Teachers College Record
The Voice of Scholarship in Education

What Do We Know About How Teachers Influence Student Performance on Standardized Tests: And Why Do We Know so Little About Other Student Outcomes?

by Thomas L. Good — 2014

Background/Context: Since the 1970s, researchers have attempted to link observational measures of instructional process to student achievement (and occasionally to other outcomes of schooling). This paper reviews extensively both historical and contemporary research to identify what is known about effective teaching.

Purpose/Objective: Good, after reviewing what is known about effective teaching, attempts to apply this to current descriptions of effective teaching and its application value for practice. Good notes that much of the "new" research on effective teaching has simply replicated what has been known since the 1980s. Although this is not unimportant (since it shows that older findings still pertain to contemporary classrooms), it is unfortunate that research has not moved beyond the relationship between general teacher behavior (those that cut across subject areas) and student achievement (as measured by standardized tests). How this information can be applied and the difficulty in using this information is examined in the paper.

Research Design: The paper is a historical analysis and reviews research on teaching from the 1960s to today.

Conclusion: Conclusion: This paper has stressed that our data base on effective teaching is limited—still it has some implications for practice. Even though the knowledge base is limited, there is no clear knowledge that teachers-in-training learn and have the opportunity to practice and use. It would seem that teacher education programs would want to assure that their graduates, in addition to possessing appropriate knowledge, would also have clear conceptual understanding and skills related to active teaching, proactive management, communication of appropriate expectations for learning, and the ability to plan and enact instruction that balances procedural and conceptual knowledge. Future research on the use of this knowledge base and its effects in teacher education programs would be informative. If done correctly, research on teaching can improve instruction. However, the research must be applied carefully if it is to have useful effects. And, as noted often in this paper, research must consider outcomes of schooling other than achievement such as creativity, adaptability, and problem finding.

INTRODUCTION

This paper describes what we know about how teachers influence student performance on standardized achievement tests. This is valuable information and every teacher should understand this knowledge base and be able to apply it in their classroom teaching. However, this knowledge base must be qualified in two important ways. First, good teaching involves much more than the knowledge of how to teach concepts, and it is unfortunate that research on teaching is silent on many important outcomes such as problem solving, creativity, originality, and problem finding. Second, most of the knowledge base linking teaching and student achievement is drawn from research in K–8 schools and much of this knowledge was collected in mathematics classrooms.

Before reviewing the research on teaching literature, I address four issues. These contextual issues or beliefs are (a) teachers need to perform many duties well and at the same time; (b) teaching is seen by many as easy; (c) teachers and schools have been a concern for a long time; and (d) teacher performance should be stable over time. I raise a few questions and provide a general context for understanding the research on teaching literature.

In October 2012, while watching snippets from both political conventions, I heard Paul Ryan, the vice presidential candidate for the Republican Party say, "If Romney is elected, he will solve the education problem." I turned to my wife, Mary, and asked, "What is the education problem?" She gave a disinterested shrug. Then I asked myself, "What is the education problem?" I could not answer. How could I not answer such a simple question? After all, I have studied education for many years. So, what did Paul Ryan mean by the "education problem"? I generated several possibilities, including:

- huge student dropout rate;
- lack of funding across states and districts;
- teaching is easy and teachers are paid too much even though they do not work in the summer;
- teachers get too little respect or compensation;
- high rate of teacher attrition;
- low rate of parent involvement;
- schools have reduced the time for physical education, art, and music;
- low rate of student commitment;
- teacher unions;
- teacher education programs;
- the soaring cost of college tuition, especially in states such as Arizona, California, and Nevada;
- student ACT scores held steady for the class of 2012—perhaps delaying the chance for test publishers and curriculum writers to prepare new tests and new curriculums;
- the poor job market for high school and college graduates;
- too many *or* too few charter schools;

- too much *or* too little support for vouchers;
- too much time on cursive writing;
- building complex, but inadequate, data systems to monitor student achievement;
- and many more.

Fortunately, Gail Collins came to my rescue in her *New York Times* op-ed on September 1, 2012 (p. A23). Her column helped me to reflect upon Ryan's comment about the education problem. She wrote that Ryan said, "But you may be interested to know that the proven methods for improving school performance are high standards, accountability . . . [and now please understand that she used direct quotes] . . . renewed focus on the constitution and the writings of the founding fathers." Research on teaching, as I understand it, does not support Congressman Ryan's perspective.

It would be inappropriate to generalize widely from the perspective of a single congressman; however, I submit that the history of school reform is largely based on the erroneous belief that at a given point in time, the "unproductive state" of educational progress is largely due to the conception of a single or very limited set of problems. In contrast, I argue that at any point in time, educational progress is influenced by many problems. Hence, the reduction of complex problems and overlapping constraints of a single issue are inappropriate as it leads to misguided reform and a waste of considerable resources.

TEACHERS ARE ASKED TO DO MANY THINGS

Part of the reason that some citizens are dissatisfied with public education is because teachers are asked to do so much and the array of duties they face guarantees that someone or some group will be dissatisfied with some aspect of their behavior. For example, teachers are expected to influence students' civic-mindedness and to prepare them for a life in a democratic society, as well as nurture student creativity and responsibility. These tasks are overwhelming, but teachers are also expected, at least in some settings, to mentor other teachers and to serve on school management leadership teams. For a list of some of the duties or outcomes that teachers are often expected to perform simultaneously, see Appendix I.

The emphasis upon teachers being responsible for so much is in direct contradiction to what other countries expect from their teachers. Sahlberg (2010) noted that one key aspect of educational success in Finnish schools is the emphasis on "less is better." Teachers in many countries are expected to plan and to teach and are not asked to do the myriad range of tasks that American teachers address. And, the high performance of Finnish students on international achievement tests is well known.

TEACHING IS EASY

A second reason that many citizens are dissatisfied with teachers and schools is because they think that teaching is easy (after all, everyone has been a student) and, thus, teachers should accomplish much more than they are perceived to do. In Appendix II, there is a list of innovations that have been advanced at one time or another, and most of these innovations were promoted with the strong belief that they would revolutionize learning. The innovations presented in Appendix II are listed to generate readers' reflection on simple-minded reform ideas. I'm confident that readers can readily add to this list. However, the expectation that these innovations were seen as having the capacity to dramatically improve education seemingly stems from a belief that teaching is easy. And, the fact that these anticipated gains never occurred suggests that the simple solutions advanced for improving education in important ways were wrong. Yet, these failed innovations are quickly forgotten and are soon replaced by new innovations and the cycle continues.

I provide two examples of innovations presumed to be effective from my undergraduate experience at the University of Illinois. (Despite these examples, I report that I received a first-rate education there!) My first experience was with the expert teacher on television. This innovation contended that by having highly knowledgeable experts on TV, many students in various locations would benefit in cost-effective ways. My only experience with this medium was through a course taught by a distinguished professor who talked about teachers as value exemplars. Unfortunately, the professor did not have advanced teaching skills and the experience was so boring that I was stressed and vowed never again to take a television course. So much for the belief that teachers' subject matter knowledge guarantees success (or that all professors on TV could serve as value exemplars).

My second experience was my assignment to a language laboratory. This "innovative panacea" involved about 50 students practicing, listening, and speaking a foreign language at the same time. Obviously, during traditional instruction, only one student speaks at a time. At first glance, this innovation appears to have merit; however, in the lab there was only one graduate assistant and most of us soon learned that we were monitored infrequently and rarely received any useful feedback. So much for the time honored educational psychology principle that immediate and appropriate feedback facilitates learning.

There are many other examples of innovations and simplistic reforms beyond my personal experiences. For example, Goals 2000 provides prima facie evidence that many policymakers believe that teaching is easy. Consider the expectations for radical improvement expressed in three of the Goals 2000 listed below. How could anyone writing such ludicrous goals not believe that teaching is easy? Apparently the authors of the report believed that "if teachers are motivated and willing to work hard, good achievement should follow."

> By the year 2000:
>
> 1. School Readiness: All children in America will start school ready to learn.
> 2. School Completion: The high school graduation rate will increase to at least 90%.
> 3. Mathematics and Science: U.S. students will be first in the world in science and mathematics achievement.

From Public Law 103-227, Goals 2000: Educate America Act (1994).

Perhaps it is possible that now we have learned to think more realistically about achievable educational goals. Unfortunately, much evidence negates this optimistic view. Consider the quote that follows, describing a promise from the District of Columbia Public Schools. **"**At least 90 percent of DCPS teachers and principals will be highly effective or effective, as determined through the IMPACT evaluation system."

It is extremely unlikely that 90% of district teachers will become highly proficient or proficient on the IMPACT evaluation by school year 2016–2017. Further, if 90% of teachers achieved an effective ranking on the observational component (whether the instrument is scored fairly or there is some fudging of scores), I am absolutely certain that higher scores on the observational instrument will not lead to important increases in student achievement. A major component of the IMPACT evaluation system is student achievement. In assessing student achievement, the school district uses value-added (VA) calculations. And, in calculating these VA scores, regression techniques are used to identify more and less effective teachers. Thus, it is mathematically impossible for 90% of the teachers to obtain such ratings (see Lavigne, 2013, Table 2 in this issue). These relative comparisons guarantee that some teachers will be relatively high or low irrespective of actual performance. (I will return to uses/misuses of observational data later in the paper).

SCHOOL REFORM

Another source of evidence that illustrates both that public schools are inadequate and that quick fixes are possible is the many reform movements in American education such as Goals 2000. The assertion that meaningful reform could be achieved quickly has often been advanced. Space prohibits a discussion of these many reforms (such as the Sputnik crisis, A Nation at Risk, Prisoners of Time, No Child Left Behind, and Race to the Top) or the many reviewers who convincingly report that these costly reforms were ineffective (Cuban, 2010; Payne, 2010; Ravitch, 2010). It would appear that at some point, the weight of the combined school reform failures would generate some skepticism when new reforms are proposed; however, intolerance for new reform has not occurred to date among policymakers, politicians, teacher educators, and citizens. New crises and new "quick" solutions continue to appear and disappear unabated.

CITIZENS HAVE LONG BEEN CONCERNED ABOUT EDUCATION

Americans have always been concerned about education and indeed many of the concerns expressed in the 1900 still linger today. Sharon Nichols and I (Nichols & Good, 2004) surveyed articles from the early 1900s that appeared in *The New York Times*, and we discovered that educational issues and expressed concerns were frequently published there. For example, in 1900, the age at which students should start school was being debated as it is today. For more examples, see Appendix III.

These historical concerns about the quality of schools continue today. For example, for a PDK/Gallup poll in 2011 (Phi Delta Kappa, 2011), in response to a question asking citizens to grade public schools, 51% gave a grade of C, 23% gave a grade of D, and 7% gave a grade of F. In the 2012 PDK/Gallup poll (Phi Delta Kappa, 2012), assigned grades were 47% C, 23% D, and 7% F. These grades clearly suggested that many Americans have concerns about public schools and that some hold acute concerns.

Why do so many citizens assign low grades to schools? Obviously, in part, citizens' judgments are a product of the continuous media reports about the "deficiencies" of public education that have been expressed since at least the early 1900s as noted. Public schools have always been of concern and of interest to Americans. Further, an important conveyer of the message that "our schools are in trouble" comes from the parade of school reforms such as No Child Left Behind and, now, Race to the Top.

UNFAVORABLE INTERNATIONAL COMPARISONS

Another source of unfavorable opinions about public schools comes from the frequent, unfavorable comparison of the performance of American students with those in other countries. Do these frequent reminders that our schools are associated with considerably less achievement than those in many other countries lead to a willingness—to invest more in schools—so that we can be more competitive with our international peers or does it dampen citizens' interest in investing in public schools? One recent study suggested that information about low American school performance on international tests leads to less interest in funding schools (Morgan & Poppe, 2012). Although the impact of "negative" news on citizens' willingness to fund schools merits more research attention, this study illustrates an important concern.

However, the media's reports of low-performing schools and teachers are also influenced by other critics including—most curiously—educators. Consider the following quote. "Teachers ignore, resist, subvert, misinterpret, selectively adopt, or otherwise distort reformers' intentions. Changes tend to be superficial, seldom penetrating the core of instructional practice" (Lefstein, 2008). It is important to note that this sharp criticism comes from an education researcher. Although I do not suggest that Lefstein's view is representative of the field, I do think that dissatisfaction with teachers is a prevalent belief held by many teacher educators who write for publication. I do not view criticism as necessarily destructive and, indeed, I believe that accurate and constructive feedback can be very helpful in improving teacher performance and student achievement. However, my point here is that the plethora of negative press can lead citizens to inappropriate assumptions and misunderstandings about the quality of our schools. For example, Morgan and Poppe (2012) after reviewing several studies detailing the low performance of American students on international comparisons contended, "These findings—often conveyed by scholars of educational reform (e.g., Darling-Hammond, 2010; Moe & Chubb, 2009)—appear to have convinced many policy proponents that the United States has fallen behind crucial international competitors in preparing its youth for the workforce and for higher education" (p. 262).

TEACHERS SHOULD HAVE STABLE EFFECTS ON STUDENTS

In the current debate—stimulated by Race to the Top—policymakers advocate using student achievement to measure teachers' effectiveness; in doing so, they make an underlying assumption that teacher effects on students' achievement should be stable over consecutive years. Researchers and policymakers are often surprised when this is not the case. Again, I believe, providing evidence that policymakers and some educators think that teaching is easy. Given that teaching is easy, why wouldn't teachers have similar effects on their students over time? In contrast, if one thought that teaching was difficult, the question might be, how can teachers achieve such similar results over time when events in their personal lives and in their teaching environment change so much?

This belief that teachers should be stable in their effects on students reflects unfavorably on the scholarship of current policymakers and researchers as we have known for some time that teachers' effects on student achievement are highly unstable from year to year (Brophy, 1973; Good & Grouws, 1975). And, we have known for some time that one reason that teachers vary in their effects on student achievement from year to year is because the characteristics of students they teach also vary from year to year (Brophy & Evertson, 1978, 1981), and there is clear evidence that the classroom composition of students in terms of the ratio of high-achieving to low-achieving students also impacts achievement (Beckerman & Good, 1981; Veldman & Sanford, 1984). As teachers well know, sometimes subtle changes can have a huge impact on classroom environment (just as teachers make a difference, so do students). I leave further discussion of this issue of instability as my colleagues Spyros Konstantopoulos (2013) and David Berliner (2013) address this more fully in their papers. However, I do want to argue that a perception that teachers' impact on achievement should be stable leads to an easy assumption that if teachers are unstable over time, they must not be putting the necessary time and effort into their teaching.

The belief that teachers should be stable in their performance effect is challenged by evidence that highly accomplished professional bowlers and golfers vary notably in their performance over time. We know that the conditions of bowling are very stable. In most instances, the bowling lane is 60 feet long and 42 inches wide. The pin is made of hard maple. It has a diameter of 2.25 inches at the base and a circumference of 15 inches at its widest point. The pin must weigh between 3 pounds 6 ounces and 3 pounds 10 ounces; the ball has a circumference of no more than 27 inches and has to weigh 10 to 16 pounds and may have 2 or 3 finger holes; and bowlers' performance is scored with 100% reliability. Bowling conditions and the ways in which success is measured remain stable, which stands in marked contrast to the highly complex, changing, and varied conditions that teachers face. Teachers would love to have such stable conditions. So, given these stable workplace conditions, wouldn't we expect bowlers, especially talented professional bowlers, to be stable in their performance over time? An examination of Table 1 shows that this is not the case.

Table 1. *Top Professional Bowling Association Tour Winners*

Top PBA Tour Winners 2008-2011			
	2008-2009	**2009-2010**	**2010-2011**
Norm Duke	1	12	4
Rhino Page	2	9	**NR**
Mike Scroggins	3	3	**NR**
Patrick Allen	4	**NR**	**NR**
Wes Malott	5	7	**NR**
Chris Barnes	6	4	2
Walter Ray Williams Jr.	7	1	**NR**
Bill O'Neill	8	2	5
Brad Angelo	9	**NR**	**NR**
Mike Fagan	10	8	17
Pete Weber	11	10	**NR**
Parker Bohn, III	12	**NR**	18
John Nolen	13	**NR**	**NR**
Tommy Jones	14	5	16
Mika Koivuniemi	15	20	1
Danny Wiseman	16	**NR**	**NR**
Steve Harman	17	**NR**	**NR**
Sean Rash	18	18	7
Jeff Carter	19	**NR**	**NR**
Ryan Shafer	20	**NR**	**NR**
Tom Smallwood	**NR**	6	3
Jason Belmonte	**NR**	11	8
Jack Jurek	**NR**	13	15
Brian Kretzer	**NR**	14	**NR**
Mike DeVaney	**NR**	15	**NR**
Ryan Ciminelli	**NR**	16	13

Jason Couch	**NR**	17	20
Brian Voss	**NR**	19	**NR**
Dick Allen	**NR**	**NR**	6
Dan MacLelland	**NR**	**NR**	9
Andres Gomez	**NR**	**NR**	10
Osku Palermaa	**NR**	**NR**	11
Tom Hess	**NR**	**NR**	12
Tom Daugherty	**NR**	**NR**	14
Michael Haugen, Jr.	**NR**	**NR**	19

As shown in Table 1, out of the top 10 bowlers in the United States in 2008–2009, 6 of these bowlers were not ranked in the top 40 in 2010–2011. Table 1 also shows that of the bowlers ranked from 11 to 20 in 2008–2009, 60% of these bowlers were not ranked at all in 2009–2010 or 2010–2011. I have also explored the stability of professional golfers' performance over time. It is very clear that even the best golfers in the world vary from year to year and often sharply so. In examining the stability of the top 200 golfers in the world in 2007, their performance in 2008, and then three years later in 2010, the variance in performance was notable. From 2007–2008, 48 of the best golfers in the world were not ranked in the following year. Eighty-eight of the top 200 golfers in the world in 2007–2008 were not ranked three years later. (At present, these data are being compiled and a full accounting will appear in Lavigne & Good, in press.)

SOME GOOD NEWS

Despite the fact that Americans generally have low expectations for public schools, it is important to differentiate their general beliefs from the more specific beliefs about schools in their own community and the school that their oldest child attends. For example, in the 2012 PDK Gallup poll, 48% of citizens gave the schools in their own community a grade of A or B. In marked contrast, 77% of citizens gave the school their oldest child attends a grade of A or B. Further, only 6% of the respondents assigned a grade of D or F. This is an important distinction; clearly, Americans think that the school that they know best (the one their child attends) does an effective job. Since citizens are much more favorable toward schools that their children attend than schools in general, it is arguable that the general negative press that teachers and schools receive lowers citizens' perceptions of schools in general, but not the ones attended by their children.

Parents often visit schools (especially when their children are younger), occasionally talk with teachers and other parents, and sometimes attend school functions such as parent–teacher conferences. However, the only information citizens receive about "schools in general" is through media accounts, national reports, and other data that are disseminated. It appears to me that media accounts, even those in professional teacher publications, are consistently negative and often present highly pejorative descriptions of teaching. These far outweigh the positive accounts. Teachers, as practicing professionals, are unique in being criticized often by the very institutions that trained them. It is rare to see articles written by faculty from business, engineering, journalism, arts and science, and medical schools describing the practice of their graduates in highly negative ways. Thus, it is not surprising that teachers in some schools have become demoralized (Payne, 2010).

CLOSING ARGUMENT

In this introductory section, I have contended that American teachers are expected to do many things well (unlike other countries where teachers are asked to focus on teaching subject matter content). Yet, there is no consensus on what teachers are to accomplish other than to raise student achievement. Some citizens want to reduce student dropout rates, others want gifted students challenged, and so forth. Despite parents' and some citizens' desires for schools to address an array of academic, personal, and social outcomes, policymakers focus on "the education problem"—inadequate student achievement, especially in contrast to students' performance in other countries. In many ways, this focus on student achievement (as a teacher problem) is naïve and simplistic because although teachers are important influences on student achievement, there are systemic effects —poverty, inequality, and inadequate resources—that are beyond individual teacher efforts. One example of forces that reduce teacher effects on student achievement is the fact that the negative effects of low SES and poverty on student achievement is much higher in the United Sates than in other developed countries (Baldi, Jin, Skemer, Green, & Herget, 2007). Berliner (2013) further identifies many factors other than teachers that can facilitate or impede student achievement.

This context (acute concern over public schools) is conducive to the spread of negative information about schools and teachers, and if citizens believe that teaching is easy, then those who advance new innovations and reforms will not face much resistance, especially when education experts proactively fail to question new movements that have little, if any, empirical evidence to support the new reform. And, of course, sometimes it is even academics who lead and advocate for the reform.

TEACHERS MAKE A DIFFERENCE

So, educational reform is fueled by faddism and advocacy, and it is based on little, if any, research evidence. It seems useful to ask, is there any research that links instruction to student achievement? Is there any credible evidence to show that some teachers have a more positive impact on student achievement than other teachers? And, if so, how do more effective teachers influence student achievement?

Importantly, as recent as 1970, it was not believed that teachers or schools had an important impact on student achievement. The

social science community "knew" that student achievement in school was largely predicted by issues of student intelligence and family conditions. Most educators also believed that teachers had little impact on student achievement. Despite this societal belief, there was some research relating teaching to student achievement. Although these data were correlational, some of the findings were very consistent, especially those reported by Rosenshine and Furst (1973). Further, in addition to the relationship between instructional behaviors and student achievement, there was clear evidence that opportunity to learn material was positively correlated with student achievement (Lundgren, 1972).

In 1975, I wrote a book entitled *Teachers Make a Difference* with two friends, Bruce Biddle and Jere Brophy (Good, Biddle, & Brophy, 1975). In our book, we laid out a logical argument that teachers impacted student achievement in important ways and we brought evidence to bear that challenged the then very popular assumptions that teaching had little impact on achievement. Further, we supported that argument by reviewing empirical evidence that was available. Our 1975 review concluded with the following statement:

> Just as data suggesting that teachers do not make a difference frequently have been overgeneralized and accepted uncritically, there is a danger that the same mistakes can be made with the kinds of data cited in this chapter. While these data do establish that certain teachers consistently outperform others and that particular kinds of teaching behavior are almost always preferable to contrasting kinds, the present state of research in the area does not allow for the use of process–product findings for accountability purposes. Ultimately, findings of this type might be usable for making decisions about teacher tenure or pay issues, but at present they are too sparse and unsophisticated to be used for such purposes legitimately. We believe that officials charged with making accountability decisions must base them on reasonable and appropriate criteria. In our belief, such criteria do not exist at this time. Thus, before prematurely committing time and money to an imperfect accountability system, time and money should be allocated to the basic research needed to identify appropriate and inappropriate teacher behavior clearly and unambiguously. (p. 85)

At the time, we noted that information about teacher effectiveness could be *overvalued* and we cautioned against the use of this information alone for evaluating teaching. There were many who followed our analysis with better empirical data to illustrate that teachers made a difference in student achievement. As one example, I highlight the work that I did with Doug Grouws to illustrate that fourth grade mathematics teachers varied in their effectiveness to advance student achievement. However, again, I want to make it clear that this is but one research paradigm and that many others contributed to research on teaching that provided a coherent and useful knowledge base. Important contributors to this literature included Barak Rosenshine, Norma Furst, Nate Gage, Bruce Biddle, Jere Brophy, David Berliner, Carolyne Evertson, Jane Stallings, Robert Soar, Ed Emmer, and many others. In describing this literature, I can but briefly allude to a few historical contributions and summarize key findings. (For more detail on research methodology and findings, see Brophy & Good, 1986; for more information on the implications of this research, see Good & Brophy, 2008, and Lavigne & Good, in press.)

GOOD AND GROUWS' NATURALISTIC STUDY

Doug Grouws and I were extremely interested in seeing if we could identify teachers who obtained higher achievement scores than other teachers teaching fourth grade mathematics to similar students under similar circumstances. In comparing teachers' effects on their students' achievement over two consecutive years, we found that it was possible to identify teachers who were consistently high or consistently low. However, importantly, some teachers (like the performance of professional bowlers mentioned above) were not stable in their achievement effects. Hence, for some time, there has been clear evidence that teacher effectiveness, as measured by student achievement scores, often varies from year to year.

After identifying more and less effective teachers, we built an observational system in order to determine how these two groups of teachers differed in their classroom behavior. Many of the variables that were included in our observational system were those that came from earlier correlational research. Our classroom observational system provided data that generally supported the earlier correlational research. Our findings appear in Table 2 (for more information, see Good & Grouws, 1975, 1979).

Table 2. Findings from Good and Grouws Naturalistic Study

- Teacher provides a review of material previously presented, collects homework, and asks students to engage in mental computation work.
- Teacher conducts development lesson where the teacher focuses on meaning and promoting student understanding through discussion and demonstration.
- Development also includes the assessment of student understanding, the asking of both process and product questions, and controlled practice where teachers elaborate upon the meaning as necessary.
- Teacher conducts homework using appropriate pace, alerting, and accountability.
- Students are given 15–20 minutes of uninterrupted time for successful practice.
- Seat work involves momentum—keeping the ball rolling—getting everyone involved.
- Seat work also involves appropriate alerting (e.g., their work will be checked) and accountability (e.g., the work is actually checked).
- Homework assignment if students have done seat work well.
- Homework should include review problems as well as current work.

EXPERIMENTAL STUDY

The data from our correlational study provided information about how teachers differed in their instructional behavior. As noted,

those teachers who achieved higher results exhibited the behaviors presented in Table 2 more frequently and consistently than teachers who achieved lower results. Our next step was to identify a school district that was willing to participate in an experimental study, allowed us to assign teachers to an experimental group that would obtain the findings from our naturalistic study right away, and assign other teachers to a control group (who would not get information until after the study was completed). Then, the experimental teachers participated in a workshop describing how more effective teachers taught in the earlier study and they were given a manual providing extensive examples of various parts of the treatment (what was meant by development, controlled practice, accountability, and so forth).

The manual basically described what we had found in our earlier study, but some new content was added from other sources. One important addition came from our knowledge of experimental work illustrating the importance of developing the meaning of the mathematic content being studied, which is often called attending to the conceptual aspects of the content or teaching for understanding. We referred to this focus on understanding as *development*. Several experiments had indicated that students learned math better when more time was placed on development (developing meaningful understanding) than was spent on practice (Dubriel, 1977; Shipp & Deer, 1960; Shuster & Pigge, 1965). Accordingly, we asked teachers to spend 20 minutes on development in each lesson.

We observed both experimental and control teachers (hence, attention to mathematics instruction was heightened in all classrooms) and collected student achievement data before and after the treatment was administered. Our results indicated that teachers in the experimental group implemented more of the recommended treatment and had more positive effects on student achievement than the control teachers (Good & Grouws, 1979). These findings established a causal link between instructional behaviors and student achievement.

Thus, our research not only provided strong evidence that teachers make a difference, but also provided information about what teachers *did* to enhance achievement. It should be noted that some types of teachers and some types of students benefited more from the treatment than other combinations of teachers and students (for more information about how teacher typologies and student typologies were measured and analyzed see Ebmeier, 1978, and Ebmeier & Good, 1979). Thus, it is important to consider that although our treatment was generally successful, the results suggested that it may be possible to adjust the treatment in ways to improve student performance for some teachers and students. And, as I will note later, observation systems used today make no attempt to consider whether or not teaching makes more or less sense for different types of students in the classroom.

The experiment Doug and I conducted was an important demonstration of teacher effects and, today, it remains arguably the strongest demonstration of an experimental link between teaching and student learning. Indeed, our study is one of the clearest examples of how teachers influence student achievement (Carpenter, Dossey, & Koehler, 2004).

SUMMARY

Brophy and Good argued in 1986 that their review refuted the critics who believed that anyone could teach ("those who can do; those who can't teach") and that teachers had little impact on student achievement. Our response to those who did not believe that teachers make a difference follows:

> Although it may be true that most adults could survive in the classroom, it is not true that most could teach effectively. Even trained and experienced teachers vary widely in how they organize the classroom and present instruction. Specifically, they differ in several respects: the expectations and achievement objectives they hold for themselves, their classes, and individual students; how they select and design academic tasks; and how actively they instruct and communicate with students about academic tasks. Those who do these things successfully produce significantly more achievement than those who do not, but doing them successfully demands a blend of knowledge, energy, motivation, and communication and decision-making skills that many teachers, let alone ordinary adults, do not possess (p. 370).

HOW DO TEACHERS INFLUENCE STUDENT PERFORMANCE?

The research that followed during the next 20 years further clarified what Bruce, Jere, and I argued, that good teaching required extensive knowledge and talented professionals. However, as noted, research on effective teaching that was focused on instructional behavior essentially stopped in the 1990s and early 2000s (see Konstantopoulos, 2013, in this issue). Research moved on to new topics and provided useful knowledge on other important topics such as teacher beliefs, teacher subject matter knowledge and performance, and project-based learning, to name but a few. However, unfortunately, this research was limited because it largely involved the use of purposeful samples (as opposed to carefully selected teachers who varied in important ways) and typically did not collect achievement data. Although these data sometimes provided thoughtful analyses and hypotheses, the failure to include achievement measurements largely rendered these results unimportant for policy decisions.

Still, some research on teaching effects continued, using large samples of teaching. Brophy (2012) and Good and Brophy (2008) integrated these findings to describe general dimensions of effective teaching. What teachers do to enhance student performance on standardized tests appears in Table 3. These are the instructional behaviors that are associated with student achievement and that tend to cut across different subjects and grade levels. Hence, these instructional behaviors can be seen as general indicators of what teachers can do to facilitate student achievement. For elaboration of these variables, and for information about subject-specific considerations of effective teaching (of which far less is known), see Good and Brophy (2008).

Table 3. General Principles of Effective Teaching

- appropriate teaching expectations
- effective use of time
- proactive classroom management
- supportive and caring classrooms
- opportunity to learn
- coherent curriculum content
- curriculum alignment
- thoughtful discourse
- scaffolding students' ideas and task involvement
- practice/application
- goal-oriented assessments

Notes. Adapted from Good & Brophy (2008).

I discuss five of these individual behaviors briefly not only to define them, but also to illustrate that these individual variables are better seen as clusters of variables. Single variables have no predictable relationship with student achievement. While discussing these variables, I briefly note the original research that identified these variables, and I show that recent research continues to illustrate the utility of these variables. Hence, my discussion combines something old with something new.

OPPORTUNITY TO LEARN

Opportunity to learn trumps everything. If students are not exposed to material, they are less likely to learn it and certainly will not understand it well. Some teachers emphasize factual information, others emphasize conceptual knowledge, and some balance factual and conceptual knowledge. Some teachers grade student essays for grammar and style, and others place more emphasis on content. Sometimes the variation in assigned content and the conceptual level at which it is addressed varies dramatically from teacher to teacher. These differences in assignments influence what students learn. Berliner and Biddle (1995) were so impressed by the research on opportunity to learn that they called it Berliner and Biddle's first law of learning. (Note they had no second law!). At first glance this may produce a "dah" reaction, as the point is subtle but powerful. Students do not receive equal opportunity in the classroom. Some students receive a diet of drill and structure; others have more choice and challenge, and so forth. As Berliner and Biddle (1995) and Good and Braden (2000) noted, one reason why students in Japan do better on algebra tests than American students is because algebra is taught considerably earlier in the curriculum in Japan than it is here. It should be no surprise that students do less well on algebra concepts when they have not had an opportunity to learn the material than those who are taught algebra. Opportunity to learn is well documented in the literature and recent articles continue to describe the robustness of this concept (Schmidt & Maier, 2009). It is useful to note, as McCaslin (2006) argued in her work, that the opportunity to learn involves not only acquisition of content, but also student dispositions and motivational outcomes. Her work comparing policymakers' conceptions of student motivation across various school reform movements suggests that these enactments of policymakers' conceptions influence how students come to know what it means to be a learner.

EFFECTIVE USE OF TIME

Opportunity to learn refers to the focus on the curriculum and on classroom climate and the opportunity to learn the formal and informal curriculum (McCaslin & Good, 1996). Effective use of time refers more specifically to how teachers use time for instruction or for classroom management. Considerable evidence illustrates that some teachers use most of their time for instruction and that others spend considerably less time on instruction.

We have known for some time that how teachers use time relates to student achievement. One excellent example of the importance of using time well was garnered in the beginning teacher evaluation study (Berliner, Fisher, Filby, & Marliave, 1978; Fisher et al., 1978). These researchers reported that 18% of the school day was spent on noninstructional topics and that, on average, 58% of allocated time was spent on academic subjects. Differences in how teachers used time were enormous. For example, in some second grade classrooms, students received 15–20 minutes of math instruction. However, in other second grade classrooms, students received 45–50 minutes of daily math instruction. These differences in allocated time were compounded by the extent of student engagement in instructional activities. In some classes, students were engaged about 50% of the time, whereas, in other classes, they were engaged about 90% of the time.

This research also illustrated that patterns of relationships between students' seatwork success and ultimate achievement varied by grade level and subject matter. Importantly, Burstein (1980) found that better student achievement was more often associated with moderate rather than high rates of seatwork success. Note that this finding is a powerful illustration of variables working in combination: more involvement in seatwork was more highly correlated with achievement when moderately difficult work was assigned. This suggests that when seatwork was too easy or too hard, involvement alone was insufficient to increase achievement.

International comparisons of how teachers use time also illustrate considerable variation in how teachers use time in interacting with students. Stigler, Lee, and Stevenson (1987) reported that in the United States in first and fifth grade classrooms, teachers spent considerably less time working with the whole class (46%) in comparison to teachers in Japan (86%) and Taiwan (77%). Also, American teachers spent less time directly providing mathematical content to students (25%) than in Japan (33%) and especially in Taiwan (63%).

Thus, time can be used as a variable to describe various instructional conditions. These distinctions include whether the teacher is spending time with the class as a whole, with groups of students, or with individual students. Further, time can be used to differentiate the amount of time that is used for instruction versus other purposes and the cognitive level of demand placed upon students.

FOCUS ON MEANING AND PRACTICE

For a long time, there has been considerable discussion about the importance of helping students understand the information, concepts, and symbols that they study. Explicitly in the area of mathematic instruction, there has been marked emphasis upon understanding mathematics (e.g., Brownell, 1947) that has continued to today as represented in various papers presented by the National Council of Teachers of Mathematics (1989) and elsewhere, including The National Mathematics National Advisory Panel (2008). However, critics have abounded both historically and currently to argue that teaching for understanding can be problematic and time consuming.

Both math and reading wars have involved debate about the extent to which instruction should involve skills and practice versus conceptual understanding (Schoenfeld & Pearson, 2009). In the area of reading, this argument has often assumed the form of decoding versus comprehension, and in mathematics, the battleground has involved conceptual understanding versus practice. Although enormous energy has been expended in either/or debates embedded in the math and reading wars, to me these arguments are silly. Who would want students who understand math problems, but cannot use the ideas in practical ways, or students who could do math, but do not understand what they did? Meaning must be embedded in a systematic approach to instruction, which involves student integration and mastery of content at both the conceptual and the application level.

Recent research continues to illustrate a positive relationship between a focus on conceptual understanding and good student achievement. For example, Rakes, Valentine, McGatha, and Ronau (2010) reviewed 82 studies of instructional improvement strategies used in algebra classes. In general, a conceptual focus was a more powerful effect on student achievement than a procedural focus. Indeed, Rakes et al. noted that the development of conceptual understanding was associated with an average effect size about double the effect size of those classrooms that focused on procedural understanding. However, I emphasize again that the point is the relative importance of conceptual knowledge (in relation to practice), not the absence of procedural information and practice. Hiebert and Grouws (2007) argued persuasively that effective math teaching requires teaching both for understanding and for skill proficiency.

TEACHER EXPECTATIONS

Teachers' beliefs about students often influence their interactions with them. There is evidence to illustrate that teachers who believe that students are capable often interact with them differently than they do with students believed to be less capable. For example, more capable students are often provided with more choice, longer time to respond to teacher questions, and more clues and other forms of help when these students initially have a difficult time in responding. Students who are believed to be less capable often receive less choice in academic assignments, fewer opportunities to respond publicly to teachers' questions, and more criticism when they respond incorrectly. In essence, these differences allow students believed to be more capable, more time to process information and to think. Some of these differences are presented in Table 4.

Table 4. How do teachers differentiate in their behavior toward students they believe to be more capable?

High Expectations
- more opportunity to respond to teacher questions
- more opportunity for follow-up questions—tell me more
- more time to respond to teacher questions
- more stay-with behavior—teacher provides clues or rephrases the questions
- more teacher praise
- more opportunity for choice

Low Expectations
- less opportunity to respond to teacher questions
- less opportunity for follow-up questions
- less time to respond to teacher questions
- more give-up behavior—teacher provides the answer or calls on another student to do so
- more teacher criticism
- less opportunity for choice

I have studied teacher expectations since my PhD dissertation at Indiana University (Good, 1968, 1970). I know that teacher beliefs about the performance potential of their students can be enormous as teachers decisions about content assignment and the responses to student work are influenced by their conceptions of student abilities. Although the research on teacher expectations is large and rich, most of this research has been descriptive and correlational; however, there is research to show a clear relationship between teacher expectation and student achievement (e.g., Brattesani, Weinstein, & Marshall, 1984; Pidgeon, 1970; Raudenbush, 1984). Much of this research has focused upon how teachers teach subsets of students differently in the classroom (e.g., boys versus girls and students believed to be high achieving versus those believed to be less capable; see Good & Brophy, 2008). However, other work (Rubie-Davies, 2007, 2010) has shown that teacher expectation of the class as a whole can impact achievement. Thus, it is useful to understand that teachers can communicate low expectations to individual students, to the whole class, or both. Unfortunately, we know much more about how to express low expectations than we do about how to communicate high expectations.

GOOD CLASSROOM MANAGEMENT

The work on proactive management was pioneered by Jacob Kounin (1970). His work involved summer camps, experimental research, and naturalistic research in classrooms. His most important work, published in 1970, explored classrooms that differed in

terms of student involvement and degree of misbehavior. In his preliminary study, he identified two sets of classrooms. In one set of classrooms students were involved most of the time and rarely ever engaged in serious misbehavior. In another set of classrooms, students were often not engaged in assigned work and sometimes engaged in serious misbehavior. Then, Kounin systematically studied these two types of classrooms using observational procedures to explain how teachers with highly involved students structured classrooms in comparison to teachers with students who were often uninvolved. One major finding from this research was that what teachers did after misbehavior occurred was typically ineffective. The difference between good managers and poor managers was the ability of good managers to prevent problems form occurring. Table 5 presents six of the ways that good classroom managers prevented problems.

Table 5. Kounin's Research: Six Proactive Ways Teachers Prevent Misbehavior

Method	Description
Alerting	Teacher indicated in advance what was expected of students. "Raise your hand if you want me to call on you." "Take 10 minutes and work on these by yourself and then we will check our work."
Accountability	Teachers followed up on their alerting. They called on only handraisers. They reviewed student seatwork after about 10 minutes. By following through on what they alerted students to do or expect, teachers maintained credibility by allowing students to see that teacher messages were consistently fulfilled.
With-it-ness	Teachers positioned themselves so they could see and monitor student actions. They let students know they were "with it" by letting students know, in unobtrusive ways, that they were aware of what students had done or not done.
Overlapping	Teachers demonstrated the ability to do two things at the same time, for example, keeping the small group the teacher is with engaged on task while dealing with a student who is not a member of the group and collecting student papers while giving clear instructions about what students are to do next.
Signal Continuity and Lesson Momentum	Teachers ignore fleeting misbehavior. They don't stop the lesson to find a video clip. They deal with sustained inattention before it leads to misbehavior.
Variety and Challenge in Seatwork	Teachers ensure that when students do seatwork, they have productive work and not just filler tasks. If some students finish work early, teachers have meaningful tasks for these students while others finish. Teachers who were effective managers assigned doable work to students, but also had enough challenge to maintain student motivation.

These data revealed exceedingly high correlations between classroom behaviors and student involvement. Indeed, these correlations were as high as they get in social science research. During whole class sessions, Kounin (1970) found that students' engagement was correlated positively and highly with teacher scores on momentum, .66; with-it-ness, .61; smoothness, .60; alerting, .60; accountability, .49; and overlapping, .46. Still, these data were only correlational. Fortunately, several researchers have demonstrated the importance of these management behaviors, and research programs (similar to what was done in the Missouri Math Program that Doug Grouws and I developed) illustrated that it was possible to teach these managerial principles to teachers. Further, in experimental studies it was found that experimental teachers in comparison to control teachers (who did not receive the training) were able to increase student involvement in important ways and reduce the occurrence of student misbehavior. Many classroom management studies illustrated the power and stability of Kounin's findings over time (Emmer, Evertson, & Anderson, 1980; Emmer & Gerwels, 2006; Evertson & Weinstein 2006). One especially important experiment was conducted by Freiberg, Huzinec, and Templeton (2009) showing that proactive management principles could be taught to teachers in ways that improved student involvement and achievement.

INTEGRATING ACROSS VARIABLES

Earlier, I stated that there were roughly a dozen variables that have consistently linked instruction to positive increases in student achievement across various contexts. Here, I have examined five generic instructional variables: opportunity to learn, time allocation, appropriate teacher expectations, balancing conceptual understanding with meaningful practice, and good classroom management. In discussing each case, I noted that we have known about the importance of the relationship between these variables and student achievement for a long time and that we continue to find evidence confirming the importance of these variables. These variables interact and collectively they provide a system of instruction. Understand that none of these variables alone is a proxy for effective teaching. It is the appropriate and sensible combination of these variables that provides a coherent instructional plan for increasing student achievement.

CURRENT EFFORTS IN EVALUATING TEACHERS

In this section, I describe three large studies that have recently replicated a knowledge base that had been established much earlier (Brophy & Good, 1986). I also comment briefly upon the current renewed use of student evaluations and observational measures to evaluate teachers. Then I briefly address the issue, "Why do we know so little about how teachers impact other student outcomes such as creativity, problem-solving, pro-social dispositions?"

NEW KNOWLEDGE

In the early 2000s, policymakers became interested in renewed efforts to identify effective teachers, to reward effective teachers, and to weed less effective teachers out of the field.

Increasingly, the low student performance in schools, especially in low-performing schools (that were primarily enrolling students who lived in poverty), grew into a renewed and vigorous call for holding individual teachers accountable for their poor performance. Race to the Top funding placed renewed emphasis upon evaluating individual teachers. This urgency for evaluating individual teachers came at a time when researchers were reporting new analytical and statistical procedures for comparing teachers' impact on student achievement. This new approach is commonly referred to as value-added assessment (for a discussion of what this method entails, see Collins & Amrein-Beardsley, 2013, and Konstantopoulos, 2013, both in this issue).

Value-added procedures are similar to the use of residual gain scores (pre- and post-achievement comparisons) that were used to assess teacher impact on student achievement in the early teacher effects literature. Although value-added procedures are somewhat more powerful (when used appropriately) than residual gain scores, this slight gain in methodology does not excuse "modern researchers" from being aware of and reporting on previous research that had established a link between instruction and achievement using residual gain scores. Yet, as we will see in the studies I review, the older established research plays no role. In Tables 6 through 9, the findings from this new knowledge are presented.

Strong (2011) attempted to explain what constituted good teaching by having different types of reviewers (differentiated by their backgrounds and roles) view videos of teachers who had a record of achieving large positive effects on student outcomes over three consecutive years and teachers who had noticeably less impact on student achievement. One of these studies indicated that classroom principals had but limited ability to identify more and less effective teachers. In the final analysis, after conducting three different studies using different samples of teachers, Strong concluded that it was possible to use observational data to distinguish high- and low-performing teachers. His findings appear in Table 6.

Table 6. Strong's Characteristics of Effective Teachers

- Lesson objective is clearly expressed to students.
- Teacher understands student background and prior knowledge.
- Teacher and/or students demonstrate consideration of the topic in more than one way.
- Teacher provides/elicits students' multiple examples.
- Teacher uses appropriate nonexamples.
- The pace of the lesson is not too slow or too fast.
- Feedback to students is about the process—not just correctness.
- Teacher uses guided practice—students practice and immediate feedback is provided.
- Teacher explains the math concepts with good clarity.
- Teachers are actively engaged—most time spent on task.

Farr (2010) studied teachers in the Teach for America program. He identified teachers who had been found to have large positive effects on students over consecutive years and studied how they taught differently than Teach for America teachers who had considerably less impact on student achievement. These data appear in Table 7. Ripley (2010) wrote a nontechnical paper summarizing the more complex technical report presented by Farr, and these results are presented in Table 8. Unsurprisingly, the dimensions in Tables 7 and 8 are highly similar.

I chose to include the conclusions from Ripley's article here, because her article includes a very positive discussion of a teacher who involves his students in mental math activities. I agree that mental calculations can provide informative and enjoyable learning experiences for students. However, I find it interesting that the use of mental math activities was part of a videotape produced and distributed by American Association of Curriculum Development to describe the Missouri Mathematics Program that Doug Grouws and I developed. Curiously, there is no mention of our earlier work in Ripley's article.

Table 7. Good Teaching Based on Amanda Ripley's (2010) analysis of Teach for America Data (2010)

- communicating key ideas
- coordinating student practice
- checking for student understanding
- tracking progress
- using organization and routine to maximize efficiency
- asserting authority by consistently following through on high expectations

Table 8. Good Teaching Based on Farr's (2010) analysis of Teach for America Data

- setting big goals for their students
- constantly reevaluating what they are doing
- using ways to involve kids in the lesson (e.g., mental math)
- demonstrating the learning that students were to engage in
- implementing an "I do, we do, you do"
- building up well-established routines

Hattie (2009) conducted a large, comprehensive meta-analysis review of studies that related instructional behavior to student achievement. He noted that his review convinced him that teachers powerfully impacted student learning (and I agree). His results are presented in Table 9.

Table 9. Visible Learning

- Teachers need to be directive and actively engaged in teaching and learning.
- Teachers need to know how students are constructing meaning.
- Teachers need knowledge of the content they teach in order to provide meaningful feedback to students.

- Teachers need to be clear about intended meaning of lessons and making these success criteria explicit.
- Teachers need to recognize and welcome student error and creating environments where students feel safe to learn and relearn.

Notes. Based on material presented by Hattie in Visible Learning, pp. 238-239

These "new" studies are but briefly presented here; however, there is sufficient information to highlight the striking similarities between these studies and the findings reported by Brophy and Good (1986). Consider the striking similarities between the data presented in Tables 6 through 9 and the data obtained by Good and Grouws (1979) that are presented in Table 10.

Table 10. Data Collected by Good and Grouws in 1979

- Teacher provides a review of material previously presented, collects homework, and asks students to engage in mental computation work.
- Teacher conducts development lesson where the teacher focuses on meaning and promoting student understanding through discussion and demonstration.
- Development also includes the assessment of student understanding by asking both process and product questions, and controlled practice where teachers elaborate upon the meaning as necessary.
- Teacher conducts homework using appropriate pace, alerting, and accountability.
- Students are given 15–20 minutes of uninterrupted time for successful practice.
- Seat work involves momentum—keeping the ball rolling—getting everyone involved.
- Seat work also involves appropriate alerting (e.g., their work will be checked) and accountability (e.g., the work is actually checked).
- Homework assignment if students have done seat work well
- Homework should include review problems

Hence, if one wants to know whether data generated roughly 40 years ago still apply in modern classrooms, the answer is a clear yes. Although this replication of established knowledge is very useful, it is disappointing that these new efforts have not provided any new information about how instructional behaviors influence student achievement or, and importantly, new types of student outcomes.

This new knowledge replicated earlier research and illustrates that many years later, previous findings are highly similar to these new data. Oddly, and unfortunately, this knowledge of teacher effects reviewed here and discussed in detail by Brophy and Good (1986), was largely ignored and/or forgotten by researchers in the 1990s and early 2000s. Since earlier research on teaching was not cited in the new research (described below) what stimulated the development of this new knowledge?

Although I confess to some personal slight (why was my and others' previous research ignored in these "modern studies"?), the major issue is that this omission was extremely costly because (had they been aware of earlier research) they could have spent their time, energy, and money expanding on this research, rather than inadvertently replicating it. Ironically, this "new" research, that teachers make more or less difference in student achievement that emerged largely post-2005, often was accompanied by statements that essentially said, "We now know that teachers affect student achievement, but we do not know *how* teachers influence student achievement." Hence, many modern researchers failed to inform school officials and policymakers that earlier research had provided hypotheses and findings that might be applicable to current classrooms. In my opinion, a more mature response would have been, "Here is what we knew some time ago linking instruction to achievement. Let's see if we can replicate and extend these previous findings to see if they are still relevant in today's instructional context."

It would be remiss for me to suggest that all researchers ignored previous process–product research. For example, Rowan, Correnti, and Miller (2002), Konstantopoulos (2006, 2011) and Konstantopoulos and Chung (2011) did acknowledge these findings—that we know teachers matter. What was missing from research after the 1990s was any attempt to explain "how much" teachers matter and other outcome variables that they might also influence. As data make it clear, teachers only account for roughly 10–20% of the variation in student achievement (see Konstantopoulos, 2013, and Lavigne, 2013 both in this issue). If the field had placed a greater emphasis on why and how teachers matter, perhaps our current teacher evaluation systems would be more focused on improving teacher practice. If research building on previous research (such as Brophy & Good, 1986; Good & Grouws, 1979) had focused on improving instruction and extending the research base, perhaps we would not be going down the road that we are presently following—holding individual teachers accountable for students' achievement when teachers account for such a small percentage of the variation that explains school learning.

APPLICATION VALUE OF THESE FINDINGS

The information presented above has implications for evaluating teachers if it is used for informative/improvement purposes. That is, this information is helpful if it can assist teachers to think about possible adjustments in instructional climates or classroom interactions that might improve student learning. This possible advantage of using these research findings is based upon the assumption that observers are well trained, that the number of observations collected is sufficient to provide representative information, and that observers can interpret the data they collect, synthesize it, and provide feedback in ways that are both informative and encourage teachers to consider using the information.

Given the present emphasis on using student achievement data and teacher observations to evaluate teaching, one might conclude that these data hold considerable importance for evaluating teachers to improve instruction. However, as noted, long ago the value of student achievement data for evaluating teachers was reduced because of the variability of many teachers' effects on students across years. Observational measures that "in theory" could be useful are limited because in design and in practice, they are believed to be measures of how teachers obtain student achievement. Hence, achievement scores and observed measures of teaching currently used are redundant—not complimentary measures of good teaching. In my opinion the earlier studies of

teaching and earlier measures of classroom observation have much more potential for effective use because they are more than checklists of isolated behaviors. As Doug Grouws and I (Good & Grouws,1975, 1979) demonstrated, considerations of "good" teaching are best placed in a system of instruction that emphasizes the integration of variables—not isolated checklists of variables.

OBSERVATIONAL MEASURES OF TEACHER EFFECTIVENESS

The current focus on effective teaching has mainly been on student achievement. However, other sources of information have been used to assess teacher effectiveness including classroom observations. However, to reiterate, for the most part, current observation systems are believed to be highly correlated with student achievement and are not used to explore other important possibilities such as fairness, interest level of lessons, and so forth.

Having noted earlier that teaching involves an integrated set of instructional behaviors, I am not optimistic that present observational systems can be used to describe and to improve teaching. Herlihy et al. (2013) provided a very comprehensive and insightful analysis of how states are proposing to use or actually using observational systems to evaluate teachers. And, my reading of their work suggests that current use of observational systems to evaluate teachers leaves much to be desired. Given the constraints of space, I can but briefly allude to the complexity of these issues. For example, many observational systems call for the demonstration of high expectations. However, as I have noted, expectations can be too high or too low, and the issue is for teachers to demonstrate appropriate expectations. How then does a classroom observer know and code if expectations are appropriate both for individual students and for the class as a whole? Although there is literature that makes it clear that teachers can express expectations that are too low for some students or for the class, most observational systems do not clarify whether the teacher expectations being measured are for individual students or for the class as a whole.

To take another example, it has been shown that good classroom management involves the demonstration of several instructional behaviors including alerting and accountability. Teachers who alert clearly (e.g., "work for 15 minutes and then we'll check your seatwork"), but who fail to hold students accountable (check the seatwork in 15 minutes), will quickly convince students not to pay attention to what they say as they do not provide credible information. Hence, the value of alerting and accountability is in their combined use. Frequent accountability without alerting is often bad classroom management and frequent alerting without accountability is almost always bad classroom management. Yet, observational systems that are presently used to evaluate teachers frequently will include alerting or accountability codes, but seldom both, and sometimes neither. Thus, it seems highly problematic that systems that do not measure the combined effect of accountability and alerting could provide any meaningful information to improve classroom management. To add further complexity, one way to communicate low expectations for individual students is to frequently alert them ("Johnny, pay attention to how I write this equation"). Thus, in some—perhaps most—observational systems used in Race to the Top accountability, teacher alerting would be coded positively by observers when, in fact, the alerting was communicating low expectations to one or more students.

STUDENT ASSESSMENTS OF THEIR TEACHERS

In addition to using standardized achievement tests to assess teacher behavior, recently, there has been considerable interest in using student evaluation to assess effective teaching. Although there are both obvious and subtle problems inherent in the use of student data to evaluate teachers, there are good arguments for including student perceptions of their teachers and classroom learning experiences (Cooper & Good, 1983; Hook & Rosenshine, 1979; Weinstein & Middlestadt, 1975). After all, students, especially in the elementary grades, spend numerous hours observing and responding to teachers.

One large study that combines the use of observation, student achievement, and student reports to study teacher effectiveness has been funded by the Bill & Melinda Gates Foundation. This research is described in detail in *Learning About Teaching: Initial Findings From the Measures of Effective Teaching project* (Bill & Melinda Gates Foundation, 2010).

Students' surveys were given to elementary and secondary school students across elementary and secondary classrooms. Students' surveys were correlated with teacher effectiveness as measured by teachers' value-added impact on students' achievement.

Five student responses were most associated with teacher's positive impact on student achievement:

- Students in this class treat the teacher with respect.
- My classmates behave the way my teacher wants them to behave.
- Our class stays busy and doesn't waste time.
- In this class, we learn a lot every day.
- In this class, we learn to correct our mistakes.

It is beyond my goal to assess fully the value of using student perceptions of teachers as part of a formal evaluation system or the limitations of the Bill & Melinda Gates study (2010); however, in the context of exploring what we do know about how teachers increase student achievement, it is useful to note that evidence linking students' positive attitudes and achievement is not new (Brophy & Evertson 1978, 1979; Good & Grouws, 1975), although a reading of the Bill & Melinda Gates Foundation report would suggest that the use of student data is new. As a case in point, consider some of the items that appeared on the School Sentiment Index as noted in Good, Biddle, and Brophy (1975):

- My teacher doesn't explain things very well.
- When I do something wrong at school, I know I will get a second chance.
- My teacher usually explains things too slowly.

- I don't do very much reading on my own.

Second, these student perceptions of good teaching are consistent with the research on teaching that I reviewed in this paper. This congruence of observational data with student perception data, I believe, is both informative and important.

Teachers who respect students are well prepared to teach and do so as soon as the class begins. Indeed, students have been known to indicate that they view teachers' evident preparation as a sign of teacher respect. Student willingness to follow teachers' instruction is typically achieved through proactive management and student perceptions that the class is productive and does not waste time. This is consistent with considerable correlational and experimental evidence from both the classroom management literature and time utilization literature. The belief that we learn a lot each day nicely fits with the literature on pacing and opportunity to learn (see Good & Grouws, 1975; Schmidt & Maier, 2009). Finally the perspective that it is "okay to make mistakes, but they should be corrected" theoretically fits with notions of constructive failure (see Rohrkemper & Corno, 1988) and the concomitant belief that failure is inevitable, but also provides a positive opportunity to see that we can learn from our mistakes.

WHY DO WE KNOW SO LITTLE ABOUT HOW TEACHERS INFLUENCE OTHER STUDENT OUTCOMES?

In this paper I do not give detailed attention to "why we know so little about how teachers affect outcomes other than achievement." However, I did include this serious deficiency as part of the title in my paper. Although I cannot give detailed attention to this complex issue, I want to emphasize that our knowledge base (although useful), is extremely limited because the research has not yielded any truly informative information about how we can achieve any outcomes that we want students to learn in school other than achievement. In brief, we know so little about teaching because the investment in educational research has been limited compared to spending in areas like health and science. When educational research occurs, it is after an innovation or reform is implemented and then the research typically shows that the new program was poorly defined, poorly implemented, and had little, if any, effect. When reforms are implemented, they tend to push normative (typical) teaching practice so far in one direction that new reforms need to push back and rebalance the older curriculum and instructional practices (Good, 1996; Good, 2010; Good & Braden, 2000; Schoenfeld & Pearson, 2009). Reforms move from too much structure to too little structure for learning and from too much teacher-directed instruction to too little and so forth.

Clearly, citizens are concerned about many outcomes other than achievement as various authors have noted (Berliner & Biddle, 1995; Good & Braden, 2000; Rothstein, 2000). Unfortunately, these concerns have not impacted policymakers' decisions for various reasons. The most central is that there are so many and varied possible outcomes (e.g., safe schools, physical fitness and health, creativity, initiative, appropriate attitudes toward work, and a plethora of others) that no consensus has emerged among educators and parents as to which of these are the most important. Given that some, if not most, desirable outcomes of schooling (dispositions for living constructively in a democracy) are difficult and expensive to measure, it is hard to generate widespread interest in systematically assessing the impact of teachers in schools on outcomes other than achievement. If a consensus was developed around one or two aspects of schooling that merit careful assessment, it would be possible to build new databases for pursuing other goals/outcomes of schooling. Thus, given that policymakers and many citizens are concerned about student achievement, in the absence of any other concentrated voices, it is likely that only student achievement will continue to be the central focus for evaluating teachers in schools.

Why do we invest so little in research and why do we tend to invest reactively rather than proactively? This is a complex issue and will be addressed more fully elsewhere (Lavigne & Good, in press). However, the "CliffsNotes™" response is that policymakers see school productivity as too low and often as an acute problem. Ironically, however, they see teaching as easy. There are many variants on the theme that teaching is easy, but cutting across this set of beliefs is the contention that if only teachers would do X or Y, we would see sharp improvements in learning. And the cycle goes on. The new "today's" solution creates the new problem, which leads to the need for reform, and so forth (Good, 1996, 2010). Researchers have also contributed to this problem in important ways. In general, research in education is not programmatic. Rather, research, like reform movements, tends to focus on the "new problem" and researchers propose new and often simplistic changes for improving teaching and learning.

CONCLUSION

In this paper, I have argued that there is an established knowledge base that relates instructional practice to student achievement as measured by standardized tests. The field has possessed much of this knowledge since at least 1986 and that knowledge base has been replicated by many different researchers. As useful as this replicated knowledge is for establishing relationships between instruction and student learning, it is disappointing that this literature has not been extended in new ways.

I have emphasized that the application of this knowledge base is better seen as patterns and combinations of instructional behaviors rather than isolated variables. Unfortunately, at present, the use of observational systems associated with Race to the Top accountability appears in most instances to be an accounting of teaching performance only in terms of the presence or absence of individual variables. This usage is not an accurate reflection of the knowledge base, in my opinion, and in many cases will result in inappropriate conclusions about teacher performance and student learning. For more information about the current state of observational systems associated with Race to the Top, see Herlihy, et al. (2013).

In addition to the issues of accurate measurement of observational systems that Herlihy and Hill address, there are also unintended effects of using observational systems (and student achievement scores) in dismissing teachers. Alyson Lavigne (2013) addressed this issue in an informative way. Her data and policy questions raise substantial doubts about the use of high-stakes testing for evaluating and tenuring teachers. It seems unlikely that many policymakers have thought about these unintended consequences, let alone given them careful thought.

At a policy level, I have reminded readers of the caveat provided years ago by Good, Biddle, and Brophy (1975). "Just as data suggesting that teachers do not make a difference frequently has been overgeneralized and accepted uncritically, there is a danger that the same mistakes can be made with the kinds of data cited in this chapter" (p. 85). I continue to caution that information about teacher effectiveness can be overvalued for various reasons. Among these constraints for using the knowledge of teacher effects on student achievement in high-stakes situations, is that this knowledge base does not include information about how teachers influence other important student outcomes such as creativity and problem finding among many other desired outcomes of schooling. Further, although individual teachers are important and do influence student achievement, they are but one factor that is important. We know, for example, that school resources and the socioeconomic status of students have a powerful impact upon student learning (Baldi et al., 2007). David Berliner (2013) has provided clear and convincing evidence that many other variables enhance or impede the effects of individual teachers on student achievement.

This paper has primarily focused on the influence of the media and policymakers in framing problems and solutions in terms that only emphasize student achievement and defining problems in ways that lead to simplistic, but expensive, reform efforts. School reform in America is characterized by constant reform and innovation such that it seems fair to describe American education as fad-driven with little respect for research and programmatic inquiry. However, I've hinted at and provided some suggestions that researchers and teacher educators also have a role in indirectly supporting, if not directly enabling, reform not based upon research.

I have mentioned several times that our database on effective teaching is limited; however, it has useful implications for practice. There is no clear knowledge that teachers learn and have the opportunity to practice using this knowledge. It would seem that teacher education programs would want to assure that their graduates, in addition to possessing appropriate content knowledge, would also have conceptual understanding and skills related to active teaching, proactive management, communication of appropriate expectations for learning, and the ability to plan and enact instruction that balances procedural and conceptual knowledge. Future research upon the use of this knowledge base in teacher education programs would be informative.

Although I have been critical of current activities that use classroom observation and student achievement information for improving teaching, I believe that these assessments, with refinement, could provide valuable information to teachers that they could use in making decisions about planning and implementing instruction. However, this is unlikely to be the case for reasons that I have already noted. After a great deal of concern and effort, considerable money has been spent on new high-stakes evaluation procedures. The winds of change are already blowing in a new direction. Rather than refining these evaluative attempts, combining the knowledge of teachers and researchers, it is likely that education will move to another set of simplistic strategies for improving education.

As I write, I see that at least five states—Colorado, Connecticut, Massachusetts, New York, and Tennessee—are experimenting with longer school days (USA Today, 2012) and although I have no opposition to a longer school day, the rationale for this policy action is not clear. Have we forgotten A Nation at Risk and its call for longer school days and longer school years? Have we overlooked the fact that many of the nations where students outperform American students have school years that are comparable to our own and sometimes shorter? I would feel better if there was a coherent statement describing how extra time would be used so that we would have a clearer independent variable (describing how more time will be used) that presumably would lead to more achievement. And, there is currently much ado about increasing teachers' knowledge of content (and of course considerable interest in assessing that knowledge with new tests) without any rationale for taking this policy action. Presently, the American Federation of Teachers is a voice among others for increasing teachers' content knowledge (2012). Clearly, there is much evidence to illustrate that teachers' knowledge of content is at best weakly related to student achievement. Obviously, there is nothing wrong with trying to enhance teachers' content knowledge, but it seems unfortunate to develop this new knowledge without also continuing to conduct research on classroom instructional behaviors as we have long known that what teachers *do* has important effects on student achievement, and likely other outcomes as well if we took the time to study them.

Still, I believe that when done correctly research on teaching can improve instruction. However, the research must be applied carefully if it is to have useful effects. And, as noted, often in this paper, research must consider other outcomes of schooling such as creativity, adaptability, and problem finding.

References

American Federation of Teachers. (2012). *Raising the Bar: Aligning and elevating teacher preparation and the teaching profession*. Washington, DC: Author.

Baldi, S., Jin, Y., Skemer, M., Green, P. J., & Herget, D. (2007). *Highlights from PISA 2006 Performance of U.S. 15-year-old students in science and mathematics literacy in an international context* (NCES 2008-016). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Beckerman, T., & Good, T. (1981). The classroom ratio of high- and low-aptitude students and its effects on achievement. *American Education Research Journal*, 18, 317–327.

Berliner, D. (2013). Exogenous variables and value-added assessments: A fatal flaw. *Teachers College Record*, *116*(1), XXX-XXX.

Berliner, D., & Biddle, B. (1995). *The manufactured crisis: Myths, fraud, and the attack on America's public schools*. New York, NY: Addison-Wesley.

Berliner, D., Fisher, C., Filby, N., & Marliave, R. (1978). *Executive summary of Beginning Teacher Evaluation Study*. San Francisco: Far West Laboratory.

Bill & Melinda Gates Foundation. (2010). *Learning about teaching: Initial findings from the Measures of Effective Teaching project*. Seattle, WA: Author.

Brattesani, K. A., Weinstein, R. S., & Marshall, H. H. (1984). Student perceptions of differential teacher treatment as moderators of teacher expectation effects. *Journal of Educational Psychology, 76*, 236–247.

Brophy, J. (1973). Stability of teacher effectiveness. *American Educational Research Journal, 10*(3), 245–252.

Brophy, J. (2012). Observational research on generic aspects of classroom teaching. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed.; pp. 755–780). New York, NY: Routledge.

Brophy, J., & Evertson, C. (1978). Context variables in teaching. *Educational Psychologist, 12,* 310–316.

Brophy J., & Evertson, C. (1981). *Student characteristics and teaching*. New York, NY: Longman.

Brophy, J., & Good, T. (1986). Teacher behavior and student achievement. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.; pp. 328–375). New York, NY: Macmillian.

Brownell, W. A. (1947). The place of meaning in the teaching of elementary school arithmetic. *Elementary School Journal, 47*, 256–265.

Burstein, L. (1980). The analysis of multilevel data in educational research and evaluation. *Review of Research in Education, 8,* 158–233.

Carpenter, T., Dossey, J., & Koehler J. (Eds.). (2004). *Classics in mathematics educational research*. Reston, VA: National Council of Teachers of Mathematics.

Collins, G. (2012, September 1). Only the good get rich [op-ed]. *The New York Times*, pp. A23.

Collins, G., & Amrein-Beardsley, A. (2013). Putting growth and value-added models on the map: A national overview. *Teachers College Record, 116*(1), XXX-XXX.

Cooper, H. M., & Good, T. L. (1983). *Pygmalion grows up: Studies in the expectation communication process*. New York, NY: Longman.

Cuban, L. (2010). *As good as it gets: What school reform brought to Austin*. Cambridge, MA: Harvard University Press.

Darling-Hammond, L. (2010). *The flat world and education. How America's commitment to equity will determine our future*. New York. NY: Teachers College Press.

Dubriel, J. B. (1977). *A study of two plans for utilization of class time in first year algebra* (Doctoral dissertation). University of Missouri, Columbia.

Ebmeier, H. (1978). *An investigation of the interactive effects among student types, teacher types, and treatment types on the mathematics achievement of fourth grade students.* (Doctoral dissertation), University of Missouri, Columbia.

Ebmeier, H., & Good, T. L. (1979). The effects of instructing teachers about good teaching on the mathematics achievement of fourth-grade students. *American Educational Research Journal*, *16*(1), 1–16.

Educate America Act. (1994). From Public Law 103-227, Goals 2000.

Emmer, E., Evertson, C., & Anderson, L. (1980). Effective classroom management at the beginning of the school year. *The Elementary School Journal*, *80*, 219–231.

Emmer, E., & Gerwels, M. C. (2006). Classroom management in middle and high school classrooms. In C. Evertson & C. Weinstein (Eds.), *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 407–437). Mahwah, NJ: Erlbaum.

Evertson, C., & Weinstein, C. (Eds.). (2006). *Handbook of classroom management: Research, practice, and contemporary issues*. Mahwah, NJ: Erlbaum.

Farr, S. (2010). *Teaching as leadership: The highly-effective teacher's guide to closing the achievement gap*. San Francisco, CA: Jossey-Bass.

Fisher, C., Filby, N., Marliave, R., Cahen, L., Dishaw, M., Moore, J., & Berliner, D. (1978). *Beginning teacher evaluation study*. San

Francisco, CA: Far West Regional Educational Laboratory.

Freiberg, H. J., Huzinec, C. A., & Templeton, S. M. (2009). Classroom management—a pathway to student achievement: A study of fourteen inner-city elementary schools. *Elementary School Journal*, *110*(1), 63–80.

Good, T. (1968). *Student achievement level and differential opportunity for classroom response*. (Doctoral dissertation). Indiana University, Bloomington, IN.

Good, T. (1970). Which pupils do teachers call on? *Elementary School Jour*nal, 70, 190–198.

Good, T. (1996). Teaching effects and teacher evaluation. In J. Sikula, T. Buttery, & E. Guyton (Eds.), *Handbook of research on teacher education* (2nd ed.; pp. 617–665). New York, NY: MacMillan.

Good, T. L. (2010). Forty years of research on teaching 1968–2008: What do we know now that we didn't know then? In R. Marzano (Ed.), *On excellence in teaching* (pp. 31–62). Bloomington, IN: Solution Tree Press.

Good, T., Biddle, B., & Brophy, J. (1975). *Teachers make a difference.* New York, NY: Holt, Rinehart, & Winston.

Good, T., & Grouws, D. (1975). Teacher rapport: Some stability data. *Journal of Educational Psychology, 67*(2), 179–182.

Good, T., & Grouws, D. (1979). The Missouri mathematics effectiveness project: an experimental study in fourth-grade classrooms. *Journal of Educational Psychology*, *71*(3), 355–362.

Good, T. L., & Braden, J. S. (2000). *The great school debate: Choice, vouchers, and charters*. Mahwah, NJ: Lawrence Erlbaum Associates.

Good, T. L., & Brophy, J. (2008). *Looking in classrooms* (10th ed.). New York, NY: Pearson.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.

Herlihy, C., Karger, E., Pollard, C., Hill, H., Kraft, M., Williams, M., & Howard, S. (2013). State and local efforts to investigate the validity and reliability of scores from teacher evaluation systems. *Teachers College Record*.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Greenwich, CT: Information Age.

Hook, C., & Rosenshine, B. (1979). Accuracy of teacher reports of their classroom behavior. *Review of Educational Researc*h, 49, 1–12.

Konstantopoulos, S. (2006). Trends of school effects on student achievement: Evidence from NLS:72, HSB:82, and NELS:92. *Teachers College Record*, *108*, 2550–2581.

Konstantopoulos, S. (2011). Teacher effects in early grades? Evidence from a randomized experiment. *Teachers College Record*, *113*, 1541–1565.

Konstantopoulos, S. (2013). Teacher effects, Value-Added Models, and Accountability. *Teachers College Record, 116*(1), XXX-XXX.

Konstantopoulos, S., & Chung, V. (2011). The persistence of teacher effects in elementary grades. *American Educational Research Journal*, *48*(2), 361–386.

Kounin, J. S. (1970). *Discipline and group management in classrooms*. New York: Holt, Rinehart and Winston.

Lavigne, A. (2013). Exploring the intended and unintended consequences of high-stakes teacher evaluation on schools, teachers, and students. *Teachers College Record, 116*(1), XXX-XXX.

Lavigne, A., & Good, T. (in press). *Evaluating teachers and students*. New York, NY: Routledge.

Lefstein, A. (2008). Changing classroom practice through the English National Literacy Strategy: A micro-interactional perspective. *American Educational Research Journal, 45*(3), 701–737.

Lundgren, U. (1972). *Frame factors and the teaching process: A contribution to curriculum theory and theory on teaching*. Stockholm, Sweden: Almqvist and Wiksell.

McCaslin, M. (2006). Student motivational dynamics in the era of school reform. *Elementary School Journal, 106,* 479–490.

McCaslin, M., & Good, T. (1996). Informal curriculum. In D. Berliner & C. Calfee (Eds.), *Handbook of educational psychology*. Washington, DC: American Psychological Association.

Moe, T. M., & Chubb, J. E. (2009). Liberating learning: Technology, politics, and the future of American education. San Francisco, CA: Jossey-Bass.

Morgan, S., & Poppe, E. (2012). The consequences of international comparisons for public support of K–12 education: Evidence from a national survey experiment. *Educational Researcher*, 41(7), 262–268.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education.

Nichols, S., & Good, T. (2004). *America's teenagers—myths and realities: Media images, schooling, and the social costs of careless indifference*. Mahwah, NJ: Erlbaum

Payne, C. (2010). *So much reform, so little change*. Cambridge, MA: Harvard Education Press.

Phi Delta Kappa. (2011). Highlights of the 2011 PDK/Gallup poll. Retrieved from http://www.pdkintl.org/poll/docs/pdkpoll43_2011.pdf

Phi Delta Kappa. (2012). Highlights of the 2012 PDK/Gallup poll. Retrieved from http://www.pdkintl.org/poll/docs/2012-Gallup-poll-full-report.pdf

Pidgeon, D. (1970). *Expectation and pupil performance* (Doctoral dissertation). Slough, Great Britain: National Foundation for Educational Research in England and Wales.

Rakes, C. R., Valentine, J. C., McGatha, M. B., & Ronau, R. N. (2010). Methods of instructional improvement in algebra: A systematic review and meta-analysis. *Review of Educational Research, 80*(3), 372–400.

Raudenbush, S. (1984). Magnitude of teacher expectancy effects on pupil IQ as a function of the credibility of expectancy induction: A synthesis of findings from 18 experiments. *Journal of Educational Psychology, 76*(1), 85–97.

Ravitch, D. (2010). *The death and life of the great American school system: How testing and choice are undermining education*. New York, NY: Basic Books.

Ripley, A. (2010, January 12). What makes a great teacher? *The Atlantic.* Retrieved from http://www.theatlantic.com/magazine/archive/2010/01/what-makes-a-great-teacher/307841/.

Rohrkemper, M., & Corno, L. (1988). Success and failure on classroom tasks: Adaptive learning and classroom teaching. *Elementary School Journal, 88*, 297–312.

Rosenshine, B., & Furst, N. (1973). Research on teacher performance criteria. In B. O. Smith (Ed.), *Research in teacher education: A symposium.* Englewood Cliffs, NJ: Prentice-Hall.

Rothstein, R. (2000). Toward a composite index of school performance. *Elementary School Journal*, *100*(5), 409–441.

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.

Rubie-Davies, C. M. (2007). Classroom interactions: Exploring the practices of high and low expectation teachers. *British Journal of Educational Psychology, 77*, 289–306.

Rubie-Davies, C. M. (2010). Teacher expectations and perceptions of student characteristics: Is there a relationship? *British Journal of Educational Psychology*, 80, 121–135.

Sahlberg, P. (2010). *Finnish lessons: What can the world learn from educational change in Finland*? New York, NY: Teachers College Press.

Schmidt, W., & Maier, A. (2009). Opportunity to learn. In G. Sykes, B. Schneider, & D. Plank (Eds.), *Handbook on educational policy research* (pp. 541–559). Washington, DC: American Educational Research Association.

Schoenfeld, A., & Pearson, P. (2009). The reading and math wars. In G. Sykes, B. Schneider, & D. Plank (Eds.), *Handbook on educational policy research* (pp. 560–580). Washington, DC: American Educational Research Association.

Shipp, D., & Deer, G. (1960). The use of class time in arithmetic. *Arithmetic Teacher*, 7, 117–121.

Shuster, A., & Pigge, F. (1965). Retention efficiency of meaningful teaching. *Arithmetic Teacher, 12*, 24–31.

Stigler, J. W., Lee, S-Y., & Stevenson, H. (1987). Mathematics classrooms in Japan, Taiwan, and the United States. *Child Development*, *58*, 1272–1285.

Strong, M. (2011). *The highly qualified teacher: What is teacher quality and how do we measure it?* New York, NY: Teachers College Press.

USA Today. (2012, December 2). 5 states to increase class time in some schools. *USA Today*. Retrieved from http://www.usatoday.com/story/news/nation/2012/12/02/longer-school-days-year/1740703/

Veldman, D. J., & Sanford, J. P. (1984). The influence of class ability level on student achievement and classroom behavior. *American Education Research Journal*, 21, 629–644.

Weinstein, R. S., & Middlestadt, S. E. (1975). Student perceptions of teacher interactions with male high and low achievers. *Journal of Educational Psychology*, *71*(4), 421.

APPENDIX I. WHAT TEACHERS SHOULD KNOW/DO TO HELP STUDENTS

Teachers should

- know the subject well.
- convey subject matter in accessible and accurate ways.
- assess degree of student subject matter learning.
- design new instruction based upon results of student assessments.
- help students to apply content knowledge to problems.
- help students to use content knowledge to find new problems.
- help students to value subject matter.
- respect student diversity and teach students to value diversity.
- promote citizens for tomorrow who value democracy.
- promote self-expression.
- teach students to be patient and take turns.
- teach students to initiate and to lead.
- create safe environments and handle disciplining.
- create pleasant environments for students.
- work in teams to improve curriculum.
- work in teams to improve school governance and functioning.
- share ideas and mentor other teachers.
- develop appropriate attitudes toward drugs.
- encourage students to become good citizens.
- enhance aesthetic appreciation for music and art.
- use creative skills.
- help students develop vocational awareness of job skills.
- help students value physical health.
- help students develop pro-social attitudes.
- help students value, understand, and use technology
- help students communicate well both orally and verbally.

APPENDIX II. ENTHUSIASTICALLY RECOMMEND INNOVATION: SOME EXAMPLES FROM THE PAST

Tape Recorder
TV
Internet
Chalk
Blackboard
Team Teaching
Expanded Class Period
Thematic Teaching
Teaching Combinations—Teach Math/Science Together
Principal as Leader
School-wide Reform
Looping
School Leadership Teams
Merit Pay
Career Ladder
Summer School
Pencil Sharpener
Peer Tutoring
Project-based Learning
Mastery Learning
Individualized Learning
Phonetics
Whole Language
Grade Retention
Grade Promotion

Magnetic Schools/Charter Schools
Gifted
Cross-grade Tutoring
Small Groups
Pair-share
Competition
Cooperation
Private
Individual Q&A
Choral Responding
Seatwork
Immersion
Distributed practice
Ability Grouping
Heterogeneous Grouping Objectives
Behavioral
Advanced Organizers
Overhead projector
Mimeograph machine
Xerox copier
Schools within schools
Magnet school

APPENDIX III. *THE NEW YORK TIMES* HEADLINES

From 1900

- Pedagogy Past and Present—Editorial: July 21
- Parasitic Luxury in College While Costly Clubhouses are Maintained, Professors are Miserably Underpaid—Criticism by President R. E. Jones of Hobart: October 14
- College Entrance Board—An Association for Examinations Organized at Columbia: November 18
- Catholics and School Money—Editorial: November 28

From 1901

- The Catholic Position on Education Defined—A Message from the Pope: April 13
- Handwriting in Schools—Authorities in Philadelphia will not Discard Vertical System: May 11
- National Educators Conclude their Work—Elementary Education a Problem the State must Solve: July 13
- Education of the Boy—High School and Early College Training Discussed: October 13

From 1902

- The Cry for Education—Letter to the Editor: March 16, 1902
- Minority's Victory in Board of Education: July 8, 1902

From 1903

- Education in the South—Needs of the States for the Welfare of the Children: January 9, 1903
- When to Begin School—Massachusetts Educators Divided as to Age When a Child Should be Put in School

Cite This Article as: *Teachers College Record* Volume 116 Number 1, 2014, p. -
http://www.tcrecord.org ID Number: 17289, Date Accessed: 8/12/2015 12:35:13 PM

Purchase Reprint Rights for this article or review