Running head: TEACHER EFFICACY RESEARCH

Teacher Self-Efficacy: Substantive Implications and Measurement Dilemmas

Robin K. Henson

University of North Texas 76203-1337

Invited keynote address given at the annual meeting of the Educational Research Exchange, January 26, 2001, Texas A&M University, College Station, Texas. Correspondence concerning this manuscript should be sent to rhenson@tac.coe.unt.edu.

Abstract

Founded in social cognitive theory, teachers' self-efficacy beliefs have been repeatedly associated with positive teaching behaviors and student outcomes. However, teacher efficacy has developed a storied history regarding construct validity and measurement integrity. Study of teacher efficacy now stands on the verge of maturity, but such developmental growth will likely be contingent on development of strong theoretical models and effective instrumentation to assess theoretical constructs. The purpose of the present article is to: (a) briefly review the theoretical foundation of teacher efficacy and critically evaluate historical attempts to measure teacher efficacy, (b) discuss important substantive implications stemming from efficacy research that may advance the field, (c) present recent measurement advances, and (d) highlight several methodologies that have been underutilized in development of teacher efficacy instruments.

Teacher Self-Efficacy: Substantive Implications and Measurement Dilemmas

Anecdotally, we all have observed others prevailing amidst adversity and trial. Consider, for example, the following:

- Louis Pasteur was only a mediocre pupil in undergraduate studies and ranked 15th out of 22 students in chemistry.
- Albert Einstein was four before he began to speak. He did not read until he was seven. His teacher described him as "mentally slow, unsociable, and adrift forever in foolish dreams."
- It has been told that a football expert once said of Vince Lombardi: "He possesses minimal football knowledge and lacks motivation."

Of course, each of us could add our own testimony of how we have fared in the world, for better or worse. Implicit in these and our own anecdotes lies the question of how people are able to face challenge, direct their actions, and somehow succeed. One answer to this question lies with the concept of self-efficacy.

What is Self-Efficacy and Why is it Important?

The theoretical foundation of self-efficacy is found in social cognitive theory, developed by former APA president (1974) and current Stanford professor Albert Bandura (1977, 1997). Social cognitive theory assumes that people are capable of human agency, or intentional pursuit of courses of action, and that such agency operates in a process called triadic reciprocal causation. Reciprocal causation is a multi-directional model suggesting that our agency results in future behavior as a function of three interrelated forces: environmental influences, our behavior, and internal personal factors such as cognitive, affective, and biological processes.

This trinity mutually impacts its members, determines what we come to believe about ourselves, and affects the choices we make and actions we take. We are not products of our environment. We are not products of our biology. Instead, we are products of the dynamic interplay between the external, the internal, and our current and past behavior. In reaction to more reductionist theories, Bandura noted: “Dualistic doctrines that regard mind and body as separate entities do not provide much enlightenment on the nature of the disembodied mental state or on how an immaterial mind and bodily events act on each other” (1986, p. 17).

Central to Bandura’s (1997) framework is his concept of self-efficacy. Bandura’s aspirations about self-efficacy were grand, as reflected in the title of his 1977 article “Self-Efficacy: Toward a Unifying Theory of Behavioral Change.” In this seminal work, Bandura defined self-efficacy as “beliefs in one’s capabilities to organize and execute the courses of action required to produce given attainments” (p. 3). Self-efficacy beliefs were characterized as the major mediators for our behavior, and importantly, behavioral change. Over the last quarter century, Bandura’s other works continued to develop and defend the idea that our beliefs in our abilities powerfully affect our behavior, motivation, and ultimately our success or failure (cf. Bandura, 1982, 1986, 1993, 1996, 1997).

Bandura (1997) proposed that because self-efficacy beliefs were explicitly self-referent in nature and directed toward perceived abilities given specific tasks, they were powerful predictors of behavior. The research literature has supported this proposition. Research has linked efficacy to a variety of clinical issues such as phobias (Bandura, 1983), addiction (Marlatt, Baer, &

Quigley, 1995), depression (Davis & Yates, 1982), and smoking behavior (Garcia, Schmitz, & Doerfler, 1990). Educationally, self-efficacy beliefs are related to academic performance and self-regulated learning (cf. Hackett, 1995; Pajares, 1996; Schunk, 1991; Zimmerman, 1995).

Importantly, efficacy beliefs help dictate motivation (cf. Maehr & Pintrich, 1997; Pintrich & Schunk, 1996). Bandura observed: "People regulate their level and distribution of effort in accordance with the effects they expect their actions to have. As a result, their behavior is better predicted from their beliefs than from the actual consequences of their actions" (1986, p. 129). From the social cognitive theory perspective, because human agency is mediated by our efficaciousness, self-efficacy beliefs influence our choices, our effort, our persistence when facing adversity, and our emotions (cf. Pajares, 1997).

In short, self-efficacy theory is a common theme in current views of motivation (Graham & Weiner, 1996), primarily because of its predictive power and application for practically any behavioral task. This article will focus on one area of self-efficacy application directly relevant to educational improvement: teacher self-efficacy. Unfortunately, teacher efficacy research has at times been theoretically confused. In effort to advance and strengthen the study of teacher efficacy, I will (a) briefly review the theoretical foundation of teacher efficacy and critically evaluate historical attempts to measure the construct, (b) discuss important substantive implications stemming from efficacy research that may advance the field, (c) present recent measurement advances, and (d) highlight several methodologies that have been underutilized in development of teacher efficacy instruments.

What is Teacher Self-efficacy and Why is it Important?

Consistent with the general formulation of self-efficacy, Tschannen-Moran and Woolfolk Hoy (in press) defined teacher efficacy as a teacher's "judgment of his or her capabilities to bring about desired outcomes of student engagement and learning, even among those students who may be difficult or unmotivated." The study of teacher efficacy is a little over two decades old and began with RAND researchers' evaluation of whether teachers believed they could control the reinforcement of their actions (Armor et al., 1976). This early work was founded on Rotter's (1966) locus of control theory, and it was assumed that student learning and motivation were the relevant reinforcers of teaching action.

Historically, the Bandura (1977) and Rotter (1966) traditions have influenced the study of teacher efficacy. Unfortunately, researchers' interpretations of these theories have significantly muddied the efficacy waters as regards the theoretical formulation of teacher efficacy and the psychometric attempts to measure the construct. Tschannen-Moran, Woolfolk Hoy, and Hoy (1998) provided a comprehensive review of these historical developments and they will not be reiterated here, with the exception of several measurement issues noted below.

In spite of the measurement confusion, teacher efficacy still emerged as a worthy variable in educational research. As Woolfolk and Hoy (1990) noted, "Researchers have found few consistent relationships between characteristics of teachers and the behavior or learning of students. Teachers' sense of efficacy . . . is an exception to this general rule" (p. 81). The idea that teachers' self-beliefs are determinants of teaching behavior is a simple, yet powerful idea.

The correlates of teacher efficacy are many when using a variety of efficacy scales and measurements. Students of efficacious teachers generally have outperformed students in other classes. Teacher efficacy was predictive of achievement on the Iowa Test of Basic Skills (Moore & Esselman, 1992), the Canadian Achievement Tests (Anderson, Greene, & Loewen, 1988), and the Ontario Assessment Instrument Pool (Ross, 1992). Watson (1991) observed greater

achievement in rural, urban, majority Black, and majority White schools for students of efficacious teachers. Teacher efficacy is also related to students' own sense of efficacy (Anderson et al., 1988) and student motivation (Midgley, Feldlaufer, & Eccles, 1989).

Regarding teacher behaviors, efficacious teachers persist with struggling students and criticize less after incorrect student answers (Gibson & Dembo, 1984). They are more likely to agree that a low SES student should be placed in a regular education setting and less likely to refer students for special education (Meijer & Foster, 1988; Podell & Soodak, 1993; Soodak & Podell, 1993). Teachers with high efficacy tend to experiment with methods of instruction, seek improved teaching methods, and experiment with instructional materials (Allinder, 1994; Guskey, 1988; Stein & Wang, 1988). Coladarci (1992) observed higher professional commitment for efficacious inservice teachers. Evans and Tribble (1986) found similar results for preservice teachers.

Clearly the study of teacher efficacy has borne much fruit. However, teacher efficacy is the subject of current debate concerning its meaning and measure (cf. Tschannen-Moran et al., 1998). The dialogue has centered on two issues. First, based on the theoretical nature of the self-efficacy construct as defined by Bandura (1977, 1997), researchers have argued that self-efficacy is most appropriately measured within context regarding specific behaviors (Pajares, 1996). Second, the construct validity of scores from the primary instruments purporting to measure teacher efficacy has been severely questioned (Coladarci & Fink, 1995; Guskey & Passaro, 1994). Accordingly, teacher efficacy is presently on the precipice of inquiry; it is ready to either move forward or fall to the wayside as a good idea that ultimately had little substance. As Tschannen-Moran et al. (1998) noted:

> This appealing idea, that teachers' beliefs about their own capacities as teachers somehow matter, enjoyed a celebrated childhood, producing compelling findings in almost every study, but it has also struggled through the difficult, if inevitable, identity crisis of adolescence. . . . teacher efficacy [now] stands on the verge of maturity. . . (p. 202)

I hope to explicate several issues in the study of teacher efficacy that will facilitate this maturation. My discussion, while not exhaustive, focuses on current issues in defining the theoretical construct and appropriate measurement strategies.

Current Issues in Teacher Efficacy

Gibson and Dembo's Teacher Efficacy Scale

As noted, social cognitive theory provides the theoretical foundation for teacher efficacy. Historically, however, Bandura's (1977) theory has been mixed with locus of control theory (Rotter, 1966). At the birth of teacher efficacy, RAND researchers (Armor et al., 1976) developed two items that were based on the locus of control orientation:

- Item 1: "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."
- Item 2: "If I really try hard, I can get through to even the most difficult or unmotivated students."

The items were intended to assess whether a teacher believed that student learning and motivation (assumed reinforcers for the teacher) were under the teacher's control. These items and this orientation guided most teacher efficacy research during the late 70s and early 80s.

Amidst concerns about construct definition and reliability of measurement with only two items, Gibson and Dembo (1984) sought to empirically develop a teacher efficacy measure.

They argued that the two items used by the RAND researchers actually corresponded to Bandura's (1977) outcome expectancy and self-efficacy dimensions of social cognitive theory, respectively. (Note: Outcome expectancy is theoretically independent from self-efficacy. Unfortunately, brevity prohibits discussion of outcome expectancy here and the reader is referred to Bandura [1997], Pajares [1997], and Tschannen-Moran et al. [1998] for thorough treatments.) RAND Item 1 was thought to assess an outcome expectancy regarding a teacher's belief whether teaching can impact student learning despite external constraints. This construct was labeled teaching efficacy (later to be called general teaching efficacy, or GTE). RAND Item 2 was thought to assess self-efficacy, or a teacher's perceived ability to positively impact student learning. This construct was named personal teaching efficacy (PTE).

In their study, Gibson and Dembo (1984) developed additional items modeled after the RAND items and retained 16 after a factor analysis (PTE: 9 items, GTE: 7 items). The PTE and GTE factors were essentially uncorrelated ($r = -.19$), a result consistent with Bandura's conceptualization of outcome expectancy and self-efficacy dimensions. A multitrait-multimethod construct validity study was conducted and the Teacher Efficacy Scale (TES) was born. The TES subsequently became the predominate instrument in the study of teacher efficacy, leading Ross (1994, p. 382) to label it a "standard" instrument in the field. The TES has also served as a launching point for the development of other similar instruments, such as the subject matter specific Science Teaching Efficacy Belief Instrument (STEBI; Riggs & Enochs, 1990).

As teacher efficacy research flourished, serious questions about the TES arose. Specifically, in a study of correlations among scores from the major instruments of teacher efficacy and related constructs, Coladarci and Fink (1995) largely found weak evidence for discriminant validity of PTE and GTE scores. Furthermore, Guskey and Passaro (1994) reported that the PTE and GTE factors correspond to not to self-efficacy and outcome expectancy dimensions, but to an internal versus external orientation, respectively. This dichotomy resembled locus of control and attributional theory orientations more than self-efficacy theory.

Importantly, the Coladarci and Fink (1995) and Guskey and Passaro (1994) studies pointed out potential theoretical confounds in the TES. We must remember that the TES was originally developed from the two RAND items which were based on locus of control theory. Gibson and Dembo (1984) later interpreted the items as reflecting self-efficacy theory. Accordingly, the TES appears to have elements of both theoretical orientations captured in its items. As might be expected from an instrument that serves two theoretical masters, the study of teacher efficacy has suffered an adolescent identity crisis as researchers have struggled to clarify the construct (see e.g., Soodak & Podell, 1996). (Note: Tschannen-Moran et al. [1998] presented a more complete discussion of these theoretical strands in teacher efficacy research, along with specific results from the Coladarci and Fink [1995] and Guskey and Passaro [1994] studies.)

Toward Construct Clarification

In effort bring some coherence to the meaning and measure of teacher efficacy, Tschannen-Moran et al. (1998, p. 227) developed a model of that "weaves together both conceptual strands" in teacher efficacy's storied history. The model attempts to take a broader, more comprehensive look at self-efficacy as it relates to teachers and explicates a cyclical feedback loop for efficacy judgments. The model, presented in Figure 1, represents an important advancement in the area that may guide future research efforts. Several theoretical issues that are particularly relevant to the current discussion are noted below.

INSERT FIGURE 1 ABOUT HERE

Sources of efficacy information. Bandura (1997) proposed that there are four general sources of efficacy building information: verbal persuasion, vicarious experiences, physiological arousal, and mastery experiences. Of these, mastery experiences are likely the most powerful influence in fostering efficacy. Prior conceptualizations of teacher efficacy have all but ignored these sources of information and their relationship to efficacy and ultimate behavior. If efficacy is a powerful influence on behavior, then investigation of factors that might influence efficacy are certainly warranted.

Task analysis and teaching competence. Tschannen-Moran et al. (1998) also argued that teacher efficacy actually is a joint, simultaneous function of a teacher's analysis of the teaching task and his or her assessment of his or her personal teaching competence or skill (concepts related to but different from the GTE and PTE scales, respectively). As described by Tschannen-Moran et al.:

> In analyzing the teaching task and its context, the relative importance of factors that make teaching difficult or act as constraints is weighed against an assessment of the resources available that facilitate learning. In assessing self-perceptions of teaching competence, the teacher judges personal capabilities such as skills, knowledge, strategies, or personality traits balanced against personal weaknesses or liabilities in this particular teaching context. . . (p. 228, italics in original)

The task analysis evaluates the specific elements of a teaching situation. Although similar to GTE, the task analysis is more specific and invokes elements that can both help and hinder teaching (GTE only deals with external constraints and ignores possible useful resources). This conceptualization is consistent with Bandura's (1997) triadic reciprocal causation, such that a teacher's efficacy belief stems from dynamic interplay of the environment, behavior, and personal factors. Assessment of resources or constraints in the teaching environment would play an important role in a teacher's efficacy judgment.

Additionally, personal teaching competence resembles the PTE scale but deals directly with perceptions of current functioning or abilities as opposed to futuristic evaluations. Self-efficacy beliefs are, by definition, judgments of expected performance at a task at some point in the future. Tschannen-Moran et al. (1998) have argued that these futuristic (self-efficacy) judgments only occur after a teacher evaluates their current skill level. Furthermore, the judgment occurs only after the teacher's current skill level is weighed against the task analysis. These two processes, task analysis and assessment of competence, are proposed to occur simultaneously and in light of each other, eventually resulting in an efficacy belief held by the teacher for the given context.

Level of specificity and content congruence. The teacher efficacy model also holds that teacher efficacy beliefs should be referenced to specific tasks. Pajares (1996) observed that self-efficacy judgments are most consistent with Bandura's (1997) theory, and most predictive of behavior, when evaluation of one's capability is matched to a specific outcome. Pajares noted:

> Omnibus tests that aim to assess general self-efficacy provide global scores that decontextualize the self-efficacy-behavior correspondence and transform self-efficacy beliefs into a generalized personality trait rather than the context-specific judgment Bandura suggests they are. . . . The problem with such assessments is that students [or teachers] must generate judgments about their . . . capabilities without a clear activity or task in mind. As a result, they generate the judgments by in some fashion mentally

> aggregating to related perceptions that they hope will be related to imagined tasks. (p. 547)

Historically, and contrary to Bandura's (1997) formulations, teacher efficacy has been measured rather globally. For example, one PTE item on the TES reads: "When I really try, I can get through to most difficult students." How much effort is necessary? How difficult are the students? Is difficulty related to behavior, motivation, and/or instruction? Are learning disabilities relevant to the situation?

Because efficacy judgments refer to one's ability to perform successfully in a task, the judgments are contextually based. Assessment of efficacy without reasonable context specificity may actually be assessment of a different construct altogether, perhaps of more general personality traits. The lack of discriminant validity for teacher efficacy measures found in the Coladarci and Fink (1995) study points to this reality.

However, Lent and Hackett (1987) cautioned that specificity and precision may reduce result generalizability. In response to this potentiality, Pajares (1997) noted:

> Judgments of competence need not be so microscopically operationalized that their assessment loses all sense of practical utility. Domain specificity should not be misconstrued as extreme situational specificity, and there is no need to reduce efficacy assessments to atomistic proportions. (p. 13)

Furthermore, Bandura (1997) argued that transfer of efficacy judgments is possible, and noted that "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). Regarding teacher efficacy, Tschannen-Moran et al. (1998) also cautioned against "developing measures so specific that they lose their predictive power for anything beyond the specific skills and contexts being measured" (p. 219). It is not very useful or predictive of other behavior if a teacher only feels efficacious for providing effective instruction if the student is in the seventh grade, does not have a learning disability, enjoys working in cooperative groups, never asks difficult questions, has a pleasant personality, and is left-handed.

Some Limited Preliminary Support

Although the Tschannen-Moran et al. (1998) model has promise for the advancement of teacher efficacy inquiry, the model must undergo empirical scrutiny before theoretical claims are proclaimed too loudly. For example, the sources of efficacy information have been largely unexamined in relation to teacher efficacy. Historically, vicarious experience and social persuasion seem to impact GTE beliefs (Watters & Ginns, 1995) and PTE may be best impacted by actual teaching experiences (Housego, 1992; Hoy & Woolfolk, 1990). However, as noted, these conceptualizations of teacher efficacy are not entirely consistent with the proposed model.

Many questions exist regarding the nature of the proposed task analysis concept. How do the sources of efficacy information contribute to the task analysis? What elements of the teaching situation are most relevant to efficacy judgments? Furthermore, it is unknown whether the task analysis varies for teachers at different developmental stages in their careers. Tschannen-Moran et al. (1998) suggested that the task analysis "will be most explicit for novice teachers and for those entering a new teaching assignment. Experienced teachers are likely to rely more heavily on memories and interpretations of similar past teaching experiences" (p. 231). However, this assumption warrants empirical evaluation.

Henson, Bennett, Sienty, and Chambers (2000) found some limited support for the model. Moving from Tschannen-Moran et al.'s (1998) assumption that novice teachers will

emphasize the task analysis in efficacy judgments, Henson et al. examined a sample of emergency certification teachers new to their teaching careers. The canonical correlation analysis revealed a relationship between teaching competence and efficacy. However, task analysis variables were not related to efficacy judgments.

In a study of collective teacher efficacy (see below) using the same theoretical framework noted in Figure 1, Goddard, Hoy, and Woolfolk Hoy (2000) noted that the task analysis and teaching competence dimensions were closely related in a factor analysis ($\underline{r} = .75$). The authors argued that the high interfactor correlation was evidence that the two dimensions are poorly discriminated by teachers and likely occur simultaneously in teachers' cognitive processing. This result and interpretation is consistent with Tschannen-Moran et al.'s (1998) model. At present, there are no other known studies explicitly evaluating the validity of this recent development in teacher efficacy. Clearly, much empirical work and model testing is needed to continue to advance the field.

Promising Substantive Implications

Recent developments in the study of teacher efficacy beg many questions regarding the function of teacher efficacy in teachers' lives. Although myriad research agendas could be developed to pursue these questions, the following represents three areas of inquiry that show great promise for the advancement of teacher efficacy. Other treatments of these and related issues can be found in Goddard et al. (2000), Tschannen-Moran and Woolfolk Hoy (in press), Tschannen-Moran et al. (1998).

Efficacy Building Information

Bandura (1986) argued that "perceived self-efficacy results from diverse sources of information conveyed vicariously and through social evaluation, as well as through direct experience" (p. 411). Furthermore, these sources of information "must be processed and weighed through self-referent thought" (p. 21). Accordingly, Bandura (1986, 1997) postulated four sources of efficacy building information: mastery experiences, vicarious experiences, social persuation, and physiological or emotional arousal. Tschannen-Moran et al. (1998) employed these sources of information in their theoretical model of teacher efficacy.

Mastery experiences are considered the most powerful influence on efficacy as they provide direct feedback regarding capabilities. However, because the feedback "must be processed and weighed through self-referent thought," all success does not lead to bolstered efficacy. Attributional analysis and causal assumptions concerning outcomes impact the interpretation of mastery experiences. Furthermore, some outcomes may be valued more than others. For example, a teacher may succeed at altering an assignment to the appropriate level of a student, but experience no increase in confidence because no special value was vested in the outcome. Social cognitive theory emphasizes that vicarious experiences can impact learning and efficacy. Social persuasion and the emotional state that one experiences during social interactions can also bolster, or weaken, self-efficacy beliefs.

Although the Tschannen-Moran et al. (1998) included these elements in their model, research examining the validity and potential impact of the sources of information on teacher efficacy is practically non-existent. If teacher efficacy is the powerful predictive construct it has been thought to be, then research examining the processes by which such efficacy is built is critical to fostering teacher efficacy and, ultimately, changing behavior. Are some mastery

experiences more salient than others? What sources if information are most effectual for preservice teachers? Do these sources differ for inservice teachers? How do teachers process and internalize vicarious experiences? What characteristics of models best facilitate vicarious efficacy growth?

Important is this discussion is the fact that most efficacy research has been self-report, survey, and correlational in nature. Henson (in press) noted that "experimental (or quasi-experimental) and/or long term designs are near absent in the literature, leaving cross-sectional snapshots of teacher perceptions of their capabilities." Such designs are unlikely to shed much light on the complex interplay between sources of efficacy information and efficacy development. Regarding student self-efficacy, Pajares (1997) recognized this methodological weakness, and noted:

> Two strategies are called for. The first is for researchers to assess both the sources and the effects of self-efficacy through direct observation rather than rely on students' self-reports; the second is to increase the use of experimental techniques so as to manipulate sources and effects. (p. 30)

Researchers of teacher efficacy would likely do well to heed this advice.

Collective Teacher Efficacy

Recently some researchers have begun to explore the construct of collective teacher efficacy. Because social cognitive theory holds that "personal agency operates within a broad network of sociostructural influences" (Bandura, 1997, p. 6) the theory "extends the analysis of mechanisms of human agency to the exercise of collective agency" (p. 7). Bandura (1997) defined collective efficacy generally as "the groups' shared belief in its conjoint capabilities to organize and execute courses of action required to produce given levels of attainments" (p. 477). Collective efficacy takes self-efficacy to the social level. Extending Bandura's theory to teachers, Goddard et al. (2000) defined collective teacher efficacy as

> a construct measuring teachers' beliefs about the collective (not individual) capability of a faculty to influence student achievement; it refers to the perceptions of teachers that the efforts of the faculty of a school will have a positive effect on student achievement. (p. 486)

Research on collective teacher efficacy as been limited (Bandura, 1993, 1997; Esselman & Moore, 1992; Goddard et al., 2000; Newmann, Rutter, & Smith, 1989), but the findings have been compelling. Using multi-level modeling, Goddard et al. (2000) reported that collective teacher efficacy was more predictive of elementary students' math and reading achievement than gender, ethnicity, and even socio-economic status. The authors indicated that "collective teacher efficacy explained 53.27% and 69.64% of the between-school variance in mathematics and reading, respectively" (p. 500).

Several studies of collective teacher efficacy have used individual teacher efficacy variables and aggregated them to the school level (cf. Goddard, 1998). This process is inconsistent with the theoretical nature of the construct, which concerns teachers' perceptions of the faculty as a whole. Collective teacher efficacy is not the simple aggregate of individual perceptions of the self; instead, it is individual perceptions of the capabilities of the entire faculty in a school organization. Therefore, collective efficacy items generally should reflect a group ("teachers can. . . ") rather than individual or self ("I can. . . ") orientation.

Bandura (1997) noted, however, that the decision to use self or group orientation in collective efficacy measurement should actually consider the level of organizational coupling

present. That is, organizations whose functions are highly interdependent (tight coupling) would be more appropriately assessed with a group orientation. Loosely coupled organizations may be more appropriately assessed with an aggregate of individual orientations. Goddard (1998) suggested that elementary schools are best characterized as tightly coupled because of shared goals and similarity of teaching roles across positions. Following the same logic, one could argue that secondary schools would be more loosely coupled, at least as compared to elementary schools.

In a study of high school teachers, Kurz (2000) noted that goal consensus, a component of organizational coupling, was more predictive ($\underline{R}^2 = .274$) of collective teacher efficacy ($\beta = .376$, $\underline{r_s}^2 = 68.89\%$) than individual personal teaching efficacy (PTE; $\beta = .229$, $\underline{r_s}^2 = 33.64\%$) and general teaching efficacy (GTE; $\beta = .183$, $\underline{r_s}^2 = 19.36\%$). This finding suggests that organizational coupling may indeed be an important variable in the study of collective teacher efficacy.

In a related study, Newmann et al. (1989) noted that the variance in teachers' own sense of efficacy in a given school can be conceptualized as a measure of efficacy consensus, and this within-group consensus may differ between schools. The authors argued that widely divergent perceptions of personal efficacy within a school may contribute to "divisiveness that may have the negative impact of reducing efficacy in the school as a whole" (p. 234). Therefore, Newmann et al. treated the within-school variance of individual efficacy as a consensus variable in the prediction of collective efficacy. Similar to Kurz's (2000) results, the group efficacy consensus yielded the strongest effect on collective teacher efficacy.

These early findings on collective teacher efficacy are compelling. If we assume that organizational systems can impact individual behavior and perceptions, and that individual behavior can impact organizational systems (as social cognitive theory does), then the role of collective teacher efficacy takes on increased importance within school systems. Can the collective belief of a faculty's capabilities foster improved teaching for the individual? What is the interrelationship between individual teacher efficacy and collective teacher efficacy? Once established, does collective teacher efficacy maintain itself with a positive feedback loop (cf. Putnam, 1993)? What are the dynamics by which organizational coupling impacts collective efficacy beliefs?

Impacting Teacher Efficacy Change

Given the current and potential educational value of the teacher efficacy construct, efforts to impact changes in teacher efficacy would be valuable in moving teacher efficacy research beyond the realm of correlational designs. Little experimental or long-term intervention research has been conducted in this area. As Ross (1994) noted, "In the absence of interventions it is difficult to tell whether teacher efficacy is a cause or a consequence of the adoption of more powerful teaching techniques" (p. 382). The limited number of studies in this area does suggest that teacher efficacy can be impacted by meaningful, active interventions (see e.g., Henson, in press; Ross, 1994).

Bandura (1997) cautioned that positive changes in self-efficacy only come through "compelling feedback that forcefully disrupts the preexisting disbelief in one's capabilities" (p. 82). Although there is consistent evidence that efficacy is most malleable in the preservice years (cf. Housego, 1990; Hoy & Woolfolk, 1990), efficacy tends to be resistant to change for experienced teachers (cf. Anderson et al., 1988; Ohmart, 1992; Tschannen-Moran et al., 1998).

Ross (1994), for example, found that general teaching efficacy increased after an eight month training on cooperative learning. Personal teaching efficacy, however, was stable across time.

Personal teaching efficacy appears to be particularly difficult to impact in experienced teachers since it is an internally held belief about oneself that solidifies with experience and time. As such, positively impacting teachers' efficacy beliefs is unlikely outside of longer-term professional development that compels teachers to think critically about their classrooms and behave actively in instructional improvement.

These professional development opportunities come in many forms. However, one particularly promising approach is via participatory teacher research, which has been suggested as one means of fostering meaningful professional development for teachers (Cochran-Smith & Lytle, 1990, 1999; Noffke, 1997). Participatory teacher research is a collaborative process by which teachers themselves critically examine their classrooms, develop and implement educational interventions, and evaluate the effectiveness of those interventions (Knight & Boudah, 1998). These activities allow teachers to actively participate in the development of practical knowledge about teaching. Teacher research models capitalize on critical thought and data-based action; social cognitive theory upholds such human agency as foundational to self-efficacy growth (Bandura, 1997).

In an academic-year long teacher researcher initiative in an alternative school, Henson (in press) reported large effects for personal and general teaching efficacy gains from pre to posttest (56.69% and 28.79%, respectively). Furthermore, collaboration was predictive of change in general teaching efficacy, which deals with teachers' perceptions of the whether teachers, in general, can overcome external barriers. As Henson explained, "The collaborative nature of teacher research may have provided otherwise unknown feedback concerning the effectiveness (or lack thereof) of other teachers."

Teacher efficacy research must grapple with the issue of efficacy change. Again, experimental studies are desperately needed. Longitudinal designs are also warranted. Current evidence suggests that teacher efficacy is indeed malleable, but that change will likely occur only via engaging and meaningful professional development opportunities, particularly activities such as teacher research initiatives that capitalize on teachers' critical thought and human agency.

Measurement Dilemmas and Advances

As has been noted, teacher efficacy research has been hampered by both construct validity and measurement problems. Some construct validity problems have been discussed above. The following discussion is intended to look more closely at the measurement weaknesses of historical teacher efficacy measures and highlight recent advances.

Historical Problems

For roughly 15 years, the TES served as the principal teacher efficacy instrument as documented by its frequency of use and its role in spawning other efficacy assessments. Unfortunately, not only were the theoretical operationalizations of the TES constructs questionable, but scores in Gibson and Dembo's (1984) original validation study were psychometrically weak. This is true despite the fact that the study was well designed in its development of items from teacher interviews, factor analytic derivation of scales, multitrait-multimethod matrix analysis, and empirical investigation using the newly obtained instrument.

Gibson and Dembo (1984) reported that two "substantial factors emerged from the factor analysis" (p. 573, emphasis added). However, substantial Factor I accounted for 18.2% of the matrix of association variance and substantial Factor II accounted for 10.6% of the variance, yielding a total variance-accounted-for of 28.8%. By most factor analytic standards (Gorsuch, 1983; Stevens, 1996), this variance-accounted-for is minimal at best, and is indicative of poor factorial validity (Thompson & Daniel, 1996). It is also well below the 53.02% average observed by Henson and Roberts (in press) in an empirical evaluation of exploratory factor analysis reporting practices.

Additionally, while most (but not all) discriminant and convergent validity coefficients in the multitrait-multimethod matrix were in expected directions, the magnitudes were not particularly indicative of strong construct validity. For example, the monotrait-heteromethod validity diagonal consisted of .30, .39, and .42, representing 9.00%, 15.21%, and 17.64% of shared variance, respectively. These coefficients would be typically characterized as low, which is problematic given that they are expected to be the largest validity coefficients in the matrix.

In a reliability generalization analysis of the PTE and GTE scales, Henson, Kogan, and Vacha-Haase (in press) observed considerable fluctuation of internal consistency estimates across studies (PTE: M = .778, SD = .057, Min. = .630, Max. = .890; GTE: M = .696, SD = .072, Min. = .550, Max. = .820). The reliability estimates were also generally low, especially for scores on the GTE scale, and none of the estimates reached the .90 mark.

In sum, the TES suffered from numerous psychometric infirmities, but found its way into, and entrenched in, the research literature nonetheless. Because of the exciting possibilities and compelling early results of teacher efficacy research, the TES was quickly adopted. It was, after all, published in a leading journal and was developed through recognized and respected methodologies. Unfortunately, the theoretical and psychometric weaknesses were overlooked, and researchers of teacher efficacy prematurely foreclosed on the instrument's developmental identity.

Importantly, Gibson and Dembo (1984) called for continued psychometric assessment of the TES. In their recommendations for future research, they noted that "construct validation should continue to be investigated across different populations and settings. Further factor analytic studies, including use of LISREL [or, confirmatory factor analysis] procedures, should be used to confirm further the trait and factor structure" (p. 579). Had researchers headed this advice, particularly the confirmatory procedures, the TES may have been revised appropriately or abandoned. Criticisms notwithstanding, the TES served an important, and perhaps necessary, role in generating research in teacher efficacy.

Ten years after its birth, the TES came under fire from both factor analytic (Guskey & Passaro, 1994) and discriminant validity (Coladarci & Fink, 1995) perspectives. These studies were discussed above and will not be revisited here. Additionally, other studies have reported divergent factor solutions for the TES (Guskey, 1988; Henson et al., 2000; Woolfolk & Hoy, 1990). These investigations served as an impetus sparking the reexamination of teacher efficacy's theoretical formulations and investigation of new ways to measure the construct.

Advances in Instrumentation

In light of the perceived weaknesses of the TES, several researchers have recently developed instruments that show promise in furthering the study of teacher efficacy.

The Ohio State Teacher Efficacy Scale (OSTES). Based on their theoretical model and the advice of Pajares (1996), Tschannen-Moran and Woolfolk Hoy (in press) sought to develop

an efficacy instrument that possessed correspondence to the tasks that teachers faced in school. They began with an unpublished teacher efficacy instrument developed by Bandura and then developed their own items as part of a graduate seminar in teacher efficacy. Focus was on inclusion of statements representative of frequent teaching activities. Initially, Tschannen-Moran (2000) argued for a one-factor solution for 32 retained items from a factor analysis. However, Roberts and Henson (2001) found poor fit for these items to one-, two-, and three-factor confirmatory factor analysis models. The three factor model demonstrated best fit, but the fit indices were not acceptable: GFI = .666, NFI = .675, TLI = .802, CFI = .814, and RMSEA = .086. Roberts and Henson recommended reduction of the number of items to a more parsimonious solution with stronger internal consistency.

In a later study, Tschannen-Moran and Woolfolk Hoy (in press) reported two rounds of principal axis factor analyses conducted on scores from independent samples. The final solution included 18 retained items that purported to measure three factors: Efficacy in Student Engagement, Efficacy in Instructional Practices, and Efficacy in Classroom Management. These factors explained 57% of the matrix of association variance and had interfactor correlations of .59, .60, and .64. Score reliabilities were .82, .81, and .72, respectively. Convergent and discriminant validity coefficients supported the construct validity of scores.

Because the interfactor correlations in the Tschannen-Moran and Woolfolk Hoy study (in press) were moderate, the authors conducted a second-order factor analysis. All three factors collapsed into one factor with pattern/structure coefficients ranging from .74 to .84. Tschannen-Moran and Woolfolk Hoy (in press) argued that the OSTES could be used for assessment of the three domains of efficacy or to yield a more generalized efficacy score. The OSTES employs a nine-point Likert scale. Sample items include: “To what extent can you influence the self-discipline of your students?” and “How much can you do to repair student misconceptions?”

Two points of caution are worth noting regarding the OSTES. First, the third factor was defined with only three items, with one item possessing a marginal pattern coefficient (.39) in one of the analyses. Second, and more importantly, the eigenvalues for the third factor were consistently borderline across analyses regarding their ability to survive a parallel analysis. Parallel analysis (Horn, 1965; Turner, 1998) has been shown to be among the most accurate methods for determining the number of factors to retain (Zwick & Velicer, 1986) and generally superior to the scree plot and eigenvalue greater than one rule, which were consulted by Tschannen-Moran and Woolfolk Hoy (in press) when making factor retention decisions. For four studies employing the 18 item OSTES, Tschannen-Moran and Woolfolk Hoy reported eigenvalues for the third factor as 1.40 (Study 2), 1.10 (Full Sample), 0.97 (Preservice teachers), and 1.24 (Inservice teachers). When comparing these eigenvalues to random eigenvalues derived from (n X 18) random data matrices for a parallel analysis (Thompson & Daniel, 1996), only the first study (Study 2) legitimately retained the third factor. The third eigenvalues in the other studies were all less than the random eigenvalues. Furthermore, had a parallel analysis been conducted, even the second factor in the Inservice teacher sample would not have been retained as the second eigenvalue (1.04) did not surpass the random eigenvalue. Therefore, caution should be exercised when using scores from this third factor and future studies should examine the factor’s invariance.

Self-Efficacy Teaching and Knowledge Instrument for Science Teachers (SETAKIST). Roberts and Henson (2000) developed a subject matter specific instrument for science teachers. Using the TES as a starting point, the authors essentially largely retained the personal teaching efficacy items, with the exception of rewording to reflect science content and elimination of past

tense verb uses. Because science education is explicitly involved with the pedagogical conversion of science information into a format meaningful for students, Roberts and Henson developed a knowledge efficacy construct, which is intended to roughly approximate efficacy for science pedagogical content knowledge (cf. Shulman, 1987). It was expected that these factors were related. An example of a teaching efficacy item is, "Even when I try very hard, I do not teach science as well as I teach most other subjects." A knowledge efficacy item reads, "I know the steps necessary to teach science concepts effectively."

Roberts and Henson (2000) tested their correlated two-factor model with a confirmatory factor analysis approach because of sufficient theoretical expectation for what factors would be present in the data. The hypothesized model and three competing models were tested, including a single-factor model, a three-factor model derived from modification indices obtained in the initial two-factor model, and an uncorrelated two-factor model. The hypothesized model yielded good fit to the data as indicated by several fit indices (CFI = .937, NFI = .876, TLI = .927, GFI = .917, RMSEA = .057). The one- and uncorrelated two-factor models did not indicate reasonable fit. Although the three-factor solution yielded estimates close to the two-factor solution, the two-factor model was deemed more plausible based on theory and for reasons of parsimony. As a caution, the SETAKIST requires additional validity evidence regarding the knowledge efficacy construct, given its attempt to assess an efficacy dimension formerly ignored in teacher efficacy research. However, the concept of assessing efficacy for pedagogical content knowledge is intriguing and worth further investigation.

Collective Teacher Efficacy Scale (CTE). As part of their research in collective teacher efficacy, Goddard and colleagues (Goddard, 1998; Goddard et al., 2000) developed a new instrument targeted at collective efficacy. Goddard et al. chose to use a group ("teachers can. . .") rather than an individual orientation ("I can. . .") to honor the social level of collective efficacy judgments. Furthermore, Goddard et al. explicitly devised the items to reflect positive and negative dimensions of the teaching task analysis and personal teaching competence, the two interrelated factors impacting efficacy beliefs in the Tschannen-Moran et al. (1998) model.

The instrument underwent several preliminary development stages, but ended up with 21 items that represented all four dimensions noted above. The authors argued for a single-factor solution that integrated the task analysis and teaching competence dimensions; the factor accounted for 57.89% of the matrix of association variance. Factor pattern/structure coefficients were quite strong, ranging from .61 to .93, and the coefficient alpha for the obtained scores was .96. Goddard et al. (2000) also conducted discriminant and convergent validity analyses and found reasonable support for their construct.

The CTE is important for at least two reasons. It is the first empirically developed collective teacher efficacy instrument, a definite advancement in the study of this emerging construct. Second, the CTE was developed on reasoned a priori theory, as opposed to a post hoc fitting of theory as was the case with the TES. Accordingly, the CTE included an integrated model of task analysis and teaching competence at the collective level. However, as a cautionary note, the CTE has not been tested across multiple samples. Confirmatory factor analytic methods would be useful in this regard.

Underutilized Methodologies

Because of the historical problems in teacher efficacy measurement and the current advances in instrumentation, there are several points of direction that developers of efficacy instruments should consider. My brief comments here are intended to facilitate stronger teacher

efficacy instruments, are not to be considered inclusive, and will focus on several methodologies that have been historically underutilized.

Comment on Factor Retention Rules

Exploratory factor analytic (EFA) strategies have dominated development of efficacy instrumentation. One critical decision, among myriad others, in EFA is how many factors to retain. This decision will affect the ultimate magnitude of the factor pattern and structure coefficients and, therefore, impacts the interpretation of extracted factors.

Many rules can be used to determine the number of factors to retain (cf. Zwick & Velicer, 1986), including the eigenvalue greater than one rule (cf. Kaiser, 1960), scree test (Cattell, 1966), minimum average partial correlation (Velicer, 1976), Bartlett's chi-square test (Bartlett, 1950, 1951), and parallel analysis (Horn, 1965; Turner, 1998). Thompson and Daniel (1996) and Zwick and Velicer (1986) elaborated these approaches. Importantly, these rules do not necessarily lead to the same conclusions regarding the number of factors present in a data set.

As with most EFA studies (cf. Henson & Roberts, in press), teacher efficacy studies have heavily depended on the eigenvalue rule and scree plot. However, Zwick and Velicer (1986) demonstrated that the eigenvalue greater than one rule almost always overestimates the number of factors. The scree plot is more accurate, but still leads to errant decisions in many cases. Parallel and minimum average partial analyses were demonstrated to be much superior, with parallel analysis getting the nod as the most accurate decision rule. Thompson and Daniel (1996) and Henson and Roberts (in press) called for consultation of multiple rules in factor retention decisions, with particular focus on parallel analysis results. Thompson and Daniel provided an SPSS program to automate parallel analyses. As evidenced in an empirical review of EFA reporting practices (Henson & Roberts, in press), parallel analysis is grossly underutilized in EFA methods; this is particularly true in teacher efficacy research.

Because the eigenvalue rule tends to overestimate the number of factors and interpreations of scree plots can be variable, it is possible that teacher efficacy research has attempted to extract too many factors, resulting in retention of weakly defined and theoretically weak factors suffering from a lack of invariance across studies. This may be true for the third factor in the OSTES (Tschannen-Moran & Woolfolk Hoy, in press) and other teacher efficacy instruments.

Using Confirmatory Methods

Because current conceptualizations of teacher efficacy are largely driven by rationale and theory, the utilization of confirmatory factor analytic (CFA) strategies are warranted. CFA allows researchers to test specific hypotheses regarding the structure of scores from instruments. Importantly, factor analysis results are a function of the scores obtained on an instrument and not the instrument only, and therefore can vary across time and place.

CFA methodology allows teacher efficacy researchers to test the factorial invariance of scores across samples. This process of repeated attempts at theory falsification is a time-honored tradition in theory development (Thompson and Daniel, 1996). As Moss (1995) noted,

> A "strong" program of construct validation requires an explicit conceptual framework, testable hypotheses deduced from it, and multiple lines of relevant evidence to test the hypotheses. Construct validation is most efficiently guided by the test of "plausible rival

> hypotheses" which suggest credible alternative explanations or meanings for the test score that are challenged and refuted by the evidence collected. . . . Essentially, test validation examines the fit between the meaning of the test score and the measurement intent, whereas construct validation entails the evaluation of an entire theoretical framework. (pp. 6-7)

CFA can be readily employed to test rival structures within data; and therefore is critical in theory development and evaluation of instrumentation used in assessing theoretical constructs.

Interesting, Gibson and Dembo (1984) called for continued validation procedures for the TES, and specifically referenced use of CFA approaches. However, this advice was not heeded in the research community and the TES moved forward without much challenge for ten years. With the exception of the studies by Roberts and Henson (2000, 2001), CFA applications in teacher efficacy research are almost non-existent.

Higher-Order Factor Analysis

Another avenue for measurement advancement lies with so-called higher-order factor analysis. Higher-order analysis examines whether first level factors can collapse into more general second- or even third-order factors. This methodology has been used in personality assessment for considerable time, noting the work surrounding the Big Five Factors of personality as an example.

Because efficacy judgments are said to be most accurate at reasonable degrees of specificity (Bandura, 1997; Pajares, 1996), teacher efficacy items should be task specific. It is possible that related tasks eventually find themselves clustered together in a factor. However, it is also possible, and almost completely ignored in the teacher efficacy literature, that these factors cluster into more parsimonious general factors. Recent work suggests that teacher efficacy can be conceptualized as a one-factor construct. Goddard et al. (2000) found one factor in collective teacher efficacy. Tschannen-Moran and Woolfolk Hoy (in press) observed three factors that collapsed into one second-order dimension (representing one exception to the lack of higher order analyses). Deemer and Minke (1999) argued for a one-factor solution to the TES after removing item wording confounds. Higher-order analyses have much to contribute to future assessment of the nature of teacher efficacy and instrument development.

Conclusion

Teacher efficacy research is roughly a quarter of a century old now. Early work suggested powerful effects from the simple idea that a teacher's beliefs in his or her ability to positively impact student learning are critical in actual success for failure in a teacher's behavior. These beliefs, called teacher self-efficacy, can be explained in Bandura's (1977) social cognitive theory, which emphasizes human agency within a multi-directional model that influences our thoughts and behaviors.

The study of teacher efficacy has suffered from poor construct validity issues. Relatedly, it has also suffered from foreclosure on instrument development before sufficient validation of scores across studies was evidenced. Recently however, there have been several important advances in the field on both substantive and measurement grounds. The teacher efficacy construct has underwent serious scrutiny and new theoretical models have emerged. Several promising instruments have been developed.

Among other issues, the advancement of teacher efficacy into the next stage of its developmental life would be fostered by empirical evaluation of the sources of efficacy building information, collective teacher efficacy, and methods for impacting efficacy change in teachers. Additionally, utilization of stronger measurement methodologies, including appropriate factor retention rules, confirmatory approaches, and higher-order analyses, would likely facilitate stronger instruments that yield more reliable and valid scores.

Given recent substantive theoretical advances and new, better founded, attempts at measuring this elusive construct, it seems apparent that teacher efficacy is ready to move beyond the adolescent angst it has been submitted to over the last few years. Assuming continued efforts are made to seek congruence between theory and measurement integrity, study of teacher efficacy just may be ready for a move into adulthood.

References

Allinder, R. M. (1994). The relationships between efficacy and the instructional practices of special education teachers and consultants. Teacher Education and Special Education, 17, 86-95.

Anderson, R., Greene, M., & Loewen, P. (1988). Relationships among teachers' and students' thinking skills, sense of efficacy, and student achievement. Alberta Journal of Educational Research, 34(2), 148-165.

Armor, D., Conroy-Oseguera, P., Cox, M., King, N., McDonnell, L., Pascal, A., Pauly, E., & Zellman, G. (1976). Analysis of the school preferred reading programs in selected Los Angeles minority schools (Rep. No. R-2007-LAUSD). Santa Monica, CA: RAND. (ERIC Document Reproduction Service No. ED 130 243)

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. Psychological Bulletin, 84, 191-215.

Bandura, A. (1982). Self-efficacy mechanism in human agency. American Psychologist, 37, 122-147.

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

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

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. Educational Psychologist, 28(2), 117-148.

Bandura, A. (1996). Self-efficacy in changing societies. New York: Cambridge University Press.

Bandura, A. (1997). Self-efficacy: The exercise of control. New York: W. H. Freeman.

Bartlett, M. S. (1950). Tests of significance in factor analysis. British Journal of Psychology, 3, 77-85.

Bartlett, M. S. (1951). A further note on tests of significance in factor analysis. British Journal of Psychology, 4, 1-2.

Cattell, R. B. (1966). The scree test for the number of factors. Multivariate Behavioral Research, 1, 245-276.

Cochran-Smith, M., & Lytle, S. (1990). Research on teaching and teacher research: The issues that divide. Educational Researcher, 19, 2-11.

Cochran-Smith, M., & Lytle, S. (1999). The teacher research movement: A decade later. Educational Researcher, 28, 15-25.

Coladarci, T. (1992). Teachers' sense of efficacy and commitment to teaching. Journal of Experimental Education, 60, 323-337.

Coladarci, T., & Fink, D. R. (1995, April). Correlations among measures of teacher efficacy: Are they measuring the same thing? Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Compeau, D. R., & Higgins, C. A. (1995). Computer self-efficacy: Development of a measure and initial test. MIS Quarterly, 19, 189-211.

Davis, F. W., & Yates, B. T. (1982). Self-efficacy expectancies versus outcome expectancies as determinants of performance deficits and depressive affect. Cognitive Therapy and Research, 6, 23-35.

Deemer, S. A., & Minke, K. M. (1999). An investigation of the factor structure of the Teacher Efficacy Scale. Journal of Educational Research, 93, 3-10.

Esselman, M. E., & Moore, W. P. (1992, April). In search of organizational variables which can be altered to promote increased sense of teacher efficacy. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Evans, E. D., & Tribble, M. (1986). Perceived teaching problems, self-efficacy and commitment to teaching among preservice teachers. Journal of Educational Research, 80, 81-85.

Garcia, M. E., Schmitz, J. M., & Doerfler, L. A. (1990). A fine-grained analysis of the role of self-efficacy in self-initiated attempts to quit smoking. Journal of Consulting and Clinical Psychology, 58, 317-322.

Gibson, S., & Dembo, M. (1984). Teacher efficacy: A construct validation. Journal of Educational Psychology, 76, 569-582.

Goddard, R. D. (1998). The effects of collective teacher efficacy on student achievement in urban public elementary schools. Unpublished doctoral dissertation, The Ohio State University, Columbus. (University Microforms No. UMI 9911194)

Goddard, R. D., Hoy, W. K., & Woolfolk Hoy, A. (2000). Collective teacher efficacy: Its meaning, measure, and impact on student achievement. American Educational Research Journal, 37, 479-507.

Gorsuch, R. L. (1983). Factor analysis (2nd ed.) Hillsdale, NJ: Erlbaum.

Graham, S., & Weiner, B. (1996). Theories and principles of motivation. In D. C. Berliner & R. C. Calfee (Eds.). Handbook of educational psychology (pp. 63-84). New York: Simon & Schuster Macmillan.

Guskey, T. R. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of instructional innovation. Teaching and Teacher Education, 4, 63-69.

Guskey, T. R., & Passaro, P. D. (1994). Teacher efficacy: A study of construct dimensions. American Educational Research Journal, 31, 627-643.

Hackett, G. (1995). Self-efficacy in career choice and development. In A. Bandura (Ed.), Self-efficacy in changing societies (pp. 232-258). New York: Cambridge University Press.

Henson, R. K. (in press). The effects of participation in teacher research on teacher efficacy. Teaching and Teacher Education.

Henson, R. K., Bennett, D. T., Sienty, S. F., & Chambers, S. M. (2000, April). The relationship between means-end task analysis and content specific and global self-efficacy in emergency certification teachers: Exploring a new model of self-efficacy. Paper presented at the annual meeting of the American Educational Research Association, New Orleans. (ERIC Document Reproduction Service No. forthcoming)

Henson, R. K., Kogan, L. R., & Vacha-Haase, T. (in press). A reliability generalization study of the Teacher Efficacy Scale and related instruments. Educational and Psychological Measurement.

Henson, R. K., & Roberts, J. K. (in press). Exploratory factor analysis reporting practices in published research. In B. Thompson (Ed.), Advances in social science methodology (Vol. 6). Stamford, CT: JAI Press.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. Psychometrika, 30, 179-185.

Housego, B. (1992). Monitoring student teachers' feelings of preparedness to teach, personal teaching efficacy, and teaching efficacy in a new secondary teacher education program. Alberta Journal of Educational Research, 38(1), 49-64.

Hoy, W. K., & Woolfolk, A. E. (1990). Socialization of student teachers. American Educational Research Journal, 27, 279-300.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. Educational and Psychological Measurement, 20, 141-151.

Knight, S., & Boudah, D. J. (1998, April) Participatory research and development. Paper presented at the annual meeting of the American Educational Research Association, San Diego.

Kurz, T. (2000). Collective teacher efficacy and organizational coupling. Manuscript in preparation. Texas A&M Univeristy.

Lent, R. W., & Hackett, G. (1987). Career self-efficacy: Empirical status and future directions. Journal of Vocational Behavior, 30, 347-382.

Maehr, M., & Pintrich, P. R. (1997). Advances in motivation and achievement (Vol. 10). Greenwich, CT: JAI Press.

Marlatt, A. A., Baer, J. S., & Quigley, A. A. (1995). Self-efficacy and addictive behavior. In A. Bandura (Ed.), Self-efficacy in changing societies (pp. 289-316). New York: Cambridge University Press.

Meijer, C., & Foster, S. (1988). The effect of teacher self-efficacy on referral chance. Journal of Special Education, 22, 378-385.

Midgley, C., Feldlaufer, H., & Eccles, J. (1989). Change in teacher efficacy and student self- and task-related beliefs in mathematics during the transition to junior high school. Journal of Educational Psychology, 81, 247-258.

Moore, W., & Esselman, M. (1992, April). Teacher efficacy, power, school climate and achievement: A desegregating district's experience. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Moss, P. A. (1995). Themes and variations in validity theory. Educational Measurement: Issues and Practice, 14(2), 5-12.

Newmann, F. M., Rutter, R. A., & Smith, M. S. (1989). Organizational factors that affect school sense of efficacy, community and expectations. Sociology of Education, 62, 221-238.

Noffke, S.E. (1997). Professional, personal, and political dimensions of action research. In M. Apple (Ed.), Review of Research in Education, Vol. 22 (pp. 305-343). Washington, DC: AERA.

Ohmart, H. (1992). The effects of an efficacy intervention on teachers' efficacy feelings. Unpublished doctoral dissertation. University of Kansas, Lawrence. (University Microfilms No. UMI 9313150)

Pajares, F. (1996). Self-efficacy beliefs in academic settings. Review of Educational Research, 66, 543-578.

Pajares, F. (1997). Current directions in self-efficacy research. In M. Maehr & P. R. Pintrich (Eds.), Advances in motivation and achievement (Vol. 10, pp. 1-49). Greenwich, CT: JAI Press.

Pintrich, P. R., & Schunk, D. H. (1996). Motivation in education: Theory, research, and applications. Englewood Cliffs, NJ: Merrill/Prentice Hall.

Podell, D., & Soodak, L. (1993). Teacher efficacy and bias in special education referrals. Journal of Educational Research, 86, 247-253.

Putnam, R. D. (1993). Making democracy work: Civic traditions in modern Italy. Princeton, NJ: Princeton University Press.

Riggs, I., & Enochs, L. (1990). Toward the development of an elementary teacher's science teaching efficacy belief instrument. Science Education, 74, 625-638.

Roberts, J. K., & Henson, R. K. (2000, November). Self-Efficacy Teaching and Knowledge Instrument for Science Teachers (SETAKIST): A proposal for a new efficacy instrument. Paper presented at the annual meeting of the Mid-South Educational Research Association, Bowling Green, KY. (ERIC Document Reproduction Service No. forthcoming)

Roberts, J. K., & Henson, R. K. (2001, April). A confirmatory factor analysis of a new measure of teacher efficacy: Ohio State Teacher Efficacy Scale. Paper presented at the annual meeting of the American Educational Research Association, Seattle. (ERIC Document Reproduction Service No. forthcoming)

Ross, J. A. (1992). Teacher efficacy and the effect of coaching on student achievement. Canadian Journal of Education, 17(1), 51-65.

Ross, J. A. (1994). The impact of an inservice to promote cooperative learning on the stability of teacher efficacy. Teaching and Teacher Education, 10, 381-394.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 80, 1-28.

Schunk, D. H. (1991). Self-efficacy and academic motivation. Educational Psychologist, 26, 207-231.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. Harvard Educational Review, 57(1), 1-22.

Soodak, L., & Podell, D. (1993). Teacher efficacy and student problems as factors in special education referral. Journal of Special Education, 27, 66-81.

Soodak, L., & Podell, D. (1996). Teacher efficacy: Toward the understanding of a multi-faceted construct. Teaching and Teacher Education, 12, 401-411.

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

Stevens, J. (1996). Applied multivariate statistics for the social sciences (3rd ed.). Mahwah, NJ: Erlbaum.

Thompson, B., & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. Educational and Psychological Measurement, 56, 197-208.

Tschannen-Moran, M. (2000, April). Developing a quantitative means for assessing teacher efficacy. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Tschannen-Moran, M., & Woolfolk Hoy, A. (in press). Teacher efficacy: Capturing an elusive construct. Teaching and Teacher Education.

Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. Review of Educational Research, 68, 202-248.

Turner, N. E. (1998). The effect of common variance and structure pattern on random data eigenvalues: Implications for the accuracy of parallel analysis. Educational and Psychological Measurement, 58, 541-568.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. Psychometrika, 41, 321-327.

Watson, S. (1991). A study of the effects of teacher efficacy on academic achievement of third-grade students in selected elementary schools in South Carolina. Unpublished doctoral dissertation, South Carolina State College, Orangebury. (University Microfilms No. UMI 9230552)

Watters, J. J., & Ginns, I. S. (1995, April). Origins of and changes in preservice teachers science teaching efficacy. Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco.

Woolfolk, A. E., & Hoy, W. K. (1990). Prospective teachers' sense of efficacy and beliefs about control. Journal of Educational Psychology, 82, 81-91.

Zimmerman, B. J. (1995). Self-efficacy and educational development. In A. Bandura (Ed.), Self-efficacy in changing societies (pp. 202-231). New York: Cambridge University Press.

Zwick, W. R., & Velicer, W. F. (1986). Factors influencing five rules for determining the number of components to retain. Psychological Bulletin, 99, 432-442.

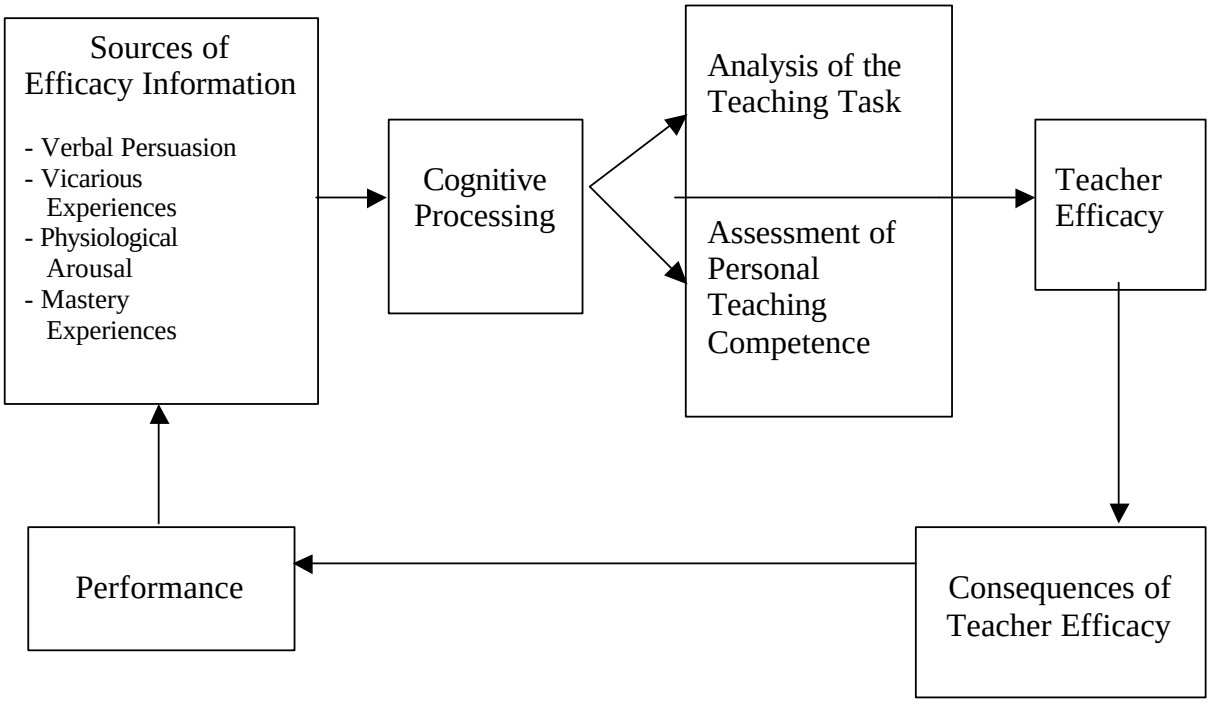

Figure 1. Multidimensional model of teacher efficacy. From "Teacher Efficacy: Its Meaning and Measure," by M. Tschannen-Moran, A. Woolfolk Hoy, and W. K. Hoy, 1998, Review of Educational Research, 68, p. 228. Copyright 1998 by the American Educational Research Association. Adapted with permission.