

Educational Researcher

<http://er.aera.net>

Attribution Error and the Quest for Teacher Quality

Mary M. Kennedy

EDUCATIONAL RESEARCHER 2010 39: 591

DOI: 10.3102/0013189X10390804

The online version of this article can be found at:

<http://edr.sagepub.com/content/39/8/591>

Published on behalf of



American Educational Research Association

and



<http://www.sagepublications.com>

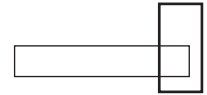
Additional services and information for *Educational Researcher* can be found at:

Email Alerts: <http://er.aera.net/alerts>

Subscriptions: <http://er.aera.net/subscriptions>

Reprints: <http://www.aera.net/reprints>

Permissions: <http://www.aera.net/permissions>



Attribution Error and the Quest for Teacher Quality

Mary M. Kennedy

Social psychologists are persuaded that researchers as well as laymen tend to overestimate the influence of personal traits and underestimate the influence of situations on observed behavior. The author of this article suggests that education researchers and policy makers may be overestimating the role of personal qualities in their quest to understand teaching quality. In their effort to understand classroom-to-classroom differences in student learning, they may focus too much on the characteristics of teachers themselves, overlooking situational factors that may have a strong bearing on the quality of the teaching practices we see. The author reviews some of these situational forces.

Keywords: teacher characteristics; teacher context; teacher knowledge

We in the United States tend to think of teachers as individual artisans. We regularly confuse *teaching* quality with *teacher* quality, as if the two were indistinguishable. We label *teachers* as caring, efficient, engaging, or boring, as if the events we see in classrooms sprang entirely from personal qualities that teachers bring with them when they enter the room. This habit of thought has a long tradition in the United States, so it should not be surprising that, as our attention has recently turned to questions about teaching quality, we have started to examine the personal characteristics of our teachers—their credentials, licensure test scores, skills, and personal values—and have overlooked aspects of their work that are outside their control, such as resources, planning time, and other aspects of the school infrastructure that might influence the quality of teaching practice.

Social psychologists are persuaded that we are all guilty of overestimating the influence of personal characteristics on behavior and underestimating the influence of the situation itself. In fact, this tendency is so widespread that it has been called the *fundamental attribution error* (Gilbert & Malone, 1995; Humphrey, 1985; Ross, 1977) and has been recognized as a fundamental component of research and theory in social psychology. In an early literature review on attribution theory, Ross (1977) argued that the fundamental attribution error was not merely the province of the man on the street but that it plagued psychology

researchers as well. In other words, researchers are just as likely as lay people to attribute behaviors to personal qualities rather than to situational influences.

In this article, I suggest that education researchers and policy makers may also have succumbed to the fundamental attribution error in their quest to understand teaching quality. I suggest that we have veered too far toward the attribution of teaching quality to the characteristics of teachers themselves, and are overlooking situational factors that may have a strong bearing on the quality of the teaching practices we see.

In the past several decades, as researchers have sought the qualities that make a good teacher, they have correlated numerous teacher characteristics with student learning outcomes; but none has demonstrated very high correlations. Despite wide recognition of these weak relationships, *teacher* quality remains a hot topic in education, as researchers continue to seek the magic ingredient that makes some teachers better than others. The recently released report by Babcock et al. (2010) on the Teacher Education and Development Study in Mathematics (TEDS-M) is a salient example: It defines and measures *teacher* quality through tests of content knowledge, and it asserts that international differences in the quality of middle school mathematics *teaching* are due largely to differences in the college-level mathematics courses that teachers in each country take prior to entering the field. Yet research on the relationship between teachers' college mathematics courses and student learning suggests that the courses have only a moderate influence on student learning (for a review of the literature that focuses specifically on math courses, see Kennedy, Ahn, & Choi, 2008).

Some researchers are now focusing on teachers' test scores rather than their prior course work (Boyd, Grossman, Lankford, Loeb, & Wycoff, 2008; Clotfelter, Vigdor, & Ladd, 2007; Goldhaber, 2007; Hill, Rowan, & Ball, 2005), but the presumption underlying all of these studies remains the same: Once we find the teacher characteristics that best forecast student learning, we can improve student learning by improving the raw material that teachers bring with them to their work. Yet, while all of these studies find some relationship between their measures and student learning, they leave a great deal of variance unaccounted for. Taken together, they remind us that the qualities teachers bring with them to their work are not enough to ensure better *teaching* practices. It is what teachers actually *do* that is most relevant to student learning.

It is time to look beyond the teacher to the teaching situation itself: the school, the classroom, the teacher's schedule, and the

teacher's resources. We need to change our tacit model of teaching from one that looks something like this:

Teacher characteristics → *Teaching practices* → *Student learning*

to one that looks something like this:

Teacher characteristics
Situation characteristics ↘
→ *Teaching practices* → *Student learning*

The problem of correctly identifying the causes of good and bad teaching practices is especially relevant in the current policy climate. Many current policies seem to presume that teaching practices follow directly from such enduring personal characteristics as credentials, knowledge, or perhaps dispositions and personality traits. For example, the notion that we should hold teachers accountable for student outcomes presumes that those outcomes are largely in teachers' hands and overlooks the role of the textbook, the physical space, and other resources. It even overlooks whether children actually attend school every day. Similarly, arguments about how best to allocate teachers among our more needy schools seem to presume that teachers who were effective last year in one setting will be equally effective next year in another setting—that good teaching arises from qualities teachers possess within themselves and carry with them from one context to another. Allocation policies assume that teachers' practices will remain constant regardless of which schools, resources, or students we place them with.

And it is not just policy makers who seek enduring personal qualities that can account for differences in student outcomes. Researchers seek better strategies for assessing teachers (see, e.g., Kennedy, 2010), assuming that such qualities have a predictable bearing on teaching practices. Advocates of all stripes seek the most effective programs for preparing teachers (Boyd, Grossman, Hammerness, et al., 2008; Constantine et al., 2009; Labaree, 2010; Stotsky, 2007; Tuttle, Anderson, & Glazerman, 2009), as if the effects of preparation will persist regardless of where teachers go upon graduation and regardless of how long they teach.

One way to estimate the importance of personal characteristics relative to situational factors lies in the year-to-year and subject-to-subject variability of teachers' effects on student achievement. A handful of studies in the 1950s and 1960s that examined the stability of teachers' effects were reviewed by Rosenshine (1970). He found that teachers' effects on student achievement were only moderately correlated from year to year, with correlations ranging from .2 to .4. And these correlations may be artificially *high* because many of the studies that Rosenshine reviewed involved highly standardized or even scripted curricula, which presumably would decrease situational variability in practice and would not be representative of the curricula in most school systems.

Questions about stability of teachers' effects are coming up again today as researchers work with more detailed and sophisticated databases and have more sophisticated statistical tools at their disposal. Some of these newer findings also raise questions about the power of teachers' enduring qualities. For instance, Aaronson, Barrow, and Sander (2003) found that, among teachers

in the lowest quartile of effectiveness during one year, only 36% remained in that quartile the next year. Of those in the top quartile, 57% remained there the following year. Rowan, Correnti, and Miller (2002) correlated elementary teachers' effectiveness in mathematics versus reading and found correlations ranging from .3 to .47 across the elementary grade levels. Papay (in press) looked at variations in value-added estimates of teachers' effectiveness across different tests and found not only variations due to test content but also variations due to differences in when the tests were administered and differences in measurement errors. Finally, McCaffrey, Sass, Lockwood, and Mihaly (2009) approached the question using variance decomposition and concluded that about 50% of elementary school variations and 70% of middle school variations in value-added estimates of teacher effects, *after measurement noise had been eliminated*, appeared to be associated with stable teacher characteristics. Their method of estimation yields a larger estimate of stability than others do, but it relies on a complex analysis that has not yet been fully vetted by scholars in the field. Moreover, even these unusually high estimates suggest that a great deal of variance is still unexplained and may be due to situational factors.

These studies raise the possibility of attribution error: They suggest that teacher effects are not as stable from year to year as we would expect them to be if they were due primarily to enduring qualities within teachers themselves. If all our ideas about teacher quality—ideas for recruiting better teachers, preparing better teachers, or better allocating the teachers we have—are based on the Fundamental Attribution Error, then our efforts to improve teachers will yield only marginal effects on teaching, and teaching will continue to vary from classroom to classroom in ways that are unaccounted for.

How does attribution error work? Social psychologists have devoted a great deal of attention to the question of how causal attributions are made and when different kinds of attributions are made. They know, for instance, that when people account for their own behavior, they place more weight on the situation than on their own personal qualities. It is mainly when accounting for other people's behaviors that they rely on personal characteristics.

One reason for the error is that we conflate behavior with personal character. That is, if we see a generous act, we assume the person doing it must be a generous person. If we hear an unkind statement, we assume the speaker must be an unkind person, overlooking entirely the range of circumstances that might have provoked the comment. Similarly, in education, if we hear teachers say things that we know to be wrong, we assume they must be ignorant.

In a review of literature on attributions, Gilbert and Malone (1995) identified four steps that are involved in drawing a causal inference about observed behavior. Each creates possibilities for misinterpretation:

- First, we *perceive* the person's situation. The possibility for error here is that we may not be aware of all relevant aspects of the situation and hence of how it might influence behavior. In the case of a teaching episode, for instance, we may not realize that one of the students in the class is emotionally disturbed and frequently throws tantrums, and that many of the teacher's actions are geared toward avoiding anything that might provoke this student.

- Next we *form expectations* for how the person should respond to the situation. The possibility for error here is that we may be overly idealistic in our expectations. In the case of teaching, virtually everyone has preformed images of what a “good” or “bad” teacher would look like, but many of these images are based on perceptions we had when we were 10 years old. Still, they contribute to our interpretation of the appropriateness of the actions we see.
- Third, we *interpret* the person’s behavior. The possibility for error here is that we may misclassify the behavior. In the case of teaching, for instance, we may think a teacher is responding to a particular student comment when in fact she is using that student’s comment to make some other point to the entire class.
- Finally, we *infer the cause* of the behavior, typically attributing it to some underlying personal characteristic. This is where errors are most visible to the social psychologists, but the errors themselves may derive from any of these stages. Thus, in the case of teaching, if a teacher says something that makes no sense *to us*, we assume that the teacher must not know what she is doing or what she is talking about.

Let’s look at an example of a teacher’s behavior in a particular situation and see how an education researcher may err during these four stages. Suppose our researcher is observing a fifth-grade math lesson offered by Ms. Katlaski, whose intention for the lesson is to introduce a strategy for multiplying whole numbers by fractions. Katlaski opens the lesson by posing the problem: We’ve learned how to multiply two fractions together and we’ve learned how to multiply two whole numbers, but what if we have a whole number and a fraction? Katlaski then gives students an example problem, $9 \times 2/3$. Although the observer does not know this, Katlaski got the example from her teachers’ guide, which also suggested that the optimal solution would be to convert the 9 to a fraction by dividing it by 1, thus creating the fraction $9/1$. The guide suggested that the problem be written on the board this way:

$$9/1 \times 2/3 = 18/3 = 6.$$

The observer is also unaware that Katlaski had almost no time to plan for this lesson. She quickly glanced over the problem before school and then headed straight into the lesson. There, she introduced the topic for the day, showed students the sample problem, and then posed the rhetorical question of how they might convert 9 to a fraction so that they could solve the problem. One of her students proposed multiplying 9 by $4/4$, a strategy that was correct but that would have been far more difficult to pursue on the board and would have required Katlaski to work through the following steps with her students:

$$36/4 \times 2/3 = 72/12 = 6.$$

Knowing that students could not compute these numbers in their heads, Katlaski knew that she could not use this example to illustrate the concept of converting whole numbers to fractions. The computations would be too distracting and students would lose the point of the exercise. Uncertain of how to proceed,

Katlaski told the student that her $4/4$ proposal would not work, and the observer in the back of the room, seeing that $4/4$ was indeed correct, concluded that Katlaski did not know the mathematical content she was teaching.

This observer might have made all the errors that Gilbert and Malone (1995) listed. First, his perception of the situation was incomplete. He did not know how much planning time Katlaski had, nor did he anticipate how the student’s solution would affect the worked example on the board as the lesson proceeded. Second, he may have had an unrealistic expectation for how the teacher should have responded. He knew that the student’s idea was correct, and he assumed, therefore, that the appropriate response was to say “yes, that is correct.” But he was unaware of the teacher’s concern about how that answer would affect the next step in the lesson—working the problem out on the board and ensuring that students could follow the logic. Third, in defining Katlaski’s behavior, he labeled it as a mathematical error when in fact she was thinking that the student’s solution would be difficult to *work out on the board* in a way that 10-year-olds would be able to follow. As a result of this combination of interpretive errors, the researcher drew the erroneous attribution that Katlaski’s behavior followed directly from a lack of mathematical knowledge on her part.

None of this discussion is intended to suggest that Katlaski’s response was a good one, or that it was excusable. She should not have told the student that the idea would not work. Rather, the point here is that the misstep was not due to a lack of mathematical knowledge but rather due to a conflating set of circumstances that momentarily stymied her. The question I raise in this essay is why all of us, as education researchers, tend to interpret events such as this as indicators of enduring qualities within the teacher rather than wondering why, for instance, her manual did not prepare her for the variety of other ways students might try to convert 9 to a fraction, or why she had so little time to plan for this lesson.

My goal in this article, then, is to introduce the possibility that the situations teachers face in their daily work provide important and underappreciated influences on teaching practice and on student learning. What might these situations be? The example of Ms. Katlaski suggests that the amount of planning time and the quality of instructional materials would be candidates. There is also a body of research that focuses on the role of school organization and climate, although that work is typically used to explain school-to-school variations. Still, even school climate is likely to be experienced differently by different teachers. What is needed is a closer examination of the situations that individual teachers experience that may provoke the practices we see in their classrooms.

In the next four sections, I examine aspects of teachers’ working conditions that are likely relevant to their effectiveness in any given classroom or school year: the parameters of the work itself, the students, institutional incursions into classroom life, and excessive reforms, or what I call “reform clutter.”

Parameters of Teachers’ Work

Education institutions structure teachers’ work in several important ways. They provide the physical space in which teachers work and the textbooks they use. Although we know that teachers work in different rooms, we know very little about how any particular arrangement might facilitate or hinder a teacher’s

effectiveness. Suppose that long, narrow rooms create problems that do not appear in square rooms, for instance. If so, a teacher assigned to a long, narrow room might routinely be perceived as less effective, and that lesser degree of effectiveness might routinely be assumed to be due to some characteristic within her rather than to a characteristic of the space she inhabits. Schools also establish rules and norms for behavior (and may differ in how consistently these are enforced), policies for rewards and sanctions, and schedules and standard operating systems that help or hinder teachers. Most districts also establish content standards, curriculum materials, learning goals, and tests to measure the goals.

Of these many aspects of teaching situations, three seem particularly relevant to teachers' effectiveness: time, materials, and work assignments.

Time

The typical teacher contract requires teachers to spend about 32 hours per week at school (Drago, et al., 1999), or a little over 6 hours per day. Of that time, teachers typically have 1 hour each day allocated to planning and the remaining 5 hours allocated to instruction. These statistics suggest that the ratio of planning time to instructional time is roughly 1:5. Presumably, Ms. Katlaski would have to use the 1-hour planning period to prepare for five other class periods, so she would have about 10 minutes to prepare for each of her lessons each day. The value of this hour may depend on what the teacher has to do during the remaining 5 hours. Because Katlaski is an elementary teacher, she must use this hour to plan her reading, mathematics, science, and social studies lessons. A secondary teacher may be able to achieve some efficiencies if he or she teaches the same course—say, introductory algebra—multiple times during the day. But this aspect of the job varies from teacher to teacher and, for any given teacher, it may vary from year to year.

Moreover, this “planning period” is also the only time during their official work day that teachers can read and respond to student work, assign grades to student products, and diagnose their learning progress. And it is during this hour that they talk to or e-mail parents, special education teachers, or any others who may share responsibility for their students.

Some teachers, of course, add to that time by working at home in the evenings (Krantz-Kent, 2008); but they do so at their own discretion, which probably contributes to differences in the total planning time that teachers actually have. Krantz-Kent's time-use survey of teachers found that 30% of teachers worked at home in the evenings, 35% worked on Saturday, and 50% worked on Sunday. If we included in our estimate the time teachers spend working at home, the average ratio of planning time to instructional time might be closer to 2:5, although there could be quite a bit of variation among teachers.

Three points can be made about teachers' planning time. One is that the planning time allocated by U.S. schools is impossibly low, and undoubtedly is one reason that the quality of teaching in the United States appears to be lower than in many other countries. In fact, my extended ratio estimate of 2:5 is still smaller than that for teachers in China and Japan, who typically spend about 16 periods per week in instruction (Stevenson & Stigler, 1992; Ma, 1999). If teachers in those countries devote 45 hours per week to their work, their ratio of planning to instruction

would be almost 2:1. The value of planning time is difficult to estimate through cross-national studies, of course, because there are so many other differences to account for. But the presence of such a large difference certainly raises the question of how important planning time is, or could be, to teaching effectiveness.

There is more: It is not uncommon for American institutions to invade teachers' planning time with other agendas: grade-level or departmental planning meetings, parent visits, or researcher visits. Institutions convene teachers to review test results, engage in professional development, plan assemblies and other special events, revise discipline policies or paperwork procedures, or engage in other institutional tasks. Professional developers invade teachers' time by encouraging them to join professional learning communities, try their hand at lesson study or action research, or spend more time examining student homework. Researchers also take up teacher time by interviewing them, surveying them, asking them to fill out logs, or asking them to obtain parental permission for research involving their students.

Ironically, many of these time-consuming noninstructional activities are intended to make teachers more thoughtful about their practice; yet they do so by reducing the time available for thinking about tomorrow's lessons. It should not be surprising, then, that, in my own research (Kennedy, 2005), I found teachers rushing to class without having even read the day's lesson in the textbook, forgetting to bring materials they had planned to use, and discovering, midlesson, unforeseen problems with their plans or materials.

Materials

Schools typically provide at least three types of curricular guides for teachers: textbooks and their associated teachers' manuals; state- and possibly district-level standards or curriculum frameworks; and one or more standardized achievement tests. The guides may or may not be coherent, internally consistent, or aligned with each other, and all have been the objects of substantial criticism. American textbooks, for instance, have been famously characterized as mile wide and inch deep (Schmidt, McKnight, Cogan, Jakwerth, & Houang, 1999). Standards have also been criticized for defining broad and vague outcomes (Cohen, 2010; Cohen & Moffitt, 2009; Shattuck, 2005), and tests have been criticized for focusing mostly on recitational knowledge. The whole set has also been criticized for lack of alignment, and there is anecdotal evidence that teachers often deviate from the materials their institutions provide for them.

One aspect of teachers' materials that has received relatively less attention is teachers' manuals. In a comparison of Chinese and American teachers, Ma (1999) found that Chinese teachers increased their depth of mathematical understanding simply through their experience teaching the content. She concluded that this growth came from their teachers' manuals, which explained the content to be taught in each day's lesson and also explained the types of misconceptions that students were likely to experience. (Of course, the value of this material would depend on teachers' having time to actually study it.) If Katlaski had had a manual like this, and the time to read it when she prepared for her lesson, it might have reminded her that there were multiple ways to convert 9 into a fraction and might even have provided a way to respond to the $\frac{4}{4}$ proposal. For instance, the manual

might have given her a set of numbers that would work better, like $3/3$, or $2/2$, fractions that might better allow students to follow the computations:

$$18/2 \times 2/3 = 36/6 = 6.$$

When curricular materials are not aligned, or when they demand more content than teachers believe they can adequately teach, teachers make their own judgments about which content is most important. When texts are cluttered with trivia, teachers decide what the main point of each unit is and what they will emphasize in their instruction. Porter and his colleagues (Porter, Polikoff, & Smithson, 2009; Porter, Smithson, Blank, & Zeidner, 2007) have shown that there are substantial variations in content coverage across schools and classrooms. What we do not know is why, or how well, teachers make their adaptations. These researchers and others have devised methods of measuring the content actually taught, measuring how well aligned it is with other content, and estimating the effect of alignment on student test scores. An important next step here would be to see how important these variations are, as compared with variations among the teachers themselves, in accounting for learning outcomes.

Work Assignments

There are also differences in teachers' individual workloads and schedules. Secondary teachers differ in the number of different preparations they need to make each day. A teacher who teaches five sections of algebra each day has an easier schedule than one who teaches one section each of algebra, geometry, trigonometry, and calculus. Workloads can also be more or less taxing depending on whether they include out-of-field assignments. That is, a teacher certified in mathematics may teach, say, four mathematics courses and one course in Spanish or history. Ingersoll (2003a, 2003b) has argued that such placements serve to "de-skill" teachers, making them interchangeable and reducing their sense of control over their work, which could, in turn, reduce the quality of their teaching.

In both elementary and secondary schools, teachers also differ in the number and variety of their extracurricular responsibilities. They monitor the halls, playgrounds, and lunchrooms; supervise after-school debate teams and science fairs; accompany students on field trips; and so forth. These additional responsibilities not only take time that might otherwise be available for instructional planning but also create additional planning obligations.

The number and variety of extracurricular responsibilities vary from teacher to teacher and may be relevant to differences in their effectiveness. Ingersoll (2003b) describes how school administrators manipulate teachers' resources and responsibilities in a variety of ways that reduce their sense of control over their work. Principals may reward or punish particular teachers by providing more or less storage space or more or fewer out-of-field assignments, better or worse classroom spaces and so forth.

If these practices are widespread, they should be investigated for their impact on apparent teaching quality. Suppose we learned, for instance, that teachers' value-added estimates of effectiveness depended on their course assignments and non-teaching duties. If this were the case, it would be easier for us to

improve teachers' effectiveness by altering their work assignments than it would be to alter their enduring personal characteristics such as their educational backgrounds or test scores.

Students

In the past couple of decades, statisticians have made a great deal of progress in their ability to measure students' rates of growth and to control for differences in their initial achievement. But student influences on teachers' instruction arise from far more than their initial test scores. Students influence teaching practices when they are restless, gregarious, or frustrated, and even when they are happy. And they interrupt the learning of other students as well.

Cohen (1988) has reminded us that, ultimately, teaching is an attempt to change other human beings, and that such enterprises cannot succeed unless the other human beings cooperate. Both American culture and American policy discourage student cooperation. American culture contains a deep anti-intellectual tendency and a strong youth orientation (Powell, 1996), both of which work against deference to teachers. At the same time, American education policy makes school attendance compulsory, thus forcing many students who have no interest in learning to attend school. In a study of secondary schools, Cusick (1983) concluded that the central problem facing those schools was that of "containing" the students who did not want to be there. Bryk, Sebring, Allensworth, Luppescu, and Easton (2010) found that stronger and weaker schools differed in part in their proportions of students who attended school, were on time, and did their homework. These are all indicators of student cooperation and willingness to be taught. They vary from one classroom to another and are not completely subject to the influence of the teacher. This is probably one reason that, for instance, Raudenbush, Rowan, and Cheong (1992) found that secondary teachers' sense of self-efficacy varied from hour to hour across the periods of the school day. Teachers' sense of efficacy depends on the particular students they teach.

But the problem is not merely one of cooperation and motivation. For even when students are willing to cooperate, they are still children. They are immature, disorganized, unreliable, and easily distractible, and they begin their schooling with little or no experience of the kind of sustained, focused, purposeful work that is required for learning. This simple fact—that teachers work with children rather than adults—has profound implications for teachers. For their task is not merely to teach other human beings about school subjects, but also to teach them how to cooperate and how to learn, neither of which comes naturally, so that teachers must also persuade students that all of this is worthwhile.

Moreover, teachers' classrooms can differ in the number of students who are pulled out during the day for special education, second language learning, or other kinds of additional instruction. These movements further complicate the teachers' work. In a qualitative study of teachers' practices (Kennedy, 2005), I observed a classroom with a severely handicapped student who moaned loudly through the entire lesson, and another in which a boy threw temper tantrums so frequently that the teacher maintained a low-key and even boring classroom atmosphere to avoid flare-ups. In a recent survey of teachers, Public Agenda (2004) found that 80% of the teachers surveyed believed that some of their students were persistent troublemakers and should be removed. Almost

half of the teachers said they had been accused of unfairly disciplining a student, and more than half said their district had backed down from assertive parents.

Even when students try to cooperate, they may still put pressure on teachers to reduce their work load. Several studies point to ways in which students try to “bargain” with teachers to reduce the cognitive demands placed on them (see, e.g., Cusick, 1983; Doyle, 1979, 1983; Metz, 1993; Sedlak, Wheeler, Pullin, & Cusick, 1986; Steinberg, 1996). This literature suggests that students themselves are a prominent influence on teaching practices and often motivate teachers to simplify assignments to maintain peaceful relations with them.

The important point here is that students can disrupt instruction in a variety of ways that are unrelated to their prior achievement levels. So when we use value-added measures of student achievement to assess teacher quality, we may still fail to account for all the ways that individual students can influence the achievement of the class as a whole.

School Incursions Into Classroom Life

Another important descriptor of teachers’ working conditions is when and how institutional practices interfere with instructional time. Some of these incursions are so commonplace that researchers may not even notice them—things like fire drills and public address announcements. But these events, even though ordinary, can disrupt the flow of a lesson and require teachers to backtrack and start their train of thought anew.

In a cross-national video study of teaching practices, Stigler and Hiebert (1999) noticed that there were more interruptions in American classrooms than in most other countries. This observation was reinforced in my study (Kennedy, 2005), which documented such things as mid-lesson telephone calls and visits from central office staff who needed signatures. Similarly, Smith (2000) outlines a number of things that interfere with instruction, including fire drills, assemblies, testing schedules, Christmas festivals, science fairs, field trips, and parent-teacher conference days. Smith concludes that, in many schools, only about 40–60% of the time allocated for instruction is actually used for instruction.

Reform Clutter

The zeal to reform American education has reached such a pitch that teachers are confronted with one or more new initiatives every year. Each new superintendent and each new building principal feels obligated to introduce a new initiative. They shift to block scheduling, then away from it; to a zero-tolerance policy, then away from that; to a new textbook or a new curriculum framework, then to something else. Each initiative requires teachers to revise their routines and strategies and pulls their attention away from their teaching and toward a new logistical problem. Often these initiatives do not alter the content of instruction, but they may nonetheless change the instructional systems within which teachers embed their instruction. Teachers devise a host of rules, routines, standard operating procedures and the like that enable them to teach with fewer interruptions and distractions. When the district or a school alters its own standard operating procedures, teachers often must alter theirs to accommodate the change. Their attention is pulled away from the content to be taught and toward how to alter their own

in-class systems to accommodate the school’s new zero tolerance policy, the new lesson length, the new public announcement schedule, the new attendance policy. Reforms can require teachers to make substantial adaptations in their annual, unit, and lesson plans; participate in additional training programs or additional staff meetings; fill out new forms; test their students more often; write longer and more complex lesson plans; post assignments online; and so forth. When the new ideas are both voluminous and time-consuming, teachers can suffer from *reform fatigue*—hence the common reference to teachers who “shut the door” so they can run the classroom as they see fit, whatever the policies and rules of the school as a whole.

Most reforms distract teachers from the core of their work, forcing them to stop thinking about science or history and to think instead about scheduling, grouping, or recordkeeping. But reforms can also have differential effects on teachers within the same building. The differences we see in teachers’ effectiveness within a school may reflect these influences as much as, or more than, they reflect the effects of the teachers’ own personal characteristics.

Discussion

Our search for enduring teacher qualities has a long history. In the 1920s, Charters and Waples (1929) surveyed parents, teachers, school administrators, and teacher educators in a quest to find an agreed-upon list of important teacher traits. On the basis of this research, the authors settled on a list of 83 traits, which they asked a panel of experts to reduce to 25. They then devised a list of “trait actions,” that is, observable behaviors that would demonstrate presence of each trait (recall that the very notion that traits can be inferred from observed actions is at the heart of the Fundamental Attribution Error). Their lists of traits and trait actions help us see what people expected from teachers at that time. For example, one important trait was “adaptability,” and a trait action deemed relevant to the trait of adaptability was that the teacher “shows a willingness to put up with a poor school system, and unfriendly community.” I received a similar comment from an anonymous reader of an earlier draft of this essay, who said that “while situations may vary from teacher to teacher and these differences in situations may make effects on students variable, teachers must somehow accommodate these variations in situation *and still be effective* [emphasis added].”

The current emphasis on teacher accountability seems to be based on a similar sentiment: Even if schools encroach on teachers’ planning time, interfere with classroom events, assign teachers to subjects outside their fields, or assign students who are disruptive to their classrooms, we nonetheless expect teachers to find a way to be effective. We measure and track their value-added test scores but we do not measure their teaching loads, planning time, student absences, proportion of difficult-to-teach or resistant students, frequency of outside interruptions, access to textbooks or equipment of good quality, or whether their instructional materials arrived before the school year began.

Readers might wonder how a research agenda could be devised that would help us learn more about these aspects of teaching. One source of knowledge about school influences on teachers is research on curriculum and test alignment, where researchers have developed measures of curriculum alignment that enable us to learn the extent to which the enacted curriculum matches the

intended curriculum and the extent to which the tested content matches the curricular content. This work was recently reviewed by Martone and Sireci (2009). Another source of knowledge is literature on aspects of school organization and climate that are relevant to student learning, well summarized and expanded upon recently by Bryk et al. (2010). For the most part, however, the school climate literature seems to assume that school climate acts uniformly on all teachers in a given building, as if its effects appeared mainly in school-to-school variations in student learning, not in classroom-to-classroom variations. This is likely not the case, for each teacher's daily experience in a school is necessarily unique to that teacher. Teachers vary in the students they serve, the extracurricular duties they take on, the amount of time they have for planning, and so forth. Nevertheless, the school climate literature is relevant in that it points to many ways in which school policies and practices can either facilitate or hinder teachers' effectiveness, and it is worth examining by any researcher interested in the relative importance of teachers' personal qualities versus their teaching situations.

A third source of knowledge about situational factors relevant to teaching is the qualitative literature on teaching. For instance, Ingersoll (2003b) examines the extent to which teachers feel they have control over a number of different types of school policy decisions; he provides numerous examples of ways in which the teachers he studied were rewarded or punished through principals' allocation of classes, materials, and students. On the basis of his findings he offers hypotheses about how these practices influence teachers and their work with students. Similarly, Johnson's (1990) early study of teachers' working conditions focuses on the effect of these circumstances on teachers' sense of satisfaction and their interest in remaining in the job. My study of teaching practices (Kennedy, 2005) provides numerous examples of classroom interruptions—telephones ringing, public address announcements, unexpected visits from central office personnel, students coming from or going to special classes, and tells how teachers respond to these events. Early research by Walter Doyle (1979, 1983, 1986a, 1986b) provides a number of insights into how students themselves can impede teachers in their work.

The question I raise is not whether teachers make a difference, nor is it about what constitutes teacher quality. My question is, To what extent is the quality of teachers' everyday *practice*—actual classroom behavior—really a function of enduring personal qualities that they bring with them, and to what extent is it a function of schedules, materials, students, institutional incursions into the classroom, and the persistent clutter of reforms that teachers must accommodate? Certainly, some teachers are better able than others to accommodate the variety of stresses and strains they may face in their work. And to the extent that this is so, these differences may be due to differences in personal qualities.

Still, social psychologists tell us that behavior tends to be more influenced by the situations we face than by our own enduring personal qualities. They are so convinced of this that they believe we are making a Fundamental Attribution Error when we attribute actions to personal qualities. If they are right, then we education researchers have succumbed to that error ourselves: We search for the teacher characteristics that can account for differences in effectiveness, and we fail to consider aspects of teaching situations that might also be relevant.

The irony is that the behavior of researchers is also affected by their research situations. Most states and districts maintain data on teachers' credentials, test scores, college majors, and so forth; they do not maintain data on how much free time teachers have, how many different subjects they teach each day, or how many extracurricular responsibilities they have. Researchers are also limited by the circumstances of their funding agencies, which may hesitate to pay the cost of gathering data on these more difficult-to-define and difficult-to-measure variables. *We study teachers' credentials because we can.*

But if we were observed by a sociologist or an organizational psychologist, perhaps that observer would make a fundamental attribution error too. Being unaware of our research situation, the observer might conclude that we were inherently short-sighted or naïve because we persisted in seeking some enduring teacher characteristic that could account for variations in student learning, and persistently overlooked the influence of teachers' situations. Such an observer might equate us with the archetypal man who has lost his watch in the alley but looks for it under the street lamp, where it is easier to see.

REFERENCES

- Aaronson, D., Barrow, L., & Sander, W. (2003). *Teachers and student achievement in Chicago public high schools*. Chicago: Federal Reserve Bank of Chicago. Retrieved from http://www.chicagofed.org/webpages/publications/working_papers/2002/wp_28.cfm
- Babcock, J., Babcock, P., Buhler, J., Cady, J., Cogan, L., Houang, R., et al. (2010). *Breaking the cycle: An international comparison of U.S. mathematics teacher preparation*. Amsterdam: International Association for the Evaluation of Educational Achievement. Retrieved from <http://www.educ.msu.edu/content/default.asp?contentID=710>
- Boyd, D., Grossman, P., Hammerness, K., Lankford, R. H., Loeb, S., McDonald, M., et al. (2008). Surveying the landscape of teacher education in New York City: Constrained variation and the challenge of innovation. *Educational Evaluation and Policy Analysis*, 30(4), 319–343.
- Boyd, D., Grossman, P., Lankford, R. H., Loeb, S., & Wycoff, J. (2008). *Measuring effect sizes: The effect of measurement error*. Paper presented at the National Conference on Value-Added Modeling.
- Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing schools for improvement*. Chicago: University of Chicago Press.
- Charters, W. W., & Waples, D. (1929). *Commonwealth Teacher Training Study*. Chicago: University of Chicago Press.
- Clotfelter, C., Vigdor, J., & Ladd, H. (2007). *How and why do teacher credentials matter for student achievement?* Cambridge MA: National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w12828>
- Cohen, D. K. (1988). Teaching practice: Plus que ça change. . . . In P. Jackson (Ed.), *Contributing to educational change: Perspectives on research and practice*. Berkeley, CA: McCutchan.
- Cohen, D. K. (2010). Teacher quality: An American educational dilemma. In M. M. Kennedy (Ed.), *Teacher assessment and the quest for teacher quality*. San Francisco: Jossey-Bass.
- Cohen, D. K., & Moffitt, S. (2009). *The ordeal of equality: Did federal regulations fix the schools?* Cambridge, MA: Harvard University Press.
- Constantine, J., Player, D., Silva, T., Hallgren, K., Grider, M., & Deke, J. (2009). *An evaluation of teachers trained through different routes to certification*. Washington, DC: U.S. Department of Education. Retrieved from <http://nces.ed.gov/Pubsearch/pubsinfo.asp?pubid=NCEE20094043>

- Cusick, P. A. (1983). *The egalitarian ideal and the American high school: Studies of three schools*. New York: Longman.
- Doyle, W. (1979). Classroom effects. *Theory Into Practice*, 18, 138–144.
- Doyle, W. (1983). Academic work. *Review of Educational Research*, 53(2), 159–199.
- Doyle, W. (1986a). Classroom organization and management. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 392–431). New York: Macmillan.
- Doyle, W. (1986b). Content representation in teachers' definitions of academic work. *Journal of Curriculum Studies*, 18(4), 365–379.
- Drago, R., Caplan, R., Costanza, D., Brubaker, T., Cloud, D., Harris, N., et al. (1999). New estimates of working time for elementary school teachers. *Monthly Labor Review*, 122(4), 31–40.
- Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, 117(1), 21–38.
- Goldhaber, D. (2007). Everyone's doing it, but what does teacher testing tell us about teacher effectiveness? *Journal of Human Resources*, 42(4), 765–794.
- Hill, H., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371–406.
- Humphrey, R. (1985). How work roles influence perception: Structural-cognitive processes and organizational behavior. *American Sociological Review*, 50(2), 242–252.
- Ingersoll, R. M. (2003a). *Out-of-field teaching and the limits of public policy*. Philadelphia: University of Pennsylvania Consortium for Policy Research in Education.
- Ingersoll, R. M. (2003b). *Who controls teachers' work? Power and accountability in America's schools*. Cambridge MA: Harvard University Press.
- Johnson, S. M. (1990). *Teachers at work: Achieving success in our schools*. Boston: Basic Books.
- Kennedy, M. M. (2005). *Inside teaching: How classroom life undermines reform*. Cambridge MA: Harvard University Press.
- Kennedy, M. M. (2010). *Teacher assessment and the quest for teacher quality: A handbook*. San Francisco: Jossey-Bass.
- Kennedy, M. M., Ahn, S., & Choi, J. (2008). The value added by teacher education. In M. Cochran-Smith, S. Feiman-Nemser, & J. McIntyre (Eds.), *Handbook of research on teacher education* (3rd ed., pp. 1249–1273). New York: Routledge.
- Krantz-Kent, R. (2008). Teachers' work patterns: When, where, and how much do U.S. teachers work? *Monthly Labor Review*, 131(3), 52–59.
- Labaree, D. F. (2010). Teach for America and teacher ed: Heads they win, tails we lose. *Journal of Teacher Education*, 61(1–2), 48–55.
- Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum.
- Martone, A., & Sireci, S. G. (2009). Evaluating alignment between curriculum, assessment, and instruction. *Review of Educational Research*, 79(4), 1332–1361.
- McCaffrey, D., Sass, T., Lockwood, J. R., & Mihaly, K. (2009). The intertemporal variability of teacher effect estimates. *Education Finance and Policy*, 4(4), 572–606.
- Metz, M. (1993). Teachers' ultimate dependence on their students. In J. Little & M. McLaughlin (Eds.), *Teachers' work: Individuals, colleagues and context*. New York: Teachers College Press.
- Papay, J. P. (in press). Different tests, different answers: The stability of teacher value-added estimates across outcome measures. *American Educational Research Journal*.
- Porter, A. C., Polikoff, M. S., & Smithson, J. (2009). Is there a de facto national intended curriculum? Evidence from state content standards. *Educational Evaluation and Policy Analysis*, 31(3), 238–268.
- Porter, A. C., Smithson, J., Blank, R., & Zeidner, T. (2007). Alignment as a teacher variable. *Applied Measurement in Education*, 20(1), 27–51.
- Powell, A. G. (1996). Motivating students to learn: An American dilemma. In S. H. Fuhrman & J. A. O'Day (Eds.), *Rewards and reform: Creating educational incentives that work* (pp. 19–59). San Francisco: Jossey-Bass.
- Public Agenda. (2004). *Teaching interrupted: Do discipline policies in today's public schools foster the common good?* New York: Author. Retrieved from <http://www.publicagenda.org/reports/teaching-interrupted>
- Raudenbush, S. W., Rowan, B., & Cheong, Y. F. (1992). Contextual effects on the self-perceived efficacy of high school teachers. *Sociology of Education*, 65(2), 150–167.
- Rosenshine, B. (1970). The stability of teacher effects upon student achievement. *Review of Educational Research*, 40(5), 647–662.
- Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173–220). New York: Academic Press.
- Rowan, B., Correnti, R., & Miller, R. J. (2002). What large-scale survey research tells us about teacher effects on student achievement: Insights from the Prospects study of elementary schools. *Teachers College Record*, 104(8), 1525–1567.
- Schmidt, W. H., McKnight, C. C., Cogan, L. S., Jakwerth, P. M., & Houang, R. T. (1999). *Facing the consequences: Using TIMSS for a closer look at U.S. mathematics and science education*. Boston: Kluwer.
- Sedlak, M., Wheeler, C., Pullin, D. C., & Cusick, P. A. (1986). *Selling students short: Classroom bargains and academic reform in the American high school*. New York: Teachers College Press.
- Shattuck, R. (2005). The shame of the schools. *New York Review of Books*, 52(6).
- Smith, B. (2000). Quantity matters: Annual instructional time in an urban school system. *Educational Administration Quarterly*, 36(5), 652–682.
- Steinberg, L. (1996). *Beyond the classroom: Why school reform has failed and what parents need to do*. New York: Simon and Schuster.
- Stevenson, H. W., & Stigler, J. W. (1992). *The learning gap: Why our schools are failing and what we can learn from Japanese and Chinese education*. New York: Summit Books.
- Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York: Free Press.
- Stotsky, S. (2007). *The negative influence of education schools on the K–12 curriculum*. Washington, DC: American Enterprise Institute. Retrieved from www.aei.org/event1595
- Tuttle, C. C., Anderson, T., & Glazerman, S. (2009). *ABCTE teachers in Florida and their effect on student performance*. Washington, DC: Mathematica. Retrieved from <http://www.mathematica-mpr.com/publications/SearchList2.aspx?jumpsrch=yes&txtSearch=ABCTE>

AUTHOR

MARY M. KENNEDY is a professor at Michigan State University, College of Education, 116 Erickson Hall, East Lansing, MI 48823; mkenedy@msu.edu. Her research focuses on teaching, teacher learning and contributions to teaching quality.

Manuscript received July 13, 2010

Revisions received September 20, 2010

Accepted October 6, 2010