Minding the Gap between Research and Policy Making

In: The SAGE Handbook for Research in Education: Pursuing Ideas as the Keystone of Exemplary Inquiry

By: David N. Plank
Edited by: Clifton F. Conrad & Ronald C. Serlin
Pub. Date: 2014
Publishing Company: SAGE Publications, Inc.
City: Thousand Oaks
Print ISBN: 9781412980005
Online ISBN: 9781483351377
DOI: http://dx.doi.org/10.4135/9781483351377

Print pages: 43-58

Subway riders in London are constantly warned to “mind the gap,” the dangerous empty space between the platform and the train. Unwary riders who fail to heed this advice may suffer a variety of unpleasant consequences, ranging from scuffed shoes to broken ankles.

In this essay, I warn readers to mind a different and vastly wider gap: the one between researchers and policy makers. Researchers often bemoan the fact that policy makers fail to take research findings into sufficient account when making policy choices (Weiss, 1977). For their part, policy makers complain that research fails to provide answers to the questions they are obliged to address. There is agreement on both sides that the current state of affairs is unsatisfactory, and there are frequent calls for change, but the gap between scholars and policy makers remains wide and apparently unbridgeable.

Why should this be the case? I argue that the gap between researchers and policy makers finds its origins not only in mutual obstinacy or misunderstanding (though these are common enough) but also in a set of dilemmas that are intrinsic to the field of policy research. These dilemmas originate in fundamental differences between the orientations and interests of the two groups. These differences almost inevitably produce disagreements about which questions merit study and which answers merit attention.

In the section that follows, I identify some of the crucial obstacles that must be overcome by those who venture into the realm of policy research, and I discuss the ways in which these obstacles serve to maintain the gap between scholars and policy makers. I then turn my attention to two policy issues—teacher quality and school choice—that have attracted a great deal of recent interest from both groups. A brief review of the scholarly and policy debates on these two issues illustrates some of the reasons why research findings seldom achieve immediate or lasting influence over policy choices. For each of these issues, however, I also provide examples of research that has successfully bridged the gap, influencing policy making while maintaining academic integrity. I conclude the essay with some reflections on whether and how the divide between researchers and policy makers might be bridged, suggesting that the creation of specialized institutions may be necessary to ensure that policy research is useful to policy makers in the field and simultaneously impressive to scholars in the academy.

I begin from the premise that policy research aims to answer policy-relevant questions and, thereby, to influence policy. I am not concerned with broader types of policy research, including research on the policy process (Baumgartner & Jones, 1993; Kingdon, 1984; Lindblom, 1959; Pressman & Wildavsky, 1984) and policy implementation (Lipsky, 1983; Mazmanian & Sabatier, 1989; Spillane, 2000). I recognize that scholarship on agenda setting and policy implementation has deep implications for the ways in which research influences policy, but a review of this literature is beyond the scope of this essay.

The Origins of the Gap

Scholars working in the field of policy research must overcome a variety of obstacles not often faced by scholars in other fields. These obstacles find their origins in the quite different orientations and interests of scholars and policy makers with respect to the research enterprise. Under many circumstances, the demands and responsibilities of academic research conflict with the needs and expectations of policymakers. Too often the consequence is acrimonious misunderstanding and mutual distrust, both of which serve to widen rather than narrow the gap between the two groups.

It is impossible to address all of the points of dissension between scholars and policy makers in a brief essay.[2] In the discussion that follows, I identify some of the most salient points under three main headings: Questions, Answers, and Arguments.

Questions

Scholars tend to be a prickly bunch, strongly resistant to the idea that anyone—federal agencies, state legislators, or deans—should tell them which research topics merit their

attention. By training and professional inclination, they are drawn to questions they find theoretically or empirically interesting, and they are rewarded for work that impresses their peers and pushes back the frontiers of knowledge. For most scholars, whether and how their work might influence policy is at best a secondary consideration.

Successful scholars focus their energies on questions that are both interesting and answerable, and they frame their questions carefully to make them tractable for research. They take great care in choosing the questions they seek to address and in designing careful research strategies to pursue the answers. Asking the right questions in the right way is the key to the success of the research enterprise.

In contrast, policy makers seek answers to questions currently featured on the policy agenda (Kingdon, 1984). Their interest in a particular question may be intense in the short run, but it generally persists only so long as the issue remains above the fold in the local newspapers. As public attention shifts, they quickly move on to other issues. Policy makers are simply not interested in the theoretical and empirical challenges that motivate scholars. They value research not because it helps to advance knowledge but only insofar as it helps them to answer the questions their constituents expect them to address.

Which Questions Are Important?

Only rarely do the questions posed by scholars coincide with those that interest policy makers. From the scholar's viewpoint, many of the questions that policy makers pose are uninteresting, whereas others have no answers. From the policy maker's viewpoint, many of the carefully framed questions posed by researchers fail to provide clear guidance on the urgent issues that they face every day.

By way of illustration, consider the divergent ways in which scholars and policy makers characteristically frame questions related to school violence. This issue appears on the policy agenda only occasionally, often in the wake of school shootings or other high-profile incidents. When it does appear, however, it generally arrives in the company of widespread public outrage and intense pressure on policy makers to do something in response. Responding to this pressure, policy makers ask urgent but intractable questions about how to prevent further incidents and express frustration at researchers' apparent lack of concern for children's safety. From the point of view of scholarship, this sudden interest looks like ill-informed hysteria.

Scholars have in fact produced a substantial body of research on the issue of school violence, but the questions that they have addressed are quite different from those that policy makers raise. Scholars have asked whether schools are dangerous environments for children and whether schools have become more or less safe over time (DeVoe et al., 2004). They

have asked about the incidence of school shootings and about the effectiveness of policy interventions aimed at preventing them (Fisher & Kettl, 2001; Moore, 2003). And they have asked questions about other forms of school violence, including bullying, that affect far more children and may be more readily susceptible to policy intervention (Moore, 2003; Smith, 2003). To policy makers seeking to prevent the next mass shooting, however, most of this research looks like irrelevant dithering.

Neither group is entirely correct in its characterization of the other, but the gap between the two groups is clearly wide. Scholars and policy makers ask different questions, and they frame their questions in different ways. One consequence is that policy research contributes little to the policy debate, leaving policy makers to adopt policies that take little account of—or sometimes fly in the face of—scholarly research findings. In the case of school violence, for example, policy makers around the country have adopted a variety of “zero tolerance” policies, which many scholars view as ineffective or even counterproductive (Dunbar & Villareul, 2002; Newman, 2004, p. 288). Competing policies that receive greater research support are dismissed because they do not answer policy makers' current questions.

Which Questions Can be Answered?

Policy makers often ask seemingly reasonable questions that scholars cannot answer. For example, policy makers may want to know whether a given policy is better than the alternatives. Scholars, quite reasonably, want a definition of *better*. Will success be judged in terms of increased student achievement or reduced dropout rates? Reduced teacher turnover or enhanced quality of new recruits? Policy makers naturally want all of these things and more; whether or not the questions they ask are amenable to research is not their concern. Without greater specificity, however, the question of which policy is better cannot be answered. Researchers may fall back on their own judgments about which inputs and outcomes are most significant, but there is no guarantee that the models they specify will satisfy the expectations of policy makers.

Even when the inputs and outcomes of interest are clearly identified, scholars may still find policy makers' questions to be intractable. Some questions cannot be answered while observing the standard canons of social science research. The variables of interest may be difficult or impossible to measure, and key contextual variables may be difficult or impossible to control. When these circumstances prevail, scholars may be reluctant to take on policy makers' questions, no matter how urgent they may be.

For policy makers, of course, the situation is rather different. Their questions demand answers. When research does not provide adequate guidance, they seek answers from other sources, as I discuss in the next section.

Faced with the gap between scholars’ and policy makers' definitions of interesting,

answerable questions, scholars need not despair. With sufficient care, questions can be framed in ways that meet the expectations of scholars and policy makers. One example is recent work in the New York City public schools that seeks to assess the relationship between measures of teachers' classroom practice and their value-added scores. This work is simultaneously interesting to academics and of immediate value to policy makers who are concerned with the measurement and enhancement of teacher effectiveness (Grossman et al., 2010).

Answers

Scholars and policy makers are likely to come to very different judgments about the usefulness of the answers they propose to their respective questions, in significant part because they hold dramatically different views of what constitutes good research. From the scholar's point of view, good research shares a number of features, including careful design, solid data, and conclusions based on cautious and responsible inference. From the policy maker's point of view, however, judgments about good research are likely to be based on criteria radically different from those valued in the academy. Compatibility with prior beliefs and political commitments, timeliness, strength and simplicity of findings, and clear implications for action all are likely to be valued above theoretical or methodological sophistication. As a result, the answers that scholars produce may be of little value from the policy maker's point of view, whereas the answers that policy makers value may not meet the standards of credible scholarship.

Ideology and Interests

Policy makers and scholars often have strongly held ideological views. Scholars, however, are rigorously enjoined to minimize the influence of their prior beliefs on their research. They are trained to observe the rules of social science and to base their conclusions on evidence and argument rather than on ideology. Policy makers have no such compunctions. For them, ideology may provide a useful filter through which to screen research findings and to decide whether they serve a larger political purpose. Findings that comport well with prior convictions or commitments may be accepted, whereas those that do not will be dismissed.

These different orientations to the role of ideology have two complementary consequences. First, policy makers may be disinclined to credit research that is not fully compatible with their ideological commitments, no matter how careful the research design or how compelling the findings. Second, policy makers' demand for research that supports or advances a particular policy agenda encourages the supply of such research from ideologically compatible but methodologically suspect think tanks and similar sources. If researchers are unable to supply the kind of information that policy makers seek, they will simply go elsewhere to find it (Weiss, 1995).

Take class size reduction, for example. This idea has immediate appeal for many constituencies, including parents, teachers, and much of the general public. For most of these groups, smaller classes are obviously better. Parents prefer schools where teachers are able to devote individual attention to every child, and teachers prefer classes that are small enough to allow them to respond to students' individual differences. Reducing class size is technically straightforward, whereas alternative policies aimed at improving teacher quality may be blindingly complex. As a result, policy makers at all points on the political spectrum are under enormous pressure to reduce class size, and they have spent vast sums of money toward this end. An increasingly solid body of research concludes that money spent on keeping classes small might better be spent on other initiatives, such as improving teacher quality (Ferguson, 1991; Ferguson & Ladd, 1996; Harris, 2002), but when research evidence conflicts with “common sense” and the opportunity to win votes, it does not carry much weight.

Timeliness

Perhaps the clearest difference between researchers and policy makers resides in their different orientations to time. As all scholars know, proper attention to the familiar steps in the research process—developing a theoretical framework, designing a study, collecting data, conducting analyses, and writing up results (not to mention seeking funding and dealing with the delays that characterize peer review and scholarly publishing)—requires a great deal of time, and shortcuts are potentially dangerous. Taking insufficient care in any of these steps can seriously if not fatally undermine the integrity of a research project. “Quick and dirty” alternatives to careful research are not likely to produce robust or persuasive findings, nor are they likely to win professional respect. As a result, research is generally slow to market; months, if not years, may pass between the initial formulation of a research question and the publication of results.

Policy makers operate on a very different schedule. The educational and other issues that come before them often arise in response to urgent public concerns, and they are expected to take action with appropriate dispatch. They rarely have months or years to wait for the completion of careful research because their constituents expect them to act immediately. They may well prefer good research to bad research, other things being equal, but adherence to the standards of the academy is of less concern than the immediate availability of useful findings. Getting it now is more important than getting it right.

Simplicity and Certainty

Social science research seldom produces simple answers. In the course of their efforts to identify relationships and measure effects, researchers are trained to navigate through a minefield of potentially confounding factors, including spurious correlation, multiple

causation, and noise in the data. The myriad choices that scholars make as they collect and analyze data depend not on rules but rather on expert judgment. As often happens, therefore, two careful researchers who make slightly different but equally justifiable decisions can easily produce dramatically different findings.

Scholars similarly recognize that their findings are context dependent; relationships that are significant under some circumstances may not necessarily hold under others. Scholars are consequently reluctant to generalize beyond the scope of the specific environments they study. Policies and practices that appear to be effective in one set of schools may have quite different effects elsewhere because of natural variation in local circumstances. In addition, schools are unusually noisy environments, both literally and figuratively, and this can make it difficult to identify or measure significant relationships even when they are present.

Social scientists recognize that their findings are almost always provisional, subject to revision on the basis of better models and better data. They are rewarded for acknowledging and exploring the complexity of the questions they investigate. Moreover, they move forward in their work through disagreement and debate (Popper, 1959). The publication of compelling new findings spurs efforts by other scholars to refine or debunk them. Ambiguity and uncertainty define the natural state of affairs in the academy; scientific progress relies on informed dissent (Kuhn, 1970).

Scholars consequently recognize that “right” answers are hard to come by. Instead of providing unequivocal support for a specific policy position, they often answer policy makers' questions with this honest answer: “It depends.” Scholars' reluctance to generalize frustrates policy makers who must make choices for schools and students across a variety of contexts. As Green (1983) observed, “Public policy is the drop-forge or the axe of social change. It is not the knife or scalpel…. Policy deals always with what is good in general, on the whole,
and for the most part” (p. 322). Policy makers are obliged to make use of the best blunt instruments that come to hand. Scholars are adept with the scalpel, but this is only rarely a good substitute for an axe.

For policy makers, the professional norms of social science research amount to little more than hand-wringing. The cautions and caveats that accompany most research findings may garner the respect of other researchers, but they simultaneously disqualify most scholarly work from serious policy influence. “It depends” rarely offers helpful guidance for the binary decisions that policy makers regularly face. Policy makers must decide whether to support or oppose bills, whether to implement new strategies or not, and whether to vote yes or no. These stark choices afford little room for the uncertainties reflected in most social science research.

Faced with these ambiguities, policy makers may adopt either of two courses of action. On

the one hand, they can simply ignore scholarly research and seek answers to their questions from other sources. Alternatively, they can adapt scholars' answers to their own purposes by ignoring the authors' cautionary notes. Research can thus be made to provide clear policy guidance, but at a potentially high cost in scholarly integrity and credibility.

Arguments

In addition to posing different questions and valuing different answers, scholars and policy makers are comfortable with different kinds of arguments. The differences are both substantive and rhetorical. First, scholars and policy makers may acknowledge different arguments as trustworthy based on their respective evaluations of different warrants. Second, the rhetorical conventions that govern most research reports convey essential information to scholars but represent barriers to understanding for nearly everyone else, including policy makers.

The Problem of Warrants

In their professional capacity, scholars are uniquely responsive to arguments supported by scientific evidence obtained on the basis of solid research; arguments that rely on other kinds of warrants are discounted. In contrast, policy makers are open to arguments supported by a variety of other warrants, including compelling personal experience and common sense, which are typically dismissed by academics. Policy makers may take scientific evidence into account in their deliberations, but other kinds of evidence are often accorded equal or even greater weight.

Scholars are generally of the view that research evidence trumps other possible warrants. They are skeptical—even dismissive—of the roles that common sense, personal experience, and compelling anecdotes regularly play in policy making. They are even cautious when considering scientific evidence. When the available research evidence is weak, scholars argue against precipitate action until more research can be conducted and often call for the careful evaluation of pilot programs in advance of full-scale program implementation. When research evidence is strong, on the other hand, scholars may call for immediate action. They are often surprised and frustrated when policy makers fail to take the strength of research findings into sufficient account when making policy choices, and they are horrified when policy makers give undue weight to other sorts of warrants.

For policy makers, in contrast, research findings merit no special reverence. They may provide a useful warrant for a policy argument, but a conversation with Aunt Eleanor may carry equal or greater weight because of her long experience on the local school board. For most people, warrants derived from experience, authority, ideology, or common sense provide more valuable guidance than do findings from academic research (Audi, 2002). When the “right” answer to a policy problem can be found as easily by talking to the people in line at

the grocery store as by reading research reports, researchers face an uphill battle to win serious attention in policy debates.

The Problem of Rhetoric

Researchers often present their findings in ways that are confusing to policy makers and in prose that is less than riveting in its rhetorical flourish. Scholars are trained to write in formats where their arguments can be fully developed and to present their results as transparently as possible. Their lengthy reports and articles typically include a review of prior theoretical and empirical work, an explanation of methodological choices, and an exploration of competing arguments. Scholars are expected to acknowledge weaknesses in the available data, to support their conclusions by presenting detailed information on their research methods and results, and to include disconfirming evidence and alternative explanations for their findings. Providing a full account of the work and how it was conducted is an essential feature of a credible and persuasive research report.

Much to the dismay of scholars, however, documents that respect the familiar conventions of academic writing are unlikely to win much attention in policy debates. Like other readers, policy makers are apt to distrust documents they do not understand and to begrudge the time it takes to sort out the rhetorical complexities and infelicities of most academic work. They are likely to find a 35-page report that bristles with statistics, footnotes, and caveats to be windy, opaque, and off-putting and to conclude that it has little of value to say on the question of what to do about the problems they face. They may find a one-paragraph summary or a set of bullet points more useful and persuasive, but only if the conclusions provide a guide for action and do not include phrases such as “It's complicated” and “It depends.” Such forthrightness does not come easily to most scholars.

The style of academic writing poses a further obstacle for policy makers. Researchers are trained to let their work speak for itself. Rather than seeking to persuade readers with rhetorical fireworks and emotional arguments, scholars present their work in sufficient depth and detail to allow their colleagues to make their own judgments. This is as it should be in academic writing, but for readers who lack the expertise needed to sort things out for themselves (including nearly all policy makers), the results are at best unpersuasive and at worst unintelligible. Other parties—including Aunt Eleanor—may be more accustomed to persuasive speaking and writing, and policy makers may be more inclined to listen to them than to scholars.

Falling into the Gap

The obstacles that scholars and policy makers must overcome to bridge the gap between them are exacerbated in specific policy debates by the ubiquity of familiar research problems.

In the absence of agreement on how to address these problems, different scholars adopt different strategies, and these inevitably produce different—and sometimes diametrically opposed—findings. From the point of view of the academy, this is entirely normal and is essential to the advancement of knowledge. From the point of view of policy makers, however, it is chaotic and dysfunctional.

In this section, I focus on two policy questions—teacher quality and school choice—and ask why research contributes so little to policy making on these issues. The answer is partly attributable to the obstacles discussed in the preceding section but is also attributable to research problems specific to the two issues. In the case of teacher quality, the lack of valid and reliable measures for key variables precludes firm responses to the questions that policy makers seek to answer. The problem of selection bias has the same consequence in the case of school choice. In the discussion of these two issues, I also present examples of policy research that has successfully bridged the gap between academics and policy makers, engaging the curiosity of the former while providing clear guidance to the latter.

Teacher Quality

Teachers matter. A lot. Despite occasional efforts to develop “teacher-proof” curricula or classroom scripts, there is general agreement among scholars, policy makers, and parents that individual differences in teacher quality have a greater influence on student achievement than any other school input. A large and growing body of research confirms this conventional wisdom (Goe, Bell, & Little, 2008; Kane & Staiger, 2008; Murnane & Phillips, 1981; Nye, Konstantopoulos, & Hedges, 2004; Rivkin, Hanushek, & Kain, 2000; Rockoff, 2004; Sanders & Horn, 1998; Wenglinsky, 2000).

Recognizing the importance of teacher quality, policy makers at the federal, state, and local levels are eager to develop policies aimed at improving the quality of teaching. In this effort, they receive unsolicited advice from unions, business organizations, philanthropists, colleges of education, and other interested constituencies hoping to influence the outcome. When policy makers turn to researchers for help in crafting their initiatives, however, they often come away empty-handed.

The questions that policy makers seek to answer appear to be straightforward. They want to know what attributes characterize effective teachers, and they want to know which policies help to ensure that more teachers display these attributes. Unfortunately, research offers few useful answers to these questions for two main reasons. First, there is disagreement among scholars (and others) about how to define and measure effective teaching. Are student test scores our primary concern, or are other outcomes also important? If we accept the primacy of test scores, how can we link students' scores to the contributions of individual teachers? Second, even if the obstacles to measuring effectiveness could be resolved, the available

evidence about the teacher attributes that are associated with effective teaching remains weak and equivocal. Many obvious indicators of teacher quality, including certification status and experience, appear to have limited effects on student performance. In the absence of clear strategies to overcome these obstacles, scholars cannot answer the questions that policy makers pose.

Can We Measure Student Learning?

Scholars and policy makers agree that effective teachers are an essential contributor to student learning. But which outcomes matter most? Much of the current research on teacher quality focuses on student test scores as the principal measure of teacher quality (Hanushek & Rivkin, 2007; Monk, 1994; Sanders & Horn, 1998). This is in large part a function of availability and ease of interpretation, as many scholars readily acknowledge. Other educational outcomes, including curiosity, morally defensible behavior, critical thinking, and civic engagement (Gutmann, 1999; Kohn, 1999; Noddings, 1995), are generally recognized as important, but they are harder to define and measure. They consequently receive less scholarly attention.

Even when researchers restrict their attention to student test scores as a measure of teacher quality, they encounter several daunting problems. First, standardized test scores are affected by many factors other than the quality of a student's teachers. Family background and prior learning have a huge influence on a student's current scores, but those factors are entirely beyond the control of the student's teacher this year. Second, even the most sophisticated assessments are subject to significant measurement error.

In an attempt to address the first of these problems, scholars and policy makers have increasingly turned to the analysis of student learning gains (Gordon, Kane, & Staiger, 2006; Rothstein, 2010; Sanders & Horn, 1998; Webster & Menro, 1997). The shift to “value-added” models is motivated by the desire to isolate the unique contributions of specific schools and teachers to student learning. In contrast to cross-sectional analyses of student achievement, these models introduce controls for family background and prior achievement by testing students at two or more points in time and analyzing performance changes between assessments. The resulting models are generally quite complex, perhaps limiting their appeal to policy makers, but they nevertheless offer a way forward on an otherwise intractable problem.

Lamentably, however, even the most sophisticated value-added models remain subject to large measurement errors stemming from three main sources. First, test scores may change from year to year for a variety of reasons, many of them unrelated to the teaching that students receive. Apart from the inevitable dispersion of observed performance around a student's true score, a flu outbreak, unusual weather, or the generosity of a helpful parent

who brings doughnuts or chewing gum on the morning of the test can all affect test scores in a specific school or classroom (Kane & Staiger, 2002; Rogosa, 1999, 2001). These problems are exacerbated when the number of observations is small, as is invariably the case in individual classrooms.

Second, test scores at best measure only a small sample of the knowledge and skills that are taught during a given school year, and tests are only administered in a limited number of grades and subjects. A genuinely comprehensive testing system would be prohibitively expensive in terms of both test development and testing time. As a result, test scores provide only a very rough measure of what children have actually learned, even in core instructional areas such as reading and mathematics. Other learning outcomes are measured with even less precision, if they are measured at all.

Third, value-added models rely on the assumption that tests can be equated across grade levels (Martineau & Plank, 2004). For student learning gains to be measured accurately, two conditions must be met. First, achievement must be measured using identical equal-interval scales on which a 5-point score difference means the same thing at all points within and across tests. Second, tests must consistently assess students on the same underlying constructs: For example, reading tests must measure the same skills (e.g., vocabulary) across grades and in the same proportions across test instruments. These assumptions are routinely violated. Scholars generally agree that the standardized tests administered by most schools do not use equal-interval scales, and it is clear that the mix of skills assessed at different grade levels should and does change as children mature and advance to each new grade level. When these assumptions are violated, even sophisticated value - added assessment models may do a poor job of measuring teacher effectiveness (Martineau & Plank, 2004).

Do Teacher Attributes Matter?

Suppose for a moment that there is a simple way to address these problems and to measure student outcomes fully and accurately. Would we then be able to extricate ourselves from the teacher quality quagmire? Sadly, no. As we lift one foot out of the swamp, the other foot continues to sink. Accurate measures of student achievement would enable us to identify teachers who successfully raise student performance, but by themselves they would provide virtually no insight into the combination of knowledge, attributes, and skills that makes these teachers successful. This, of course, is precisely what policy makers need to know as they seek to design policies to recruit, reward, develop, and retain better teachers for American classrooms.

When it comes to identifying the attributes of effective teachers, researchers have barely reached the starting gate. According to Goldhaber (2001), easily measured attributes,

including experience and advanced degrees, explain only about 13% of a teacher's contribution to student achievement, leaving 97% of the variance in teacher performance unaccounted for. There is some evidence that teachers with high scores on tests of verbal ability perform well in the classroom (Ballou & Podgursky, 1997; Ehrenberg & Brewer, 1995). The selectivity of the undergraduate institution that a teacher attended (Ehrenberg & Brewer, 1994; Summers & Wolfe, 1977) and the possession of a major in the subject area taught (Monk, 1994) may also be associated with teacher effectiveness, but the explanatory power of these variables remains extremely small. As Woods (1994) noted, "Teachers' personal qualities are considered to be the decisive factor in effective teaching. However, it is never quite clear what these qualities are" (p. 84).

Differences in classroom practices may account for a substantial share of the unexplained variation in teacher performance, but specifying different practices and linking them to differences in student achievement is a messy, time-consuming task (Grossman et al., 2010; Wenglinsky, 2000). Sound data on classroom practices are hard to come by, especially on a scale that would allow meaningful comparison and generalization. In short, much of what seems to matter most—teacher practices—is difficult and expensive to measure. Information regarding the other piece of the puzzle—teacher characteristics—is often unavailable, and when available, it rarely proves to be particularly helpful.

Where does this leave policy makers? They pose a seemingly straightforward question, and scholars answer with a resounding, "Hmmm…." The reason research does not have much impact in the teacher quality debate is not because policy makers are ignoring academic work or because academics are failing to address urgent policy questions; rather, it is because research cannot yet provide unequivocal responses to the questions that policy makers seek to answer. It is therefore little cause for wonder when policy makers throw up their hands in despair and seek guidance from more helpful sources, including Aunt Eleanor and the president of the local Chamber of Commerce.

There are nevertheless some grounds for optimism that research on teacher quality may bear fruit in terms of policy. For example, recent work by Pam Grossman and her colleagues (2010) in the New York City public schools shows that two very different approaches to identifying effective teachers—students' value-added scores and classroom observations of instructional practice—are significantly correlated, which suggests that the search for valid and reliable measures of teacher effectiveness is not entirely quixotic. The Gates Foundation and the federal government are financing large-scale policy experiments that support the development and implementation of new approaches to teacher evaluation and compensation. Each of these initiatives includes a substantial commitment of resources to fund systematic research on program impact, which may ultimately point the way toward more helpful guidance for policy makers as they seek to recruit, support, retain, and reward talented teachers.

School Choice

Another issue that has attracted significant interest from both researchers and policy makers is school choice. The debate over vouchers, charter schools, and other policies that expand the array of choices available to households has captivated many scholars, and policy makers on both sides of the issue have marshaled an impressive array of research findings to support their positions (Henig, 2009). On balance, however, it is uncertain whether research has had any influence in the policy debate beyond reinforcing the convictions of the already persuaded. Despite the proliferation of sophisticated studies conducted by admirable scholars, the findings from research on school choice have proved to be too fragile and contested to persuade key actors in the policy debate. One key reason for this failure is the familiar research problem of selection bias.

Much of the policy debate on school choice has focused on the question of whether and how student achievement in “choice” schools—Catholic schools, private schools, voucher schools, and charter schools—differs from student achievement in traditional public schools. The answer to this question turns decisively on whether the students who make nontraditional choices are comparable to the children who remain in traditional public schools. If the two groups of students are indeed similar, then achievement differences across schools may rightly be attributed to differences in the schools themselves. If the two groups of students differ in significant ways, however, then achievement differences across schools may be at least partially attributable to selection bias.

Not surprisingly, therefore, the scholarly debate on school choice has featured sharp disagreements over sample selection and the constitution of appropriate control groups. The early furor over Milwaukee's voucher program originated in disagreement about which public school students were comparable to students receiving vouchers (Rouse, 1998; Witte, 1998). Recent efforts to design voucher experiments in Florida, Arizona, Washington, DC, and elsewhere have been similarly vexed by the problem of selection bias. Small shifts in the composition of control groups can produce substantially different conclusions about the efficacy of voucher schools (Hoxby, Murarka, & Kang, 2009; Krueger & Zhu, 2004; Myers, Peterson, Mayer, Chou, & Howell, 2000). In addition, students who participate in choice programs may differ from those who do not in unmeasured but important ways, including religious commitment and parental support. It has even been argued that the very act of receiving a voucher may distinguish choice students from other students in ways that are conceptually independent of the effectiveness of the schools they attend (Floden, 2000).

On the one hand, there is a large and persuasive body of evidence suggesting that students in Catholic schools perform better than similar students in public schools (Bryk, Lee, & Holland, 1993; Coleman, Hoffer, & Kilgore, 1982), and there is a small but growing literature

suggesting that at least some of the students who participate in school choice programs also experience gains (Hoxby, Murarka, & Kang, 2009; Peterson & Howell, 2003; Witte, Weimer, Schlomer, & Shober, 2007). On the other hand, there is also an extensive literature calling all of these results into question on the grounds that students who make nontraditional choices are systematically different from students who remain in traditional public schools (Fuller, Elmore, & Orfield, 1996; Krueger & Zhu, 2004). The consequence is an impasse; disputes over selection bias and the identification of appropriate control groups cloud the debate over school choice, leaving uncontested findings in extremely short supply.

That the scholarly debate on the impact of school choice policies remains stalled by the intractable problem of selection bias hardly means that research findings have been excluded from the policy debate. In fact, the policy debate over choice, charter schools, and vouchers rages on with undiminished intensity, and research continues to be deployed prominently on both sides of the issue. Nevertheless, the policy debate is driven not by research findings but rather by ideology, political commitments, and other considerations (Henig, 2009; Plank & Sykes, 2000). Advocates of school choice policies seize on data showing achievement gains to argue that “choice works,” especially for disadvantaged students, and that publicly funded choice programs should be expanded. Opponents embrace competing research and argue that the gains associated with school choice are either illusory or too small to justify the losses by students left behind. For now, therefore, the question of whether school choice policies lead to improved student performance remains unanswered.

This, however, is neither the only nor necessarily the most fruitful way in which to frame the debate on school choice (Plank & Sykes, 2000). Scholars pursuing other questions have begun to produce less contested findings and to generate more useful guidance for policy makers. For example, work on the different incentives created by policy choices reflected in charter school legislation (Arsen, Plank, & Sykes, 1999), on the nature and extent of innovation in charter schools (Mintrom, 2000, 2001), and on the critical roles played by authorizers in ensuring accountability and supporting improvement in charter schools (Vergari, 2009) has identified specific ways in which state policies can be strengthened to increase the likelihood that charter schools will deliver on their initial promise. These and similar studies have contributed to the emergence of more nuanced views of charter schools and to a better-informed and ultimately more productive conversation about policies to realize the promise and avoid the pitfalls of charter schools. Credulity about charter schools as a reform strategy nevertheless continues to hold great sway among policy makers, in the Obama administration and beyond (Plank & Sykes, 2003). The gap between researchers and policy makers remains deep and wide.

Can the Gap be Bridged?

Where does this leave us? Can scholars and policy makers overcome their differences? The

long-standing tradition is one of mutual incomprehension. Researchers complain that policy makers do not seek out or value their advice (“what we know”), whereas policy makers complain that research findings bear little connection to the problems they are obliged to address (“what we need”). This is mutually advantageous, at least in the short run. Neither group is obliged to change its behavior, and each is able to cast the blame elsewhere.

As researchers, we may find this tradition to be quite comfortable. We can win the praise of our colleagues and enjoy the rewards of academic success without worrying about the disregard of policy makers. If they choose to ignore our work, the loss is theirs, not ours.

Comfort with the prevailing situation may carry a price, however, to the extent that scholars care about what goes on in schools. When policy makers ignore research findings, they may well repeat past mistakes or adopt misguided policies. For example, policy makers' undiminished enthusiasm for class size reduction comes at a huge cost, not only in financial terms but, more important, in terms of failure to invest in alternative policies that promise greater impact on student learning. Minding the gap between scholars and policy makers is a poor substitute for action; if research is to contribute to the development of policies and the improvement of schools, we will have to find a way to build a durable bridge that will more effectively connect the two groups.

If the gap is to be bridged, the responsibility will fall mainly on scholars, rather than on policy makers. Even policy makers who are eager to do the right thing have no real reason to believe that scholars have the answers they seek, and they are regularly confirmed in this judgment. Thus, the key question is how researchers can make their work more accessible and useful to policy audiences.

A relatively cheerful answer to this question comes from Weiss (1977). In her view, scholars need only to continue doing their academic work according to the canons of scholarly practice. Over time, useful findings will seep into the policy conversation, shifting the terms and boundaries of the debate. This has occurred in some important instances. For example, research findings from the Perry Preschool Project (Barnett, 1985) demonstrated that the benefits of early educational intervention far outweighed the costs, drawing the attention of policy makers to the issue of prekindergarten education. Findings from the Third International Mathematics and Science Study and other comparative assessments of student performance have strengthened efforts to raise standards and expectations in American schools (Schmidt et al., 2001). In many other cases, however, the policy debate moves forward without any attention paid to policy research, while potentially useful findings languish in scholarly obscurity.

This may be the best we can do, but most policy researchers aspire to inform policy debates more directly. Individual scholars may focus their energies on questions that are of immediate

policy concern, but this only begins to address the problem. They must also gain the attention of relevant policy makers, often by producing multiple representations of their work tailored for audiences that include the media and legislative aides. They must compete with lobbyists and well-funded partisan think tanks to gain recognition. They must simultaneously tend to their academic obligations and produce work that meets scholarly expectations for rigor and theoretical reach. Some exceptionally gifted and energetic scholars can pull this off, but most cannot. Most researchers need institutional support to ensure that research findings reach policy makers in a timely and accessible way.

The rise of university-affiliated policy centers offers one promising way forward. Centers such as the Consortium for Policy Research in Education, Policy Analysis for California Education, and the Education Policy Center at Michigan State University reflect efforts to institutionalize stronger connections between the worlds of research and policy. These centers strive to address issues of concern to policy audiences on reasonable timelines and to translate research findings into terms and formats that are accessible and useful to policy makers. They build institutional, as well as individual, relationships between universities and a variety of policy audiences. These institutional relationships endure beyond the terms of particular policy makers and allow both established and emerging scholars to quickly gain credibility with key constituencies.

Rather than viewing policy makers' needs as a burden, these policy centers recognize those needs as their raison d'être. They have supported excellent scholarly work while acknowledging the needs of policy audiences at every step in the research process, from formulating questions to disseminating findings. This orientation has increased their credibility to the point that policy makers may actively seek their guidance, as in the teacher compensation example described earlier. Over time, therefore, some of these centers have managed to construct bridges—still fragile and lightly traveled—across the divide that separates scholars and policy makers. Sadly, though, the number of similar centers that have died for lack of institutional commitment and support is far larger.

Some of the reasons policy research has so little to say to policy makers are intrinsic to the research enterprise, but others can be changed. Scholars cannot change the political environment within which policy is made, nor can they change policy makers' preferred sources of information. All they can change is their own behavior. Policy researchers who seek to affect ongoing policy debates may therefore find it worthwhile to devote considerably more attention and energy to solving the problems of relevance, timing, translation, and dissemination that restrict the current impact of their work. This is a challenging task, but the reward of increased policy relevance might make it worthwhile.

In closing, it is important to acknowledge that deeper engagement in policy research entails significant risks, both for individual scholars and for the broader scholarly enterprise.

Engagement in policy research almost inevitably requires attention to research topics and publication in formats and media that generally earn relatively little credit in the academy. Producing the kind of work that is valued by policy makers may divert time and energy away from work that would be admired by other scholars. Writing targeted to policy audiences might not always measure up to the expectations of academic audiences.

In addition, the intensity of contemporary policy debates and the voracious appetite of policy audiences for politically useful research findings may have pernicious consequences for scholarship. Those who publish most of their work on editorial pages and in think tank broadsides may find that their work serves mainly to provide policy makers with rhetorical ammunition that reinforces prior support for or opposition to specific policy initiatives, rather than with a deeper understanding of the issues. They may also find their own scholarly credibility severely impaired.

In the realm of policy research, it is essential to strike a balance that protects the integrity of academic research, even as it turns scholarly work to more practical ends. Given the urgency of the issues facing American schools and the commitment of many policy researchers to provide assistance in addressing those issues, a move toward the production of scholarship that is useful and accessible to policy makers may yield substantial rewards for schools and their students. Many researchers, therefore, may find that sacrificing some degree of scholarly autonomy and methodological elegance in exchange for greater policy influence is a worthwhile trade. It is nevertheless essential for scholars to be wary of the dilemmas and difficulties that characterize policy research and to mind their step as they stride across the gap.

Notes

1. A previous version of this chapter was coauthored by Debbi Harris.

2. For further discussion of the issues raised in this essay, see for example Kirst (2000), Moss (2007), and Hess (2008).

References

References
Arsen, D., Plank, D.N., & Sykes, W.G. (1999). School choice in Michigan: The rules matter. East Lansing: Michigan State University, Education Policy Center.

Audi, R. (2002). Epistemology: A contemporary introduction to the theory of knowledge. New York: Routledge.

Ballou, D., & Podgursky, M. (1997). Teacher pay and teacher quality. Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Barnett, W.S. (1985). Benefit-cost analysis of the Perry Preschool Program and its policy implications. Educational Evaluation and Policy Analysis, 7, 333–342.

Baumgartner, R., & Jones, B.D. (1993). Punctuated equilibria in politics. In R.Baumgartner & B.D.Jones (eds.), Agendas and instability in American politics (pp. 3–24). Chicago: University of Chicago Press.

Bryk, A.S., Lee, V., & Holland, P. (1993). Catholic schools and the common good. Cambridge, MA: Harvard University Press.

Coleman, J., Hoffer, T., & Kilgore, S. (1982). High school achievement: Public, Catholic, and private schools compared. New York: Basic Books.

DeVoe, J., Peter, K., Kaufman, P., Miller, A., Noonan, M., Snyder, T., et al. (2004). Indicators of school crime and safety: 2004. Washington, DC: National Center for Education Statistics.

Dunbar, C., & Villareul, F.A. (2002). Urban school leaders and the implementation of zero tolerance policy: An examination of its implications. Peabody Journal of Education, 77, 82–104. http://dx.doi.org/10.1207/S15327930PJE7701_5

Ehrenberg, R.G., & Brewer, D.J. (1994). Do school and teacher characteristics matter? Evidence from high school and beyond. Economics of Education Review, 13 (1), 1–17. http://dx.doi.org/10.1016/0272-7757%2894%2990019-1

Ehrenberg, R.G., & Brewer, D.J. (1995). Did teachers' verbal ability and race matter in the 1960s? “Coleman” revisited. Economics of Education Review, 14 (1), 1–21. http://dx.doi.org/10.1016/0272-7757%2894%2900031-Z

Ferguson, R.F. (1991). Paying for public education: New evidence on how and why money matters. Harvard Journal on Legislation, 28, 465–498.

Ferguson, R.F., & Ladd, H.F. (1996). How and why money matters: An analysis of Alabama schools. In H.F.Ladd (ed.), Holding schools accountable (pp. 265–298). Washington, DC: Brookings Institution.

Fisher, K.M., & Kettl, P. (2001). Trends in school violence: Are our schools safe? In M.Shafii & S.L.Shafii (eds.), School violence: Assessment, management, prevention (pp. 73–83). Washington, DC: American Psychiatric Press.

Floden, R. (2000). Does choice produce gains in student achievement outcomes?East Lansing: Michigan State University, Education Policy Center.

Fuller, B., Elmore, R., & Orfield, G. (1996). Who chooses, who loses? Culture, institutions, and the unequal effects of school choice. New York: Columbia University, Teachers College Press.

Goe, L., Bell, C., & Little, O. (2008). Approaches to evaluating teacher effectiveness: A research synthesis. Washington, DC: National Comprehensive Center for Teacher Quality.

Goldhaber, D.D. (2001). How has teacher compensation changed? In W.J.Fowler, Jr. (Ed.), Selected papers in school finance, 2000–2001 (pp. 11–30). Washington, DC: National Center for Educational Statistics.

Gordon, R., Kane, T.J., & Staiger, D.O. (2006). Identifying effective teachers using performance on the lob. Washington, DC: The Brookings Institution.

Green, T.F. (1983). Excellence, equity, and equality. In L.S.Shulman & G.Sykes (eds.), Handbook of teaching and policy (pp. 318–341). New York: Longman.

Grossman, P., Loeb, S., Cohen, J., Hammerness, K., Wyckoff, J., Boyd, D., et al. (2010). Measure for measure: The relationship between measures of instructional practice in middle school English Language Arts and teachers' value-added scores (Working paper 16015). Cambridge, MA: National Bureau of Economic Research. http://dx.doi.org/10.3386/w16015

Gutmann, A. (1999). Democratic education. Princeton, NJ: Princeton University Press.

Hanushek, E.A., & Rivkin, S.G. (2007). Pay, working conditions, and teacher quality. The Future of Children, 17, 69–86. http://dx.doi.org/10.1353/foc.2007.0002

Harris, D.N. (2002). Identifying optimal class sizes and teacher salaries. In H.M.Levin & P.McKewan (eds.), Cost effectiveness analysis in education (pp. 177–191). Larchmont, NY: American Education Finance Association.

Henig, J.R. (2009). Spin cycle: How research is used in policy debates: The case of charter schools. New York: Russell Sage Foundation.

Hess, F.M. (2008). When research matters: How scholarship influences education policy. Cambridge, MA: Harvard Education Press.

Hoxby, C.M., Murarka, S., & Kang, J. (2009). How New York City's charter schools affect achievement, August 2009 report. Cambridge, MA: New York City Charter Schools Evaluation Project. http://dx.doi.org/10.3386/w14852

Kane, T.J., & Staiger, D.O. (2002). Volatility in school test scores: Implications for test-based accountability systems. In D.Ravitch (ed.), Brookings papers on education policy, 2002 (pp. 235–283). Washington, DC: Brookings Institution.

Kane, T.J., & Staiger, D.O. (2008). Estimating teacher impacts on student achievement: An experimental evaluation (Working paper 14607). Cambridge, MA: National Bureau of Economic Research Working Paper. http://dx.doi.org/10.3386/w14607

Kingdon, J.W. (1984). Agendas, alternatives, and public policies. New York: HarperCollins.

Kirst, M.W. (2000). Bridging education research and education policymaking. Oxford Review of Education, 26, 379–391. http://dx.doi.org/10.1080/713688533

Kohn, A. (1999). The schools our children deserve: Moving beyond traditional classrooms and “tougher standards.”Boston: Houghton Mifflin.

Krueger, A.B., & Zhu, P. (2004). Another look at the New York City school voucher experiment. American Behavioral Scientist, 47, 658–698. http://dx.doi.org/10.1177/0002764203260152
Kuhn, T.S. (1970). The structure of scientific revolutions (2nd ed.). Chicago: University of Chicago Press.

Lindblom, C.E. (1959). The science of muddling through. Public Administration Review, 19, 79–88. http://dx.doi.org/10.2307/973677

Lipsky, M. (1983). Street-level bureaucracy: Dilemmas of the individual in public services. New York: Russell Sage.

Martineau, J.A., & Plank, D.N. (2004). Fairness in accountability policy: Is value-added assessment the answer?East Lansing: Michigan State University, Education Policy Center.

Mazmanian, D.A., & Sabatier, P.A. (1989). Implementation and public policy. New York: University Press of America.

Mintrom, M. (2000). Leveraging local innovation: The case of Michigan's charter schools (Working Paper No. 6.). East Lansing: The Education Policy Center at Michigan State University
Mintrom, M. (2001). Policy design for local innovation: The effects of competition in public schooling. State Politics and Policy Quarterly, 1, 343–363.
http://dx.doi.org/10.1177/153244000100100401

Monk, D. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. Economics of Education Review, 13 (2), 125–145.
http://dx.doi.org/10.1016/0272-7757%2894%2990003-5

Moore, M.H. (2003). Deadly lessons: Understanding lethal school violence—Case studies of School Violence Committee. Washington, DC: National Research Council School Violence Committee; Committee on Law and Justice; National Research Council; Board on Children, Youth, and Families; and Institute of Medicine.

Moss, P. (2007). Evidence and decision-making. Yearbook of the National Society for the Study of Education (Part 1). Hoboken, NJ: Wiley-Blackwell.

Murnane, R.J., & Phillips, B.R. (1981). Learning by doing, vintage, and selection: Three pieces of the puzzle relating teaching experience and teaching performance. Economics of Education Review, 1, 453–465. http://dx.doi.org/10.1016/0272-7757%2881%2990015-7

Myers, D., Peterson, P., Mayer, D., Chou, J., & Howell, W.G. (2000). School choice in New York City after two years: An evaluation of the school choice scholarships program (Interim Report No. MPR 8404-036). Princeton, NJ: Mathematica Policy Research.

Newman, K.S. (2004). Rampage: The social roots of school shootings. New York: Basic Books.
Noddings, N. (1995). A morally defensible mission for schools in the 21st century. Phi Delta Kappan, 76, 365–368.

Nye, B., Konstantopoulos, S., & Hedges, L.V. (2004). How large are teacher effects?Educational Evaluation and Policy Analysis, 26, 237–257. http://dx.doi.org/10.3102/01623737026003237

Peterson, P., & Howell, W.G. (2003). Efficiency, bias, and classification schemes: Estimating private-school impact (No. PEPG 03-01). Cambridge, MA: Harvard University, Program on Education Policy and Governance.

Plank, D.N., & Sykes, G. (2000). The school choice debate: Framing the issues. East Lansing: Michigan State University, Education Policy Center.

Plank, D.N., & Sykes, G. (2003). Lighting out for the territory: Charter schools and school reform strategy. In K.E.Bulkley & P.Wohlstetter (eds), Taking account of charter schools: What's happened and what's next. New York: Teachers College Press.

Popper, K.R. (1959). The logic of scientific discovery. New York: Basic Books.

Pressman, J.L., & Wildavsky, A.B. (1984). Implementation: How great expectations in Washington are dashed in Oakland—Or, why it's amazing that federal programs work at all, this being a saga of the Economic Development Administration as told by two sympathetic observers who seek to build morals on a foundation of ruined hopes (3rd ed.). Berkeley: University of California Press.

Rivkin, S.G., Hanushek, E.A., & Kain, J.F. (2005). Teachers, schools, and academic achievement. Econometrica, 73, 417–458. http://dx.doi.org/10.1111/j.1468-0262.2005.00584.x

Rockoff, J.E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. American Economic Review, 94, 247–252. http://dx.doi.org/10.1257/0002828041302244

Rogosa, D. (1999). Accuracy of individual scores expressed in percentile ranks: Classical test theory calculations (No. R305B60002). Washington, DC: National Center for Research on Evaluation, Standards, and Student Testing.
Rogosa, D. (2001). Shoe shopping and the reliability coefficient. Educational Assessment, 7, 255–258. http://dx.doi.org/10.1207/S15326977EA0704_01

Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. Quarterly Journal of Economics, 125, 175–214. http://dx.doi.org/10.1162/qjec.2010.125.1.175

Rouse, C. (1998). Private school vouchers and student achievement: Evidence from the Milwaukee choice program. Quarterly Journal of Economics, 113, 553–602. http://dx.doi.org/10.1162/003355398555685

Sanders, W.L., & Horn, S.P. (1998). Research findings from the Tennessee value-added assessment system (TVAAS) database: Implications for educational evaluation and research. Journal of Personnel Evaluation in Education, 12, 247–256. http://dx.doi.org/10.1023/A:1008067210518

Schmidt, W.H., McKnight, C.C., Houang, R.T., Wang, H., Wiley, D.E., Cogan, L.S., et al. (2001). Why schools matter: A cross-national comparison of curriculum and learning. San Francisco: Jossey-Bass.

Smith, P.K. (2003). Violence in schools: The response in Europe. London: Routledge Falmer.
Spillane, J.P. (2000). Cognition and policy implementation: District policy makers and the reform of mathematics education. Cognition and Instruction, 18 (2), 141–179. http://dx.doi.org/10.1207/S1532690XCI1802_01

Summers, A.A., & Wolfe, B.L. (1977). Do schools make a difference?American Economic Review, 67, 639–652.

Vergari, S. (2009). Charter school policy issues and research questions. In W.G.Sykes, B.Schneider, & D.N.Plank (eds.), Handbook of education policy research (pp. 478–490). New York: Routledge.

Webster, W.J., & Menro, R.L. (1997). The Dallas value-added accountability system. In J.Millman (ed.), Grading teachers, grading schools: Is student achievement a valid evaluation measure? (pp. 81–99). Thousand Oaks, CA: Corwin.

Weiss, C.H. (1977). Research for policy's sake: The enlightenment function of social research. Policy Analysis, 3, 531–545.

Weiss, C.H. (1995). The four “I's” of school reform: How interests, ideology, information, and institution affect teachers and principals. Harvard Educational Review, 65, 571–592.

Wenglinsky, H. (2000). How teaching matters: Bringing the classroom back into discussions of teacher quality. Princeton, NJ: Educational Testing Service.

Witte, J.F. (1998). The Milwaukee voucher experiment. Educational Evaluation and Policy Analysis, 20, 229–251.

Witte, J.F., Weimer, D.L., Shober, A.F., & Schlomer, P.A. (2007). The Performance of charter schools in Wisconsin. Journal of Policy Analysis and Management, 26, 557–573. http://dx.doi.org/10.1002/pam.20265

Woods, P. (1994). The conditions for teacher development. In P.P.Grimmett & J.Neufeld (eds.), Teacher development and the struggle for authenticity: Professional growth and restructuring in the context of change (pp. 83–100). New York: Columbia University, Teachers College Press.