft (1962) tailored and broadened the LBDQ with items that described further the behavior and actions of a school principal. Factor analysis of their innovative survey showed four factors of principal behavior emerge: aloofness, production emphasis, thrust, and consideration. Eventually, Halpin and Croft supplemented their principal behavior scales with teacher factors to form a climate inventory for schools. These teacher factors consisted of disengagement, hindrance, esprit, and intimacy. The confluence of teacher and principal behavior scales formed six
possible elementary school climates: open, autonomous, controlled, familiar, potential, and closed. Thus, the Organizational Climate Description Questionnaire (OCDQ) was conceived and has been revised by Hoy and his colleagues (Hoy, Tarter, & Kottkamp, 1991; Hoy, Hoffman, Sabo, & Bliss, 1996; Hoy & Sabo, 1998). Bass (1990) notes that in comparison to the original initiating structure and consideration factors, the OCDQ yields “more insight into the dynamic interplay among the climate of schools, the behavior of leaders, and the response of teachers than could be produced by the use of two factors …” (p. 516).

Oldham (1976) augmented the LBDQ to enhance its ability to gauge organizational leadership. His additional scales included items that measure subordinate perceptions of their leaders’ propensity to personally reward, personally punishments, goal set, design feedback systems, place personnel, and design job systems. The results of Oldham’s work demonstrate scales with higher correlations to leadership effectiveness than the factors of the LBDQ. In effect, Oldham’s (1976) research showed that enhanced transactional practices, (i.e., leaders giving praise and recognition along with material rewards) improved employee performance and leadership influence (Bass, 1990).

In short, while the LBDQ may have been limiting in some ways, it, nonetheless, sparked several lines of research that still stand today. One contribution of the LBDQ to modern day understanding of superior-subordinate relationships is via transformational leadership, a focus of this study. Another leadership behavior model utilized the behavior stream that helped bridge the research into another branch of leadership inquiry,
the contingency approach of leadership theory. A review Blake and Mouton’s (1964) managerial grid precedes an explanation of the contingency model of leadership.

**The Managerial Grid.** Two factors corresponding closely to initiating structure and consideration form the basis of Blake and Mouton’s (1964) managerial grid. This leadership construct conceives managers as manifesting leadership styles that vary in the importance they place on people (consideration-like) concerns and the emphasis they place on production matters (initiating structure-like). Aligning each of these dimensions along chart axes, Blake and Mouton theorize five separate leadership styles from the various intersections of the dimension strengths (measured from 1 to 9). Leaders’ scores for the people-oriented and task-oriented factors come from their endorsements of various factor-relevant statements. The five leadership styles resulting from crossing of the factor scores are as follows: (a) Authority-obedience (high production-low people), (b) “Country Club” (low production-high people), (c) Impoverished management (low production-low people), (d) “Organization Man” (middle production – middle people), and (e) Team management (high production-high people).

The authors contend that a high-high leadership orientation (team management) has constantly shown to facilitate performance criteria throughout various studies (Blake & Mouton, 1978). One study, in particular, highlights this point. Two comparable self-contained units of a larger company participated in a long-range study spanning ten years. The results of this study rested on a pre-post comparison of the two smaller organizations after one experienced experimental treatment. The treatment consisted of company-wide development consistent with the high-high leadership managerial philosophy. The non-experimental unit received no comparable training. The results of the study showed that
the treated organization realized a 400 percent gain in profits over the control unit. As compelling as high-highs might seem as a universal leadership behavior, Blake and Mouton (1982) realized the different situations may call for different leadership behaviors.

Yukl (1998) explains that while Blake and Mouton advocate an optimum leadership behavior (high-high), they “clearly recognize the need for leaders to select specific forms of behavior that are appropriate for a particular time or situation” (p. 56). Effective leaders should maximize their people and task concerns, but, in doing so, their behaviors reflecting such considerations will change from underling to underling and from situation to situation. Consequently, Blake and Mouton’s (1982) managerial grid leadership model is both a universal theory and a situational one. The goal of maintaining high-high behavior dictates a leader framework for selecting from a range of behaviors. The situational aspect of the theory manifests in its suggestion to choose behaviors that match the characteristics of the situations. Blake and Mouton’s (1982) work never specifies which situations enhance or impede which behaviors. As such, their model falls short of being classified a full-fledged contingency model of leadership.

Contingency Theory

Yukl (1998) characterizes a contingency model of leadership as one that specifies certain situational variables as enabling or hindering the effectiveness of certain leader behaviors. Bass (1990) suggests contingency models answer the question of whether leaders should behave consistently or vary their behaviors to fit the demands of the situation with the latter alternative. Two models worth mentioning couple traits or
behaviors with situations to maximize the leader’s effectiveness: Fielder’s (1967) least preferred coworker theory and House & Mitchell’s (1974) path-goal framework. The theories will be addressed in the order of their chronology.

_Least Preferred Coworker_. Bass (1990) contends that Fiedler’s (1967) least preferred coworker theory stands as the most studied leadership contingency model. It also represents the first major theory to assert that precise situational factors influence the effectiveness of certain types of leadership behaviors or traits (Hoy & Miskel, 2001). In Fiedler’s (1967) work, his contingency model matched a leadership trait, derived from a least preferred coworker score, with certain situational factors. The least preferred coworker score comes from the leader’s rating of his or her most incompatible worker along 16 bipolar adjective scales (e.g., pleasant – unpleasant, friendly – unfriendly, efficient – inefficient). A high LPC score represented a consideration-oriented leader and a low LPC score represented a task-oriented leader. Fielder’s (1978) piece explained that these qualitatively different leaders led best in different contexts, differentiated by the strength of certain situational factors.

Fielder (1978) maintained that effective task and people-oriented leaderships were contingent on certain situational factors. He termed these contextual characteristics as being task structure, positional power, and leader-member relations -- elements of leadership control of subordinates (Yukl, 1998). Task structure describes the degree to which the targeted work has defined operating procedures, end product specifications, and quality standards. Positional power refers to the extent the leader has the legitimate authority to evaluate, reward, and sanction subordinate work. Leader-member relations gauge the extent leaders have close, supportive, and loyal rapports with their
subordinates. When all three of these environmental elements indicate very favorable or very unfavorable subordinate control, then task-oriented leaders tend to be more effective. When the three situational factors measure inconsistently, people-oriented leaders tend to be more effective (Fiedler, 1967).

Although this model has received much study and support (Peters, Hartke, & Pohlman, 1985; Strube & Garcia, 1981) and has been credited with being the first systemic empirical work on leadership theory (Hunt, 1999), it does have its weaknesses. One shortcoming of this theory is its changing assertions about what the LPC trait means. Yukl (1998) states that the latest meaning infers a leader’s hierarchical work orientation. A low-LPC score indicates a leader is motivated to attain achievement objectives before realizing close, interpersonal relationships. If performance goals are achieved, the close relationships can be pursued. The high-LPC leader has the opposite order of work motivation. Rice’s (1978) review of LPC research over 25 years challenges Fielder’s (1978) latest interpretation of the LPC score. He found that the LPC trait acted more like an attitude or a value than a hierarchy of motivation. Schriesheim and Kerr (1977) found no evidence for a motivation hierarchy conceptualization. Offerman (1984) found evidence that the LPC operates more like a non-committal attitude. Thus, much debate lingers over the conceptualization of the LPC.

Nonetheless, Fiedler (1967) broadened the field of leadership inquiry by considering situational variables as moderators of leadership effectiveness. As such, Fiedler (1978) argued that by manipulating the situational factors, a leader stood a better chance of success than by manipulating a leader’s qualities. This notion was a novel
approach to organizational thinking. The next model under consideration also expands leadership theory by studying contingency ideas, but it focuses on leadership behavior. House and Mitchell’s (1974) path-goal leadership model is reviewed next.

Path-Goal theory. The path-goal leadership model theorizes that leader behavior provides, clarifies, and facilitates subordinate means to employee satisfaction and achievement. House (1971) explains the theory this way:

The motivational function of the leader consists of increasing personal payoffs to subordinates for work-goal attainment and making the path to those payoffs easier to travel by clarifying it, reducing roadblocks and pitfalls, and increasing the opportunities for personal satisfaction en route. (p. 324)

Yukl (1998) points out that leader behavior influences employee satisfaction and achievement differentially. Situational factors help determine the effects leadership activities have on subordinate satisfaction and performance. Bass (1990) succinctly captures the essence of this model by noting that the leader needs to furnish only those elements of the situation missing for employees to realize both contentment and the targeted organizational objectives. The employees’ abilities and motivations, the organizational tasks, and the work environment dictate those elements absent and present for achieving the dual aims of leadership effectiveness: organizational productivity and employee satisfaction. Thus, those dictates determine the proper leadership behavior for attaining the same aims.

The original model (Evans, 1970) defined only two leadership behaviors: supportive leadership (akin to consideration) and directive leadership (akin to initiating structure). House and Mitchell (1974) broadened these behaviors to number four, thereby adding participative leadership and achievement-oriented leadership. The effectiveness
of these behaviors depends on the situational conditions, as outlined above. According to Yukl (1998), requisite conditions have been clearly delineated for directive and supportive behavior, only. The model of motivation on which this theory rests is consistent with expectancy theory. Leaders need to enhance subordinates’ beliefs that the instrumentality, outcomes, and valence of outcomes are attractive and within reach.

Directive leadership thrives when the work detail is vague and amorphous, the subordinates are unproven, and the job specifications are lacking. Unclear directions, lacking relevant experiences, and uncertain role functions cause subordinates to doubt their prospects of attaining the task’s finished product. By enhancing each of these doubt-producing factors, leaders can improve the outcome expectations and valences subordinates will generate. Yukl (1998) points out that role ambiguity can be upsetting, and as such, leaders can improve job satisfaction by clarifying the demands of the job. The corrective behaviors for such situational ailments are directive leadership, as called for the path-goal model of leadership.

Demanding, monotonous, and hazardous tasks call for supportive leadership actions. Such direction facilitates employees’ satisfaction and effort by improving their self-confidence, decreasing their concerns, and minimizing the dislikable features of their jobs. Using expectancy theory verbiage, supportive leadership enhances the intrinsic value of performing task-related duties (valence) and the beliefs that such jobs can be completed (expectancy). Competent employees and intrinsically motivating tasks provide contexts that render supportive leadership unnecessary and relatively futile. The theorized situations calling for participative and achievement-oriented leadership lack the
detail of the suggested directive and supportive leaderships' scenarios. However, even
the directive and supportive leadership styles haven’t yet passed empirical muster.

Wofford and Liska (1993) and Podsakoff, Mackenzie, Ahearne, and Bommer (1995) conducted extensive reviews of path-goal research and situational moderator variables, respectively. Wofford and Liska focused on the directive and supportive leadership forms, in particular. The combined conclusions of the two meta-analyses suggest that any decisions about the effectiveness of the path-goal model are premature, despite the plentiful studies. Yukl (1998) claims that “not enough studies were available to provide an adequate test of hypotheses about situational moderators of participative and achievement-oriented leadership” to form any determinations. Wofford and Liska (1993) and Yukl (1989) blame methodological shortcomings for the confounded results stemming from the path-goal studies. Put another way, the only conclusion emerging from the meta-analyses indicates that the path-goal model has not been adequately tested yet.

Yukl (1998) credits the path-goal leadership theory as contributing the conceptual paradigm for including situational variables in the study of leadership behavior. Although the theory lacks consideration of self-concepts in its conceptualization of employee motivation, an asset of transformational leadership (Shamir, House & Arthur, 1993), it provides a framework for linking various leadership behaviors to employee effectiveness and satisfaction in different task and subordinate contexts. Bass (1985) considers these situational factors in conceiving the optimum contexts for transformational leadership behavior, the leader behavior under study (Bass, 1985, p. 158).
The situational moderators investigated during this time amassed to such numbers that confusion and complexity began to emerge (Hunt, 1999). The involved research investigating such complex leadership models yielded little “conceptual capital” (Hoy & Miskel, 2001). Leadership theory scholars expressed doubts about the fruits of their labor. Leadership theories and models lacked much explanation for the performance of organizational members of the early 1980’s revolutionary corporations (i.e., high-tech entrepreneurial companies), organizations revolutionized by spirited executives, and for the epic historical feats by the followers of many famous leaders (Bass, 1985). The leadership models advocated through the first three-quarters of the 20th century explained a modest amount of variance in leadership behavior (1985). Something more than material or psychic exchanges must take place between leaders and subordinates in order to account for the supreme achievements of organized groups throughout history. A new leadership paradigm evolved during the late 1970’s that began to theorize other relationships between leaders and followers. These new possibilities tried to clarify the more potent motivations and behaviors of followers. This new leadership framework, based much on the concept of charisma, will be reviewed next.

New Leadership Paradigm

Roughly 30 years transpired from when Weber’s (1947) charismatic authority notions were documented to their utilization in leadership theory. Bass (1985) feels that Weber’s charismatic leadership ideas didn’t receive much theoretical attention during the intervening years because of organizational sciences’ focus on leader-follower exchange models. The methods and techniques of organizational research lent themselves to the study of material and psychological transactions more than to the study of higher-order
dealings between leader and follower. “[Material and psychological] exchanges are easier to sense, observe, record, and measure. They are logically compelling as long as we can posit that man is a rational and economic being” (p. 6). However, Weber (1947) maintains that some of the most influential leaders in history operated with charismatic authority, “the direct antithesis of [rational authority]” (p. 361). The early use of charisma in leadership models will be explored by reviewing Weber’s charismatic authority conceptions, House’s (1977) charismatic leadership model, and Burns’ (1978) transformational leadership theory.

Charismatic authority. According to Weber (1947), a German sociologist, societies experience a cycle of leadership that includes charismatic authority, rational-legal authority, and traditional authority. The cycle might start with an influential figure that captures the attention of many followers based on a quality that separates him from ordinary people. These traits or abilities endow him with qualities “not accessible to the ordinary person, but are regarded as of divine origin or as exemplary” and, as such, accords him status as a leader (p. 358). The virtuous or moral interpretation of the quality, or charisma, is immaterial, according to Weber, as the important aspect of charismatic authority rests on how the individual is actually seen by those under the influence of his or her charismatic leadership. Thus, good and evil charismatic leaders exist.

In time, the leadership beliefs of the charismatic authority become “routinized” by having its non-rational qualities rationalized (p. 363). The practices espoused by the leader need to have economic and, otherwise, practical importance for the charismatic leader to have any legacy. Those practices deemed to have utility become
institutionalized and protected by policies, laws, and regulations, otherwise known as rational-legal authority. As time passes, the people following in this society continue to practice the beliefs espoused by the charismatic leader through customs and traditions, hence the nomenclature “traditional authority.” Although charismatic authority captures the imagination and spirit of followers, its legitimacy and recognition lasts only as long as its ability to fulfill the needs of its followers.

The vagueness of Weber’s charismatic leadership ideas aren’t too dissimilar to Carlyle's (1847) “Great Man” leadership stories previously mentioned, in terms of leader identification. An element of mystery surrounds both theories of leadership, in that a magical, spiritual, or supernatural power belongs to the leader. However, Weber offered a framework for how the practices under the leader became institutionalized. The multidimensional theory of House (1977), including leader traits, behaviors, and situational contexts, distinguishes his charismatic leadership theory from the previous ones.

Charismatic leadership theory. Although strictly theoretical and empirically untested, House’s (1977) charismatic leadership model offers a series of observable manifestations regarding the behavior and effects of charismatic leaders, representing progress from the mysterious and amorphous descriptions of earlier models. House distinguishes between charismatic and noncharismatic leaders via their behaviors, traits, situational conditions, and effects on followers. The effects of charismatic leaders include subordinates’ sense that their superior’s perceptions are accurate, follow the leader unconditionally, comply with the leader freely, have positive emotional feelings for the leader, passionately work toward the realization of the group’s goals, hold beliefs they can further the mission of the group, and maintain ambitious work standards (Yukl, 91
The likely behaviors of the charismatic leader include articulating an attractive future, distributing power, acting in exceptional ways, creating impressions, subjecting self to risks, managing impressions, performing self-sacrifices, and behaving in ways for followers to emulate (Hoy & Miskel, 2001).

The implications of House’s theory manifest with subordinates’ self-concepts and emotions. This feature of the paradigm is emphasized and further explicated in a subsequent writing of House (Shamir, House, & Arthur, 1993). The crux of their model, again untested, suggests leader behaviors engage subordinates’ self-concepts that, in turn, influence their motivational mechanisms. An example of such dynamics is as follows:

Charismatic leaders increase effort-accomplishment expectancies by enhancing the followers’ self-esteem and self-worth. They enhance self-esteem by expressing high expectations of the followers and confidence in the followers’ ability to meet such expectations (Yukl, 1989; Eden, 1990). By so doing, they enhance followers’ perceived self-efficacy, defined as a judgment of one’s capability to accomplish a certain level of performance. Self-efficacy is a strong source of motivation (Bandura, 1986, p. 351). (Shamir, House, & Arthur, 1993, p. 582).

It is interesting to note that although Shamir, House, and Arthur (1993) reference Bandura’s (1986) social cognitive theory, of which self-efficacy is the central motivational mechanism, they do not explicitly explain how leader behaviors foster any of self-efficacy’s sources of information. Bandura (1977, 1986) states rather explicitly what the sources of self-efficacy are and how they can be manipulated to foster stronger competency perceptions, so their absence in Shamir et al.’s model is noticeable. It follows then that if leadership enhances subordinates’ motivation through their developed self-efficacy beliefs, the mechanisms by which self-efficacy beliefs develop warrant mention. Shamir et al. (1993) consider their model rare in that it “[shows] how
charismatic leaders activate self-concept related motivations, and how these motivations can explain the effects that are not well explained by current theories” (p. 579). Such claims may be more indicative of prior leadership theories’ generalizations than demonstrative of the detail inherent in their theory’s alleged psychosocial blueprint of leadership effects. This void hopes to be addressed with the present study.

Shea and Howell (1999) tested charismatic leadership’s effect on subordinate performance while also gauging subordinate self-efficacy’s role in mediating task feedback. Ninety-nine graduate students manufactured electrical wiring harnesses under a charismatic and a noncharismatic leader. Interestingly, performance quality of harnesses was not affected by leadership style. However, those subjects creating harnesses under the direction of a noncharismatic leader were more sensitive to no task feedback than those working under a charismatic leader. In other words, charismatic-led subjects making harnesses and receiving no task feedback performed not significantly better than other charismatic-led subjects receiving task-feedback. Feedback meant more for non-charismatic led workers than charismatic-led employees.

Furthermore, the interaction between task feedback and leadership style was mediated by the self-efficacy perceptions of the workers. When self-efficacy was considered in the regression equation, the interactive relationship between leadership style and task feedback became nonsignificant, hence its mediating effect. Self-efficacy also mediated the effect of feedback, regardless of leadership style.

Shea and Howell’s (1999) results concerning the relative effectiveness of charismatic behaviors aren’t surprising. Yukl (1998) contends charismatic leadership theories are still relatively new and the research testing their propositions is insufficient,
so individual research results have little to which compare. This is especially true for charismatic paradigms emphasizing the relationships between leaders and the collective, since the early notions of charismatic leadership stressed the dyad and group relationship with the leader. In particular, Yukl credits charismatic leadership theory with providing an explanation, as untested it might be, for the exceptional influence some leaders have on their employees. These models underscore the importance symbols and emotions have in the work motivations of subordinates. Charismatic leadership highlights the importance of leaders in conveying the meaning and purpose of the work to subordinates. Taken together, these attributes of charismatic leadership, the consideration of the collective and employee emotion, and the conveyance of work meaning contribute much to the study at hand.

Shea and Howell’s (1999) findings are relevant not only to the empirical validation of charismatic leadership, but also to the variables considered in this study. Self-efficacy’s role as a mediating variable between leadership and performance provides initial support for exploring the relationship between transformational leadership, collective efficacy, and school achievement. As noted by Shamir et al. (1993), charismatic leadership is very similar to transformational leadership. The emotional, collective, and symbolic considerations mentioned above begin to convey the similarities between the two. More similarities, and the few distinctions, between these post-exchange leadership designs, emerge as this review moves to the focus of this study, transformational leadership.
Transformational Leadership

Frustration with the limited amount leadership performance explained by contemporary models led to the development of new paradigms of leadership. Ironically, historical notions regarding great leaders inspired much of the new leadership thinking. A review of this latest leadership theory's origins, underlying factors, constructs, and effects follow.

Theoretical beginnings. Transformational leadership evolved from the work of James Burns (1978). Burns reviewed leadership practices of rulers and politicians over hundreds of years, distilling the characteristics and elements of influential leadership in his seminal work, titled Leadership. He determined that motives and resources were the key components of power. The essence of the power is found in the relationship between motives, resources, leaders and followers. At any one time, resources and motives can be found with either player. Thus, Burns' conception of power, or leadership, was rooted in the relationship between and among all members of an organization. He was the first to suggest such origins of leadership.

Burns conceived of a transactional leadership style that rested on the exchange of valued items or actions during a bargaining process. Each member involved in the process is interested because his motives may be advanced through the transaction. Although the purposes of both parties can be furthered through the exchange, they are not changed because of it.

Unlike transactional leadership, transformational leadership changes the motives and resources of those involved. It elevates both to higher levels. Motives are lifted from the material into the moral domain, thus the level of concern between leaders and
followers is heightened. Resources are expanded to achieve this enhanced purpose. The nature of transactional exchanges maintains the current status of motives and resources whereas the spirit of transformational dynamics lifts the behavior, resources, and motives of all. Proponents of transformational leadership argue that participants within an organization need more than contingent rewards for agreed-upon work. Organization members will achieve more as they sense a higher meaning and purpose for their work.

Scholars and practitioners welcomed the transformational leadership theory in the late 1970’s. Burns’ ideas came at a time when leadership theory was under fire, as noted above. The contemporary thinking of the day struggled to generate any new knowledge and questioned the usefulness of doing so (Hoy & Miskel, 2001). However, as exciting as Burns’ ideas were concerning the relationships between higher-level effort, resources, and motivation, scholars noted limitations. The perspective lacked an empirically tested model operationalizing the ideas. However, another scholar came along to provide the needed empirical model to keep the transformational ideas alive.

*Transformational leadership and empirical support.* Bass (1985) picked up on Burns' (1978) ideas and developed scales to measure transformational leadership, for he agreed with much of what Burn's conceived. "Much of the [theorized transformational changes in subordinate performance expectations] will be found in Burns (1978)" (Bass, 1985, p. 20). Explicitly, Bass was intrigued by Burns' model depicting leadership influences that went beyond transactional exchanges between the leader and the led. However, the two scholars differed in their conception of transformational behaviors in three distinct ways. First, Bass viewed a subordinates' expansion of needs, at any level of Maslow's hierarchy, as a potential for increased performance whereas Burns held that the
expansion had to be for a higher level. Secondly, Burns qualified transformational leadership as only that which benefits the subordinates or society. Bass (1985) argues that transformational leadership can have a good or bad effect on subordinates or society; the only qualifier is that the "followers' attitudes and behaviors were transformed by the leader's performance" (p. 21). Finally, Burns viewed transactional and transformational behavior as polar ends of a continuum. The more a leader behaved transactionally, the less he could act transformationally. Bass perceived leadership to be a two-factor structure. Leaders can act in varying degrees from either perspective.

In all, though, Bass (1985) complemented much of Burns' (1978) theoretical work with instrumentation and empirical validation. He provided characteristic behaviors and the statistical support for a transformational leadership model, a paradigm closely related to charismatic and visionary leadership conceptions (Fiol, Harris, & House, 1999). Bass was involved with the Ohio State Leadership studies of the 1950's when the LBDQ and the concepts of consideration and initiating structure were developed. One might notice these elements from his background emerge in his transformational leadership scale. His conceptualization of leadership further stemmed from not only Burn's ideas, but he also borrowed House's (1977) notions describing charismatic leadership (Hunt, 1999). Charisma is the most important behavior of transformational leadership (Bass, 1985), for its description, as well the description of the other transformational behaviors follows.

Various forms of transformational leadership differ modestly according to the number of distinguishable administrative functions (e.g., Bass, 1985; Leithwood, 1994). Bass and Avolio's (1993) commonly identified four “I’s,” the common moniker for the operationalized leadership behaviors of this model in their transformational leadership
scales (Bass & Steidlmeier, 1999; Hoy & Miskel, 2001; Yukl, 1998). However, the four "I's" have evolved into five "I's," as Bass and Avolio (2000) responded to criticism from various scholars (Hunt, 1991; Yukl, 1994) regarding the problems of distinguishing between behaviorally-based charisma and attribution-based charisma. Idealized influence, formerly known as charisma, has factored into behaviors and attributed dimensions (Bass & Avolio, 2000). In addition to the two forms of idealized influence (behaviors and attributed), transformational leaders display inspirational motivation, intellectual stimulation, and individual consideration. These four will be serve as operational definitions within this transformational leadership platform.

Transformational leadership behaviors. The first “I,” idealized influence, describes behaviors and impacts that develop followers’ trust. According to Bass (1996), participants accept changes in how and what work gets done as leaders demonstrate idealized influence, or charisma. Without trust, change must be motivated through its merits alone - perhaps a much more difficult inducement. In order for leaders to show integrity they must be consistent in word and deed. They must show that they live what they speak. Transformational leadership displays congruence between spoken thoughts and behaviors. Bass (1996) reports that transformational leaders are consistent rather than random in their behavior. Simons (1999) argues that transformational leadership cultivates subordinate trust and loyalty through behavioral integrity. Integrity manifests in those viewed as role models. Avolio (1999) reports Federal Express qualifies its Best Leaders as those displaying integrity by serving as role models for others. Hoy and Miskel (2001) explain idealized influence results from serving as a role model for
followers. In sum, Avolio (1999) cites integrity as one of the features of the idealized influence transformational dimensions. The following leader behaviors are some manifestations of idealized influence (Hoy and Miskel, 2001):

- demonstrating high standards of ethical and moral conduct; sharing risks with followers in setting and attaining goals, considering the needs of others over their own desires; using power to move individuals or groups toward accomplishing their mission; conveying a vision and cause, but never for personal gain. (p. 415)

Inspirational motivation, the second “I,” describes spirited appeals and images to increase participants’ focus on a vision. Bass and Steidlmeier (1999) state these appeals and images encourage commitment to mutual goals and responsibilities. A mutually constructed future vision stands as the prerequisite to all such inspirational appeals. Inspirational motivation behavior utilizes verbal, symbolic, and body language to inspire followers to work toward a common vision. The feedback component of this open communication process takes the form of follower suggestions, if any, about very different and desirable alternative goals or methods. Barnett, McCormick, and Connors (2001) portray inspirational motivation as leader behavior that motivates and stimulates organizational members by "providing meaning and challenge to their followers' work" (p. 26).

Intellectual stimulation prompts creative thinking about school problems, especially related to teaching and learning. The evaluation, formulation, and implementation processes of problem solving are all open for participant input. Leaders encourage the abandonment of traditional thinking. Stimulation of new ideas is promoted by questioning current strategies, posing problems from different perspectives, and supporting new procedures for work (Hoy & Miskel, 2001). Leithwood (1994)
reports high levels of problem solving expertise support transformational leadership practices. The thought processes behind the problem-solving expertise include deliberate use of problem-solving processes, varied information relevant to the problem, and focused values. Intellectual stimulation promotes intelligence, rationality, and careful problem solving (Bass, 1996). The value often associated with problem-solving behavior concerns students' welfare. Competence in transformational leaders manifests in their attitudes, too. These leaders consider their staff a valuable source of ideas, maintain openness to new ideas, and exert self-confidence borne from successful experience.

“Expert thinking lies behind effective [transformational] leadership practices and is not contingent.” (p. 515). By solving complex, difficult tasks, leaders model cerebral skill and galvanize staff trust. Bass (1996) defines intellectual stimulation as promoting intelligence, rationality, and careful problem solving. It seems clear that transformational leadership demonstrates and promotes competence.

Individualized consideration behavior, the fourth “I,” gives organizational members personal attention. The leader focuses attention on each teacher and staff person’s need for achievement and growth. Bass and Steidlmeier view individualized consideration as service to others, underscoring "… the necessity of altruism if leadership is to be anything more than authoritarian control (1999, p. 189). Transformational leaders take time to know members individually and help them develop their capacity to grow and develop as educators. Leadership concern for followers narrows the often-perceived expanse between subordinates and superiors. As noted earlier, transformational leaders focus on individual needs for progress and development by serving as advisors and mentors (Bass, 1996). When Avolio (1999) asked
transformational workshop respondents to describe their ideal leaders, they shared that superiors answered subordinate questions quickly, expressed concern for subordinates’ well-being, linked followers to custom training opportunities, were available when needed, and repeatedly encouraged through mentoring, coaching, and counseling – an approximate description for individualized consideration. Although idealized influence and inspirational motivation are the most effective leaderships for stimulating organizational performance (Avolio, 1999), individual consideration’s role should not be minimized. Not only does individual consideration still have a significant effect on follower behavior, its absence has serious implications for subordinate morale. Interestingly, a lack of individual consideration strongly relates to teacher burnout (Leithwood, Menzies, Jantzi, & Leithwood, 1999). Both empirically proven and purposefully designed, transformational leadership includes behaviors for meeting the individual needs of teachers.

Bass's transformational behaviors are measured in an instrument labeled the Multi-Factor Leadership Questionnaire (MLQ) that gauges employee perceptions of their leaders' full range of behaviors, including transactional and non-leadership actions (Bass & Avolio, 2000). The scales measuring transformational behavior and transactional evolved from factor analyses of the original questionnaire and later editions of it. Some modifications have occurred to scale items (i.e., they became more behavior-oriented) on intellectual stimulation and idealized influence. Moreover, idealized influence has differentiated into behavior- and attributed-based dimensions, as noted above.
Inspirational motivation was added later. Research testing transformational leadership behaviors' relationship with different manifestations of effectiveness has often used the MLQ (Yukl, 1998). A review of relevant research follows.

Research on transformational leadership. Lowe, Kroek, and Sivasubramaniam (1996) performed a meta-analysis of research studying transformational leadership to synthesize the varied findings. The studies considered had to employ the MLQ, establish a variable for leader effectiveness, document the sample size, report a Pearson correlation coefficient for leadership effectiveness and leadership style, and operationalize leader as being a direct supervisor of the respondent, not a hypothetical or idealized leader. Thirty-nine studies met these criteria. Because of inspirational motivation's later addition to the MLQ, it was not part of this meta-analysis.

Results of Lowe et al.'s analysis indicate that charisma, individualized consideration, and intellectual stimulation correlate with leader effectiveness. The researchers boldly underscore the certainty of this finding by declaring, "Those who have asserted that the transformational construct has been embraced because of the affective lure of its implications - rather than on empirical, practical, or rational grounds - are impeached by the consistency of this result across studies" (Lowe et al., 1996, p. 415).

Moderator relationships were revealed in their analysis, as well. Organizational type moderated the correlation strength between leadership effectiveness and leadership behavior. Public organizations provided settings in which transformational leadership behaviors more strongly correlated with leadership effectiveness than private concerns. Subordinate perceptions of leader effectiveness displayed stronger correlations with transformational behaviors than objective measures of leader effectiveness. Yukl (1998)
states this phenomenon is typical of survey research of leadership behavior. As theorized by Bass (1985), charisma steadily ranked as the strongest predictor of leader effectiveness among the three transformational actions, regardless of organization type, level of leader, and how effectiveness was measured.

The authors' find that intellectual stimulation more strongly correlates with leader effectiveness in public organizations than in private organizations. The surprising nature of this finding rests on the popular notion that public institutions typically follow more bureaucratic processes and procedures, such that status quo thinking holds sway. Indeed, Merton (1957) notes that bureaucracies may often curb innovation and achievement with seniority-oriented career paths. However, Lowe et al. reason that it may be the very rigid nature of many public organizations that causes subordinates to relate intellectual stimulation behavior to their supervisors' effectiveness. Private firms have a competitive need to be more innovative and risky than publicly funded concerns.

To the extent that intellectual stimulation involves cognitive reappraisal of the status quo and the questioning of long held assumptions, it appears that this construct is more highly associated with performance in the public sector than it is in private industry" (Lowe et al., 1996, p. 415).

Finally, the authors reiterate their support for this rationale by noting intellectual stimulation behaviors' fit for military and educational organizations, a thought worth noting for the present study. A consideration of transformational leadership in educational organizations, in fact, marks the next junction in our literature review.

*Transformational Leadership in Schools*

Leithwood and his colleagues have largely done the study of transformational leadership in educational institutions since 1994. Leithwood et al. devised their own
transformational leadership instrument to measure leadership behaviors in elementary and secondary school settings. They focused primarily on the relationship between school leadership behaviors, school culture, and teacher and student outcomes. The impetus for their leadership studies, their initial research, the development of their scale, and a recent studies outlines this review of Kenneth Leithwood's work with transformational leadership in school settings.

School reform and leadership. Leithwood (1994) endeavored to explicate the requisite leadership behaviors for leading schools through necessary school restructuring efforts. He felt that the leadership style popular at that time, instructional leadership (Greenfield, 1987), was wholly inadequate for effecting the changes needed for improving schools. Instructional leadership behaviors lead schools through what Leithwood terms first-order changes, or modifications to the technical aspects of teaching. These changes manifest in such instructional initiatives as switching focus to constructivist forms of learning or to teaching for understanding. However, instructional leadership does not focus on second-order changes, or alterations to the psychosocial aspects of the school. Such changes include fostering a collaborative vision, facilitating an industrious and fertile work culture, and sharing leadership with others, among other transformations. Leithwood, Jantzi, and Steinbach (1999) state that substantial evidence indicates a sole focus on first-order changes in school restructuring efforts lead to failure for most change initiatives (Fullan, 1993; Furhman, 1993). Thus, Leithwood and colleagues endorse leadership behaviors facilitating second-order change initiatives.
They believe transformational leadership activities are such behaviors. Leithwood and his colleagues developed a school-specific transformational leadership model based on the aforementioned premise as described below.

*School-specific transformational leadership scale.* The scale developed by Leithwood (1994) derived from two sources: the results of an extensive qualitative study of educational leaders especially successful in reforming schools and “theory-testing / theory modification” research based on Bass’s (1985) model of transformational leadership (p. 502). The modifications Leithwood made to the Bass-oriented scale items reflect the realities of leading public schools. The first major publication of Leithwood’s (1994) transformational leadership model revealed six behavioral dimensions: vision articulation, goal consensus building, high-expectation conveyance, modeling, intellectual stimulation, and individualized support. Despite the conceptual distinctions of the instrument’s factors, items frequently fall onto one factor when data sets are factor analyzed (Leithwood & Jantzi, 1999; Leithwood & Jantzi, 2000). Number of factors notwithstanding, the conceptual similarities between Leithwood’s transformational leadership model and Bass’s (1985) framework are apparent. Moreover, as reported by Leithwood (1994), the effects found for his school-based model parallel the effects found for nonschool transformational scales. An example of such effects follows.

*Early results.* Leithwood’s (1994) early work on transformational leadership spanned four years before constructing, testing, and publishing his research. His findings suggest transformational leadership has significant direct an indirect effects on school progress with reform programs and on subjective measures of student outcomes. In particular, transformational behavior has statistically significant influence on three
sources of teacher commitment. In descending order of influence, transformational behaviors affected teachers’ goals, teachers’ context beliefs, and teachers’ capacity beliefs. Additionally, in-school conditions, a concept comprised of school goals, culture, programs and instruction, policies and organization, and resources, related most strongly to many of the dependent variables, namely seven school restructuring programs and two teacher-perceived student outcomes. Finally, Leithwood (1994) found that leadership behavior has a moderately strong relationship with in-school conditions, a construct significantly correlated with several outcome variables, as previously noted.

Leithwood’s (1994) work sets the stage for the present study in several ways. First, he provides initial support for transformational leadership relating to teacher commitment to change, a construct conceptually similar to collective efficacy in that competency beliefs and context beliefs are considered. Teacher commitment, though, measures teachers’ perceptions of their individual competencies and contexts, not the collective’s. Secondly, Leithwood demonstrated a significant and positive relationship between transformational leadership and a proxy for school climate, labeled in-school conditions. This relationship is important because collective efficacy serves as one manifestation of school climate (Goddard et al., 2000). Thirdly, his 1994 study laid the groundwork for one of his later studies testing the relationship between family background and transformational leadership, a dynamic tested in the present study. A review of Leithwood’s later study, Leithwood and Jantzi (2000), follows.

Transformational Leadership and School Conditions. Leithwood and Jantzi (2000) based their research on the assumptions most school restructuring programs make about teachers’ capacity development for successful reform to occur. The success of
these initiatives also depends on significant levels of teacher motivation and commitment for effectively handling the challenges attributed to school reform projects. In pursuit of validating the ripeness of reform conditions for transformational leadership, Leithwood and Jantzi tested a model mapping relevant variables, presumably drawn from the aforementioned assumptions, along with measures for student outcomes and external organizational controls. Measures of student engagement were included to test the notion that transformational leaders influence organizational outcomes (see Shea and Howell, 1999). Family educational culture served in the place of socio-economic status as a control for student background, on the basis it provides twice the statistical power for predicting “academic learning” (Walberg, 1984).

Leithwood and Jantzi tested the validity of their model using LISREL, a statistical procedure to gauge the fit of one’s conceptual framework with the quantified relationships drawn from the data. The results of this analysis provide both expected and surprising findings. Not surprising was the strength of the relationship between family educational culture and student school engagement. Given Coleman et al.’s (1966) seminal study in which family background clearly overwhelmed any effect schools might have on student school performance, this finding was expected.

Transformational leadership had a nominal, yet statistically significant, effect on student engagement. This finding, too, finds support in the literature, as Ogawa and Hart (1985) defended such relationships. They noted scant student achievement variance is available for between school differences, such as principal leadership perceptions. This contention warrants reconciliation with Goddard et al. (200). They found that collective
efficacy beliefs explained more than 50% of the between school variance in student achievement scores, as noted previously. This 50% represents half of the variance available to between-school factors.

Surprisingly, family educational culture had an insignificant effect on transformational leadership, another dynamic worth noting for the present study. Intuition suggests teacher perceptions of transformational leadership behavior should relate to the academic preparedness of the students, often indicated by their family background.

Still, more surprise stems from the modest relationship between school conditions and student engagement. Although the fairly small relationship between school conditions and student engagement seems shocking, given other reported relationships of school climate and student outcomes (Hoy & Sabo, 1998), more surprising is the absence of teacher commitment, motivation, and capacity measures in the construct. After all, these concepts formed the basis of their study’s purpose, as noted above. Although the authors state, “For purposes of this study, organizational culture was defined as the norms, values, beliefs, and assumptions that shape members’ decisions and practices,” their school climate variable didn’t take into account beliefs as they related to motivation, capacity, or commitment (Leithwood & Jantzi, 2000, p. 115). School conditions, as conceived by the authors, comprised of teachers’ perceptions of school goals, culture, information collection, policies and procedures, shared planning, and structure and organization. To the extent these concepts invoke notions of school climate, the findings are surprising.
Leithwood and Jantzi’s (2000) choice to conceive school organization in the manner they did still leaves the study of transformational leadership’s effect on teachers’ beliefs and student outcomes a relatively untested and timely proposition for two reasons. Evidence supporting school reform assumptions (i.e., teacher capacity development) still is very limited (Leithwood & Jantzi, 2000), warranting a study of transformational leadership, collective efficacy, and student achievement in schools undergoing mandated change. Moreover, the recent reauthorization of the Elementary and Secondary Education Act mandating public school reform adds further urgency to testing these reform assumptions with schools affected by ESEA. Thus, this is where the present study "jumps off" from the relevant literature (Pajares, 2002. p. 3).

To this point, the literature reviewed teacher efficacy and transformational leadership. The last major concept to be studied for this research is student achievement. Student achievement serves as the school outcome variable of this study. A focused review isolating attention on the specific operationalizations of student achievement comprises this section.

Student Achievement

The importance of student achievement in our country assumed new levels of political and legislative significance with the presidential signing of the No Child Left Behind Act of 2001: Reauthorization of the Elementary and Secondary Education Act on January 8, 2002 (No Child Left Behind, 2002). The passing of this law begins the most extensive federal role in public school education since the inception of the ESEA in 1965. The Act is based on four major changes for local and state interests in public education: stronger accountability for student performance, more options for local control, more
choices for parents, and provisions for developing teaching methods that have been proven to work (No Child Left Behind, 2002). The accountability provision of the law carries special relevance for this study, as it requires states to implement annual standardized assessments in reading and math no later than the 2005-2006 school year for students in grades 3 through 8. The reasons for mandating such requirements are two-fold.

Two Reasons for Public Demand of Improved School Achievement

Much of the interest for state assessments of school performance comes from public demand for school accountability. The plight of many youth not connected with schools concerns political leaders, scholars, and other careful observers of education and society. One-quarter of all adolescents leave school before graduating (Steinberg, 1999). Furthermore, 40% of high school graduates discontinue their formal education. The importance of school attainment is as follows. Because continuation of formal education is often contingent on prior school success, student achievement warrants serious study for improving school attainment rates. Moreover, because school attainment is the one of the best predictors of future economic gain (Steinberg, 1999), these figures are cause for concern.

The problems of youth disengagement from school go beyond economic and financial hardships, however. The achievement gap between majority and minority classes has been well documented since Coleman et al. (1966), and little has changed since to close it, for the subtitle of the ESEA bill sent to the President called for closing the achievement gap. Whites continue to achieve at higher levels in school than do those students of minority status (Campbell, Hombo, & Mazzeo, 2000). The minority class
students under performing whites are most closely linked with academic failure and poverty (Steinberg, 1999) and will comprise close to 36% of the US population by the year 2040 (US Census Bureau, 2002). If more progress is not made in closing the achievement gap as has been realized since the Coleman Report, a sizeable portion of the US population will comprise the underclass. Certainly, a disenfranchised class of this size makes for potential political instability, if not worse (e.g., French Revolution, Argentina Crisis, etc.). Thus, the reasons justifying public demand of school accountability rest on political stability and economic interests.

State assessments of school student performance manifest as one measure of school accountability. Annual reporting of student assessment test scores, disaggregated by social economic class and race, brings regular public scrutiny to the achievement gap problem. This scrutiny has not been unanticipated, though.

Ohio's Role in Preparing for School Accountability

Many states have been preparing annual assessments during the years leading up to the "No Child Left Behind" law. Only three states did not have some form of statewide assessment during the 1997-1998 school year (Wolk, 1998). In 2001, Ohio, the state in which data will be gathered, passed a bill titled "Senate Bill 1" requiring its department of education to implement achievement and diagnostic tests yearly to children, kindergarten through the eighth grade (Stump, 2001). For much of the 1990's, Ohio had been administering yearly proficiency tests in grades four, six, nine, and twelve (Stump, 2001), so the tenor of Senate Bill 1's mandates are not completely new. However, the yearly reporting of assessment scores to the public, as required by the No Child Left Behind law, causes the state to make further changes to meet federal
compliance. As Senate Bill 1 reads now, school reporting of student performance only takes place for those grades in which achievement tests are taken, not for the grades in which diagnostic exams are administered (Stump, 2001). Thus, school reform continues in Ohio, as well as in the rest of the nation's schools.

School achievement in this study will be operationalized as either the percentage of students passing the fourth-grade math proficiency exam or as the percentage of students passing the fourth-grade reading exam. Both measures will be considered as dependent variables. The validity of the proficiency exams serving as measures of student achievement rests on the state's purposes for employing their use, the process of their creation, and their correlation with other national standardized achievement tests.

*Validity of Ohio's Fourth Grade Proficiency Tests*

The fourth-grade proficiency tests serve to measure a fourth grade student's level of literacy and competency (Ohio Department of Education, 1995). A student's level of competency and literacy would seem to proxy one's level of achievement. Our nation's main educational progress program, the National Assessment of Educational Progress, reports reading progress of public school students under the heading "Achievement/Proficiency" (Mullis, 1992, p. 200). Thus, the proficiency-achievement distinction between the dependent variable and the instrument operationalizing it appears to be inconsequential.

Ohio's fourth-grade proficiency exams are products of a multi-stage development process. Initially, a committee of 25 people convened for each content area targeted for proficiency testing to insure the items comprising such exams are valid. Teachers made up half the committee members and various representatives of relevant professional
educational organizations comprised the other membership. Steps were taken to insure proper committee representation of races, cultures, genders, and geographies. Content area committees decided on learning outcomes for grades via discussion of and reference to the State Board of Education adopted model courses of study.

Once the learning outcomes were established for a grade-level content area, bids were drawn up specifying the activities and products contractors would need to satisfy in order to complete the construction of a proficiency exam. The successful vendors chosen to develop the exam then presented their ideas for test and item specifications to the content review committee for discussion. After the test and item specifications were approved, the vendor developed items that met those specifications. Drafted items were then subjected to a five-step evaluation process including bias review, content review, field-test, bias review with field test results, and content review with field-test results. "Only items that go through all five steps successfully are eligible to be included in the item bank and used in the operational form of the test or in the practice test" (Ohio Department of Education, 1996, p. 3). This test development process validates the test, insofar, that educators spanning from classroom to college and from government to professional settings had input into the construction procedure. A standardized procedure reviewed each item for content and bias-free satisfaction.

Finally, the validity of the proficiency exams may be implicated by their correlations with other standardized tests. A sample of 158 seventh grade students took the Ohio sixth grade proficiency exam one year before taking the Stanford Achievement Test - Form 9, a nationally reputable assessment. Both sets of test scores were correlated for comparison. The math sections of both tests correlated moderately high ($r = .779$, $p <$
and the reading tests correlated slightly lower ($r = .627$, $p < .01$). Although these Pearson correlation coefficients don’t indicate the tests are perfect substitutions for one another, the quantified relationships indicate similar concepts are being tested. Given that the sixth grade and fourth grade proficiency tests are created using the same standard procedures, it is assumed the fourth grade proficiency test would correlate with a fifth grade Stanford Achievement Test similarly.

**Administration and Scoring of Proficiency Tests**

The fourth-grade proficiency tests follow administration procedures like most other standardized tests (Ohio Department of Education, 1996). Students are not permitted to use reference materials of any kind during each of the two and one-half hour exams. They may only use the writing instruments they bring to the test and the test materials provided by the teacher. Test administration manuals suggest teachers remind students to read over their choices carefully as the multiple choice sections of the test often use frequently made mistakes as response choices. Both the reading and the math tests have objective, short-answer, and extended-answer questions. Teachers are encouraged to administer the half-length practice test to help prepare student for the actual proficiency exams.

An outside scoring company secured through a competitive bidding process initiated by the Ohio Department of Education scores the proficiency tests. The scoring company works in conjunction with a Rangefinder Committee to establish benchmarks for consistency in scoring the short and extended response test items. The Rangefinder committee is made up of teachers and ODE representatives who read, score, and discuss hundreds of field-test papers to arrive at common notions of performance standards.
The mechanism for insuring consistency during the scoring process works this way. Tests are scored by groups of raters gathered at tables. Each table of scorers mark a calibration packet as part of their work grading tests. Deviant scorers are identified by their abnormal reactions to the calibration packets and have their tests re-scored for that session. Consistent instances of scoring abnormalities give reason for scorer removal (Ohio Department of Education, 1996).

*Rationale for Using Ohio Proficiency Tests*

There were many features of the proficiency tests that made their inclusion in this study a valuable asset for this research. Every elementary school in Ohio must administer the same fourth-grade reading and writing proficiency exams during the same one-week period of the year following a standard set of administration directions. Thus, the state fourth-grade reading and math assessments provide common measurements of student achievement following approximate protocol of quantitative research practices. The scoring procedure, as described above, also follows strict standardized procedures to insure consistent rating of the achievement scales. Having the assessments channel back through the Department of Education makes for a central station in which to collect the measures. The availability and feasibility of these assessments were, perhaps, the most appealing features. At no expense to the schools or to this study, all subjects were queried for achievement in the same way, with the same instrument, and scored uniformly.
Theoretical Rationale, Hypotheses, and Path Model

This final section of the chapter brings together the empirical and theoretical findings previously reviewed to develop a framework for the proposed variable relationships in this study.

Rationale for Hypotheses

Research and theory suggest a positive relationship between transformational leadership behaviors and efficacy beliefs (Fuller & Izu, 1986; House & Shamir, 1993; Shamir et al., 1993; Shea & Howell, 1999). To understand this dynamic, Bandura's (1997a) social cognitive theory needs to be reviewed. Bandura conceives triadic reciprocal causation as the theoretical design linking behavior, socio-contextual features, and personal factors for learning and motivation. All three dimensions mutually determine each other in an interactive manner. This broader framework contextualizes the relationships affecting individual's motivations for psychosocial functioning. Self-efficacy stands as one of the central mechanisms in forming motivation for agency (Bandura, 1977, 1986, 1997a).

The sources of self-efficacy beliefs determine one's level of belief that he or she can attain certain performance standards. A consideration of one's mastery experiences, vicarious learning, social persuasion, and emotional states pertaining to a given situation forms the competency belief about successfully performing the behavior in question. Transformational leadership behaviors, the four "I's," serve as influences affecting personal and cognitive factors. More specifically, the four "I's" each influence at least one source of efficacy information. Consider the chart below for examples of theorized linkages between transformational leadership behaviors and a subordinate's self-efficacy.
beliefs. The chart suggests that to the extent a principal manifests transformational
leadership behaviors, his or her teachers will perceive stronger beliefs about faculty
competence.

<table>
<thead>
<tr>
<th>Transformational Leadership Behaviors</th>
<th>Sources of Collective Efficacy Beliefs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idealized Influence (e.g., modeling)</td>
<td>Mastery Experiences</td>
</tr>
<tr>
<td>Inspirational Motivation</td>
<td>Vicarious Learning</td>
</tr>
<tr>
<td>Intellectual Stimulation</td>
<td>Social Persuasion</td>
</tr>
<tr>
<td>Individualized Consideration</td>
<td>Emotional States</td>
</tr>
</tbody>
</table>

Figure 2.2: Correspondence between Transformational Leadership Behaviors and Sources of Efficacy Information

Recalling Bandura's (1997a) assertion that self-efficacy beliefs and collective efficacy beliefs differ only in their unit of agency, these theorized relationships between principal behavior and efficacy information apply the collective level of psychosocial dynamics. For example, to the extent a principal models for the staff at team meetings how to analyze student data and apply the results to instruction, he or she has produced vicarious learning experiences for all the teachers on that topic. To the extent that the principal provides time and the activity for teachers to work with one another on student data issues, he or she is facilitating the staff's mastery experiences. To the extent the principal expects and motivates teachers to apply his or her knowledge of data analysis to his or her teaching, he or she is providing their persuasion. Finally, a principal can reinforce the staff's sense of mastery by recognizing proximal gains in their real advances.
on that issue. Although specific relationships between the four "I's" and particular sources of self-efficacy beliefs will not be estimated in this study, a positive relationship is expected between transformational leadership and collective efficacy. Thus, the following hypothesis is advanced:

H1: All dimensions of transformational leadership behavior are positively related to teachers’ collective efficacy; that is, idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration are all positively related to collective efficacy.

Goddard et al. (2000) and Bandura (1993) demonstrate positive relationships between collective efficacy and school achievement. The theoretical underpinnings supporting this relationship use Coleman's (1985, 1987) notion about the normative effectives of climate on group behavior. Beliefs emerge within groups to dictate expected behaviors of members. These beliefs act as norms providing some influence over the actions of group participants. In terms of schools, Coleman's theory suggests that deviant behavior by any teacher risks corrective reactions by other teachers. These reactions likely take the form of verbal remarks proportionate to the norm violation.

Goddard and Goddard (2001) found that mean teacher efficacy scores across schools fluctuated more with schools' collective efficacy beliefs than with the schools' socioeconomic characteristics. Although causation cannot be determined in such cross-sectional studies, certainly one can theorize teachers' beliefs and behaviors about teaching stem, in part, from their perceptions of the staff's ability as a whole. Therefore, perceptions about staff competence influences what one believes he or she can accomplish. Bandura (1977) advocates that such efficacy perceptions relate to one's propensity to choose more challenging tasks, resilience in the face of setbacks,
persistence in the face of challenge, effort investments in given endeavors. In sum, collective efficacy beliefs facilitate teacher behaviors and attitudes that foster student learning. As such, the following hypothesis emerges:

H2: Collective efficacy is positively related to student achievement, in particular reading and math achievement.

As noted from the Tschannen-Moran et al. (1998) and Goddard et al. (2000), teacher and collective efficacy beliefs stem from teachers' perceptions of the task and the task's context. Since Coleman et al.'s (1966) study of educational equality, socioeconomic status reigned as one of the strongest influences on student achievement. Certainly, Tschannen-Moran et al.'s integrated model of teacher efficacy and Goddard et al.'s model of collective efficacy mesh with Coleman's observation. Socioeconomic status (SES) measures imply much about the context in which schools operate. For instance, the resources available to schools and families, the academic attainment of students' parents and guardians, and the community's expectations for schools all play roles in supporting the educational process. Context perceptions need to be considered for assessing the likelihood of completing the task at hand. Despite the context perceptions of teachers folding together with their task perceptions in forming collective efficacy notions (Goddard et al., 2000), SES significantly affected the collective perceptions of teachers. Thus, context has an independent affect on staff competency belief, as replicated in Hoy, Sweetland, and Smith (2002) and Bandura (1993). The same relationship is expected in this study.

H3: Socioeconomic status is positively related to collective efficacy.
Liethwood and Jantzi (2000) did not find any relationship between school socioeconomic context and transformational leadership behavior in their study. However, they conceived family background differently than how SES is usually conceived (2000). Many of the collective efficacy studies to date (Goddard et al., 2000; Bandura, 1993; Goddard & Goddard, 2001; Hoy et al., 2002) operationalized SES in a more objective and fiscal manner than how Leithwood and Jantzi (2000) measured family background, which was conceived as student subjective measures of their school and career aspirations, their parent involvement, the academic and work structure of their homes, among other family considerations. Students may have overestimated their actual behaviors and attitudes on these scales, as social desirability may have been at work. Subjective and objective measures of the same concept may yield different results, as noticed in Lowe et al. (1996). It seems reasonable to suggest that context matters when perceiving the actions of a leader. When the context of a school drowns in financial woe, subordinates may believe lacking environmental resources overshadow the effects of a principal, no matter how charismatic, inspirational, or personable he or she is. By the same token, when the environment of a school sets amidst readily available resources, subordinates may feel leadership has much at his or her disposal to effect any needed change. For these reasons, the following hypothesis is forwarded:

H4: Socioeconomic status is positively related to transformational leadership.

Ogawa and Hart (1985) demonstrate through their research that principal behaviors explain very little of between school variation in student performance. Leithwood and Jantzi's (2000) work confirmed this observation: transformational leadership's limited but significant effect on student outcomes. Jencks et al. (1972)
contend that only 15 percent of between school achievement variation is explained by school characteristic differences, making any significant leadership effects limited at best. The reason for this particular dynamic lies in the extent to which school leaders meaningfully interact with the students. The small achievement variation explained by school leader actions may be attributed to the overall tone and attitudes they infuse into the school climate which affect students as well as teachers.

However, principals' effects on climate relates strongly to teachers' efficacy perceptions (Bandura, 1997a; Lee, Dedrick, & Smith, 1991). Leithwood and Jantzi's (2000) study, described earlier, showed that leadership behavior related strongly to school conditions, a concept very similar to climate. One main influence of climate stands as the degree principals inspire, stimulate, model for, and care about the teachers. To the extent leaders practice these behaviors, they foster the efficacy beliefs of teachers, a proposition explained earlier. These beliefs, components of school climate (Hoy & Sabo, 1998), much more strongly predict student performance measures than leader behaviors do (Bandura, 1997a; Goddard et al., 2000). Teachers interact daily with students, so it is their behavior that more strongly affects student achievement, behavior theoretically driven by their beliefs in their ability, as per self-efficacy theory. As such, the following hypothesis is advanced:

H5: Transformational leadership behavior is directly related to student achievement and indirectly related to student achievement through collective efficacy.

Socio-economic status has long been one of the strongest predictors of student achievement (Coleman et al., 1966). The reason for such effects lies in the community and school resources available for student learning. In school communities where SES is
relatively high, material and intellectual resources provide aid to students in their learning. Conversely, in low-SES areas, additional academic challenges, represented by lacking material and intellectual capital, impede student learning. Although such challenges may be overcome by teacher resilience, perseverance, and dedication (Brookover et al., 1979), the direct relationship between SES and student achievement remains a potent factor.

Goddard et al. (2000) and Bandura (1993) show that SES has a positive relationship on teacher collective efficacy beliefs. Recalling that task analysis perceptions influence teacher competency beliefs, the role of SES emerges during times of such determinations. To the degree teachers view their instructional roles as being helped by available material and parental resources, direct manifestations of SES, their efficacy beliefs probably improve. Conversely, to the degree teachers find their tasks negatively affected by deficient teaching materials and community academic support, their efficacy beliefs likely decline. Because teacher beliefs affect their attitudes, perseverance, resiliency, and choice of instructional techniques (Ross, 1998), SES should also indirectly affect student achievement through collective efficacy. Thus, the following relationship is hypothesized:

**H6:** Socioeconomic status is directly related to student achievement and indirectly related to student achievement through collective efficacy.

*Path Model*

The hypothesized relationships come together to form a causal model of student achievement. The model moves the analysis from a bivariate one to a multivariate one, where the relative strengths of each variable's influence on each other can be estimated.
Furthermore, modeling the hypotheses in this fashion permits a potentially cohesive and systematic explanation for student achievement. Before proceeding with the explanation of the causal model, consider its illustration below:

![Diagram of a theoretical model of school achievement](image)

**Figure 2.3: A Theoretical Model of School Achievement**

The two-way arrows noted in the path diagram indicate reciprocal influences between concepts. These relationships are consistent with Bandura's (1986) social cognitive theory. Triadic reciprocal causation explains psychosocial functioning, motivation, and human agency within an interactive and mutually deterministic network of environmental, behavioral, and cognitive influences. In the model developed above, both transformational leadership behaviors and collective efficacy beliefs have direct
influences on teachers' sources of self-efficacy beliefs. Collective efficacy affects, at least in part, teachers' instructional behaviors and attitudes (Goddard & Goddard, 2001), which in turn are assumed to influence student achievement scores (Ashton & Webb, 1986; Midgley et al., 1989). In the opposite direction, student achievement scores shape the cognitive processes of principals and other teachers. For instance, to the degree a transformational principal views the effects of his behaviors to be positive (i.e., student scores are high), a principal will act in transforming ways. Relatedly, to the degree student scores indicate success, teachers will perceive their abilities to be strong.

Similarly, transformational leadership behaviors persist in accordance with the perceived abilities of the staff. Leaders who sense their behaviors as having a positive effect on the staff will continue looking for opportunities to behave in that manner. To the degree leaders' behave in transformational ways, teachers' sources of collective efficacy beliefs will be enhanced, as per the previous linkages made between the four "I's" and efficacy determinants.
CHAPTER 3

METHODOLOGY

To test the hypotheses and causal path model developed in the last chapter, data was collected from a set of 146 elementary schools in Ohio. The sampling, data collection, instrumentation, and data analysis procedures are described below.

Sample

The sample for this research comprised of 146 elementary schools in Ohio. Steps were taken to solicit participation of urban, suburban, and rural schools. Schools having 21 or more certified teachers and containing grade 4 were considered for the study. According to data files downloaded from the Ohio Department of Education website, roughly 1,095 schools qualify for this study, out of 2,689 elementary schools in the state. Most of the qualifying schools have a K-4 or K-5 structure (67%). Using measures of urban-rural density and SES, a distribution of participating schools was monitored to maintain balance according to these indices. To solicit participation, a phone script was constructed to standardize the recruiting process.

Research Instrument

Four variables are of primary concern in this study: transformational leadership, collective efficacy, socioeconomic status, and student achievement. Two of the variables are measured using constructs developed elsewhere; a construct for transformational
leadership is designed specifically for this study. Socioeconomic status is operationalized using a statistic describing the number students participating in the schools’ federal lunch programs. Before describing the existing constructs, the process used to develop the transformational leadership instrument will be explained.

Development of the Transformational Leadership Instrument

A pilot study was conducted to develop a set of reliable and valid items to measure transformational leadership. The theoretical base from which transformational leadership items were conceived belonged to Bass (1985) and Leithwood (1994). Additional items included in the pilot instrument stemmed from discussions with fellow researchers collaborating on an encompassing larger school study.

Bass and Avolio's (2000) most recent construct of transformational leadership provided a five-dimension framework from which to build our pilot instrument. Twenty items represented the transformational dimensions idealized influence (attributed), idealized influence (behavioral), inspirational motivation, intellectual stimulation, and individual consideration. Eight items originating from Leithwood and Jantzi (1999) were added, where appropriate, to further describe the framework's theorized dimensions. Four final questions were added based on discussions with research colleagues about noticeable gaps in the items describing the five dimensions. In all, 32 items made up the survey used to pilot a transformational leadership instrument. The items were as follows:

Idealized Influence (attributed)
1. Instills pride in me for being associated with him/her
2. Goes beyond self-interest for the good of the group
3. Acts in ways that build my self-respect
4. Displays a sense of power and confidence
5. Symbolizes success and accomplishment within our profession
Idealized Influence (behavioral)
6. Talks about his/her most important values and beliefs
7. Specifies the importance of having a strong sense of purpose
8. Considers the moral and ethical consequences of his/her decision
9. Emphasizes the importance of having a collective sense of mission
10. Promotes an atmosphere of caring and trust among staff
11. Models problem solving techniques that I can readily adapt for my work
12. Practices what he preaches
13. Leads by example

Inspirational Motivation
14. Talks optimistically about the future
15. Expresses his/her confidence that we will achieve our goals
16. Talks enthusiastically about what needs to be accomplished
17. Articulates a compelling vision of the future
18. Gives me a sense of overall purpose
19. Gives me a sense of being important

Intellectual Stimulation
20. Re-examines critical assumptions to question whether they are appropriate
21. Seeks differing perspectives when solving problems
22. Suggests new ways of looking at how we do our jobs
23. Gets me to look at problems from many different angles
24. Facilitates opportunities for staff to learn from each other
25. Is a source of new ideas for my professional learning
26. Stimulates me to think about what I am doing for my students

Individualized Consideration
27. Treats me as an individual rather than just a member of a group
28. Focuses me on developing my strengths
29. Spends time teaching and coaching me
30. Treats each of us as individuals with different needs, abilities, and aspirations
31. Provides moral support by making me feel appreciated
32. Is concerned about my personal problems

Scale Justification

A 6-point Likert scale is employed to follow the format set by Bass and Avolio (2000). Because their scale was the source for most of the items, following their response system seemed most appropriate. Bass and Avolio's (2000) Multifactor Leadership Questionnaire (5x) served as the basis of our transformational leadership
instrument for the following reasons. First, there have been meta-analyses performed on wide ranging applications of the MLQ confirming its validity and utility, and demonstrating transformational leadership’s effects on employee and organizational outcomes (e.g., Lowe et al., 1996). No such meta-analyses exist concerning Leithwood's instruments. Secondly, Bass and Avolio (2000) make available technical information supporting the construction, reliability, and validity of the instrument (Bass & Avolio, 2000). No such report exists for Leithwood's instrument (Jantzi, Personal Communication, 2001). Finally, although Leithwood et al. have conducted the most extensive research in schools concerning transformational leadership (Leithwood & Duke, 1999), other studies have done the same using the MLQ (e.g., Arends, 1993; Barnett et al., 2001; Catanyag, 1985; Gorham, 1992; Ingram, 1997).

Preliminary Review

Before testing the instrument with a sample of practicing teachers, the instrument was reviewed one last time by a panel of three experts (a distinguished professor of educational administration, an instructional leadership researcher and former elementary teacher, and an efficacy researcher and former school administrator). The panel reviewed the scale for format; dimension representation; instruction coherency; and item clarity, appropriateness, and orientation. A few changes were made concerning syntax and grammar, but the scale remained largely intact before being subjected to a pilot test.

The Pilot Test

The data set against which the transformational leadership instrument is tested came from 188 teachers in central Ohio, western New York, and Virginia. Classes at the Ohio State University and The College of William and Mary supplied completed surveys
in Ohio and Virginia. Various teacher center in-services, BOCES workshops, Buffalo State College and school district training programs provided completed teacher surveys from Western New York. Practicing public elementary teachers completed the 106 surveys from New York. The surveys provided by the colleges came from elementary and secondary, public and private teachers.

Results of the Pilot Study

An initial factor analysis, utilizing principal axis extraction and varimax rotation fitting, of all 32 transformational items revealed one underlying dimension. The five dimensions theorized in the most recent Bass literature (Bass & Avolio, 2000) were not realized in our analysis. All items loaded on one factor with loadings of at least .49. Thus, efforts were made to make a more parsimonious instrument.

After sorting the 32 items from the first factor analysis, two items with the highest factor loadings for each of the theorized dimensions (the five “I’s”) were maintained in the model. In one instance, two items of the same class and of equal loadings were analyzed for deletion consideration. “Lead by example” was chosen over “Considers the moral and ethical consequences of his/her actions” because the former item characterized the underlying factor (idealized influence – behavior) better than the latter. Consistency between word and deed and modeling seemed more descriptive of idealized influence than the morality of one’s behavior. Also, it is easier for teachers to perceive a principal “Leading” than it is to see him or her “Considering.” As a result of this balanced selection of the five “I” items, 22 prompts were dropped from the analysis, leaving the following items.
<table>
<thead>
<tr>
<th>Factor Loading</th>
<th>Theorized Dimension</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.80</td>
<td>individualized consideration</td>
<td>Treats me as an individual rather than just a member of a group</td>
</tr>
<tr>
<td>0.80</td>
<td>idealized influence (attributed)</td>
<td>Instills pride in me for being associated with him/her</td>
</tr>
<tr>
<td>0.80</td>
<td>individualized consideration</td>
<td>Provides moral support by making me feel appreciated</td>
</tr>
<tr>
<td>0.79</td>
<td>idealized influence (attributed)</td>
<td>Acts in ways that build my self-respect</td>
</tr>
<tr>
<td>0.76</td>
<td>inspirational motivation</td>
<td>Gives me a sense of being important</td>
</tr>
<tr>
<td>0.75</td>
<td>idealized influence (behavioral)</td>
<td>Practices what he preaches</td>
</tr>
<tr>
<td>0.74</td>
<td>idealized influence (behavioral)</td>
<td>Leads by example</td>
</tr>
<tr>
<td>0.73</td>
<td>intellectual stimulation</td>
<td>Seeks differing perspectives when solving problems</td>
</tr>
<tr>
<td>0.67</td>
<td>intellectual stimulation</td>
<td>Helps staff to learn from each other</td>
</tr>
<tr>
<td>0.59</td>
<td>inspirational motivation</td>
<td>Gives me a sense of overall purpose</td>
</tr>
</tbody>
</table>

Table 3.1: Factor Loadings of Items Representing the Four Dimensions of Transformational Leadership
In preparing the instrument for the major study, thought was given to the
differences between the pilot study and the larger one for which the pilot was employed.
For the most part, very few respondents reflected perceptions of the same school leader.
At one point, 75 pilot surveys were collected representing 42 school leaders in New
York. It is assumed a very similar ratio held true for the whole pilot sample. The nature
and location of participating graduate classes provide little reason to suspect otherwise.
The accuracy of such few observations concerning leader behavior may not be the same
as when more observations are made of the same principal, is planned for the larger
study. The conceptual dimensions of transformational leadership could still unfold under
larger study conditions, in which a ratio of at least seven randomly selected teachers per
principal would be maintained. Therefore, more items were added to the survey
instrument to help measure the five theorized dimensions.

Conceptual and empirical reasoning guided the supplemental items added to the
existing ten. Also, conceptual considerations caused a few changes in the dimensional
classification of some items. For example, Bass and Avolio (2000) consider "Specifies
the importance of having a strong sense of purpose" as an idealized influence behavior,
whereas it could be argued that conveying a strong sense of purpose could be
inspirationally motivating, as well. Such alterations were made in a few cases.
Additionally, the difference between attributed and behavioral idealized influence did not
appear critical, especially in light of the pilot study results. Hence, both theoretical
dimensions folded together in this study. A chart depicting the final items and their
factor loadings are displayed in Appendix A. The organization of
transformational leadership items after factor and conceptual analysis is as follows:
<table>
<thead>
<tr>
<th>Transformational Leadership Items</th>
<th>Description</th>
</tr>
</thead>
</table>
| Idealized Influence              | Instills pride in me for being associated with him/her  
                                  | Acts in ways that build my self-respect  
                                  | Practices what he preaches  
                                  | Leads by example  
                                  | Goes beyond self-interest for the good of the group  
                                  | Promotes an atmosphere of caring and trust among staff |
| Inspirational Motivation         | Specifies the importance of having a strong purpose  
                                  | Gives me a sense of being important  
                                  | Gives me a sense of overall purpose  
                                  | Stimulates me to think about what I am doing for students |
| Individualized Consideration      | Is concerned about my personal problems  
                                  | Provides moral support by making me feel appreciated  
                                  | Treats me as an individual than just a member of a group |
| Intellectual Stimulation         | Seeks differing perspectives when solving problems  
                                  | Is a source of new ideas for my professional learning  
                                  | Models problem solving techniques that I can readily adapt to my work  
                                  | Helps staff to learn from each other  
                                  | Suggests new ways of looking at how we do our jobs. |

Table 3.2: Transformational Leadership Item Description
Scale Validity and Reliability Indicators

The grade- and public- school status for more than half of the sample respondents was tracked. For comparison purposes, factor analyses were performed on the total sample set (n=188) and a subset of just known elementary teachers (n=106) to gauge any differences, if any, concerning the perceptions of strictly elementary teachers. The comparison between respondent groups follows for each transformation scale developed throughout the pilot study:
### Table 3.3: Respondent Group Comparison

<table>
<thead>
<tr>
<th>Scale</th>
<th>Total Variance explained</th>
<th>Reliability Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>32-Item Scale</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All respondents (n=188)</td>
<td>71.9%</td>
<td>.987</td>
</tr>
<tr>
<td>Known Elementary Teachers (n=106)</td>
<td>72.0%</td>
<td>.987</td>
</tr>
<tr>
<td><strong>18-Item Scale</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All respondents (n=188)</td>
<td>76.65%</td>
<td>.981</td>
</tr>
<tr>
<td>Known Elementary Teachers (n=106)</td>
<td>77.60%</td>
<td>.982</td>
</tr>
<tr>
<td><strong>10-Item Scale</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All respondents (n=188)</td>
<td>80.63%</td>
<td>.973</td>
</tr>
<tr>
<td>Known Elementary Teachers (n=106)</td>
<td>82.05%</td>
<td>.976</td>
</tr>
</tbody>
</table>

As noted in the charts, trimming the transformational scale to 18 items from 32 does not sacrifice the reliability of the instrument in any meaningful manner. Importantly, the reliability coefficient exceeds the .80 criterion (Bryman & Cramer, 134).
1997) for multiple item scales, suggesting the scale items are consistently referring to the same underlying concept. Moreover, the elementary teacher perceptions of their principal's behavior did not seem to vary considerably from the observations of teachers in general.

The only modification made to the transformational leadership scale after the pilot study was the addition of another individualized consideration item. Believing the individualized consideration scale might be light on items and lacking a reference to the principal's knowledge about subordinate interests and talents, the following item was added: "Doesn't know my talents and interests" (reverse scored). Leithwood and Jantzi (1999) has a similar item - "Is aware of my unique needs and expertise" (p. 475), validating our notion. The 18 other items maintained their standing on the scale, creating a 4-6-item subscale for each of the 4 "I's."

Criterion-Related Validity

Empirical validation of the transformational leadership scale is limited to the other constructs measured at the time of the pilot study. An instructional leadership scale developed by a colleague served as the criterion against which the transformational scale was validated. Instructional leadership conceptions developed by Hallinger and Murphy (1985) and Murphy (1990) served as the instrument's theoretical foundation. The scale was tested with the same sample of teachers whom responded to the transformational leadership pilot survey. Factor analysis revealed three dimensions emerging from the instructional leadership survey pilot that were consistent with the findings of Hallinger and Murphy (1985) and Murphy (1990): developing mission and goals (alpha=.92), managing the educational production function (alpha=.93), and
promoting professional development (alpha=.90). It seems reasonable to suggest that a transformational leader who endeavors to improve school performance tends to the instructional behaviors of teachers. Indeed, Leithwood (1994) states "[instructional leadership behaviors] are necessary elements of any reform strategy likely to have payoff for students" (p. 500). However, a limited focus on instructional behaviors jeopardizes change programs. Thus, instructional leadership is a necessary, but not a sufficient behavior for real organizational growth (1994). Consequently, one would expect to see some positive correlation between transformational and leadership behaviors. To be sure, a high correlation existed between the transformational behavior scale and the three instructional leadership factors (r > .75 for each correlation, p < .01).

Socioeconomic Status Construct

A measure of SES exists for the schools by way of their federal school lunch program participation rates. Students qualify for free and reduced meals according to their families' household income, a reasonable approximation of a student's socioeconomic status. With the school being the unit of analysis, the proportion of students entitled to free and reduced lunch determines the aggregate SES measure per school participating in the research. One assumption implied by using this SES measure is that most students who qualify for the federal lunch program come forward with completed applications.

Collective Efficacy Construct

A 12-item likert-type scale developed by Goddard (2001a) serves as the collective efficacy instrument in this study. The construct's conceptual factors, group competency beliefs and perceptions of task analyses, fold into one into one dimension whose value
derives from the mean of its item scores. Examples of items found on each of the
generously- and negatively-oriented conceptual factors include "Teachers in this school
are able to get through to difficult students," "If a child doesn’t want to learn teachers
here give up," "These students come to school ready to learn," and "Students here just
aren’t motivated to learn" (p. 22). The reported reliability score for this instrument is
alpha=.94 with an explained variance of 64%. Goddard (2001a) provides support for the
criterion- and conceptual validity of this scale.

Math and Reading Achievement Constructs

The Ohio 4th grade proficiency exams measure math and reading achievement in
this study. Experts involved in the multi-stage, committee-oriented review process
determine the tests’ content validity by verifying the developmental appropriateness,
content-suitability, and bias-free status of each test item. Reliability scores for the 2001
administration of the math and reading proficiency exams were 0.87 and 0.84
(Cronbach’s alpha), respectively. The math exam contains 40 questions - 30 multiple
choice, 8 short answer, and two extended response. The reading exam consists of 30
questions – roughly 22 multiple choice, 6 short answer, and 2 two extended response.
Each type of question is worth a set number of points and the sum of the points earned is
the student’s raw score. A scaled score for each student is determined given the student’s
raw score. The scaled score is compared against a predetermined proficiency level to
establish a student’s proficiency status for that exam. The unit of analysis for this study
is the school, so consequently a school’s mean scaled score for each exam provides the
achievement measures for a given school.
Data Collection

Transformational leadership and collective efficacy data were collected from each school during school-site visits scheduled during regular monthly faculty meetings, after securing permission from principals for our brief on-site study. A skilled researcher disseminated the survey instruments in an alternating fashion so that one set of teachers chosen at random responded to the leadership items while another independent group of teachers answered collective efficacy questions. Following this protocol insures practical and methodological independence between constructs. Because the unit of analysis was the school, all data were aggregated to the school level, making this procedure possible. All the teachers were assured anonymity. School SES and achievement data were secured from the Ohio Department of Education. As noted earlier, SES’s construct derived from the proportion of students whom participated in the federal lunch program at each school, while achievement measures stemmed from the percentage of students passing each test at the schools.

Data Analysis

Teachers’ responses were aggregated to the school level using SPSS to calculate the means, standard deviations, and reliability coefficients (Cronbach’s alpha) for all the scales. First, intercorrelations were generated to test the bivariate relationships, and then the model was tested using multivariate statistics. Exploratory and confirmatory factor analyses were employed to verify the number of dimensions observed and to assess the accuracy of our conceptual model under these conditions.

LISREL 8.5 (Scientific Software International, 2000) provided the technical means for employing Structural Equation Modeling (SEM) techniques in this study.
This statistical procedure, for example, estimates the direct effects of leadership behavior on student achievement and the indirect effects of leadership behavior on student achievement through collective efficacy, while controlling for SES and prior achievement. SEM allows theoretical models describing conceptual relationships, represented by implied covariance matrices, to be compared against the covariance matrices calculated from the relationships found in the data. The default estimation algorithm used to gauge the strength and direction of relationships between variables is maxim likelihood.

Several scholarly pieces call for this type of analysis. Leithwood and Jantzi (2000) report that the study of leadership’s indirect effects on student outcomes requires researchers to “identify those alterable conditions likely to have direct effects on students, and to inquire about the nature and strength of the relationship between them and leadership” (p. 114). They inform, from a reading of Hallinger and Heck (1996), that only one mediating variable, school goals, reliably provided leaders with indirect effects on student achievement. Indeed, SEM analysis provides a means for determining if collective efficacy supplies another venue for leaders to influence student achievement.

Hallinger and Heck (1996) cite a lack of research exploring antecedent influences on leadership while studying its effect on school outcomes. They claim SES and prior achievement antecedents were often used to control effects on student achievement, but not on leadership. Again, SEM and relevant data allow this study to help nourish that line of research, as Hallinger and Heck (1996) proclaim,
The greatest progress in the field will yield from research that places the principal in the context of school and its environment ... The theoretical groundwork has been laid for such studies, and analytical methods appropriate to studying this type of comprehensive framework (e.g., structural equation modeling) are available for use in this domain. (p. 34)

Finally, Heck and Hallinger (1999) restate their calls for more comprehensive models of leadership inquiry. Most studies of leadership have relied on finding direct effects using simple statistical methods, such as descriptive measures, correlational analyses, and tests of mean differences between principal leadership types, perhaps explaining some of the inconclusive or weak findings (Leithwood & Jantzi, 2000). However, even regression analyses, a more advanced technique, lacks the sophistication to entirely expose the indirect effects of leadership on school outcomes (Heck & Hallinger, 1999). SEM stands as one of the few technologies thus far to adequately estimate direct and indirect relationships school leadership may have on school achievement. Heck and Hallinger assert,

SEM can examine a variety of theoretical models including those with latent (unobserved) variables, direct and indirect effects, ... reciprocal effects, data collected at several points in time, and data that are hierarchical in nature. SEM techniques, when combined with a comprehensive conceptualization of leadership, have produced consistent findings across countries. (1999, p. 152)
CHAPTER 4
RESULTS

This chapter conveys the analysis of the data for this study. The sample of schools is described, using demographic data provided by the Ohio Department of Education, the transformational leadership variables was verified using factor analysis techniques, and the research hypotheses are tested using correlational analyses and structural equation modeling.

School Sample Description

One hundred and forty-six elementary schools participated in the study. Although their composition is not a random one, efforts were made to draw schools from urban, suburban, and rural settings. As such, relevant elementary schools from 33 counties comprise the sample. Ohio has 44%, 23%, and 33% of its elementary schools in urban, suburban, and rural settings, respectively. Correspondingly, the study’s schools distribute across respective urban descriptions at a rate of 36%, 37%, and 27%. Further comparisons to the state utilize mean wealth, student enrollment, teacher experience, teacher salary, and teaching staff measures given in Table 4.1. Schools that administer grade four proficiency tests constitute relevant state schools.
The grade organization can be another characteristic differentiating elementary schools. The grade configuration of the schools in the study heavily favored a K-5 format, as seen in Table 4.2. Percentages for the state will not sum to 100 because several other grade configurations exist in relevant elementary schools.
Table 4.2: Grade Configurations

<table>
<thead>
<tr>
<th>Grade Format</th>
<th>Sample</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grades 1-5</td>
<td>2.05%</td>
<td>2.25%</td>
</tr>
<tr>
<td>Grades 1-6</td>
<td>0.68%</td>
<td>1.38%</td>
</tr>
<tr>
<td>Grades 4-6</td>
<td>0.68%</td>
<td>1.69%</td>
</tr>
<tr>
<td>Grades K-4</td>
<td>7.53%</td>
<td>13.90%</td>
</tr>
<tr>
<td>Grades K-5</td>
<td>59.90%</td>
<td>38.78%</td>
</tr>
<tr>
<td>Grades K-6</td>
<td>30.14%</td>
<td>27.54%</td>
</tr>
</tbody>
</table>

Although certain characteristics of the sample may resemble those of the state’s elementary schools, inferences made from our study should be cautiously applied to Ohio elementary schools in general.

Four thousand, sixty-nine teachers responded to the surveys. The vast majority (95.79%) of teachers gave indication of their gender, and females largely comprised that group (92.56%).

Transformational Leadership Construct Development

Variable determination comprises the next section of this study. Factor analyses, both confirmatory and exploratory, guided the construction of the transformational
leadership variable. The existence of the transformational structure manifested in this study's data can be detected using a combination of literature and exploratory and confirmatory factor analyses.

Exploratory factor analysis (EFA) aids in the detection of possible factor dimensions inherent to the data. Although such analysis may lead researchers to possible construct structures, exploratory factor analysis makes use of little to no previous theory about the relationships between the observed variables and the latent variables. As such, EFA is a theory generating strategy. Once relationships between and among observed and latent variables are theorized, confirmatory factor analysis (CFA) can be utilized to test the degree to which this theory holds up against empirical data. Thus, CFA is a theory testing strategy (Stevens, 1996).

CFA is one application of structural equation modeling. By employing structural equation modeling techniques using LISREL 8.50 (Joreskog & Sorbom, 2000), hypotheses can be tested against a set of data for their empirical fit. Constructing a structural equation model infers that one wishes to relate various concepts, latent and/or observable, to test the direction and strength of their associations. Sound theory should drive the construction of the SEM, giving reason for the relationships to be tested. In this case, the four hypothesized dimensions, as manifested by their respective set of indicators, are represented by a covariance matrix and are fitted to the collected data, which is also represented by a covariance matrix. The null hypothesis in such testing is $S = \Sigma(\theta)$, where sigma is the data matrix and sigma theta is the hypothesized relationships matrix. As one might suspect, the test statistic, minimum fit chi-square, needs to be a value such that it fails to reject the null hypothesis. One advantage of using LISREL
rests with its abundance of model fit indices (Byrne, 1998). Several will be utilized, in addition to the minimum fit chi-square statistic, to gain a comprehensive assessment of the empirical validity of several models.

The following measures will be considered when assessing the compatibility of theoretical factor and path model structures to the collected data: a) the chi-square minimum fit function test statistic and accompanying probability value, b) supplemental, and more “practical” global model fit indices (Bollen, 1989; Byrne, 1998), and c) the signs, magnitudes, and statistical significances of model components. The specifics of such reporting elements will be explained in more detail upon the first application of confirmatory factor analysis, which is an assessment of Bass’s (1985) four-factor structure of transformational leadership.

Four-Factor Solution

Although much research contends transformational leadership behaviors exhibit a different number of dimensions (e.g., Barnett, McCormick, & Conners, 1999; Leithwood, 1994, Lowe et al., 1996), this study hypothesized a four-factor structure, based largely on Bass's (1985) work. As such, the testing of this structure’s existence in the data was first consideration. Perceptions of school leadership came from 19 items describing the four theoretical dimensions of transformational leadership: idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration.

Confirmatory factor analysis (CFA) tests the extent to which observations of leadership behaviors describes the four "I's," intellectual stimulation, inspirational motivation, individualized consideration, and idealized influence. In this first CFA, the 19 indicators describing the four I's comprise the first transformational model to be
tested. The school-level averages of teacher responses to the nineteen indicators form the data set. Thus, 146 observations of elementary school leadership constitute the n. Appendix B illustrates a conceptual diagram depicting the four-factor model. An explanation of the model fit measures follows.

LISREL 8.5 defaults to the maximum likelihood fitting function for estimating a model’s parameters (Joreskog & Sorbom, 2000). Maximum likelihood generates an estimated chi-square statistic, and as such provides a probability value indicative of the overall fit of the model since the distribution of chi-square values is known (Bollen, 1989). The value of the chi-square statistic, in relation to the model’s degrees of freedom, determines whether the fitted residuals, resulting from $\Sigma - \Sigma(\theta)$, deviate from the population values of zero. The differences between the corresponding elements of the implied covariance matrix (theory) and the observed covariance matrix (reality) constitute the matrix of residuals. The lower the chi-square value gained from the accumulation of these residuals, the more likely the theory represents the reality.

The four-factor solution yields a relatively large chi-square value $x^2(146, N = 146) = 551.66$, $p = .000$, thus rejecting the null hypothesis, $s = \Sigma(\theta)$. This construct of transformational leadership as a four-factor model was not supported; however, the chi-square test is a very rigorous standard for model fit testing. Bollen (1989) describes the supposition undergirding the chi-square test:
[One Assumption] affecting the chi-square approximation is that it assumes that 
\( \sigma = \sigma_{\theta} \) is exactly true. In virtually all cases, we do not expect to have 
a completely accurate description of reality. The goal is more modest. If the 
model that leads to \( \sigma_{\theta} \) helps us to understand the relation between 
variables and does a "reasonable" job of matching the data, we may judge it as 
partially validated. The assumption that we have identified the exact process 
generating the data would be accepted. Yet the chi-square test derives from a 
comparison of the hypothesized model \( H_0 \) to \( H_1 \), a model of perfect fit. A perfect 
fit may be an inappropriate standard, and a high chi-square estimate may indicate 
what we already know - that \( H_0 \) holds approximately, not perfectly. (p. 268)

Hence, practical measures of model fit are considered in a comprehensive 
assessment of this factor structure. Sample size and model complexity are known to 
impact differentially on certain fit statistics (Browne and Cudeck, 1993; Gerbing and 
Anderson, 1993; Tanaka, 1993). Thus, if a model adequately represents the data, the 
thresholds of several fit indices should be satisfied (Tallon, Kraemer, Gurbaxani, & 
Mooney, 1997). The Root Mean Squared Error of Approximation (RMSEA), Root Mean 
Squared Residual (RMR), Adjusted Goodness of Fit Index (AGFI), Expected Cross-
Validation Index (ECVI), Hoelter's Critical N, and the Incremental Fit Index (IFI), 
representing different families of fit statistics, will all be employed for their model 
assessments, the preponderance of which should lend satisfactory model determination 
more clarity and validity.

Byrne (1998) claims that the RMSEA is one of the most instructive criteria in 
theoretical model testing. This model fit measure gauges the degree to which the 
theoretical model deviates from the same model with optimally chosen parameter values. 
The extent to which the theoretical model deviates from a parallel model with ideal 
parameters is conveyed in \textit{per degrees of freedom}. Therefore, this measurement of fit 
reflects consideration for model complexity. Byrne (1998) suggests RMSEA values less
than 0.05 signifies a good fitting model, as values closer to zero approach better fits. The RMSEA of the four-factor transformational leadership model is 0.15, denoting a subpar fit.

The RMR is a measurement of the average residual from fitting the theoretical model to the sample covariance matrix. The smaller the average standardized residual size, the better the model fit the data. In order to bring uniformity to residual interpretation, differences between measured and hypothetical correlations may be calculated (Bollen, 1989). The residuals comprising a standardized RMR’s matrix are the measured deviations between the corresponding elements of the sample’s intercorrelation and the hypothesized intercorrelation matrices. The standardized RMR measure represents the average quantity across the intercorrelation residuals. Standardized RMR values less than 0.05 indicate well-fitting models, as the range of possible values for the standardized RMR is from zero to one (Byrne, 1998). The four-factor transformational leadership model yielded a 0.045, suggesting satisfactory fit.

The AGFI yields the relative amount of variance explained by the model (Byrne, 1998). Pillai, Schriesheim, and Williams (1999) describe the AGFI as a fit measurement that “reflects the relative amount of variances and covariances in the sample matrix to the variances predicted by the model matrix,” adjusted for the degrees of freedom relative to number of variables employed (p. 912). The range of possible values for the AGFI is from zero to one, with scores greater than 0.90 indicative of good model fits (Byrne, 1998). The AGFI generated from this sample of data does not support a four-factor solution of transformational leadership (AGFI=0.60).
Fit measures exist that gauge the extent to which the hypothetical model structure cross-validates across similar-sized samples. The ECVI is one such fit measure, as it indexes the relative quality of the hypothetical model against quality indices for fully saturated and fully independent representations using the same hypothesized constructs and similar sized samples. Because the ECVI coefficients can take on any one of an infinite array of values, an anticipated range of acceptable values is not applicable to this goodness of fit measure (Byrne, 1998). However, the lowest ECVI index corresponds with the most likely model type to be replicated in future like-sized samples (1998). In the four-factor model, the transformational leadership construct fails to exhibit the lowest value ECVI (i.e. 4.86 as opposed to 2.62 for the saturated model).

Hoelter’s Critical N is the next gauge of model quality generated via inspection of sample size adequacy. This qualitatively different goodness-of-fit measure stems from efforts to calculate a model fit indicator not dependent on sample size. More precisely, the statistic reflects the minimum number of observations necessary to realize a satisfactory model fit using the chi-square test. Conventional guidelines suggest a Critical N value of 200 or more implies satisfactory model fit (Hoelter, 1983). Transformational leadership, as our tested four-factor solution, generated a Critical N of 50.59, falling short of the 200 N guideline.

The last model fit gauge utilized in this study is the IFI. The family of incremental fit indices compares the hypothetical model against a baseline model, often being the null model where all parameters are set to zero (Byrne, 1998). Bollen (1989) noticed that the "classical" practical fit indice of the 1980's, the incremental NFI (Bentler & Bonett, 1980), had been insensitive to model complexity and overly dependent on
sample size. Thus, he incorporated the number of free parameters in the model, represented by the degrees of freedom associated, and the sample size into the NFI, creating a new IFI (Bollen, 1989). Tallon, Kraemer, Gurbaxani, & Mooney (1997) advocate an IFI greater than 0.95 for the model to be satisfactorily representing the data. The four-factor solution yields an IFI of 0.91, falling short of the suggested guideline.

The preponderance of model fit measurements indicates a bad correspondence between the four-factor model and the sample data. As such, a review of the model components is unnecessary. However, when global fit indices suggest a good fit, the quality of the model can still be questioned by statistically insignificant parameter estimates, negative variances, minimal squared multiple correlations ($R^2$'s) and inexplicable parameter sign directions (Bollen, 1989; Breckler, 1990). Although these component qualities will not be reviewed for this model, other models revealing good fit, as indicated by global fit measures, will necessitate a review of the components for further validation of model quality.

With ample evidence supporting lack of fit for the four-factor solution, other leadership model structures warrant consideration. Recalling the pilot study of the transformational leadership realized a one-factor solution, attention now turns to the fit between a one-factor transformational leadership model and the data of this study.

*One Factor Solution*

A one-factor solution is not without precedent in the literature of transformational leadership constructs using the MLQ. Arends (1993) observed a one-factor model and Pillai et al. (1999) utilized a composite measure of transformational leadership, as well. Bass and Avolio (2000) recognize that one of the criticisms of Bass's (1985) leadership
model is that the four dimensions theoretically supporting the construct have not been consistently realized in research and when they have been, their intercorrelations are extremely high. Such high correlations can sway researchers to collapse the dimensions into one global construct to avoid problems with discriminant validity between the 4 I's. Previous experiences by other researchers with the MLQ justify subjecting a one-factor model to a confirmatory factor analysis. A conceptual diagram of this model is shown in Appendix C. The structure’s factor loadings appear in Appendix D.

The chi-square minimum fit function for the one-factor solution returned a statistically significant test value, signifying the model does not fit the data. Moreover, five of the six more practical global fit measurements failed to support this model structure. The standardized RMR was the only gauge to realize a satisfactory value, and that value did not have much wriggle room (0.047). Again, with such overwhelming evidence not supporting the one-factor structure, little reason exists for analyzing the component values. Despite the fact this structure yields an alpha reliability of 0.9831, the discrepancy between internal consistency measures and factor structure indices has been noted before in scholarship.

Pillai et al. (1999) witnessed comparably high reliability indices in their factorization of transformational leadership, yet also observed the same unfavorable global fit indices in their confirmatory factor analysis. Thus, these researchers settled for a global measure of transformational leadership, citing that "in the past, the MLQ has consistently yielded high reliabilities but its factor structure has also been problematic (Bycio, Hackett, & Allen, 1995; Howell & Avolio, 1993; Yammarino & Dubinsky, 1994; Yammarino, Spangler, & Bass, 1993)" (p. 910).
At this point, a firm factor structure based on theoretical foundations (i.e., four dimensions, Bass, 1985) or on direction implied via the pilot study (i.e., one dimension) bears little empirical support. As such, exploratory factor analysis (EFA) offers potential for revealing a leadership construct supported by the data. Stevens (2002) distinguishes exploratory factor analysis from confirmatory factor analysis by the a priori status of factor structure and indicator-factor relationship determinations. With EFA, both determinations are left open-ended for the sake of having the data guide these considerations.

Two Factor Solution

One of the most widely used criteria for determining the number of factors evident in a construct is the eigenvalue-greater-than-one rule, or Kaiser's criterion (Bryman, 1997; Stevens, 2002). Such strategy will help discern the salient dimensions within the research data. Another consideration beyond the factor determination criterion is the factor rotation decision. Analytical tension exists in reconciling the need to orient the factors so as best to represent reality or to fix them so as to ease interpretation. Rennie (1997) asserts that rotating factors orthogonally or obliquely makes little difference in realized factor loadings but makes appreciable difference in factor interpretation. Her reasoning is as follows:

It is suggested that regardless of the magnitude of the correlation among the factors that it is unnecessary to do oblique rotation in addition to orthogonal rotation. The correlation among the factors yields results that are more difficult to interpret than results of an orthogonal rotation. Moreover, the slight difference between results of an oblique rotation and an orthogonal is virtually insignificant. Thus, it becomes much easier and simpler to rotate the factors orthogonally. (p. 10)
EFA revealed two factors that had eigenvalues greater than 1. The first factor generated an eigenvalue of 14.645 and explained 76.169%. The second factor had an eigenvalue of 1.086 and explained 4.413%. The first nine indicators loaded onto the first factor and the last ten loaded onto the second dimension. Focusing attention on to the first factor, the nine indicators, in order of loading magnitude, are as follows:
<table>
<thead>
<tr>
<th>Factor Loading</th>
<th>Item</th>
<th>Theorized Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.881</td>
<td>Suggests new ways of looking at how we do our jobs.</td>
<td>Intellectual Stimulation</td>
</tr>
<tr>
<td>0.818</td>
<td>Specifies the importance of having a strong sense of purpose.</td>
<td>Inspirational Motivation</td>
</tr>
<tr>
<td>0.801</td>
<td>Is a source of new ideas for my professional learning.</td>
<td>Intellectual Stimulation</td>
</tr>
<tr>
<td>0.782</td>
<td>Stimulates me to think about what I am doing for my students.</td>
<td>Inspirational Motivation</td>
</tr>
<tr>
<td>0.754</td>
<td>Models problem solving techniques that I can readily adapt.</td>
<td>Intellectual Stimulation</td>
</tr>
<tr>
<td>0.740</td>
<td>Leads by example.</td>
<td>Indealized Influence</td>
</tr>
<tr>
<td>0.708</td>
<td>Helps staff to learn from each other.</td>
<td>Intellectual Stimulation</td>
</tr>
<tr>
<td>0.694</td>
<td>Gives me a sense of overall purpose.</td>
<td>Inspirational Motivation</td>
</tr>
<tr>
<td>0.665</td>
<td>Practices what he or she preaches.</td>
<td>Idealized Influence</td>
</tr>
</tbody>
</table>

Table 4.3: Items Constituting the First Factor of the Exploratory Factor Analysis

The highest loading items on each factor provide the strongest notions about the underlying concepts constituting the dimension (Bryman, 1997). Inspection of the highest-ranking indicators suggests that intellectual stimulation behaviors may heavily...
influence the first factor. At first glance, it appears that only four of the top seven indicators qualify as intellectual stimulation items. However, careful consideration of the indicators "Stimulates me to think about what I am doing for my students" and "Leads by example" may sway the balance to six out of the top seven.

Although the item "Stimulates me to think about what I am doing for my students" falls under the Inspirational Motivation dimension, two reasons suggest why it may fit better within the intellectual stimulation classification. First of all, the word "stimulate" stands as the active verb, evoking compelling consideration of its reclassification to the category bearing the same word. Second, the item was moved from the intellectual stimulation dimension during the pilot study because conceptually it appeared to fit within the motivation domain, as stimulation is synonymous with motivation. Such a move was congruent with efforts to balance the item counts within each of the four "I's." Leithwood and Jantzi (1999) originally cast this item as an intellectual stimulation prompt. The empirical evidence at hand indicates they were correct in doing so.

"Leads by example" was an item borne out of discussion among the research team closely associated with this project. Originally conceived as a behavior common to charismatic leaders, the team felt "Leads by example," or modeling, best described the idealized influence dimension of transformational leaders. However, modeling has also long been advocated as an effective means for influencing the behavior of others (e.g., social learning theory, Bandura, 1971). In organizational contexts, modeling is a most salient form of influence among staff (Scarborough, 2001). As compared to other forms of academic efficacy information, women responded to the modeling of others with a
much greater sensitivity than men do in rigorous academic concerns, such as math, science, and technology (Zeldon & Parjares, 2000). Because most of the subjects in this sample are female, respondents may be more apt to view "Leads by example" as a form of intellectual stimulation than as a kind of idealized influence. The first factor was called intellectual stimulation and explained 76% of the variance in this factor analysis. A review of the second factor follows its item loadings shown in Table 4.4.
<table>
<thead>
<tr>
<th>Factor Loading</th>
<th>Item</th>
<th>Theorized Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.850</td>
<td>Treats me as an individual rather than just a member of a group</td>
<td>Individualized Consideration</td>
</tr>
<tr>
<td>0.814</td>
<td>Provides moral support by making me feel appreciated</td>
<td>Individualized Consideration</td>
</tr>
<tr>
<td>0.754</td>
<td>Promotes an atmosphere of caring and trust among staff.</td>
<td>Idealized influence</td>
</tr>
<tr>
<td>0.748</td>
<td>Acts in ways that build my self-respect</td>
<td>Idealized influence</td>
</tr>
<tr>
<td>0.748</td>
<td>Is concerned about my personal problems.</td>
<td>Individualized Consideration</td>
</tr>
<tr>
<td>0.711</td>
<td>Instills pride in me for being associated with him/her.</td>
<td>Idealized influence</td>
</tr>
<tr>
<td>0.703</td>
<td>Goes beyond self-interest for the good of the group.</td>
<td>Idealized influence</td>
</tr>
<tr>
<td>0.683</td>
<td>Gives me a sense of being unimportant.</td>
<td>Individualized Consideration</td>
</tr>
<tr>
<td>0.649</td>
<td>Seeks differing perspectives when solving problems.</td>
<td>Intellectual Stimulation.</td>
</tr>
<tr>
<td>0.588</td>
<td>Doesn't know my talents and interests.</td>
<td>Individualized Consideration</td>
</tr>
</tbody>
</table>

Table 4.4: Items Constituting the Second Factor of the Exploratory Factor Analysis
Inspection of the second factor items suggests that the convergence of like
dimension prompts is not like the first factor’s analysis, given the spread of theorized
dimensions and cross loadings. Certainly, the first two items indicate an individualized
consideration orientation; however an interpretation of the other eight items is
confounded by the presence of three other types of factors. The dual loadings of many of
the items during this factor analysis further complicate the interpretation of these
dimensions. Ho (2000) and Bryman and Cramer (1997) offer variable elimination as one
way to handle significant cross loadings, an option congruent with one major purpose of
factor analysis – to decrease the number of variables considered in a study and achieve
parsimony (Rennie, 1997).

Appendix E exhibits a principal axis factoring, with a varimax rotation, of the
items remaining after the indicators that loaded on both factors at weights greater than
.500 were eliminated. Setting the criterion standard at a level at which no indicators
loaded onto more than one factor follows the direction of Bryman and Cramer (1997).
Removing ten items that cross-loaded yielded a factor structure displaying dimensions
with much more homogeneity. Individualized consideration and intellectual stimulation
are the respective factors in this more orderly solution. However, attaining such a clean
factor structure using a rotation technique that assumes no factor intercorrelations causes
one to search for ways to validate these findings, for few dimensions constituting a
higher-order factor have such a null interrelationship (Raven, 1994).

Some researchers show skepticism about the validity of varimax-like rotation
techniques. Pedhazur and Schmelkin (1991) express such doubt, writing, “…we believe
that such solutions are, in most instances, naïve, unrealistic, portrayals of behavioral
Therefore, it may be wise to validate the transformational leadership factor structure found thus far by performing factor analysis with another rotation technique that allows the factors to correlate. An oblique rotation of factors would support the MLQ literature that contends the four “I” subscales highly correlate (Bass & Avolio, 2000; Yukl, 1998). Furthermore, such careful thought would help stem the tide of many researchers’ rotation strategies being limited to the default option of their particular statistical software package, as suggested by Hair, Anderson, Tatham, and Black (1995). Hence, the nine-items remaining after the elimination of the cross-loaded indicators were subjected to factor analysis using oblique rotation.

Appendix F displays the results of the new analysis. Two factor matrices need review when performing oblique rotations: the pattern matrix and the structure matrix (Bryman & Cramer, 1997). The pattern matrix reveals the individual item relationships, in terms of a “beta weight,” with their corresponding factors, controlling for all the other indicators. The structure matrix displays the bivariate relationships between each indicator and factor (Stevens, 2002). The data provided by the oblique factor analysis validates the factor structure suggested by the orthogonal factor examination. The “beta weights” provided in the pattern matrix, a set of values often used for figuring factor scores, relate the indicators to the same factors as in the factor analysis with varimax rotation. The Pearson correlation coefficients exhibited in the structure matrix indicate the stronger coefficients for each variable are linked with their respective factors suggested previously. Finally, the two factors are shown to correlate at $r = .757$. The first factor, individualized consideration, explains 68.27% of the variance, and the second factor, intellectual stimulation, explains 8.25 % of the variance.
Explicitly, the items comprising each subscale are as follows.

<table>
<thead>
<tr>
<th>Factor Loading</th>
<th>Item</th>
<th>Theorized Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.851</td>
<td>Treats me as an individual rather than just a member of a group.</td>
<td>Individualized consid.</td>
</tr>
<tr>
<td>0.813</td>
<td>Provides moral support by making me feel appreciated.</td>
<td>Individualized consid.</td>
</tr>
<tr>
<td>0.760</td>
<td>Is concerned about my personal problems.</td>
<td>Individualized consid.</td>
</tr>
<tr>
<td>0.691</td>
<td>Gives me a sense of being unimportant. (Rev. scored)</td>
<td>Inspirational Motivat.</td>
</tr>
<tr>
<td>0.617</td>
<td>Doesn't know my talents and interests. (Rev. scored)</td>
<td>Individualized consid.</td>
</tr>
</tbody>
</table>

Factor 2:

<table>
<thead>
<tr>
<th>Factor Loading</th>
<th>Item</th>
<th>Theorized Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.873</td>
<td>Suggests new ways of looking at how we do our jobs.</td>
<td>Intellectual Stimulat.</td>
</tr>
<tr>
<td>0.813</td>
<td>Is a source of new ideas for my professional learning.</td>
<td>Intellectual Stimulat.</td>
</tr>
<tr>
<td>0.806</td>
<td>Specifies the importance of having a strong sense of purpose.</td>
<td>Inspirational Motivat.</td>
</tr>
<tr>
<td>0.806</td>
<td>Stimulates me to think about what I am doing for my students.</td>
<td>Intellectual Stimulat.</td>
</tr>
</tbody>
</table>

Table 4.5: Items Comprising the Two Factors of Transformational Leadership
All the items load on their theorized dimension, except for two. “Gives me a sense of being unimportant” was another product of the research team’s discussions. It was originally classified as an inspirational motivation indicator, based on the notion that personal motivation may come from feelings of personal importance. Upon reflection, this item connotes a clear image of leader attention to the personal subordinate. The word “me” is clearly the object of the leader’s provision of importance, indicating, perhaps, a bias toward individualized consideration. Inspirational motivation items may need to bring the subordinate’s focus outside of oneself, beyond matters of focused personal attention, and towards more spirited images, goals, and futures.

The reason for the second indicator, “Specifies the importance of having a strong sense of purpose,” to not load as theorized is more elusive. However, the item’s classification as an intellectual stimulation indicator may not be totally inexplicable. Though it was originally framed as an idealized influence indicator (Bass & Avolio, 2000), perhaps specifying a critical reason for beholding purpose in teaching is cerebrally uplifting to the respondents. Citing specific and meaningful grounds for maintaining purpose may take cognitively charged insight amidst a profession that is often described with clichés and grandiose purposes for its existence. Moreover, the standards and assessments movement now afoot may cause educators to believe uniformity colors practice, possibly dulling one’s purpose for teaching. A principal who is able to sharpen a teacher’s focus through highlighting specific significances of the teaching task may be intellectually stimulating.

Before establishing this factor structure for transformational leadership as the construct of choice, this model needs to be subjected to a confirmatory factory analysis to
note how well it fits the data. Roberts and Henson (2001) call for more rigorous testing of construct instrumentation via confirmatory factor analysis. Henson (2001) states CFA is critical for the testing of scales used for evaluating theoretical models. Accordingly, this two-factor leadership model submits to a CFA.

_Parsimonious Two-Factor Solution_

Appendix G displays the conceptual design of the leadership model being tested. The chi-square minimum fit function yielded a statistically non-significant value $x^2(25, N = 146) = 31.05$, $p=0.19$, meaning the chi-square test failed to reject the null hypothesis. Moreover, all the more practical, albeit less rigorous, global fit indices satisfied their thresholds for model fit. To insure the integrity of the model matches the quality suggested by the model fit indices, its components demand review.

Bollen and Long (1993) suggest attention be paid to the squared multiple correlations of the indicators, the strength of the parameters, and the sign of the estimates. The statistical significance of the estimates should be such that the estimates are reliably different than zero, meaning that there is in fact a relationship between the variables in question (Breckler, 1990; Byrne, 1998). The squared multiple correlations, noted as alternative reliability measures above, for all nine indicator variables average 0.758. This means the underlying concept explains 75.8% of the indicators’ variances, on average.

Appendix H exhibits the standardized path estimates, or factor loadings, linking the indicators with their respective factors. As can be seen, the path standardized path coefficients indicate strong, positive relationships between indicators and underlying factors. The measurement errors associated with the indicators are disturbances that upset the relationship between the indicators and their underlying factor. The closer the
measurement terms are to zero, the better the indicator measures the latent variable. Not
unexpectedly, the magnitude of the measurement terms is inversely associated with the
strength of the path coefficients linking the indicators to their dimensions. The two
weakest path coefficients, “Doesn’t know my talents and strengths” and “Gives me a
sense of being unimportant,” have the two highest measurement errors. They were also
the only two indicators that posed a leadership behavior with a negative orientation.
Other than these two items, the health of the coefficients is strong. All coefficients are
statistically significant. Nothing about the component analysis brings the contention of
the global fit indices into question. Appendix I reviews all confirmatory factor analysis
model fit indicators across the four transformational leadership models considered.

Further validation of this model comes from prior research following similar
principal leadership measurement methodology, using survey scales. Hoover, Petrosko,
and Schultz (1991) found individualized consideration and intellectual stimulation to be
the only distinct factors emerging in their analysis of school leaders using the MLQ.
Barnett et al. (2001) discerned two subfactors of the MLQ in their study of Australian
school leadership: an individualized consideration factor and a hybrid of the other three
“I’s” in the other leadership factor. Lowe et al. (1996), in their meta-analytical review of
MLQ literature noted that intellectual stimulation observations of the leader were
particular salient in mechanistic organizations, like public service agencies. “[It] may be
the very nature of the mechanistic organization that propels transformational leadership
in the form of intellectual stimulation leader behavior to be highly salient to individuals”
(p. 415).
Similar logic may be applied to the finding of individualized consideration in school leaders. Although organizational theory suggests a supervisor-subordinate ratio of 1:5 to 1:10 for optimal administrative effectiveness (Hoy & Miskel, 2001), the span of control for most principals clearly surpasses that guideline. The typical elementary school principal supervises a staff of 30 professional educators and an average of 14 support staff (Profiling the principalship, 2002). The challenge to be individually considerate to thirty teachers while tending to the needs of 425 students, on average, as well as to the demands of the broader school community, is profound (Profiling the Principalship, 2002). Hence, a principal who is sensitive to the personal needs and desires of his or her teachers acts conspicuously, given the organizational context and demands of the job. Empirically, teachers notice.

The most fundamental theoretical support for the factor structure found in this study may come from the contentions about leadership behavior made by Halpin and Winer (1957). Their study of military officers and industrial supervisors, via subordinates’ responses to the LBDQ leadership scale, produced a leadership model that contained two factors, as previously described: initiating structure and consideration (Bass, 1990). Given that subsequent leadership researchers built upon the foundation set by the LBDQ (i.e., Halpin & Croft, 1962; Oldham, 1976; Seltzer & Bass, 1987), with Bass being among those scholars, it is not surprising to find constructs conceptually similar to the LBDQ’s original dimensions.

The two-factor structure model captures transformational leadership perceptions via mean teacher scores generated from their observations of their school principals. Intellectual stimulation and individualized consideration stand as the two dimensions of
transformational leadership. Although they correlate highly \( r=0.726, \ p < 0.01 \), CFA testing discourages folding the two factors into a unified dimension. When the parameter linking the two factors together is set to one, representing factor convergence, the model fit measures fail to meet satisfactory thresholds – thus the unified model structure falls short of confirmation. Maintaining a dual factor organization of transformational leadership realizes internal consistency as Cronbach’s alpha reliability coefficients for the individualized consideration and intellectual stimulation factors are .9169 and .9529, respectively.

Collective Efficacy

Goddard’s (2001) 12-item collective efficacy scale, a shorter version of the Goddard’s original collective efficacy scale, constituted the school climate variable in this study. Its construction corrected the imbalance among the theoretical dimensions, task analysis and group competence, of the previous 21-item collective efficacy scale (Goddard et al., 2000). Goddard (2001a) reports his shorter scale, a one factor solution, has high internal reliability \( \alpha = 0.94 \), contains items that load at 0.67 or higher, explains 64.10% of the variance extracted, and has an eigenvalue of 7.69.

A one-factor solution of the short-form proved reasonable in the current sample. The instrument’s internal reliability remained strong \( \alpha = 0.93 \) and the one-factor construct explains 56.36% of the variance. All the items loaded strongly on the first factor; in fact, three-quarters of the items load greater than 0.67.

In sum, transformational leadership was operationalized as two separate variables, intellectual stimulation and individualized consideration. Collective efficacy was measured as a single construct. A review of the construct relationships apparent in the
data can now be reviewed for testing the research hypotheses now that measurements for
the major independent variables have been established. After noting the descriptive
statistics for each of the variables in the hypotheses, the bivariate correlations and path
models describing the multivariate relations will be discussed.

Descriptive and Correlational Statistics of Research Variables

Means, minimums, maximums, and standard deviations were calculated to gain a
broad overview of the measures in question. Prior school achievement is included to
answer the call by researchers for school leadership models to reflect antecedent effects
on principal influence (Hallinger & Heck, 1996). The results of the descriptive
calculations are shared in Table 4.3.
<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individualized Consideration</td>
<td>146</td>
<td>.64</td>
<td>3.80</td>
<td>2.75</td>
<td>.56</td>
</tr>
<tr>
<td>Intellectual Stimulation</td>
<td>146</td>
<td>.25</td>
<td>3.87</td>
<td>2.53</td>
<td>.63</td>
</tr>
<tr>
<td>Collective Efficacy</td>
<td>146</td>
<td>3.26</td>
<td>5.47</td>
<td>4.50</td>
<td>.52</td>
</tr>
<tr>
<td>2001 Mean Math Scaled Score</td>
<td>139</td>
<td>169.00</td>
<td>254.00</td>
<td>223.22</td>
<td>15.10</td>
</tr>
<tr>
<td>2001 Mean Reading Scaled Score</td>
<td>139</td>
<td>167.00</td>
<td>236.00</td>
<td>217.77</td>
<td>9.16</td>
</tr>
<tr>
<td>Free/Reduced Lunch Rate 2002</td>
<td>141</td>
<td>.00</td>
<td>.84</td>
<td>.28</td>
<td>.23</td>
</tr>
<tr>
<td>2002 Mean Math Scaled Score</td>
<td>146</td>
<td>185.00</td>
<td>252.00</td>
<td>225.14</td>
<td>13.71</td>
</tr>
<tr>
<td>2002 Mean Reading Scaled Score</td>
<td>146</td>
<td>200.00</td>
<td>235.00</td>
<td>220.80</td>
<td>7.12</td>
</tr>
</tbody>
</table>

Table 4.6: Research Variables’ Descriptive Statistics

The first hypothesis predicted that all dimensions of transformational leadership positively related to collective efficacy. Correlational analysis does not support this contention, as evidenced by Table 4.4. Neither intellectual stimulation nor individualized consideration relate to collective efficacy. It should be noted, though, that the relationship linking collective efficacy and intellectual stimulation hovers close to statistical significance (r=.16, p=.059).
The second hypothesis predicted that collective efficacy was positively related to student achievement, or more specifically to math and reading achievement. This hypothesis was supported: collective efficacy correlated with the mean school math scaled score ($r = 0.719, p < 0.01$) and with the mean school reading scaled score ($r = 0.745, p < 0.01$). A partial correlation was computed to analyze the relationship between collective efficacy beliefs and student achievement controlling for SES, because student family background is often related to achievement (Coleman et al., 1966). Collective efficacy still independently and significantly related to math achievement ($r = .29, p < 0.01$) and reading achievement ($r = .33, p < 0.01$). The stronger the teachers perceived the collective competency of the staff, the higher the school’s level of achievement.

The third hypothesis proposed that the school’s socioeconomic status was positively related to collective efficacy. Adjustments were made to transform the negative orientation of SES, as it was measured in this study, to a positive one. More specifically, SES was measured by noting the proportion of a participating school’s student body enrolled in the federal lunch program; in this case, the more students participating in the free and reduced lunch program, the lower the socioeconomic status (SES). The empirical evidence of the study supports this hypothesis; collective efficacy was strongly related to SES ($r = .79, p < 0.01$). The greater a school’s proportion of students not on the federal lunch program, the greater the staff’s collective competence.

The fourth hypothesis suggested that SES was positively related to transformational leadership. The findings show that SES was positively related to individualized consideration ($r = .19, p < 0.05$) but was unrelated to intellectual
stimulation. The more comfortable economic circumstances in which a school resides, the more likely a principal will manifest individualized consideration.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Individ. Consider.</td>
<td>0.92(^{a})</td>
<td>.726(^{**})</td>
<td>0.137</td>
<td>0.161</td>
<td>0.104</td>
<td>.192((^{*}))</td>
<td>.206((^{*}))</td>
<td>.0139</td>
</tr>
<tr>
<td>Intellect Stimul.</td>
<td>0.95(^{a})</td>
<td>0.156</td>
<td>0.091</td>
<td>0.054</td>
<td>-0.109</td>
<td>.216(^{**})</td>
<td>.155</td>
<td></td>
</tr>
<tr>
<td>Collect. Efficacy</td>
<td>0.93(^{a})</td>
<td>.605(^{**})</td>
<td>.591(^{**})</td>
<td>.794(^{**})</td>
<td>.719(^{**})</td>
<td>.745(^{**})</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Math 2001</td>
<td>0.87(^{a})</td>
<td>.900(^{**})</td>
<td>.629(^{**})</td>
<td>.798(^{**})</td>
<td>.759(^{**})</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading 2001</td>
<td>0.84(^{a})</td>
<td>.597(^{**})</td>
<td>.704(^{**})</td>
<td>.702(^{**})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Socio-economic</td>
<td>b</td>
<td>.750(^{**})</td>
<td>.794(^{**})</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Math 2002</td>
<td>0.86(^{a})</td>
<td>.901(^{**})</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading 2002</td>
<td></td>
<td>0.84(^{a})</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: **Correlation is significant at the 0.01 level (2-tailed). *Correlation is significant at the 0.05 level (2-tailed)

\(^{a}\) Reliability coefficient (Cronbach’s alpha)

\(^{b}\) Construct measured with single-item indicator, thus alpha can’t be measured.

Table 4.7: Correlations among Research Variables
Path Models of Student Achievement

The fifth hypothesis proposed that transformational leadership was related to student achievement both directly, and indirectly through collective efficacy. Because bivariate analysis reveals that neither dimension of transformational leadership related to collective efficacy, further empirical testing for the indirect effects of transformational leadership on student achievement through collective efficacy is unwarranted. Furthermore, when the direct relationship of transformational leadership with student achievement is studied, consideration must be given to contextual influences, for leadership behavior has been shown to vary according to socioeconomic surroundings (Hallinger, Bickman, & Davis, 1996). When relating transformational leadership behaviors to student achievement, controlling for SES, only intellectual stimulation was statistically significant ($r=.20$, $p<.05$). Individualized consideration remained statistically insignificant. Thus, intellectual stimulation is the transformational leadership behavior employed as a principal influence in the proposed model of student achievement.
Consider two models of student achievement illustrated in figures 4.1 and 4.2.

Figure 4.1: Model of School Math Achievement

Model’s Explained Variance of Math Achievement = 0.76

Bold path coefficients are statistically significant (p<0.05)
Both models offer respectable predictions about achievement, as the math model explains 76% of the variation in the 2002 math scaled scores and the reading model explains 72% of the 2002 reading scaled scores. Please refer to Appendix J for a comprehensive look at the models’ goodness-of-fit statistics. All but one of the model fit statistics yielded satisfactory correspondence between the models and the data, thus implying the proposed theory is reasonable. In the model including math achievement, intellectual stimulation
does directly relate to student achievement (beta=.118, p<.05), even after controlling for SES, prior school achievement, and collective efficacy. However, in the reading achievement model, intellectual stimulation does not relate significantly to reading achievement (r=0.06, p=.259), when SES, prior school achievement, and collective efficacy are taken into account. Thus, the fifth hypothesis is true only for math achievement. Moreover, it is only true when considering the direct effects of leadership on math achievement.

The sixth hypothesis is that socioeconomic status is directly related to student achievement and indirectly related to student achievement through collective efficacy. Both models of achievement support this hypothesis. Socioeconomic status relates to math achievement directly (r=.29, p<.01), indicating that as the SES of a school’s student body increases, the achievement of the school likely increases, too. SES’s direct effect on reading achievement is statistically significant, as well (beta=.42, p<.01).

As shown by the path models above, SES, for both models, works through collective efficacy, as well as prior school achievement, to affect school achievement. LISREL 8.50 (Joreskog & Sorbom, 2000) calculates standardized total and indirect effects to quantify the influence of any variable on any other one, provided paths between the relevant combinations of variables exist. SES’s indirect effects on math achievement (beta=0.46, p<0.01) and reading achievement (beta=0.37, p<.01) are substantial.

Noteworthy, though, SES has its largest influence on collective efficacy, not student achievement. SES’s total standardized effects, that is its indirect and direct effects, on math achievement measures .76 (p<.01), whereas its total effects on the collective efficacy beliefs of teachers total .81 (p<.01). Similarly, the socioeconomic
background of the school’s student body constitutes total effects equaling .79 (p<.01) on reading achievement, still less than SES’s estimated total influence on the competence perceptions of the staff.

This study would be remiss in not reporting what parameter emerged as the marked path for improving the model’s fit with the data. Although this research model passed all but one of the fit statistic standards utilized in this study (the models’ correspondence with the data didn’t quite reach quality level for the standardized RMR test), a path parameter was flagged, signaling the model’s fit with the data could be improved by freeing the parameter identified. The parameter in question was the relationship from collective efficacy to intellectual stimulation, the direction of which is critical. Interestingly, although the modification indices suggested freeing the path from collective efficacy to intellectual stimulation, the result of doing so did not realize a statistically significant path parameter. For both the math and reading models, the t-test used for testing the probability the path parameter is indicative of a systematic, not random, relationship between the two variables generated a test statistic of 1.63, below conventional standards for statistical significance.

Conclusion

This chapter comprised two major sections. First, the factor structure of transformational leadership was considered. Systematic attention was given to a one-factor, a four-factor, and a two-factor structure leadership variable to address the possibilities suggested by existing scholarship, the pilot study, and the data at-hand. A two-factor structure of leadership emerged, meeting the rigor of confirmatory factor analyses, to measure intellectual stimulation and individualized consideration.
Second, the hypotheses were tested. Correlational analyses and structural equation modeling provided insight regarding the accuracy of the research hypotheses. The results of this research fail to support the hypothesis that transformational leadership influences collective efficacy. The findings show collective efficacy is positively related to student achievement even controlling for SES and prior achievement. Socioeconomic status is positively related to collective efficacy and is positively related to individualized consideration, but it is not related to intellectual stimulation. Transformational leadership behavior, via intellectual stimulation, is directly related to student achievement, but is not indirectly related to student achievement through collective efficacy. Finally, socioeconomic status is directly related to student achievement and indirectly related to student achievement through collective efficacy and prior school achievement. Reasons for these findings will be considered in the following chapter.
CHAPTER 5
DISCUSSION OF RESULTS

For the past 15 years, public schools have been experiencing a wave of standards and accountability reform pressures like that of no other previous time (Elmore, 2000). As the demands for better public schooling coincide with demands for more school funding, calls for school accountability ring louder. Tennessee’s experience in the early 1990’s serves as a model for understanding what is taking place nationally and in other states, with regard to the ramping up of standards, funding, and accountability (cf: Sanders, 1998; and Horn, 1998). Although various reform efforts may have disparate ideas about how schools should change and should be evaluated, ultimately, most hold the expectation that all students should achieve to a certain standard. Moreover, schools are asked to educate all students to reach these standards of achievement despite the fragmenting of families, the availability of drugs, and the increasing needs of an ever-diversifying student population that also challenge schools.

Teacher beliefs about their staff skill level, or collective efficacy, inform educators of their likely persistence, effort, resilience, and stress thresholds for specified tasks. Collective efficacy beliefs serve as social norms, behaving like normative expectations and support for teaching behavior. Finding factors related to collective efficacy would contribute to the tools needed for effective school reform. Despite the growing research demonstrating the positive effects of collective efficacy beliefs (ex., Bandura, 1993, 1997; Campion, Medsker, & Higgs, 1993; Esselman & Moore, 1992;
This study investigated the relationship between leadership behaviors and collective efficacy. A model was developed for predicting and explaining school achievement, which built upon the theoretical correspondence between transformational leadership and collective efficacy. Transformational leadership and its positive effect on organizational functioning is supported by educational research (e.g., Leithwood, 1994; Leithwood, Jantzi, & Steinbach, 1999; Leithwood & Jantzi, 1999; Leithwood & Jantzi, 2000) and by organizational theory in general (e.g., Bass, 1985; Bass & Avolio, 2000; Howell & Avolio, 1993). Transformational leadership’s conceptual fit with the determinants of collective efficacy and its documented link with notable organizational performances made it an attractive factor to include in this study of collective efficacy antecedents.

Leaders who cultivate robust staff beliefs develop capacity for increased student achievement in their schools. From this theoretical assertion, a set of research questions evolved into research hypotheses. Subsequently, transformational leadership, collective efficacy, student achievement, socioeconomic status, and prior achievement variables were operationalized and measured from survey responses, test data, and federal lunch program rates from 146 public elementary schools. The hypotheses were tested, and the results are summarized as follows:
Summary of Findings

1. A combination of exploratory factors analyses and confirmatory factor analyses performed on the sample data resulted in a two-factor structure of transformational leadership. Intellectual stimulation and individualized consideration constitute the two factors.

2. This research failed to support the hypothesis that transformational leadership related to collective efficacy. Neither intellectual stimulation nor individualized consideration was related to collective efficacy.

3. Collective efficacy was positively related to student achievement. Indeed, collective efficacy was directly related with math ($r=.719$, $p<.01$) and reading ($r=.745$, $p<.01$) achievement. Collective efficacy maintained its statistical significance even after school SES and prior achievement were introduced into the achievement model; that is, collective efficacy was related to student achievement even after controlling for SES and prior achievement.

4. Socioeconomic status was positively related to collective efficacy ($r=.794$, $p<.01$).

5. Socioeconomic status was positively related to individualized consideration, but did not relate to intellectual stimulation.
6. Intellectual stimulation was related directly to student achievement, but it was not
related to it indirectly, through collective efficacy.

7. Socioeconomic status was directly related to student achievement and indirectly
related to student achievement through collective efficacy and prior school
achievement.

Discussion

The hypotheses of this study had some empirical support. Of the theorized
relationships supported in this study, collective efficacy was positively related to student
achievement; socioeconomic status was positively related to collective efficacy;
transformational leadership was positively related to student achievement; and
socioeconomic status was directly and indirectly related to student achievement.
However, the findings did not support the hypothesis that transformational leadership
related to collective efficacy and offered only partial support to the theory that
socioeconomic was status related to transformational leadership. Each of the
relationships is discussed in light of current research.

Collective efficacy and student achievement. These findings support prior
research that claims the importance of collective efficacy as a school variable predictive
of school performance (Bandura, 1993; Hoy, Sweetland, & Smith, 2001; Goddard, 2001).
Noteworthy, however, is that the strength of the relationship between collective efficacy
and achievement diminishes appreciably when prior achievement and SES are considered
in the models. The standardized path coefficients realized in the structural equation
models (estimates similar to beta weights in traditional regression analyses) were .18 and .22 for math and reading, respectively. The decreased effect sizes of collective efficacy for both models are to be expected; prior achievement and SES have long been known to strongly influence academic performance (see Coleman et al., 1966; Parker, 1994). Hence, the variance accounted for in models involving just collective efficacy and student achievement spuriously inflate the true relationship between the two variables.

Social cognitive theory suggests that the relationship between collective efficacy and student achievement is mutually deterministic (Bandura, 1993, 1997). The bi-directional association linking collective efficacy and achievement highlights two important components of the tripartite: responsive environment and effective behavior. First, students demonstrating high achievement serves as a responsive environment for the school staff, validating the staff’s efforts as constructive and efficacious. Such feedback to the staff develops their perceptions of competence and facilitates future instructional interventions for improving student achievement. In turn, effective behavior stemming from confident teachers likely causes improved student achievement to continue a cycle of mutual improvement. A receptive social context and high teacher efficacy beliefs make for a potent combination. “A high sense of [collective] efficacy in a responsive environment that rewards valued accomplishments fosters aspirations, productive engagement in activities, and a sense of fulfillment” (Bandura, 1997, p. 20).

Socioeconomic status and collective efficacy. Goddard’s (2001) theoretical design of collective efficacy offers a plausible reason for the positive relationship between these two school achievement predictors. Two dimensions comprise collective efficacy: task analysis and group competence. Both factors are implicated by the social
context of the schools, operationalized in this study as the SES variable. Conceivably, teachers working in lower socioeconomic settings may encounter less environmental responsiveness to their efforts, as students living in surroundings with fewer resources tend to come to school less ready to learn (Coleman et al., 1966). Confidence in one’s skills may waver if, over time, the application of instructional skills realizes minimal gain in expected results.

The task analysis dimension of collective efficacy contributes to the explanation of this relationship, too. Wilson (1996) applies social cognitive theory to the plight of the unemployed in new poverty neighborhoods, communities resulting from the disappearance of work. He posits that because most of the adult population in these neighborhoods suffers from unemployment, the jobless adults’ low efficacy beliefs likely stem from environmental restriction considerations. Such thoughts stem from task analysis reviews, rather than from low competency beliefs. Likewise, teachers facing comparable environmental responsiveness may attribute their efficacy beliefs to task analysis factors, rather than competency thoughts. That is, they may lower their efficacy judgments because the task is so difficult in the unresponsive environment, not because they question their competence.

*Intellectual stimulation and student achievement.* Although the indirect effects of transformational leadership on student achievement did not materialize in this study, the direct effects of intellectual stimulation on student math achievement merit focused discussion.

The magnitude of the intellectual stimulation – math achievement relationship is weak ($r=.12, p<.05$). However, several reasons exist for not discounting the significance
of its effects or its implications. First, this model of achievement considers three influences most closely associated with student performance: the students themselves (prior achievement), the family (SES), and the teachers (collective competency beliefs of the staff). Given that, the direct effects of leadership behavior on achievement are noteworthy, for principals are often not viewed as such direct influences (ex., Wiseman, 2002). Second, the very fact that intellectual stimulation is recognized as a distinct leadership behavior prompts consideration of elementary school environments for adult learning. Perhaps, the mechanistic nature and bureaucratic organization of schools cause innovative and challenging leadership ideas to be conspicuous (Lowe et al., 1996; Podsakoff et al., 1990). Also, such school leader action may be contrary to the teachers’ receptiveness to learning. Indeed, the resistance to novel ideas may be viewed in the socialization of teachers into the profession. The flow of new ideas into schools by new teachers is often discouraged (National Research Council, 2000). “Cognitive reappraisal of the status quo” gives meaning to the essence of intellectual stimulation (Lowe et al., 1996). In sum, intellectual stimulation’s direct effect is statistically significant, even when strong influences on school performance are considered. Also, its role as a distinct behavior suggests leaders who practice it make themselves noticeable, perhaps in the face of norms that may not always initially embrace such behavior.

There are other reasons intellectual stimulation deserves more careful consideration than might be typically given to other variables with similar magnitudes of influence. Hallinger and Heck (1996) contend, with support from Ogawa and Hart (1985), that only about 15% of the total variance of student achievement outcomes can be attributed to between school differences. Intellectual stimulation, in this study, accounts
for about 1.4% of the variation in student math achievement – a number given to cursory attention, at first. However, given that the fourth-grade reading test is a high-stakes test, influence over 1% of the score distribution represents an important opportunity to make a difference. For example, 125,827 students took the 4th grade proficiency tests in 2002 (Ohio Department of Education, 2002). If principal behavior, such as intellectual stimulation, can influence the scores of 1.4% of these students, then 1,762 children’s academic futures can be positively affected.

Podsakoff et al. (1990) underscored the importance of intellectual stimulation in their research on transformational leadership effects on trust, satisfaction, and organizational citizenship behaviors. Their concluding remarks regarding intellectual stimulation provide an engaging summary of this behavior’s contribution to school staff:

It now seems quite clear that the leader who is able to intellectually stimulate subordinates will not only foster the perception of effectiveness among subordinates, but will also amplify performance itself as gauged by independent measures of productivity. Inducing employees to appreciate, dissect, ponder, and discover what they would not otherwise discern is perhaps the basis of behavior that comes closest to our prototypical abstractions of ‘true leadership.’ (p. 415)

Hallinger et al. (1996) state that of all the school level administrators, elementary principals have the best opportunity to directly influence students. The size of elementary schools allows more principal access to students, based on their smaller enrollments. Also, the nature of elementary schools is more conducive to principal influence of students, perhaps due to their less rigid schedules, and to their more focused attention on fewer areas of cognitive development (i.e., the four R’s). Moreover, elementary principals may impact their pupils greater because of the more impressionistic minds of grade-school students.
*Socioeconomic status and student achievement.* Of little surprise is SES’s direct effect on student achievement. However, the magnitude of its direct effects deserves elaboration. Like the case illustrated with collective efficacy, SES’s effects on achievement are overstated when consideration of other variables likely contributing to achievement is lacking. When collective efficacy, prior achievement, and school leadership are factored into the achievement models, SES’s direct relationship with achievement falls to .29 in math and .42 in reading. These effects represent roughly a .40 standardized unit drop from SES’s bivariate relationship with achievement.

Nonetheless, SES relates directly to student achievement, and has been known to for decades (Coleman et al. 1966). SES implicates the social and economic backgrounds of a school’s students; hence, both resources and social capital are important for students. Both “possessions” of the students, the economic and social resources, serve as important facilitators for developing children’s intellectual capacities. Economic resources can provide students with the materials, experiences, conditions, and health aids that facilitate well-being and positive cognitive development. Because economic resources help people gain access to mainstream society, social assets often come with financial ones. Social resources provide networks necessary for important peer and adult support, provisions for important assistance and advantageous opportunities in life. SES’s indirect effects on school achievement through prior achievement can be explained by the reasoning used for its effect on current achievement; the temporal status stands as the only difference between SES’s effect on the two school performances.

The indirect effects of SES on school achievement through collective efficacy implicate teachers’ beliefs. Adverse student body characteristics associate more strongly
with teacher beliefs about staff competence than with the actual achievement of the school. In the math and reading models of school achievement, SES’s standardized direct effects on collective efficacy were 0.71. SES’s standardized direct effects on math and reading school achievement were 0.29 and 0.42, respectively. These sets of relationships illustrate a consistent psychological phenomenon. “People's level of motivation, affective states, and actions are based more on what they believe than on what is objectively the case” (Bandura, 1997b, p. 2).

*Transformational leadership and collective efficacy.* Although transformational leadership was not found to relate to collective efficacy, there may be plausible explanations for the lack of association.

Elmore (2000), in recalling Newman et al.’s (1989) research of organizational factors affecting school sense of efficacy, contends that there is no independent effect of teachers' perceptions of principals' leadership on teachers' sense of efficacy, unless it is deliberately connected to tangible and immediate problems of practice. The items used in this study’s transformational scales did not specify any particular instructional practice. Perhaps, if common challenges in practice had been identified – those that seem to pose problems for most schools – they could have been infused into the leadership survey items. For example, if leadership behaviors had been expressed in terms of providing vision, motivation, and ideas related to differentiated instruction (Tomlinson, 2000), perhaps the teachers could have related better to how leadership behavior influenced their efficacy beliefs. However, there may be another reason why transformational leadership behavior didn’t relate to collective efficacy beliefs.
Two dimensions of transformational leadership, idealized influence and inspirational motivation, were not discernible by the teachers participating in this study. This was not novel, as other researchers have found these dimensions of transformational leadership indiscernible to their research subjects (e.g., Barnett, McCormick, & Connors, 2001; Hoover, Petrosko, and Schultz, 1991). However, these two may have been the more influential behaviors for bolstering collecting efficacy perceptions. In fact, Avolio (1999) notes that idealized influence and inspirational motivation are the most effective leaderships for stimulating organizational performance, an ends for which collective efficacy is deemed a vital means in this study. Idealized influence behaviors and charismatic behaviors have been termed the “primary leadership mechanisms for influencing organizational cultures” (Hipp & Bredeson, 1995), in a reference to Schein’s (1985) view of such actions. Podsakoff et al. (1990) describe idealized influence as tapping the “core” of transformational leadership, a view shared by other researchers (e.g., Kouzes & Posner, 1987). Lowe et al. (1996), based on their meta-analysis of transformational leadership studies, considered idealized influence, or charisma, to be the scale most closely linked with a general impression of transformational leadership, “due to the nature of the items which make up the scale and because the construct itself tends to represent affective reactions of subordinates to the leader” (p. 414). Without the primary dimension of transformational leadership, not finding a statistically significant relationship between principal behavior and collective efficacy may not be surprising.

_Socioeconomic status and individualized consideration._ Principal behavior correlating with school socioeconomic context has been reported previously. Hallinger, Bickman, and Davis (1996) reported that the nature of principals’ leadership differed
systematically in relation to the schools’ student body’s socioeconomic status. Hallinger et al.’s work (1996), although offers little explanation for the reason between SES and school leadership, recognizes that several other researchers have found a similar relationship (Goldring, 1990, 1993; Goldring & Pasternak, 1994; Heck et al., 1990; Scott & Tedlie, 1987).

Larson, Sweeney, Christian, and Olson (2002) synthesize assertions made by Baum (1998) in reporting, “Differences in [the context’s] socioeconomic status can lead to differing leadership styles. These [differences] can be counterproductive if not managed effectively and may hinder … leadership” (p. 39). In quoting Baum (1998) further, Larson et al. (2002) write, “… no mental stratagem really eliminates racial differences” (p. 39), implying that cultural differences may be working behind the SES differences observed. To the extent that cultural differences grow between the principal and the teachers as the SES of the school decreases, the personable behavior of the principal may be challenged. Certainly, cultural differences among people can delay or discourage bonds formed between them.

Perhaps a more immediate explanation might exist for the individualized consideration – SES link. The sheer weight of demands placed upon the school for educating students with backgrounds not likely conducive for school learning might limit the principal’s opportunities for engaging in personal exchanges with the staff. Certainly, an overwhelmed principal may be hard-pressed to be a people-oriented leader.

Practical, Research, and Theoretical Implications

Practical implications. In an age when increased accountability and demands for improved achievement challenge public schools, the performance of all students on
observable measures becomes central. This study highlights two factors that directly relate to school achievement and stand as qualities more susceptible to school influence than SES and prior achievement. Intellectual stimulation and collective efficacy may hold promise for advancing the aims of elementary schools.

Intellectual stimulation behaviors challenge the means and content of teacher instruction to insure the outcomes sought are being pursued by the most defensible methods. The problems facing education may stem, in part, from relatively poor and underutilized observations of teaching and learning by the schools, themselves. The dearth of data collection and data analysis by school administrators supports this contention (Creighton, 2001), as data stand as measured observations.

School leaders who stimulate staff by conveying intellectually compelling thoughts operate from and generate the need for observations of teaching and learning. Quality observations require a command of learning theories, pedagogical competence, a working familiarity of human development, management skills, a reflective disposition, alertness, and openness (Bulterman-Bos, Terwell, Verloop, & Wardekker, 2002). Such an assertion implies that educators who hold solid command of disciplines related to education will make better observations of learning. These same educators will likely form better interventions to realize student achievement. Thus, to the degree intellectually stimulating leaders cause teachers to sharpen their observations of teaching and learning, they probably increase student performance. Although intellectual stimulation did not relate to collective efficacy, its potential for improving student
performance is documented in this study. Challenging the teaching practices and content likely keeps the staff alert and mindful (Langer, 1990), steering them away from the blind haze of routine and the paralysis of rigidity.

Although no leadership behaviors were found to influence collective efficacy in this research, consideration of principal actions likely to elevate staff competence beliefs should continue in research and in practice. Because it is amply documented that people's beliefs are among the most powerful predictors of behaviors, motivations, and affective states (Bandura, 1993, 1997), attempts to enhance the collective efficacy perceptions of staff are well-placed efforts. Thought might be directed at highlighting the positive results of teaching efforts, so staff could witness the responsiveness of their environment. Stated differently, a savvy principal might conceive ways to sharpen the observations of his or her teachers so that they are more sensitive to the positive effects of their craft. Relatedly, a leader endeavoring to reverse low staff efficacy beliefs requires "explicit, compelling feedback that forcefully disputes … pre-existing" notions (Bandura, 1997, p. 82). Data, a resource likely at the disposal of most intellectually stimulating leaders, may pose as an important means to both of these collective efficacy-enhancing strategies.

Theoretical Implications. Perhaps one of the most theoretically intriguing observations to emerge from this study stems from the structural equation model of school achievement. The modification index suggested a path be drawn from collective efficacy to intellectual stimulation, the direction of which is critical. Although the parameter estimate resulting from that relationship was not significant, its t-value was not drastically distant from significant standards. The close failing of this estimate
notwithstanding, this path value suggests a leader’s behavior may be influenced by the collective beliefs of his or her subordinates. The theoretical implications of such a possibility could cause some leadership theorists to reconsider where school reform begins – with the actions of the principal or the beliefs of the staff? Although much contemporary school reform literature suggests educational change needs to include the psychosocial aspects of the school (Leithwood, 1994; Senge, 1990; Sergiovanni, 2000), many models of school leadership do not consider the possibility that the staff influences leadership. Research variables used historically to study leadership antecedent factors have included community type and homogeneity, school size, student SES, school grade level, and principal personal characteristics (Hallinger & Heck, 1996; Hallinger et al., 1996). Any research exploring the effects of staff beliefs on leader behaviors that exists may be novel, sparse, or both.

Support for such theory can be found in the literature on effective teaching. If learning is conceived as a relatively permanent change in behavior and knowledge (Woolfolk Hoy & Hoy, 2003), then the teachers’ charge is to change the behavior and knowledge of his or her students. Effective teaching strategists suggest that teachers may achieve this aim by, among other practices, providing scaffolding opportunities to students so that the pupils can develop their understanding based on what their knowledge base was at the time the concept is introduced (Gillespie, 2002; Vygotsky, 1978). The same might be said of leaders, if their aim is to change the behavior and knowledge of his or her teachers in order to effect significant reform. Effective leadership may entail providing scaffolding opportunities to teachers so that they can build their understandings of and competence with initiatives based on their previous
knowledge. In both cases, leaders and teachers construct proper scaffolding based on where the "learner" is starting. Such prompts for change, then, may need to start from subordinate cues. Teacher perceptions may influence leadership actions, in the same way that student perceptions may influence teacher actions.

Research implications. Although transformational leadership did not relate to collective efficacy in the current study, it may be premature to determine that this style of leadership does not relate to the collective efficacy. Recall that the four dimensions of transformational leadership failed to materialize in this research. Only intellectual stimulation and idealized influence were identified as leadership characteristics of school principals. These two characteristics are basic dimensions of leadership not unlike the task-oriented (initiating structure) and people-oriented (consideration) dimensions of the early LBDQ (Halpin & Winer, 1957) and still found in contemporary scholarship on leadership (Chen & Bliese, 2002). Inspirational motivation and idealized influence were not involved in the relationship tested in this study. Acknowledging that factors comprising transformational leadership have been difficult to identify in research (Bass & Avolio, 2000; Yukl, 1998), efforts to construct transformational items distinguishable by elementary teachers should still continue. The empirical importance of the two dimensions not found in this study has been demonstrated elsewhere (Avolio, 1999, Lowe et al., 1995; Podsakoff et al., 1990) and the existence of such data suggests that the task of discerning them is not insurmountable. Given the premium placed on school leadership, scholarship aimed at identifying productive principal behaviors is effort well spent. Moreover, there may be direction to take in attaining such factor items.
Future research might do well to construct survey leadership items that characterize particular problems of practice, as Elmore (2000) observed teacher efficacy perceptions only relate to leadership behavior when specific teaching issues are directly implicated. Similarly, researchers have called for efficacy scale items to be more context-sensitive and specific to instructional behaviors (Bandura, 1997; Roberts & Henson, 2001; Pajares, 1997). Thus, the prospect of framing transformational leadership and collective efficacy items around specific instructional issues and testing for the strength of their relationship holds potential for future scholarship.

Inquiry into collective efficacy may continue regarding teachers’ attributions for their competency beliefs. Wilson (1996) concedes his thoughts are untested about the efficacy belief attributions of the unemployed. He makes, though, a reasonable conjecture that task analysis attributions of the unemployed may result from their neighbors’ shared and relevant experiences concerning the task at hand (a survey of the context). Likewise, the collective efficacy belief attributions of teachers are largely untested, if the call for more studies to research the antecedents of efficacy beliefs (Goddard & Goddard, 2001) may include study of collective efficacy belief attributions.

Do the collective efficacy beliefs of teachers stem more from task analysis considerations or from competency perceptions? Might collective efficacy perceptions be driven more by task analysis evaluations than collective competency beliefs? If this is so, do teachers pay too much heed to the SES background of the students in shaping their collective efficacy judgments, rather than focusing more on their talents and skills at hand?

One other area that prompts research questions centers on the SES –leadership behavior relationship. Are the principal behavior adjustments to the school context, as
advocated by Hallinger et al. (1996), proactive or reactive, or, as social cognitive theory might suggest, mutually deterministic? Why was individualized consideration and not intellectual stimulation related to SES, especially since instructional leadership behaviors, conceptually similar to intellectual stimulation actions, related to SES (1996)? Is the link between leadership behavior and school SES culturally driven (ethnicity or race constraints), resource-driven (time and effort constraints), both, or caused by some other influence?

In their own study, Hallinger et al. (1996) conceive school leadership as those behaviors that manage curriculum and instruction and create an academically focused climate in the school. Interestingly, though, the researchers attempt to sum up this leadership – SES relationship as further evidence that principals adapt their behavior to the community context in which they work. Like the individualized consideration link with SES in this study, questions abound about why communities with lower SES warrant such purportedly healthy leadership behaviors any less than higher SES communities. Is this a conscious and proper adaptation by school leaders? Perhaps ethnographic or other qualitative studies may reveal why certain leadership behaviors vary with the SES context of their schools.

Conclusion

Present-day Americans owe a debt of gratitude to our forbearers for jointly struggling to overturn inhumanities and creating social conditions that allowed a better life (Bandura, 1997b). Though conditions today in our country may look nothing like the hardships witnessed by colonists, nineteenth century Unionists and industrialists, or this
past century’s “Greatest Generation” (Brokaw, 1998), the challenges facing our country now may be just as pivotal. Future generations of Americans may look back on this time as the context for challenges that fortunately resulted in their favor, due, in part, to the collective efforts of educators. Or they may not.

As referenced to earlier, there is a crisis brewing in our country that just recently received focused federal attention (No Child Left Behind Act, 2002). The likely perpetrator and probable redeemer of this developing urgency is public education, for an increasing number of under-prepared citizens are leaving our schools. “The time bomb is the failing American system of education that is producing a growing underclass of citizens who are not equipped to deal with modern life and the future job market” (Fenton, 2001, p. 1). To the extent we cultivate our own sense of collective efficacy to correct the inequities perpetuated by schools, we will provide for our children and grandchildren the domestic political stability we currently enjoy and an improved system of education.

Collective efficacy serves as one of the few school level variables that influences school achievement, regardless of the socioeconomic context and previous attainments of the school. Though no link was found between principal behavior and collective efficacy, this study showed that school leadership contributes directly, albeit weakly, to school performance of students. The practical applications of these findings, and the implications drawn for further theoretical consideration and research, hold important potential for school improvement. In all, this study adds to an important line of research. What is known about collective efficacy to date may help schools improve the lives of students and the nation’s social and political future.
APPENDIX A

FACTOR ANALYSIS OF TRANSFORMATIONAL LEADERSHIP SCALE

<table>
<thead>
<tr>
<th>Scale Item</th>
<th>Factor Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instills pride in me for being associated with him/her</td>
<td>.927</td>
</tr>
<tr>
<td>Leads by example</td>
<td>.924</td>
</tr>
<tr>
<td>Acts in ways that build my self-respect</td>
<td>.918</td>
</tr>
<tr>
<td>Practice what he preaches</td>
<td>.914</td>
</tr>
<tr>
<td>Gives me a sense of being important</td>
<td>.906</td>
</tr>
<tr>
<td>Provides moral support by making me feel appreciated</td>
<td>.906</td>
</tr>
<tr>
<td>Goes beyond self-interest for the good of the group</td>
<td>.895</td>
</tr>
<tr>
<td>Promotes an atmosphere of caring and trust among staff</td>
<td>.883</td>
</tr>
<tr>
<td>Stimulates me to think about what I am doing for my students</td>
<td>.877</td>
</tr>
<tr>
<td>Treats me as an individual rather than just a member of a group</td>
<td>.864</td>
</tr>
<tr>
<td>Is a source of new ideas for my professional learning</td>
<td>.859</td>
</tr>
<tr>
<td>Suggests new ways of looking at how we do our jobs</td>
<td>.858</td>
</tr>
<tr>
<td>Gives me a sense of overall purpose</td>
<td>.858</td>
</tr>
<tr>
<td>Seeks differing perspectives when solving problems</td>
<td>.849</td>
</tr>
<tr>
<td>Helps staff to learn from each other</td>
<td>.842</td>
</tr>
<tr>
<td>Specifies the importance of having a strong sense of purpose</td>
<td>.833</td>
</tr>
<tr>
<td>Models problem solving techniques that I can readily adapt for my work</td>
<td>.831</td>
</tr>
<tr>
<td>Is concerned about my personal problems</td>
<td>.743</td>
</tr>
</tbody>
</table>
APPENDIX B

4 FACTOR STRUCTURE OF TRANSFORMATIONAL LEADERSHIP SCALE
APPENDIX C

1 FACTOR STRUCTURE OF TRANSFORMATIONAL LEADERSHIP SCALE
APPENDIX D

EXPLORATORY FACTOR ANALYSIS OF TRANSFORMATIONAL LEADERSHIP ITEMS

<table>
<thead>
<tr>
<th>Rotated Factor Matrix(a)</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>C50TL_IS</td>
<td>.881</td>
<td>.315</td>
</tr>
<tr>
<td>C12TL_IM</td>
<td>.818</td>
<td>.393</td>
</tr>
<tr>
<td>C30TL_IS</td>
<td>.801</td>
<td>.405</td>
</tr>
<tr>
<td>C20TL_IM</td>
<td>.782</td>
<td>.444</td>
</tr>
<tr>
<td>C43TL_IS</td>
<td>.754</td>
<td>.511</td>
</tr>
<tr>
<td>C38TL_I1</td>
<td>.740</td>
<td>.583</td>
</tr>
<tr>
<td>C49TL_IS</td>
<td>.708</td>
<td>.575</td>
</tr>
<tr>
<td>C4TL_IM</td>
<td>.694</td>
<td>.604</td>
</tr>
<tr>
<td>C45TL_I1</td>
<td>.665</td>
<td>.655</td>
</tr>
<tr>
<td>C44TL_IC</td>
<td>.414</td>
<td>.850</td>
</tr>
<tr>
<td>C28TL_IC</td>
<td>.486</td>
<td>.814</td>
</tr>
<tr>
<td>C46TL_I1</td>
<td>.529</td>
<td>.754</td>
</tr>
<tr>
<td>C42TL_I1</td>
<td>.589</td>
<td>.748</td>
</tr>
<tr>
<td>C16TL_IC</td>
<td>.381</td>
<td>.748</td>
</tr>
<tr>
<td>C32TL_I1</td>
<td>.639</td>
<td>.711</td>
</tr>
<tr>
<td>C36TL_I1</td>
<td>.617</td>
<td>.703</td>
</tr>
<tr>
<td>R_C8TL_</td>
<td></td>
<td>.683</td>
</tr>
<tr>
<td>C24TL_IS</td>
<td>.558</td>
<td>.649</td>
</tr>
<tr>
<td>R_C47TL_</td>
<td></td>
<td>.588</td>
</tr>
</tbody>
</table>
### Rotated Factor Matrix(a)

<table>
<thead>
<tr>
<th>Factor</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treats me as an individual rather than just a member of a group.</td>
<td>.851</td>
<td></td>
</tr>
<tr>
<td>Provides moral support by making me feel appreciated.</td>
<td></td>
<td>.813</td>
</tr>
<tr>
<td>Is concerned about my personal problems.</td>
<td></td>
<td>.760</td>
</tr>
<tr>
<td>Gives me a sense of being unimportant. (Rev. Scored)</td>
<td></td>
<td>.691</td>
</tr>
<tr>
<td>Doesn't know my talents and interests. (Rev. Scored)</td>
<td></td>
<td>.617</td>
</tr>
<tr>
<td>Suggests new ways of looking at how we do our jobs.</td>
<td>.873</td>
<td></td>
</tr>
<tr>
<td>Specifies the importance of having a strong sense of purpose.</td>
<td>.813</td>
<td></td>
</tr>
<tr>
<td>Is a source of new ideas for my professional learning.</td>
<td>.806</td>
<td></td>
</tr>
<tr>
<td>Stimulates me to think about what I am doing for my students.</td>
<td>.806</td>
<td></td>
</tr>
</tbody>
</table>

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.

* a Rotation converged in 3 iterations.
APPENDIX F

EXPLORATORY FACTOR ANALYSIS OF TRANSFORMATIONAL LEADERSHIP ITEMS USING OBLIQUE ROTATION

| Pattern Matrix(a) | Factor  \\
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Treats me as an individual rather than just a member of a group.</td>
<td>.922</td>
</tr>
<tr>
<td>Provides moral support by making me feel appreciated.</td>
<td>.820</td>
</tr>
<tr>
<td>Is concerned about my personal problems.</td>
<td>.814</td>
</tr>
<tr>
<td>Gives me a sense of being unimportant. (Rev. Scored)</td>
<td>.782</td>
</tr>
<tr>
<td>Doesn't know my talents and interests. (Rev. Scored)</td>
<td>.697</td>
</tr>
<tr>
<td>Suggests new ways of looking at how we do our jobs.</td>
<td>-1.007</td>
</tr>
<tr>
<td>Specifies the importance of having a strong sense of purpose.</td>
<td>-.867</td>
</tr>
<tr>
<td>Is a source of new ideas for my professional learning.</td>
<td>-.866</td>
</tr>
<tr>
<td>Stimulates me to think about what I am doing for my students.</td>
<td>-.829</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Axis Factoring.
Rotation Method: Oblimin with Kaiser Normalization.

a Rotation converged in 6 iterations.
EXPLORATORY FACTOR ANALYSIS OF TRANSFORMATIONAL LEADERSHIP ITEMS USING OBLIMIN ROTATION

<table>
<thead>
<tr>
<th>Structure Matrix</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Treats me as an individual rather than just a member of a group.</td>
<td>.947</td>
</tr>
<tr>
<td>Provides moral support by making me feel appreciated.</td>
<td>.944</td>
</tr>
<tr>
<td>Is concerned about my personal problems.</td>
<td>.851</td>
</tr>
<tr>
<td>Gives me a sense of being unimportant. (Rev. Scored)</td>
<td>.746</td>
</tr>
<tr>
<td>Doesn't know my talents and interests. (Rev. Scored)</td>
<td>.667</td>
</tr>
<tr>
<td>Suggests new ways of looking at how we do our jobs.</td>
<td>.658</td>
</tr>
<tr>
<td>Stimulates me to think about what I am doing for my students.</td>
<td>.750</td>
</tr>
<tr>
<td>Is a source of new ideas for my professional learning</td>
<td>.715</td>
</tr>
<tr>
<td>Specifies the importance of having a strong sense of purpose.</td>
<td>.696</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Axis Factoring.
Rotation Method: Oblimin with Kaiser Normalization.
APPENDIX G

CONCEPTUAL DESIGN OF A TWO-FACTOR MODEL OF TRANSFORMATIONAL LEADERSHIP
APPENDIX H

STANDARDIZED PATH COEFFICIENTS OF A TWO-FACTOR MODEL OF TRANSFORMATIONAL LEADERSHIP
# APPENDIX I

## A REVIEW OF CFA GOODNESS-OF-FIT MEASURES ACROSS FOUR TL MODELS

<table>
<thead>
<tr>
<th>Fit Measures</th>
<th>4 Factor (Bass, 1985)</th>
<th>1 Factor (Pilot Study)</th>
<th>2 Factor (EFA all)</th>
<th>2 Factor (9 indicators)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-Square Min. Fit Function P-val (only test of statistical diff.). P should be greater than .05 so that the model fails to reject the null hypothesis, $\Sigma = \Sigma(\theta)$.</td>
<td>P=0.0</td>
<td>P=0.0</td>
<td>P=0.0</td>
<td>P=.17</td>
</tr>
<tr>
<td>Root Mean Square Error of Approximation RMSEA (&lt; .05)</td>
<td>0.15</td>
<td>0.20</td>
<td>0.16</td>
<td>0.040</td>
</tr>
<tr>
<td>Parsimony of fit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standardized RMR (Root Mean Squared Residual) (&lt; .05) Absolute Test of Fit (average standardized residual values derived from fitting $S= \Sigma(\theta)$)</td>
<td>0.045</td>
<td>0.047</td>
<td>0.042</td>
<td>0.025</td>
</tr>
<tr>
<td>AGFI (Adjusted Goodness of Fit Test) (&gt; .90) Measures quality of model against no model at all. (Bollen, p. 1989)</td>
<td>0.60</td>
<td>0.47</td>
<td>0.57</td>
<td>0.92</td>
</tr>
</tbody>
</table>
A REVIEW OF CFA GOODNESS-OF-FIT MEASURES ACROSS FOUR TL MODELS

<table>
<thead>
<tr>
<th>Fit Measures</th>
<th>4 Factor (Bass, 1985)</th>
<th>1 Factor (Pilot Study)</th>
<th>2 Factor (EFA all)</th>
<th>2 Factor (9 indicators)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECVI (Expected Cross Validation Index) An ECVI is obtained for an independence model, the hypothesized model, and a saturated model. The hypothesized model’s index should be the lowest one.</td>
<td>ECVI mod = 4.86 ECVI ind = 30.78 ECVI sat = 2.62</td>
<td>ECVI mod = 7.55 ECVI ind = 30.78 ECVI sat = 2.62</td>
<td>ECVI mod = 5.53 ECVI ind = 30.78 ECVI sat = 2.62</td>
<td>ECVI mod = 0.48 ECVI ind = 9.56 ECVI sat = 0.62</td>
</tr>
<tr>
<td>Hoelter’s Critical N (CN) Model Fit Test Independent of Sample Size. CN estimates a sample size that would be sufficient to yield an adequate model fit for $\chi^2$ test. (&gt; 200)</td>
<td>50.59</td>
<td>39.23</td>
<td>45.53</td>
<td>203.19</td>
</tr>
<tr>
<td>IFI (Incremental Fit Index) Measures quality of parsimony with regard to sample size and degrees of freedom. Higher values indicate superior fit. (Byrne, 1998)</td>
<td>.91</td>
<td>.86</td>
<td>.89</td>
<td>0.99</td>
</tr>
</tbody>
</table>
## APPENDIX J

A REVIEW OF GOODNESS-OF-FIT MEASURES ACROSS BOTH SCHOOL ACHIEVEMENT MODELS

<table>
<thead>
<tr>
<th>Fit Measures</th>
<th>Math Model</th>
<th>Reading Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-Square Min. Fit Function P-val (only test of statistical diff.). P should be greater than .05 so that the model fails to reject the null hypothesis, $\Sigma = \Sigma(\theta)$.</td>
<td>P=0.44</td>
<td>P=0.42</td>
</tr>
<tr>
<td>Root Mean Square Error of Approximation RMSEA (&lt; .05) Parsimony of fit</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Standardized RMR (Root Mean Squared Residual) (&lt; .05) Absolute Test of Fit (average standardized residual values derived from fitting $S= \Sigma(\theta)$)</td>
<td>0.058</td>
<td>0.053</td>
</tr>
<tr>
<td>AGFI (Adjusted Goodness of Fit Test) (&gt; .90) Measures quality of model against no model at all. (Bollen, p. 1989)</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>ECVI (Expected Cross Validation Index) An ECVI is obtained for an independence model, the hypothesized model, and a saturated model. The hypothesized model’s index should be the lowest one.</td>
<td>ECVI mod = 0.20 ECVI ind = 3.18 ECVI sat = 0.22</td>
<td>ECVI mod = 0.20 ECVI ind = 2.95 ECVI sat = 0.22</td>
</tr>
<tr>
<td>Hoelter’s Critical N (CN) Model Fit Test Independent of Sample Size. CN estimates a sample size that would be sufficient to yield an adequate model fit for $x^2$ test. (&gt; 200)</td>
<td>582.39</td>
<td>558.83</td>
</tr>
<tr>
<td>IFI (Incremental Fit Index) Measures quality of parsimony with regard to sample size and degrees of freedom. Higher values indicate superior fit. (Byrne, 1998)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
REFERENCES


207


