Models and Predictors of Teacher Effectiveness: A Review of the Literature with Lessons from (and for) Other Occupations

Douglas N. Harris
Stacey A. Rutledge

January 1, 2007
Abstract:

This study compares research on the theoretical models and predictors of teacher effectiveness with those of other occupations, focusing on three specific predictors of worker effectiveness: cognitive ability, personality, and education. The comparison of the teacher and other worker studies yields a variety of ways in which research on teacher effectiveness might be improved and expanded: First, the worker literature illustrates specific theoretical models, such as job-organization fit, that complement existing models regarding the work of teachers. The potential value of extending worker models to teaching in this way is reinforced by the fact that the three teacher characteristics mentioned above predict effectiveness in similar ways among teachers and other workers. Second, by outlining multiple models of effectiveness, it is possible to identify the important dimensions on which they vary, such as the unit of analysis and the assumed roles of the individual worker in relation to the organization. Third, research on other workers highlights some ways to improve the measurement of the three predictors and teacher effectiveness, going beyond the use of student test scores. While we advocate no particular theoretical model of teacher effectiveness, these findings provide guidance for how future research might build on current knowledge.

Acknowledgements: This work is supported by a Teacher Quality Research Grant from the United States Department of Education, Institute for Education Sciences. The authors wish to thank William Anthony, Robert Floden, David Ketchen, David Monk, Andrew Shouse, and Tim Sass for their valuable comments and William Ingle and Cynthia Thompson for their able research assistance. The statements made and views expressed are solely the responsibility of the authors.

Corresponding Author: Douglas N. Harris, Department of Educational Policy Studies, 213 Education Building, 1000 Bascom Mall, Madison, WI 53706. Phone: 608-262-4406; Email: dharris3@wisc.edu.
Introduction

There is wide agreement among researchers and educators that teachers play an important role in the lives of students—that students will fair better if placed in a classroom led by Teacher A than in one led by Teacher B. But what characteristics distinguish the more effective teachers from others? And why do these characteristics matter? In other words, what is it about specific characteristics that lead to greater effectiveness? These are some of the most fundamental questions in educational research and are related to nearly every major aspect of educational systems, including teacher training, certification, hiring, and evaluation. They are also the basic questions addressed in the present study.

We review the research literature regarding the characteristics of effective teachers, using and building on other recent reviews. These include a wide variety of perspectives (from economics, sociology, psychology, and others), as well as different theories and conceptualizations of effectiveness, measures of teacher effectiveness and teacher characteristics, and empirical methods. Unlike previous reviews, we integrate our review of empirical evidence with a theoretical discussion of the modeling of teacher effectiveness, identifying prevailing models and considering what the empirical evidence implies about model validity.

But perhaps the more important contribution of this review is that it juxtaposes the teacher literature with that of worker effectiveness in other occupations. One reason for doing so is that the literature on non-teaching occupations is quite extensive and incorporates a wider range of research perspectives. Second, the comparison highlights the distinctive features of teaching as an occupation, such as the level of job complexity and the different types of personal skills that teachers must master in order to interact effectively with students, parents, fellow teachers, and administrators. Third, there is a storied history of trying to make schools run more like businesses, but little evidence about how the effectiveness of teachers actually compares to workers in private sector industries. While other studies have used evidence from other occupations as background for empirical analyses (see, for example, Rowan, Chiang, & Miller, 1997), this is, to our
knowledge, this is the only study to consider non-teaching as part of a comprehensive effort to develop future directions for teacher research.

From this review and synthesis, we identify three main directions for teacher research. First, the worker literature illustrates models that complement existing theories regarding the work of teachers. For example, the job-organization fit model locates the worker within the organization and focuses on the important relationships between them. The potential value of extending worker models to teaching in this way is reinforced by the fact that some worker characteristics seem to predict effectiveness equally well in teaching compared with other occupations. We focus particularly on three potential predictors of effectiveness—cognitive ability, personality, and education—which are perhaps the only three factors for which considerable evidence has been paid in both sets of research literature. While the similarity in empirical findings may reflect a general underlying model of effectiveness that applies across occupations, we also argue that there is likely a set of occupation- or domain-specific predictors that such general models cannot incorporate.

Second, by outlining multiple models of effectiveness, it is possible to identify the important dimensions on which they vary and therefore to be more precise about the assumptions the models make. For example, in describing the models of worker effectiveness below, we find that models vary in their unit of analysis and the implied importance given to organization-level differences. Few models of teaching give much attention to the organizational context as a mediator of teacher effectiveness even though there is evidence that administrators emphasize the organization in their personnel decisions (Harris, Rutledge, Ingle & Thompson, 2007). The choices made along these dimensions, in both teaching and other occupations, partly reflect the different disciplinary perspectives of the researchers.

Third, research on other workers highlights some ways to improve the measurement of the three predictors and teacher effectiveness. As we will show, research during the 1960s focused on principal evaluations of teachers as the measure of effectiveness. More
recently, studies have focused almost entirely on student test scores which, while useful, also constrain the types of models of effectiveness that can be tested. The principal evaluation remains an alternative instrument that, with some improvements in design, could broaden the types of models that can be tested.

While we advocate no particular model of teacher effectiveness, these directions may also address certain criticisms of existing research. In particular, it has long been argued that there is a disconnect between theory and evidence in research on teachers (e.g., Getzels & Jackson, 1963; Wallen & Travers, 1963), a contention we find some evidence to support. By showing how models from other occupations might be extended and focusing attention more clearly on the dimensions of the teaching models and the measures of predictors and teaching outcomes, this study helps to address this theory-evidence disconnect.

In section II, we define our key terminology—effectiveness, predictors, and models—and show their relationships to one another. We also identify dimensions that differentiate the different models in the research on worker effectiveness. This sets up, in section III, a review of the common models of worker effectiveness, measures of effectiveness and commonly considered predictors. A similarly structured review of research regarding teachers is provided in section IV. A comparison and synthesis of the two separate literatures is in section V and includes further discussion of the above directions for future research.

**Terminology and Typology of Research Approaches**

*Key Terms*

There are three central and interrelated terms that guide this comparison between the teacher and non-teacher literatures: effectiveness, worker characteristics, and models of effectiveness. We define effectiveness as the degree to which workers produce outcomes related to the objectives of their organizations. This term minimizes implied biases toward a particular research perspective, compared with terms like “productivity” that
are closely associated with the perspective of economists. Also, we deliberately avoid the term “performance” because this is sometimes defined as effectiveness but, in other cases, as the specific worker practices that are associated with effectiveness. While practice and effectiveness are obviously related to another, this is a somewhat different topic and is beyond the scope of the present study.

The definition and conceptualization of effectiveness is also closely related to how it is measured. One common distinction is between objective and subjective measures. The findings here show that recent research on teachers relies almost entirely on an objective measure, the student test score, while research on other workers relies on a subjective measure, the supervisor evaluation. Moreover, the focus on student test scores in research is almost a complete reversal from past decades of teacher research, suggesting that the models of teacher effectiveness used by researchers have changed in recent years, in tune with the larger shift toward student achievement as the central focus of education and educational decision-making.

Research on effectiveness has also sought to identify characteristics of workers that are associated with the above measures of effectiveness. Here, we focus on three specific worker characteristics—cognitive ability, personality, and educational background—because they are frequently used in studies of both teachers and other workers. Following the existing occupational literature, we refer to these characteristics as “predictors” of effectiveness.

A model of effectiveness, the third and overarching term, establishes a theory-driven relationship between effectiveness and worker characteristics. Such models begin with a general theory of work and then illustrate the theory’s practical implications, including hypothesized relationships between specific types of effectiveness and worker characteristics. These hypothesized relationships can then be tested empirically and the model can be adapted accordingly. Or put differently, one can test whether the hypothesized “predictors” actually predict effectiveness and thereby test the validity of the models. As we will see below, the models of effectiveness in teaching are lacking in
some important ways.

Because we are synthesizing research on models of effectiveness from different research literatures and sources, it was necessary to perform multiple parallel searches for studies. We performed standard searches of research databases as well as manual searches of journals that were most likely to include studies on the respective topics. We also solicited studies from scholars of effectiveness in non-teaching occupations. Because the methods varied somewhat depending on which research literature was being considered, the more specific search methods and inclusion/exclusion criteria are discussed in footnotes at the beginning of each section. Also, because the purpose of the study is to characterize and compare the two literatures, we focus especially on many excellent reviews in the existing literature.

**Typology of Models**

In reviewing the studies, we compared models of effectiveness and identified their similarities and differences. We find only a few studies that have tried to distinguish the different general types of models (Rowan, 1990; Scott, 2004; Stewart & Carson, 1997), although we find important limitations in these previous attempts. Thus, the categories in table 1 are not based on previous research, but rather on the authors’ own observations and synthesis.

| TABLE 1 |

As suggested by the first row of table 1, the research approaches are naturally determined by the backgrounds of the researchers. In economics, the individual person is the unit of analysis and effectiveness is defined by worker productivity, such as dollar value of output per hour. We consider this a “final outcome” in the sense that it relates closely to the ultimate objective of the organization (in this case, profit). Psychology-based studies also focus on individual workers and often use final outcomes, but the more
distinctive feature of these studies is that they break down work into sets of well-defined tasks or job requirements for individual workers. In this sense, both economics and psychology focus on the individual worker and emphasize the importance of technical or “hard” skills.

Sociology and organizational behavior focus instead on processes and the relationships among individuals that facilitate effectiveness. Effectiveness is more a by-product of positive relationships and positive work environments and these studies almost always define effectiveness in terms of “intermediate outcomes” that are indirectly related to the organization’s main objective (e.g., the level of worker turnover) and indicative of the quality of the work environment. By focusing on relationships, these models tend to emphasize the importance of workers’ soft skills, such as communication, leadership, and ability to work with others. They also frequently discuss the instability and dynamic nature of the external work environment, emphasizing the importance of another soft skill, worker adaptability.

The above dimensions serve four specific purposes in the analysis that follows: (1) to categorize the specific models discussed in the teaching and non-teaching literatures and therefore to unify the discussion; (2) to illustrate the distinctive features of teaching as an occupation and as an object of research; (3) to clarify the conceptual models underlying the empirical analyses of the predictors of teacher effectiveness and therefore to help interpret the meaning of the empirical results; and (4) to classify different models according to the specific dimensions and therefore to use the dimensions to provide possible new directions for models of teacher effectiveness. To achieve these purposes, we now turn to a discussion of models of worker and teacher effectiveness, respectively, and describe these models along the same dimensions shown in table 1.

Worker Effectiveness: Models, Measures and Predictors

We began this study by providing evidence that student outcomes vary depending
on whether students are assigned to Teacher A versus Teacher B. The evidence supporting this conclusion is arguably even stronger when considering research on Worker A and Worker B. For instance, Schmidt and Hunter (1983) find that the average worker produces 40 percent more output (per dollar of salary) compared with a worker who is one standard deviation below the mean effectiveness. Next, we discuss how this variation in worker effectiveness might be explained by specific models and predictors of effectiveness.

Models of Effectiveness

Our review revealed three specific models of worker effectiveness that were both frequently cited and which seemed to have particular potential for learning about teacher effectiveness. The “classic” model of worker effectiveness from the industrial psychology literature is arguably the most dominant in studies of non-teaching occupations. This model posits that the worker effectiveness is determined by job knowledge (Schmidt & Hunter, 1983), which, in turn, is a function of cognitive ability and, to a lesser extent, on job experience. The logic of the model is that cognitive ability allows workers to learn necessary information and skills more quickly and accurately, while job experience gives workers more time for such learning to take place (Schmidt, Hunter, & Outerbridge, 1986).

This industrial psychology model is clearly task-oriented and places little emphasis on the role of the work environment. As Cohen and Pfeffer (1984) write, “there is an implicit assumption in much of the literature on jobs and job structures that there is some underlying technical imperative that causes the employment relationship and occupational structures to look the way they do” (p.2). In other words, there is something inherent about each occupation that requires carrying out particular tasks, irrespective of organizational characteristics or properties of the work environment. Other researchers have echoed this interpretation of the classic model. Thus, in the case of nurses, the industrial psychology model would suggest that the ability to take a
patient’s blood pressure and correctly identify dangerously high levels, is more important than the amount of empathy shown to patients or to the ability to communicate and coordinate with doctors to solve the patient’s problems.

Two additional models of effectiveness focus on the “fit” between the person and the job. The “person-job fit” (P-J) or “job compatibility” model assumes worker effectiveness depends mainly on how well the specific strengths of the worker match the job requirements. An implicit assumption behind this is that there are no “good” and “bad” worker characteristics per se, but only those that match the needs of the job. Thus, some workers who have high cognitive skills may still not be a good fit for specific jobs, especially those jobs that require little job knowledge. In these cases, cognitively skilled workers are more likely to become disengaged from their work and to have greater opportunities for compensation and other rewards by switching jobs.

“Person-organization fit” (P-O) models take this one step further by positing that organizations have cultures and values and that the degree to which these match the preferences and values of specific workers can affect workers’ success and likelihood of staying on the job. P-O fit is also important because the job of each worker is likely to evolve and change over time, as the needs of the organizations change. From this perspective, it is more important that each worker fit the personality of the organization, which is unlikely to change quickly, rather than fit the job, which is likely to change more rapidly.

Werbel and Gilliland (1999) argue that the importance of P-J fit and P-O fit depends on the job. Specifically, P-J fit is relatively more important when technical job requirements are substantial and P-O fit is relatively more important when the organization has a distinctive culture, the career ladder is lengthy, the interaction among organization members is great, and work flexibility is high. Put differently, P-O fit matters more than P-J fit when the worker is in an organization that involves more interaction between people or when the workers themselves will have a variety of jobs over time.
Table 2 applies the general approaches and methods discussed in table 1 to the three specific models discussed above. It is important to note that the assumptions and implications listed in the first column vary in the degree to which they capture the key aspects of each model. The P-J fit model is especially poorly captured, as indicated by the large number of “unclear” results, meaning that the models are not defined in such a way that a clear determination can be made. The reason is that fit models make few assumptions about the person and job characteristics that define a good fit. Thus, a model that compares a job’s required interactions with other workers and the worker’s extroversion represents as much a model of fit as one that compares a job’s level of technical skill with a worker’s cognitive skills. This notion of fit therefore may be too generic for the development of theory and empirical evidence.

[TABLE 2]

**Effectiveness Measures**

Very little attention is paid to defining effectiveness and we suspect that most choices of effectiveness measures are chosen for reasons of convention and convenience. The supervisor evaluation is the most common measure of effectiveness in research studies on other occupations. Borman, Ferstl, Kaufman, Farmer, & Bearden write that this represents “arguably the most comprehensive measure of overall performance” (2003, p.290). These authors also indicate, however, that too little attention has been paid by researchers of worker effectiveness to identifying and understanding components of the supervisor evaluation or the evaluation rubric.

Suggesting a move toward more objective measures of effectiveness, Borman et al. also point out that “a recent trend has been to study job performance in its own right in an attempt to develop substantive models of performance” (2003, p.299). Some of the objective measures used in the literature are organizational productivity, customer satisfaction, team/business performance, and firm profit (Ployhart, 2004). Studies using
the “fit” models tend to focus especially on turnover intention and actual turnover, as well as worker satisfaction and commitment. Schmitt, Gooding, Noe, & Kirsch (1984) find in a review of 99 studies that empirical studies using objective measures have better explanatory power than those using subjective measures. But, again, the supervisor evaluation remains the most common measure of worker effectiveness.

Worker Characteristics as Predictors

We focus our review on three predictors of worker effectiveness that are commonly studied in the research literatures on both teachers and other workers: cognitive ability, personality, and educational background.

Cognitive Ability

Cognitive ability is both the most commonly studied factor and the one most widely believed to be an important factor affecting job effectiveness. There is, of course, considerable debate about the conceptualization and measurement of cognitive ability, referred to alternatively as intelligence, “g,” general mental ability and general cognitive ability. Measures of cognitive ability include questions regarding verbal analogies, figure classification and number series (Efklides, Papadaki, Papantoniou, & Kiosseoglou, 1997). Some of this debate centers on the level of generality; that is, the degree to which certain abilities are evident across knowledge domains and tasks, as opposed to being domain- or task-specific (Demetriou & Efklides, 1994). Several recent studies have found that general cognitive ability is a better predictor of worker effectiveness than measures of domain- and task-specific abilities (Borman et al., 2003; Prediger, 1989). Also, while cognitive ability is sometimes assumed to be fixed or innate, there is substantial evidence that it is influenced by environmental factors and can change over time (Petrill et al., 1998).

Results of studies of cognitive ability find that it is a consistent and relatively strong predictor of worker effectiveness. Schmidt and Hunter (1998) review studies that use one or more of 18 possible predictors, including integrity tests, peer ratings, and
others. They also include studies that have a wide variety of effectiveness measures, including wages and measured output. Their results suggest that cognitive ability is the worker characteristic that best predicts effectiveness, consistent with the hypothesis of the classic industrial psychology model.

Schmitt et al. (1984) find very similar results and are more explicit about the effectiveness measures used in the 99 studies they reviewed. Performance ratings, turnover, achievement/grades, productivity, wages, and status change (presumably promotions and demotions) are all considered as effectiveness measures. They too find that cognitive ability is the best predictor of effectiveness.

We found two articles that focused on how the role of cognitive ability and other worker characteristics varied by job type. Hunter (1983) performed a meta-analysis of more than 500 studies. After dividing the occupational spectrum into five levels of “complexity,” he shows that cognitive ability is the best predictor for three occupational levels with the highest complexity and still a significant predictor for the lower two levels. Second, Schmitt et al. (1984) report the average predictive ability for all predictors and effectiveness measures across occupational groupings, finding that effectiveness in clerical, managerial and professional occupations is more predictable than unskilled labor, skilled labor, and sales.

**Personality**

Personality is another common predictor of job effectiveness and interest in the topic is growing. Day, Bedeian, & Conte (1998) write that “a common problem with personality research, in general, and with predicting job-related criteria [effectiveness] in particular, is dealing with the vast array of available personality measures” (p.2070). One framework that is now widely accepted in the literature is the “Big Five” (Wright & Boswell, 2002, p.255), which measures conscientiousness, emotional stability, extroversion, agreeableness, and openness (Goldberg, 1990). Borman et al. (1991) and Wright and Boswell (2002) find that personality factors add considerably to the
explanatory power in studies of worker effectiveness. Barrick, Mount, and Judge (2001) find that conscientiousness is the best predictor of the five.

Results from other studies suggest a weaker and less well-defined role for personality. Schmitt and Hunter conclude that personality is a weak predictor, although they also find that conscientiousness is still the best of the personality predictors (1998, p.272). Others researchers, such as Wright and Boswell (2002) have argued that the role of conscientiousness is “equivocal” (p.256), although they also find that it predicts outcomes such as task effectiveness, career success, and employment status.

Research on P-O fit focuses on the relationship between characteristics (cognitive ability, personality and others) and organizational characteristics. The fit between the individual and the organization is then correlated with measures of organizational characteristics and organization-level effectiveness. Much of this literature, while relatively new, focuses on the relationship between worker personality and the types of organizations they are drawn to (Borman et al., 2003). For example, Chatman (1991) studies P-O fit model and measures organizational characteristics using the Organizational Culture Profile (p.465), which includes 54 items organized into nine categories. She finds that job candidates in accounting firms whose values are more similar to the organizations’ values adjust to the job more quickly. These workers also “feel most satisfied and intend to actually remain with [the organization] longer” (p.459).

The importance of studying effectiveness at the organization level, as in the P-O model, is highlighted by the work of Ployhart (2004) who argues that different levels of aggregation are associated with different theories of effectiveness. He finds that it is not necessarily the case, for instance, that organizations whose workers have higher average levels of cognitive ability will necessarily be more effective as organizations, even if the individuals themselves are more effective in their narrowly defined jobs. In another study, Page (2001) simulates decisions made by hypothetical computer-generated “participants” that have varying types of skills. He finds that groups of these participants
that have diverse sets of skills make better decisions than groups with less diversity, even when the latter group has higher levels of average cognitive ability. This suggests that having a diversity of certain characteristics may be a more important goal than having workers who all have a specified characteristic (e.g., extroversion) that is desirable “on the average.” Indeed, it may even be the case that some worker characteristics are unimportant on the average, but that organizations need a certain small proportion of workers with certain characteristics. A similar point is made by Muchinski and Monohan (1987) who distinguish between “complementary fit” between workers and organizations in which it is desirable to have a more homogenous set of workers, from “supplementary fit” in which it is beneficial to have diversity. In short, characteristics that predict individual effectiveness may not predict organization effectiveness, even if the assumptions about individual predictors are valid.

**Educational Background**

In contrast to cognitive ability and personality, education is one of the factors that can be readily used as an instrument for worker selection and as a means of improving effectiveness for workers already hired. Here, we focus on educational degrees and do not consider forms of licensure and certification forms of credentials that are less prevalent, and therefore less often studied, outside of teaching.

There seems to be almost universal agreement in the industrial psychology and related literatures that educational credentials have little predictive ability. Hunter and Hunter (1994) indicate that background credentials such as education have little predictive power. Schick and Kunnecke (1982) find that personality is a better predictor of future success on the job than academic achievement or school attended. Also, Ariss and Timmins (1989), in their public sector study, find no effect of education level or type of degree on supervisor ratings. However, compared with cognitive ability and personality, studies of educational background have been rare (Ariss & Timmins, 1989). Indeed, the extensive reviews by Borman et al. (2003), Schmidt and Hunter (1998),
and Schmitt et al. (1984) do not even consider education among the lists of possible predictors.

Economists, on the other hand, have focused a great deal of attention on the role of education, through the vast “return to schooling” literature. These studies use worker wages as the measure of effectiveness (productivity) and education levels represent one of the key explanatory variables. Such studies generally find that an additional year of schooling is associated with a 5-10 percent increase in annual earnings (Krueger & Lindahl, 2001). This economics research would seem to contradict the industrial psychology and related studies that either find no relationship between education and effectiveness or simply assume that no relationship exists.

The meaning of this evidence is ambiguous, however. In particular, there is debate among economists about whether the economic return to schooling reflects actual learning (human capital) or whether those people with more education simply have traits that led them to get more education—traits that are also positively associated with worker effectiveness. The latter “sorting” interpretation is supported by Weiss (1995) who finds, for example, that workers who have more education also have greater “perseverance.” If people with more perseverance obtain more education, then it becomes unclear which is the real underlying cause of the economic return to education.

Cognitive ability, more so than perseverance, is the worker characteristic most commonly considered as one that confounds the interpretation of the effects of education. Krueger and Lindahl (2001) address this by reviewing studies that include both educational background and the SAT score, a measure of pre-college ability and one that is highly correlated with cognitive ability (Efklides et al., 1997). While the SAT score does appear to play a statistically significant role, this reduces the economic return to education by only about 10 percent (e.g., a return of 8 percent decreases to 7.2 percent). The fact that controlling for this important worker characteristic has little effect on the measured return to education suggests, in contrast to Weiss, that the human capital interpretation is the most valid.
Does the interpretation matter for the purposes here? At one level, we can say that the distinction between the human capital and sorting explanations is irrelevant. Of course, education is correlated with other important worker characteristics and the same could be said of cognitive ability and personality. However, if education is mainly reflecting other worker characteristics then this would present a more fundamental problem. This suggests a need to more clearly theorize about what workers need to know to be effective and to try and measure these specific types of education more clearly.

We have now introduced several models of effectiveness as well as the measures of effectiveness and three worker characteristics that are often considered to predict effectiveness. The most important observation we wish to make at this point is that the models of worker effectiveness are usually (with the exception of the above discussion of education) closely related to the empirical evidence. For example, the classic model from industrial psychology, still the most widely cited, starts with a clear and concrete theory of worker effectiveness, namely, that knowledge is an important aspect of work and that workers with greater cognitive ability learn the necessary job knowledge faster and better than other workers. This theory is associated with a hypothesis that workers with high cognitive ability are more effective and a substantial body of empirical research supports this hypothesis. As we will see below, the same type of connection between theory and evidence is rarely found in research on teachers.

**Teacher Effectiveness: Models, Measures and Predictors**

In this section we provide a similarly structured review of evidence related to teachers, beginning with a discussion of teacher effectiveness models, followed by teacher effectiveness measures and predictors. So, whereas the earlier section provided macro-level findings across occupations, this section delves into specific conceptions and measures of the occupational characteristics of teaching.

Because teaching is a specific occupation, it is useful to start by considering its
distinctive features, which need to be accounted for in models of effectiveness. This is not an easy task as conceptions of teaching, as will be discussed in the models section, differ. Teaching is largely an occupation in which teachers function both within their classroom and as a member of the larger school organization, something that has been identified for teachers as a dual allegiance to both the school and students (Jackson, 1968). Models of teaching place different emphasis on the relationship between students, teacher, and the school administration. Central to these conceptions are questions of the types of decisions that teachers make in the classroom, the nature of their knowledge base, and finally, the degree of control that the administration has over the teaching task. All emphasize what Lortie (1975, p.117) calls the “interpersonal transactions and states which teachers realize with their students.”

There is evidence that the nature of the teacher-school linkage differs by individual school (Rowan, Chiang, & Miller, 1997), grade level (Firestone & Herriot, 1982; Herriot & Firestone, 1984), school size (Bidwell, Frank & Quiroz, 1997) and policy environment (Coburn, 2004). So while discussions of models emphasize different aspects of teaching, this additional evidence suggests that the school—and, in turn, the organizational context—are important factors as well. This implies further that a person-job or a person-organization fit could be found in teaching, though this line of inquiry has not been actively pursued by researchers on teacher effectiveness.

Rowan (1994) compares the complexity of teaching to other occupations using data from the Department of Labor and finds that teaching ranks second highest on a seven-point scale in the level of complexity involved when dealing with data and people, which, given the previously discussed evidence on workers in general, suggests that cognitive ability is likely to be an important predictor of teacher effectiveness. Teaching also ranks second highest on a scale of five in using reasoning and language and in the middle on the same scale for mathematics. Taken together, these findings provide support for the importance of cognitive ability, personality, and education in handing the complexity associated with teaching (Rowan, 1994).
While the work of teaching is complex, there is no consensus on what constitutes “best” teaching, but rather multiple and sometimes conflicting strategies reflecting broad curricular and instructional debates as well as the beliefs and values of the individual teacher (Johnson, 1990). Further, differences between the age, academic level, and needs of students mean that teaching requires different skills and knowledge in different contexts. These multiple contexts underscore that effective teaching is neither simple nor fixed.

Models of Teacher Effectiveness

There are at least four prevailing models, or conceptions, of teachers’ work found in the research literature: teaching as labor, teaching as craft, teaching as profession, and teaching as art (Darling-Hammond, Wise & Pease, 1983; Mitchell & Kirchner, 1983; Rowan, 1994). These conceptualizations focus on different characteristics of the school as an organization and the role of the teacher within that organization. Each has an implied model of effectiveness based on its particular characterization of teachers’ work. None consider how the external environment shapes the model. Still, they offer important conceptions showing different priorities and philosophies of teaching. Here we discuss these four models and associated empirical research.

In teaching as labor conceptualization, teachers’ work is understood to occur within the school system hierarchy and is subject to oversight by administrators. At the classroom level, teachers adhere to external demands as well as school-level routines and procedures that can be measured and evaluated. Teaching is seen as a rational activity and teachers are seen as adaptable to new theories and external circumstances (Firestone & Bader, 1991).

There are at least two branches of educational research that utilize—and in some cases empirically support—the teaching as labor view. Responding to organizational theory that assumes members of organizations engage in behavior and decision-making to maximize organizational productivity, early studies of school organization focused on
the bureaucratic components of the work (Anderson, 1968; Bidwell, 1965; Eddy, 1969; Lortie, 1969; Metz, 1978; Rogers, 1968). These studies identified bureaucratic aspects of schools including a division of labor with staff assigned to specific instructional and managerial tasks, job recruitment based on competence and merit, a hierarchy in the status of jobs, and routinized procedures and a rationalization of tasks for administrators and teachers (Bidwell, 1965). At the same time, they described schools as characterized by a structural looseness or a loose coupling with the teaching technology semi-autonomous from the administrative structure (Weick, 1976). Many of the policy efforts of the last twenty years including standards and assessments, school sanctioning, and merit pay have both drawn and relied on this conceptualization in policy implementation. School sanctioning, for example, is predicated on assumptions that the policy approach will lead to organizational alignment by principals, and teachers around the high stakes assessment (Debray, Parson, & Avila, 2003; Mintrop, 2003).

The resource allocation research (outside of economics) provides an additional example of the teaching as labor model. This approach contends that different levels of the school system have a “conditional and contributionary relation to events and outcomes occurring at adjacent ones” (Barr & Dreeben, 1983, p.7) and focuses, among other things, on the way in which teachers make decisions about the allocation of time, materials and students within classrooms. Empirically, these studies demonstrate that some teachers are more likely than others to allocate similar resources in ways that raise student achievement scores (Barr & Dreeben, 1983; Gamoran & Dreeben, 1986). For example, in a study of fifteen first grade classrooms, Barr and Dreeben find that teachers vary in their decisions about how to use the same curricular and instructional materials and how to group students and that these decisions seem to be associated with higher student achievement gains (controlling for initial student achievement). Rowan and Miracle (1983) also find that the way that teachers implemented their curriculum affected student achievement. In all of these studies, there is an implicit assumption, consistent with the teaching as labor view, that some decisions are “better” than others.
and that more effective teachers are the ones who make the better decisions most often. While these studies make a clear connection between theory and evidence, they do not try to identify the specific characteristics of the individual teachers that made better decisions about student grouping and reading instruction.

By contrast, in the conception of teaching as a profession, teachers’ work is understood to include a technical knowledge as well as an exercise of judgment as to how to apply this knowledge with students. As teaching is highly uncertain, it requires professional judgment in which teachers draw on their formal knowledge base as well as their knowledge of learners (Firestone & Bader, 1991; Shulman, 1987). Teachers’ knowledge base is gained through education, experience, and meaningful professional development (Darling-Hammond, 1998). Further, teachers’ work occurs within a community supported by administrators and peers who provide tools and support for effective work (Talbert & McLaughlin, 1994). An effective teacher under this model draws on a strong knowledge base, as well as in collegial relations with peers, to deal with the uncertain environment of the classroom.

Again, there are many branches of research that use and support the teaching as profession view. Teachers in schools with high levels of professional involvement are more likely to participate in on-going continuous improvement (Little, 1982). Schools have networks of teachers leading to different patterns of communication and interaction (Siskin, 1995). Teachers make policy implementation decisions through formal and informal interactions with networks of colleagues within the school (Coburn, 2001). A series of studies have also examined how subject matter identification shapes teachers’ work in high schools. This research has found that teacher networks shape, in large part, their support networks and their collegial interactions (Stodolsky, 1984; Stodolsky & Grossman, 1995). At the high school level, departments serve as a primary locus of organization for teachers defining their professional community and their values (Siskin and Little, 1995). The subject matter field also shapes, in large part, teachers’ attitudes about the technology of teaching, particularly conceptions of how material should be
presented and the types of instructional activities that should be used.

Rowan (1994) lends support to the teaching as profession model in his previously cited study comparing the complexity of teachers’ work to other occupations. Given findings on the relative complexity of teaching compared to other occupations, Rowan writes that “the relatively high level of ‘language development’ required for teaching… appears to have an objective, functional value in the education workplace and is therefore more than an expression of false ‘professionalism’” (1994, 12).

Less commonly studied and drawn on by researchers are understandings of teaching as craft and as art. In the teaching as craft understanding, teachers maintain discretion over their classroom decisions, applying their knowledge base of curricular and instructional techniques as appropriate. In addition, there is a focus on the interaction between a teacher and his or her students. In this conceptualization, an effective teacher employs “pedagogical procedural information useful in enhancing learner-focused teaching in the dailiness of classroom action” (Grimmett & MacKinnon, 1992, 387). Teachers are expected to draw on both pedagogical content knowledge as well as knowledge of each student to optimize student learning. The teaching as craft view is therefore similar to teaching as a profession, except with a greater focus on the classroom and the student-teacher relationship, rather than the professional focus on relationships with peers and administrators. Further, teaching is seen less as a professional endeavor than as one in which each teacher hones their skills and individual ability.

Finally, teaching as an art focuses on the individualized nature of the teaching task and the unpredictable nature of the classroom environment. While teachers draw on their professional repertoire and knowledge, their teaching reflects their unique personality as well as the distinctive interactions with students. An effective teacher under this categorization has a special gift and talent with students but also thinks critically about the pedagogical implications of external influence (Gage & Berliner, 1989). Teaching as art, as with the craft and profession views, also appears to be relationship oriented, although for somewhat different reasons. Learner-centered teaching implies that the
outcomes of interest are defined by the student, not the organization, thus there are no "final outcomes," the way we have defined them. Likewise, the teaching tasks arise out of the relationships teachers have with each student, not vice versa.

Table 3 summarizes the above discussion and expands on the initial modeling in Table 2 for research on teachers. The four models of teacher effectiveness are described along the six dimensions. These four models of teacher effectiveness highlight not just the various conceptualizations of teaching, but also its inherently distinct features. In particular, the four models highlight the different kinds of interactions that teachers must take part in on a daily basis. Of the four, the labor and craft approaches are more task-oriented than the professional and art models with the former two placing importance on the management of classroom resources such as curriculum and instructional strategies and the latter two emphasizing the importance of the individual teachers’ knowledge base and personality. The labor—and, even more so, the professional—model focus on teacher interactions with other teachers and administrators (with the implication that these affect pedagogical interactions with students). The craft and art versions, in contrast, focus exclusively on interactions with students. These multiple interactions in teaching suggest that no single conception of teaching is likely to capture all elements of the teaching occupation.

We have called the four views of “models” of teacher effectiveness, although they might be better viewed as first steps towards such models. As shown in table 3, it is difficult to describe the four models along the dimensions shown in table 1 and 2. One reason is that the evidence described above, while it certainly informs models of effectiveness, are not direct tests of any particular model. For this reason, it is impossible to say whether studies of particular models use particular definitions or measures of effectiveness. We have therefore listed all of these as “unclear.”

[TABLE 3]
Consider the rows in table 3 related to the nature of the work and work standardization. We discussed above the different types of relationships that teaching involves (relations with colleagues, students, and so on) and how they relate to the teaching as profession and art views. The task-relationship orientation of the craft view is unclear because the source of teacher’s knowledge (the focal point of the craft view) is unstated. The teaching as labor view, because of its association with bureaucratic rule-based school structures, is implicitly focused on tasks. For the same reason, we argue that the teaching as labor view assumes that teaching is standardized (and focused on technical skills). In contrast, teaching is extremely varied in the teaching as art view because each situation is considered to be different and the appropriate practices are based on teacher judgments in assessing and responding to interactions with students within the classroom.

The work environment is central to the teaching as labor because the work environment includes the bureaucratic structure that defines this view. The work environment is also central in the teacher as profession view although for a different reason: in this case, the work environment is defined in terms of relations with colleagues. The work environment (at least outside the classroom) plays a smaller role in the teaching as craft and art views, although in both cases one could argue that the situational decisions made by teachers are based on knowledge about students gained beyond the classroom and are informed by interactions with students, teachers, parents, and other members of the community.

Many of the assumptions and implications of the four models are “unclear” and even some of those we have labeled otherwise can still be viewed as somewhat uncertain. One possible reason is that these dimensions are not the most relevant in teaching compared with other occupations. More convincingly, the above views of teaching combine in ambiguous ways what Wallen and Travers call a “theory of ethics” and a “theory of behavior” (1963). One can have a philosophy about the objectives (a theory of ethics) that is completely separate from how particular teaching practices affect those outcomes.
(theory of behavior or action). But, a model of effectiveness, the way we define it, requires both. As shown in table 3, the definition and measurement of outcomes are central components, thus necessitating the theory of ethics. Other dimensions require establishing a theory of behavior that connects teacher characteristics (and associated practices) to these outcomes.

The results in table 3 therefore suggest that the four models fall short of models of teacher effectiveness because they do not clearly distinguish the theories of ethics and behavior and they tend to emphasize the former and give to little consideration to the latter. One can therefore view teaching as an art, for example, and never require any evidence that this type of teaching leads to more student achievement. As we will argue later, these models could be improved—that is, a clearer connection between theory and evidence can be made—by providing both types of theories and by explicitly stating and testing the assumptions and resulting hypotheses.

**Effectiveness Measures in Teaching**

The most common measure of teacher effectiveness in the literature is the student test score gain, followed by evaluations of teachers by school principals, college faculty, peers, students, parents, and research observers. The focus on student test scores in research is driven more by the increased availability of the scores resulting from the accountability movement (Linn, 2000). Because achievement is widely considered to be an important educational outcome, researchers have embraced these easily accessible measures for in empirical studies. Because the tests are administered frequently (usually annually), they have also introduced the possibility of attributing objective measures of student outcomes to students’ individual teachers, something that previously was almost impossible on a large scale. Recall the earlier discussion indicating that teaching is “loosely coupled” and that teachers’ work with students occurs largely in isolation. In theory, this means that the effect of the teacher can also be isolated—when outcome measures are available on a frequent basis.
There are of course a variety of objections to this when one considers the specific steps necessary to estimate the effects of teachers. In the context of the debate regarding “value-added” models, concerns have been raised regarding the reliability of such estimates (Ballou, Sanders, & Wright, 2004), the psychometric properties of the tests for such purposes (Koretz, 1992), and their general meaning and interpretation (Raudenbush, 2004). Perhaps more importantly, even though teaching is loosely coupled, it is still not clear whether the effects of teachers and schools can be truly separated from those of home and community factors. Harris and Sass (2006) discuss and test some of the underlying assumptions of value-added models. Harris (forthcoming) discusses a wide variety of evidence about the strengths and weaknesses of value-added measures of teacher effectiveness.

Another readily available effectiveness measure is the principal evaluation. One advantage these have over student achievement scores is that the principal has direct knowledge about each teacher’s practices and behaviors, leaving no need to rely on sophisticated statistical methods. In addition, the principal evaluation can incorporate many different measures of effectiveness. However, even after extensive searches of standard databases and queries of scholars on this topic, we have found very few recent articles in peer-reviewed journals that use the teacher evaluation to study the characteristics of effective teachers (see below).

This is not to say that evaluations themselves are not a frequent subject of study. Indeed, significant questions have been raised regarding their validity and utility (Darling-Hammond, Wise, & Pease, 1983; Epstein, 1985; Medley & Coker, 1987; Ostrander, 1996; Peterson, 2005; Stiggins & Duke, 1988; Stodolsky, 1984). It is worth considering the two main shortcomings of this measure. First, there is a lack of consistency across studies regarding the evaluation rubric. For example, McNergney and Satterstrom (1984) use data from an evaluation designed to measure the student-teacher’s quality of expression, teaching procedures, classroom atmosphere, control, organization/planning, resourcefulness, initiative, and cooperation. In contrast, the evaluation considered in
Lucas and Schmitz (1991) focuses on instructional technique, classroom management, and communication.

Second, there is evidence of an upward bias in the evaluation scores. Lucas and Schmitz (1991) found that 81.7 percent of teachers in their sample received perfect scores on their overall student teaching evaluation. In addition, Ostrander (1996) compared evaluations of individual teachers by principals, parents, students, and the teachers themselves (self-assessment), each using the same evaluation rubric. Principals gave the highest evaluations of any of the four groups.

A third characteristic of evaluation might be viewed as both a strength and a weakness: the subjectivity of the principal. The issue of subjectivity is fundamental to the debate about teacher quality and to the conclusions of the present study. On the positive side, some objectives of teaching, such as student engagement, are difficult to measure objectively and therefore may be best measured subjectively. In addition, the objectives of teaching are somewhat undefined and vary depending on one’s perspective and priorities—students, parents, teachers, voters, and policymakers all have somewhat different views (Darling-Hammond, Wise & Pease, 1983). The problem, of course, is that subjective ratings are influenced by a variety factors irrelevant to the teachers’ actual performance on the specified evaluation criteria. For example, there is evidence that ratings of student-teachers by their supervising teachers are higher when the two teachers have similar personalities (Sprague, 1997).

As an empirical matter, there is at least some evidence of a positive correlation between evaluation scores and student achievement gains (Harris & Sass, 2007b; Jacob & Lefgren, 2005; Medley & Coker, 1987; Milanowski, 2004), suggesting that objective and subjective measures may both have some degree of validity in teaching. It is therefore possible to justify consideration of both objective and subjective measures when trying to understand and predict teacher effectiveness.
Teacher Characteristics as Predictors

The results of empirical studies of teacher effectiveness also depend on which teacher characteristics are considered and how these are measured. We consider teacher personality, educational background, and cognitive skills (respectively). Note that the order of discussion is different than in the discussion of workers, reflecting in part the different amount of emphasis the factors have received in past research.

Teacher Personality

The potential importance of teacher personality has long been of interest to education researchers (e.g., Barr, 1952; Del Popolo, 1965; Tyler, 1960). Most of the research on personality focuses on the types of people who enter the teaching profession, rather than their effectiveness. Of the studies focusing on effectiveness, all use the teacher evaluation as the measure of effectiveness and nearly all focus on student-teachers. The 16 Personality Factors Inventory (16PF) is used to measure personality in most older studies on the topic, while the Meyers-Briggs Type Indicator (MBTI) has been the most common in recent studies. Guddemi, Swick, and Brown (1986) cite six U.S.-based studies that use the 16PF as a predictor of teacher evaluations. Three of the six studies find that teacher effectiveness is positively related to conscientiousness (Barr, 1961; Davis & Satterly, 1969; McClain, 1968) experiment-minded/venturesome (Hanjam, 1969; McClain, 1968; Swick & Ross, 1972) and outgoing/extroverted (Hanjam, 1969; Lamke, 1951; McClain, 1968). Two studies find that teacher effectiveness is related to self-control (Barr, 1961; McClain, 1968). These results are partially consistent with some newer studies using 16 PF. For example, Sparks and Lipka (1992) find that, compared with other teachers, master teachers are more warm-hearted, socially outgoing, attentive, driven, respectful, generous in personal relations, hard to fool, and able to maintain interpersonal contacts. Shaughnessy et al. (1995) administered the 16 PF to 56 teachers identified by principals as “outstanding.” They found that these teachers scored high on emotional maturity, self-sufficiency, and self-control.
More recent studies have used the MBTI, which categorizes personality along four dimensions: extroversion-introversion, sensing-intuition, thinking-feeling, and judgment-perception. Sachs (2004) surveyed elementary teachers with five or more years of experience in an urban school district. No differences were found between effective and ineffective teachers. The Sachs study, however, is apparently the only one in recent years to use a final outcome measure of teacher effectiveness. Others have focused on the relationship between MBTI and teacher efficacy (Henson & Chambers, 2003), teacher practices (Houtz et al., 1994; McCutcheon, Schmidt, & Bolden, 1991) and the likelihood of staying in the profession (Sears, Kennedy, & Kaye, 1997).

The change in direction in more recent studies—to the MBTI and away from the 16PF—is difficult to explain. The older 16PF literature provides some relatively consistent patterns in results, yet only one of the criteria found to be related to effectiveness in older studies (extroversion) is even measured in the MBTI. Moreover, growing access to both supervisor evaluations and student test scores makes carrying out studies of effectiveness using final outcome measures much easier than in the past. One possible explanation is that the personality measures available are not sufficiently targeted to individual occupations (Byrnes, Kiger, & Schectman, 2003; Guion & Gottier, 1965; Murphy & Dziewczynski, 2005).

Teacher Education

We begin with studies of the relationship between teacher education and teacher evaluations as the measure of effectiveness. As with teacher personality, these focus on student-teachers (unless otherwise indicated). McNergney and Satterstrom (1984) find no relationship between student-teacher evaluations and scholastic aptitude, cognitive ability, or interpersonal maturity. Lucas and Schmitz (1991), in apparent contrast, find a significant relationship between student-teacher evaluations and both the cumulative GPA and scores on the English portion of a college basic academic subject exam. However, no relationship is found between evaluations and ACT scores or the non-
English parts of the academic subject exam, which included math, science, and social studies.

Similarly, Boger, Boger, and Huarng (1987) find no significant relationships between the teacher evaluation and any of the academic predictors: undergraduate GPA, instructor appraisals of the student during their first education course, and scores on tests of general ability in reading, language, and math. Guyton and Farokhi (1987) study teachers who were evaluated by a peer teacher, principal or assistant principal, and an outside evaluator. They find that GPA is correlated with the evaluation scores, but that a test of basic skills scores is not. Ferguson and Womack (1993) study the effectiveness of secondary student teachers using a researcher-developed instrument administered by four types of evaluators during student teaching. They find that GPA in introductory education courses, GPA in student teachers’ majors, and NTE specialty scores are all related to effectiveness, with coursework being the best predictor.

Other studies focus especially on teacher certifications. Guyton and Farokhi (1987) find that certification scores were uncorrelated with evaluations by principals. Taylor and Worden (1986) examined the effectiveness of student-teachers and find no relationship between evaluations and any of the following: GPA, scores on the ACT, tests of reading, writing, and listening, or a Basic Professional Studies Examination (PSE). The only factor that was related to effectiveness was a subject matter test.

The Praxis Pre-Professional Skills Test is commonly used to assess teacher skills for the purpose of certification. A small number of studies have been done that find no statistically significant relationship between student teacher Praxis scores and student teaching evaluations (Dybdahl, Shaw, & Edwards, 1997; Hicken, 1992).

The above discussion of studies using teacher evaluations as the measure of effectiveness strongly suggests that there are no clear patterns in the literature regarding predictors considered or those found to be associated with teacher effectiveness. These inconsistencies are reduced, albeit only slightly, in the discussion below of studies using student achievement scores as the measure of effectiveness. This portion of the review

Wayne and Youngs (2003) identify 21 value-added studies of the relationship between student test score gains and teacher characteristics that meet certain criteria. Their main conclusions are: (1) “some relationship exists between college ratings and student achievement gains” (p.97); (2) there is a relationship between teachers’ mathematics coursework and student mathematics gains in high school, but no such effects are apparent in elementary grades or other subjects; and (3) mathematics certification is related to students’ math scores gains, but there is insufficient evidence with regard to other subjects and grades.

Rice focuses on the same set of teacher characteristics, except that she considers a larger number of studies that use a wider variety of research methods and outcomes that go beyond student achievement. Her conclusions are also similar with only a few minor differences in conclusions that are relevant here: (1) Wayne and Youngs expressed strong hesitation about drawing conclusions about teacher experience, while Rice is more confident in the positive effects of this factor. (2) Rice finds that teacher coursework in pedagogy and subject area is positively related to education outcomes, especially when the two are combined together (content-based pedagogy) focused on the subject being taught. The effect of subject knowledge is most pronounced at the high school level. This distinction between different types of coursework (content versus pedagogy) is not considered in the studies reviewed by Wayne and Youngs.

Wilson, Floden, and Ferrini-Mundy (2001) consider more than 300 studies and identify 57 that have been published in peer-reviewed journals within the previous two decades and which achieve a high degree rigor based on several criteria. They also consider a variety of effectiveness measures. Like Rice, they find that both pedagogical and content knowledge are important, although they argue that little is known about the specific types of knowledge within these categories that is most important. In their follow up study, Floden and Wilson (2003) add studies that were not published in peer reviewed
journals and consider some additional questions. The main new conclusion, for the purposes here, is that they find teacher experience to be a factor that “might” be related to teacher effectiveness, similar to the finding of Rice.

**Teacher Cognitive Ability**

The issue of teacher cognitive ability in educational research dates back nearly a century. For example, Boyce (1912) asked a small group of administrators to rate teachers according to their overall effectiveness and their intellectual capacity. Getzels and Jackson (1963) in their review argue that measures of cognitive ability were inconsistently related to principal evaluations of teachers (the measure of effectiveness most common at the time) and that researchers therefore gave up on it. The topic received renewed interest with the publication of the Coleman Report (1966). Coleman’s study administered a 30-item test of verbal ability to teachers and found that this was the best predictor of those school factors measured. But this was not enough to spark renewed interest among researchers. Indeed, few if any studies can be found since the publication of the Coleman report that use direct measures of cognitive ability, such as those used in industrial psychology. Aside from the small number of studies of college entrance exams and related test scores, relatively few studies have been published that use even indirect measures. One reason may be that researchers of teacher effectiveness have tended not to focus on cognitive ability, perhaps in part out of a fear of suggesting that teachers are “born not made.” This choice is unfortunate, given the evidence that this is a consistent predictor of effectiveness and the evidence that cognitive is ability is not innate. Rather, like personality and education, cognitive ability appears to be significantly influenced by environmental factors and can change over time (e.g., Petrill et al., 1998).

Returning to the earlier reviews, Rice (2003) and Wayne and Youngs (2003) both find that the teacher test score effect is only positive and significant when college selectivity/prestige is excluded. Rice also goes further and argues that this might be interpreted as an effect of cognitive ability. This conclusion is based on the consistent finding, especially
in large-scale studies, that teacher effectiveness is correlated with measures of teacher verbal skills, college entrance exams, and post-college certification exams (Harris & Sass, 2007a). There is evidence that some of these measures, especially the SAT, are correlated with cognitive ability (Efklides et al., 1997) and the same is likely true of the other tests as well. This effect of test scores only arises when college selectivity is not controlled. Because college entrance exams play a central role in college entrance, college selectivity and the pre-college abilities of teachers are likely to be highly correlated. Thus, the inclusion of college selectivity appears to absorb the effect of cognitive ability, making the correlations with other teacher test scores insignificant. Unfortunately, understanding the role of cognitive ability in teaching is complicated by the fact that studies on the subject rely on indirect measures, such as college entrance exam scores.

In this section, we have considered the models of teacher effectiveness, common measures of teacher effectiveness, and predictors of effectiveness. The results provide evidence that cognitive ability and personality, and to a lesser degree education, have been found to be positively related to teacher effectiveness. However, there have also been significant changes over time in both the way these studies have been carried out and in the results they have found. In the next section, we discuss how the comparison of teacher research with that of other workers can help to address this problem in future research.

**Discussion and Future Directions**

In this section, we provide further discussion of the collective meaning of the above review and describe our main conclusions. First, the worker literature illustrates models that complement existing theories regarding the work of teachers. Perhaps the most important example is the job-organization fit model which locates the worker within the organization and focuses on the important relationships between them. This is contrasted with the more common job-person fit which implicitly assumes that organizations are all essentially the same and that it is most important for the person to match the type of work.
The potential value of extending worker models to teaching in this way is reinforced by the fact that some worker characteristics seem to predict effectiveness in teaching, equally well as in other occupations. In the research on workers in general, we saw that cognitive ability is the worker characteristic most commonly included in studies of worker effectiveness and the factor most consistently associated with effectiveness. Also, recall the findings of Hunter (1983) and Schmitt et al. (1984) which suggest that the role of cognitive ability is strongest in complex occupations—occupations such as teaching.

To confirm that teaching would be in the upper level of job complexity, we consider the work of Rowan (1994) who compares teaching to other occupations using the Department of Labor’s Dictionary of Occupational Titles. He finds that 75 percent of occupations were rated less complex than teaching in elementary and secondary schools in terms of worker functions, general educational reasoning, and vocational preparation.

Considering that the Hunter findings hold for occupations in the top 60 percent in terms of complexity, this provides evidence that cognitive ability is also an important predictor of teacher effectiveness. This conclusion is also reinforced by research that finds consistent, positive correlation between worker wages and college entrance exams (Krueger & Lindahl, 2001).

A similar positive role for cognitive ability is found in recent studies, although these are limited by the use of indirect measures such as college entrance exams scores. Older studies tend to find no relationship between cognitive ability and teacher effectiveness (except for Coleman, 1966), but this might be explained by limitations in research methodology and reliance on principal evaluations as the measure of effectiveness. As with Rice (2003) and others, we interpret this as evidence that cognitive ability has a positive influence on teacher effectiveness.

We also find some similarities between teachers and other workers regarding personality. Specifically, conscientiousness and extroversion, the only two personality traits for which there is substantial evidence for both teacher and other workers, are consistently related
to effectiveness for both groups. The role of personality is somewhat clearer in studies outside of teaching, but even these suggest that personality is a secondary predictor compared with cognitive ability. Finally, we find substantial methodological difficulties in estimating the effect of education and, consequently, neither group of studies finds a consistent role for it in teacher effectiveness.

The implication of these similar findings across occupational groups (mediated by occupational complexity) is that there may be a general underlying model of effectiveness. Further, a theory-driven model of effectiveness for teaching might begin with models of worker effectiveness. Of course, this does not in any way preclude occupation- or domain-specific models and predictors that such general models cannot incorporate.

Second, by outlining multiple models of effectiveness, it is possible to identify the important dimensions on which they vary and therefore to be more precise about the assumptions the models make. Empirical research on teacher effectiveness has focused on the individual as the unit of analysis while models of teachers work situate teachers within the organizational context of the school. While the models we review discuss the teachers’ role in the organization, few studies have linked teacher effectiveness in the organizational context in which teaching occurs.

While linking the individual and organizational level is perhaps useful, it is also essential that the distinctions between them be kept clear. Recall the earlier finding that groups make better decisions when the members have a diversity of skills, even when sacrificing some level of cognitive ability. Likewise, it would not be desirable, even if it were possible, for schools to have a group of teachers who are identical, even if they are identical in ways that would seem desirable. Put differently, it is probably important to have a mix of teachers in a school who collectively work to achieve the school’s objectives.

Third, research on other workers highlights some ways to improve the measurement of the three predictors and teacher effectiveness. The teacher literature is more reliant on an objective measure of effectiveness, namely student achievement, while non-teaching occupations are much more reliant on the subjective supervisor evaluation. In this
respect, it is ironic that teacher research is criticized for being too subjective. The arguments for using student achievement are that the measures relate to a final outcome and teaching is loosely coupled, so that student achievement gains might reasonably be attributed to individual teachers. The arguments in favor of evaluations are that they provide more direct evidence of each teacher’s effectiveness and allow one to consider outcomes other than student achievement. There are good reasons to believe that the use of such evaluations of teachers, especially with some improvement in evaluation design, could help broaden the types of hypotheses that can be tested and therefore the types of models of teacher effectiveness that can be empirically supported. In any event, the measures chosen by researchers should not be determined solely by the availability and popularity of student test scores.

Measures of predictors could also be improved. Studies of workers often use direct measures of cognitive ability. Whereas, the conclusions about teachers rely on evidence from college entrance exams, which are, at best, very noisy measures of cognitive ability and, at worst, measures of different and unrelated constructs. There are also differences in how personality is measured across the two groups of studies. Research on workers in general increasingly relies on the “Big Five,” which has only small overlap (conscientiousness and extroversion) with either the 16PF or MBTI used to study teachers.

**Lessons for Research on Other Occupations**

We do not intend to leave the impression that research on worker effectiveness is without flaws. Indeed, this review also provides lessons for research on other occupations. First, it seems likely that models of worker effectiveness will have greater predictive ability if they are contextualized through specific occupational and organizational characteristics. Research on teaching has perhaps taken this too far, sometimes giving the impression that teaching is completely different than every other occupation and that every teaching situation is unique. But there is also a sense from the industrial psychology model that
something important is missing. One way that research on workers could maintain the strength of generality, while improving contextualization, is by developing models that account more concretely for different types of work. We found (and discussed) only a few studies that considered how the effects of the various worker characteristics varied by occupational type.

Second, Borman et al. argue studies of other occupations could also be improved by relying less on supervisor evaluations and, when they are used, taking greater care in considering what these actually measure. We agree and, as we argued earlier, a variety of outcome measures are necessary to test the variety of possible models.

Third, we discussed some of the problems in identifying the effects of educational background. Given that education can be influenced more easily by policy and leadership compared with personality and cognitive ability, researchers of worker effectiveness should pay more attention to this topic and carry out studies in which effectiveness can be judged before and after training takes place. This recommendation is also closely related to the previous one in that the primary benefits of education may arise not in the supervisor evaluation, but in intermediate outcomes such as worker turnover.

**Conclusion**

The main purpose of this study is to determine what can be learned from research on other occupations scholars move forward in addressing the many challenges to establishing clear models of teaching and to testing the validity of such models. In addition to these, summarized above, we provide concluding thoughts on the theory-evidence disconnect and two particular challenges that underlie that problem.

More so than other occupations, there is substantial debate about what Wallen and Travers called the “theory of ethics” or philosophy of teaching. Even among educators, there is disagreement about what teachers are trying to achieve. In contrast, it is really only the theory of behavior that is really at issue in other occupations. But issues of ethics and philosophy alone cannot provide a full explanation. Every philosophy of
teaching can define its own objectives and these in turn can be associated with theories of behavior. Philosophical issues therefore make the task more difficult, but far from impossible. Despite these philosophical issues, we do not share the concern, described (but not endorsed) by Rowan (1994) about whether teaching “ever can be . . . guided by an abstract and scientific body of knowledge” (p.4). Instead, we have suggested, based on a comparison of research on other occupations, some clear paths for future research that can help address this problem.

The second issue goes beyond the relationship between theory and evidence to include a disconnect between theory and practice. While this study has not addressed the topic directly, we speculate that there is a substantial disconnect between the findings in this article and the actual approaches to training, hiring, and evaluating teachers. The importance of having an appropriate mix of workers has already been noted. In addition, there is evidence that schools appear to place value on strong interpersonal communication skills and good classroom management and instructional skills, leaving as secondary factors such as university background and experiences (e.g., Ralph, Kesten, Lang & Smith, 1998). This suggests a disconnect not only between theory and practice, but also between research and practice. To improve practice, these issues having to be part of the starting point of research, rather than trying to infer conclusions about policy and practice after the research has been conducted.

Again, we have not tried to identify a specific model of teacher effectiveness that addresses all of these issues, nor do we downplay the level of difficulty involved is achieving such a task. The fact that such a fundamental issue remains with us after a half-century stands as a testament to the challenge involved. But the issue can be addressed and, with this study, we have tried to provide some concrete steps toward that end.
References


Schick, G. J., & Kunnecke, B. F. (1982). Do high school grades, top schools or an advanced degree lead to job security and extraordinary salary progression? Interfaces, 12(1), 9-18.


Table 1: Models of Worker Effectiveness and Research Approaches

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Common Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General Approach</strong></td>
<td></td>
</tr>
<tr>
<td>Research fields</td>
<td>Psychology, Economics, Sociology, Organizational Behavior</td>
</tr>
<tr>
<td>Unit of analysis</td>
<td>Individual worker, Team, Organization</td>
</tr>
<tr>
<td>Definition of effectiveness</td>
<td>Intermediate outcomes, Final outcomes,</td>
</tr>
<tr>
<td>Measure of effectiveness</td>
<td>Objective measures, Subjective measures</td>
</tr>
<tr>
<td><strong>Assumptions/Implications</strong></td>
<td></td>
</tr>
<tr>
<td>Nature of production/work</td>
<td>Task-oriented, Relationship-oriented</td>
</tr>
<tr>
<td>Degree of work standardization/need for worker judgment</td>
<td>Standardized, Varied/unpredictable</td>
</tr>
<tr>
<td>Importance of the work environment</td>
<td>Central, secondary</td>
</tr>
<tr>
<td>Types of worker skills hypothesized to be important</td>
<td>Technical; Interpersonal and related skills</td>
</tr>
</tbody>
</table>
Table 2: Models of Worker Effectiveness

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Models of work:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Industrial Psych.</td>
</tr>
<tr>
<td>General Approach</td>
<td>Psychology</td>
</tr>
<tr>
<td>Primary research fields</td>
<td>Worker, Task</td>
</tr>
<tr>
<td>Primary units of analysis</td>
<td>Mostly intermediate outcomes</td>
</tr>
<tr>
<td>Effectiveness measure</td>
<td>Subjective</td>
</tr>
<tr>
<td>Nature of production/ work</td>
<td>Task-Oriented</td>
</tr>
<tr>
<td>Standardization/need for judgment</td>
<td>Standardized</td>
</tr>
<tr>
<td>Significance of the work environment</td>
<td>Secondary</td>
</tr>
<tr>
<td>Types of worker skills hypothesized to be important</td>
<td>Technical skills</td>
</tr>
</tbody>
</table>
### Table 5: Models of Teacher Effectiveness

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Teaching as:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Labor</td>
</tr>
<tr>
<td>General Approach</td>
<td></td>
</tr>
<tr>
<td>Primary research fields</td>
<td>Psychology, Sociology</td>
</tr>
<tr>
<td>Primary unit of analysis</td>
<td>Teacher, organization</td>
</tr>
<tr>
<td>Effectiveness definition</td>
<td>(unclear)</td>
</tr>
<tr>
<td>Effectiveness measure</td>
<td>(unclear)</td>
</tr>
<tr>
<td>Assumptions/Implications</td>
<td></td>
</tr>
<tr>
<td>Nature of production/work</td>
<td>Task-oriented, mediated by formal school structures</td>
</tr>
<tr>
<td>Degree of work standardization/need for worker judgment</td>
<td>Generally standardized</td>
</tr>
<tr>
<td>Assumed significance of the work environment</td>
<td>Central</td>
</tr>
<tr>
<td>Types of worker skills hypothesized to be important</td>
<td>Technical skills</td>
</tr>
</tbody>
</table>
Notes

1. This may sound obvious, but some branches of research have suggested just the opposite. In particular, beginning with the Coleman Report (1966), and continuing with the education production function literature through the early 1990s (Hanushek, 1986, 1996), studies consistently found that students’ home and community environment were the strongest predictors of student learning, apparently implying that there was little that teachers or schools could do that had a measurable effect. More recent research has used a “value-added” approach to modeling education production, making it possible to estimate overall effects for individual teachers. Some studies have found significant variation in these teacher effects (Rivkin, Hanushek, & Kain, 2005; Sanders and Rivers, 1996), although there remains some question about whether this variation is the result of statistical bias and/or noise. See additional discussion of this literature in section IV.

2. In their discussion of teaching, Getzels and Jackson (1963) tried to integrate their discussion of theory with empirical evidence, but they wrote that “the greater number of [empirical] studies were so atheoretical that we were forced to abandon this plan” (p.507). Similar statements have been made regarding the identification and measurement of both dependent and independent variables is somewhat ad hoc. As Wallen and Travers write, “these [teaching] methods are derived only to the most limited degree from scientific research . . . [and] these variables tend to be intuitively derived rather than empirically derived” (1963, p.466).

3. Dunkin (1997) makes similar distinctions among “competencies” (knowledge/skills), “performance” (behavior), and “effectiveness.”

4. Objections are sometimes raised to the subjective-objective distinction because all measures involve some degree of subjectivity. We do not dispute the criticism, but we do believe a useful distinction can be made between those measures in which the subjectivity arises mainly from the perspective of the researcher (e.g., in the construction of test scores) and those arising from both the subjectivity of the researchers and the subjectivity of a person answering a question (e.g., in the construction and answering of interview questions by school principals).
5. We especially thank Drs. William Anthony and David Ketchen of College of Business at Florida State University who suggested studies regarding worker effectiveness.

6. Previous studies have attempted to create dichotomies that encompass these and related dimensions. For instance, Stewart and Carson who refer to the “traditional” versus “emergent” models and to the “technical” and “social” inputs that are associated with them (1997, p.168). Scott (2004) refers to “organic” models and considers “relational” models to be a specific subset that focuses on the relationship between the workers and the working environment. Others refer to “rational” models, although we find this term inadequate in the dichotomy because it implies, falsely in our view, that the relational model is somehow “irrational.” More generally, we find none of these dichotomies adequately captures all of the substance of the dimensions in table 1.

7. As Keeley (1980) writes, from this perspective jobs are defined by their “patterns of interaction” with workers in other positions.

8. The studies in this section were identified from two main sources. First, after inquiring among colleagues who have studied workers in other occupations, we identified several recent literature reviews (e.g., Borman et al., 2003; Ployhart, 2004; Schmidt & Hunter, 1983). Many of our conclusions are drawn directly from these reviews. We also identified some specific studies cited in these reviews for further consideration. Finally, we performed searches in PsycINFO using keywords “personnel,” “worker,” or “employee” combined with other words, such as “evaluation,” “performance,” “productivity,” and “effectiveness.”

9. Cohen and Pfeffer write that, from the industrial psychology perspective, “it is not sensible to speak of organizational factors” related to worker effectiveness (1984, p.2). Stewart and Carson (1997) write that “commonly accepted ideas about staffing are grounded in traditional practices of industrial psychology and are often incompatible with an emphasis on social relationships among organization members” (p.158). Finally, Chatman (1991) writes that “industrial psychologists have looked at personnel selection almost exclusively from the perspective of organizations selecting individuals for particular jobs” (p.461).

10. Stewart and Carson write that research on organizational fit is relatively new (1997,
p.174). Also, Sheridan (1992) describes seven variables that describe organizational culture: “detail, stability, innovation, team orientation, respect for people, outcome, aggressiveness.”

11. There are also two types of person-organization fit: supplementary fit is directed at filling gaps and implies a need for worker diversity while complementary fit is directed towards worker homogeneity (Borman et al., 2003).

12. The idea of a P-O match in organizations with “distinctive cultures” is problematic because it presumes that there is an unstated “normal culture” and that there are “normal workers” who are a good fit with that culture.

13. Ployhart et al. (2004) provide a table with a similar structure.

14. Note that only some of these 18 instruments would be considered worker characteristics, which the primary interest here. Others, such as work samples, relate more directly to worker practices. In addition to these predictors, Borman et al. (2003) discuss vocational interests, biodata, and situational judgment tests.

15. The use of wages as a measure of performance is justified on the economic theory that workers are paid according to their contribution to production (marginal product). Beyond wages and output, it is unclear which effectiveness measures were used in the studies they reviewed.

16. Schmitt et al. (1984), unlike most of the literature, also consider “concurrent” studies of effectiveness, which focus on the effectiveness of workers already on the job. As a result, they treat the supervisor evaluation as a (concurrent) predictor rather than the effectiveness measure. Indeed, they find that the supervisor evaluation is the best predictor of effectiveness, which could be viewed as support for using the supervisor evaluation instead as the dependent variable. In any event, because the supervisor evaluation is generally, and we would say appropriately, used as the outcome, our discussion excludes it as a predictor.

17. Schmitt et al. (1984) also consider predictors that cannot be described as “characteristics.” Indeed, the “work sample” and scores from “assessment centers” are better able to predict worker effectiveness, but these are based on actual work observations.
18. Huffcut et al. (2001) provide a taxonomy of personality traits (p.899-900). One reason for interest in personality as a predictor is that tests of cognitive skills tend to result in the disproportionate hiring of whites over minorities, resulting in an “adverse impact.” Personality tests are considered less prone to this phenomenon, but yet still reasonable predictors of job effectiveness.

19. Mount, Barrick, and Strauss (1999) also provide a review of the literature.


21. The studies regarding teacher effectiveness were identified as follows: First, we identified the four commonly cited reviews of the literature discussed in the text (Rice, 2003; Wayne & Youngs, 2003; Wilson, Floden, & Ferrini-Mundy, 2001; Wilson & Floden, 2003). In some cases, we relied on the conclusions from their reviews. In other cases, we identified studies cited within these reviews. Also, some of the topics of this paper are not considered in the reviews. Articles regarding teacher personality were identified using combinations of the search terms “teacher,” “personality,” “effectiveness,” and “affective” in the databases Education Abstracts and JSTOR. Regarding teacher education, we searched the same databases with terms such as “teacher,” “student teaching,” “education,” and “effectiveness.” From these searches, we focused on studies that correlated measures of teacher characteristics (predictors) with measures of teacher effectiveness. With few exceptions, discussion is limited to articles in peer-reviewed journals.

22. There has been research in education suggesting that when there is a person-organization fit, there are increased student test scores. Research on school effectiveness has found that schools in which the principal and teachers share common views on “the internal school processes that are directly linked to student learning” (Hallinger and Heck, 1996, 38) there is higher student achievement (Hallinger and Heck, 1996). Newman, Smith, Allensworth and Bryk (2001) come to similar conclusions in their study of instructional program coherence. It should be noted that both lines of research seek to identify organizational factors leading to higher student achievement rather than the relationship between teacher and organizational effectiveness.

23. A critical interpretation of this conception is that effectiveness is defined by the
ability of the teacher to comply with organizational expectations and external demands, suggesting a deskilling and deprofessionalizing of teachers (Apple, 1988).

24. Darling-Hammond, Wise and Pease make this point more generally about the four models when they write that “the conceptions of teaching signal different definitions of success” (1983, p.292)

25. There is also evidence of upward bias when student-teachers are evaluated by supervising teachers (Sprague, 1997).

26. We chose to explore the broader category of personality; however, Rowan, Chiang and Miller (1997) make a persuasive case for using motivation as a predictor.

27. Six of the studies using the 16 PF sought to identify personality characteristics of successful pre-service teachers. One looked at successful teachers in the field (Lamke, 1951). A variety of others were excluded because they focused on non-U.S. contexts (Kyriacou (1985), Start (1966), Start & Laundy, 1973; Warburton, Butcher, & Frost, 1963).

28. The measure of effectiveness if not defined in this study.

29. The 16PF and MBTI are not the only personality tests used. Most noteworthy is Hill (1968) who finds, consistent with the above research, that successful teachers are more likely to demonstrate conscientiousness and control. Other studies of personality using different measures include Lucas and Schmitz (1991), McNerney and Satterstrom (1984) and Manning and Payne (1984), and Veldman and Kelley (1965).

30. The four types are evaluators are: cooperating teachers, content specialist student teaching supervisors, school of education student teaching supervisors, and student teacher self-evaluations.

31. We exclude studies of alternative certification because the term is defined quite loosely which makes it somewhat difficult to interpret this literature.

32. The four criteria in Wayne and Youngs (2003) are: (1) use of student test scores as the outcome and include teacher characteristics as independent variables; (2) use of test scores accounts for prior achievement; (3) includes measure of student SES; and (4) focus on students in the United States.

33. Wilson, Floden, and Ferrini-Mundy (2001) considered studies that used the
following methods: (1) experimental and quasi-experimental studies that used random assignment or matched samples; (2) non-experimental multivariate studies that controlled for relevant differences among students; and (3) simple comparisons of credentialed and non-credentialed teachers were acceptable as long as consideration was also given to student control variables.

34. We could not obtain a copy of the Boyce (1912) study. Therefore our brief description is paraphrased from Getzels & Jackson (1963).

35. Of course, the main conclusion of the Coleman Report is that school factors matter relatively little overall compared with home factors. Thus, to be clear, Coleman concludes that teacher verbal skills represent the best predictor among a set of measures that are not very strong predictors.

36. Ployhart (2004) appropriately distinguishes the level/unit of the theory from the level/unit of the analysis. While we do not make this distinction in table 1, we do consider this below.