Teaching, Rather Than Teachers, As a Path Toward Improving Classroom Instruction

James Hiebert and Anne K. Morris

Abstract
For several historical and cultural reasons, the United States has long pursued a strategy of improving teaching by improving teachers. The rarely questioned logic underlying this choice says that by improving the right characteristics of teachers, they will teach more effectively. The authors expose the assumptions on which this logic is built, propose an alternative approach to improving teaching that engages teachers (and researchers) directly in the work of improving teaching, present some indirect evidence to support this approach, and examine the cultural traditions and beliefs that have kept the conventional approach in place for so long.

Keywords
school/teacher effectiveness, teacher education/development, teacher learning

The thesis of this article is that working directly on improving teaching—the methods used to interact with students about content—is the most productive option for improving classroom instruction. To support the thesis, we identify three ways that educators and policy makers have worked to improve teaching: (a) recruit more talented people, (b) improve the qualifications of teachers entering or already in the profession, and (c) improve the instructional methods that are implemented in classrooms. These approaches are not mutually exclusive. Indeed, the combination of all three might be optimal. But our goal in this article is to highlight the value of the third approach because it is often overshadowed in policy discussions by the first approach and treated by educators as a by-product of the second approach. We argue that these priorities should be reversed.

Why the Familiar Should Look Strange
The history of U.S. education is filled with efforts to improve schools and classrooms (Tyack & Cuban, 1995). What approaches have been tried to improve the quality of classroom instruction? The U.S. approach frequently has focused on improving the quality of the teachers (Kennedy, 2010). This approach includes recruiting people with the right characteristics (e.g., distinguished academic records, strong content knowledge, high motivation, desirable personality traits; e.g., teachforamerica.org), training preservice and inservice teachers to acquire these characteristics, and removing teachers who are presumed not to exhibit them as evidenced by their poor performance (usually measured by students’ achievement scores; Klein et al., 2010). As will be seen later, the United States persists with this approach even though all, of the problems we identify. We then contrast this updated second approach with directly studying and improving the methods of teaching. Our description includes a potpourri of inconclusive evidence that we believe supports increasing the prominence of this third approach. Finally, we examine a series of reasons that explain why attention to the people who teach, despite the evidence, remains the U.S. preferred strategy for improving teaching.

Corresponding Author:
James Hiebert, School of Education, University of Delaware, Newark, DE 19716, USA
Email: hiebert@udel.edu
the data linking any of the characteristics mentioned above with students’ learning are weak, and the historical record shows no lasting impact of this approach (Cuban, 1993; Hoetker & Ahlbrand, 1969).

What makes the persistence of focusing on teachers as the path to improving teaching especially strange is the logic that spawned its existence in the first place. As described by David Cohen (2010), the decentralized educational system in the United States discouraged the adoption of a common curriculum.

Absent a common curriculum, teachers could not learn how to teach it, let alone how to teach it well. Hence, teacher education consisted of efforts to teach teachers to teach no particular curriculum. This was very strange, since to teach is always to teach something, but the governance structure of U.S. education forbad the specification of what that something would be. Teacher education was accommodating: teachers would be taught how to teach no particular version of their subjects. (Cohen, 2010, p. 377)

How does one train teachers to teach a generic curriculum? The preferred strategy was to help teachers acquire general pedagogical skills (e.g., managing classroom discipline) and to instill in teachers the characteristics mentioned earlier (e.g., strong content knowledge, desirable personality traits). These characteristics “were criteria not of teaching quality but of teacher quality; the tacit assumption was that the latter was a proxy for the former” (Cohen, 2010, p. 380). Later examples will illustrate how pervasive the assumption remains that teacher quality translates into teaching quality.

An equally strange assumption underlying the belief that teaching can be improved by improving the quality of teachers is that knowledge for teaching should be held in the heads of individual people rather than in artifacts. Artifacts, or knowledge products, survive individuals and can be shared and improved over time (Bereteir, 2002; Morris & Hiebert, 2011). Because of these advantages, professions that work to improve their performance over time and on a wide scale find ways to preserve their constantly updated knowledge in artifacts (Argyris & Schön, 1996; Gawande, 2009; Morris & Hiebert, 2009). But the U.S. education system has chosen a different path by investing in people and mostly ignoring the option of building instructional products that are improved over time. Assuming that people are the best vehicle for holding knowledge leads to Dewey’s (1929) lament about the U.S. educational system—teachers take their best ideas with them when they retire. Everyone needs to start over when they enter the profession.

The third approach—engaging teachers and researchers in improving the methods of teaching—has been proposed for some time (e.g., Dewey, 1929; Gallimore, Ermeling, Saunders, & Goldenberg, 2009; Gravemeijer & van Eerde, 2009; Lampert & Graziani, 2009; Richards, 1938; Rothkopf, 2010; Stigler & Thompson, 2009; Wilson & Daviss, 1994) but has never really gained serious traction in the United States. Rather than (a) assume that people can acquire the characteristics educators associate with good teachers and then translate these into improved methods and (b) assume that teachers should hold all the knowledge for teaching, this approach asks teachers (and researchers) to engage directly in the work of improving teaching methods and to preserve the knowledge they acquire in artifacts or products that can be shared and continuously improved. To set the stage for describing what this approach might look like in practice, we first review recent and promising adaptations of the second approach to improve teaching.

Training Teachers to Acquire Teaching Skills—Revisited

Recently, a number of researchers (Ball & Forzani, 2009; Ball, Sleep, Boerst, & Bass, 2009; Grossman, Hammerness, & McDonald, 2009; Grossman & McDonald, 2008; Lampert, Beasley, Ghoussinei, Kazemi, & Franke, 2010) have argued that the previously reported weak relationships between teacher characteristics and teacher effectiveness (e.g., Buddin & Zamarro, 2009; Goe, 2007; Solomon, Bigler, Hanushek, Shulman, & Walberg, 2004) are due to the distance between the characteristics measured and the actual work of teaching. If one examines up close the nature of teachers’ practice, it is possible, claim these researchers, to identify instructional routines that could be taught to teachers and that significantly affect the quality of teaching and the subsequent learning of students. These instructional routines could replace some of the content of traditional teacher preparation and professional development programs and become measurable characteristics that predict teachers’ effectiveness in the classroom.

Examples of high-impact instructional routines include posing problems, providing explanations, responding to student thinking, and leading class discussions (Ball et al., 2009; Grossman & McDonald, 2008). Common to these routines is that they occur with frequency, are central to the practice of teaching, facilitate the learning of all students, and are learnable by novices in the laboratory-type settings in teacher preparation programs (Ball et al., 2009; Grossman et al., 2009). That these routines, or teacher moves, are “approximations of practice” (Grossman et al., 2009) means they are closer to classroom teaching than the skills and knowledge embedded in more conventional characteristics (e.g., general content knowledge) and thus more likely to be used during actual teaching.

This updated approach to training teachers identifies more promising characteristics and could complement the approach we will highlight. But, we want to draw the contrasts between these two approaches as sharply as possible to raise the visibility of the third approach. So, we note that this more carefully developed second approach still faces the two fundamental problems identified earlier. The approach still
treats people, rather than artifacts, as the best repositories of knowledge, and it still requires the transfer of acquired characteristics (in this case, instructional routines) into real classroom settings. The second of these problems, acknowledged by the researchers pursuing it, can be examined as two subproblems: (a) identifying the instructional routines that link most closely to students’ learning and (b) teaching the routines in laboratory settings and then hoping each teacher can transfer the routines to actual teaching situations and use them flexibly and appropriately.

Solving the first subproblem—choosing the right instructional routines—is enabled by the analyses of teachers’ practices. And there is emerging empirical support linking teacher knowledge and behaviors closer to classroom practice with students’ learning (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009; Hill, Rowan, & Ball, 2005). But the eventual selection of the right routines depends on a richer and more reliable knowledge base for teaching than is currently available.

The second subproblem—transferring complex skills learned in one setting to another—is likely to be a harder nut to crack. According to the researchers pursuing this approach (e.g., Grossman & McDonald, 2008), it requires teacher educators to decompose the more complex routines into parts and then requires each teacher to study and rehearse the parts, recombine them into meaningful wholes, and acquire a conceptual knowledge of these routines that enables their users to adapt them to local contexts (i.e., subject matter, learning goals, students’ achievement levels, etc.). Transferring knowledge from one setting to another is a notoriously challenging problem for people (National Research Council, 1999). Transferring professional training from laboratory to the field is especially difficult (March, 2002). March (2002) notes several professional training problems that are “endemic to personnel systems.” One of these is that training for a professional role in a laboratory setting rarely captures the real activity. Reviewing other professions, as well as teaching, March says,

The generality of the phenomenon [poor transfer of professional skills from the laboratory to the field] suggests that it is not simply a problem with training but also stems from some fundamental difficulties both in communicating general knowledge and integrating that knowledge with contextual knowledge. (p. xiii)

The third approach we favor deals with these longstanding problems by beginning at a different place, with a different set of assumptions.

An Alternative: Improving Teaching by Improving Teaching

What does it mean to improve teaching by working directly on the methods used to teach? Of the variety of forms this work could take, we favor a form that can resolve, to some degree, the two problems identified earlier: store knowledge in a way that can be shared among teachers and passed along to support continuous, lasting improvement; and, acquire skills of teaching in the contexts in which they will be used to minimize the transfer problem. We believe both these resolutions are enabled by the development of artifacts or instructional products: public, shareable materials for teaching that can accumulate improvements over time as researchers and teachers create them, try them out, and feed the information from the trials back into the products, both in the form of actual changes to the products and rationales for the changes that constitute a growing base of knowledge for teaching (Morris & Hiebert, 2011).

In our opinion, two kinds of instructional products are especially useful: specially annotated lesson plans and common assessments. Annotated lesson plans can store the knowledge acquired for improved methods of teaching, and common assessments can ensure that changes to lessons are improvements, not just changes. Before elaborating on the value of these products, we want to acknowledge that other instructional products have been described and are compatible with those we emphasize.

Some researchers engaged in identifying instructional routines that can be taught to teachers (see above) have also reported on the development of instructional products (parenthetically, this shows that working within one approach does not preclude working within the other). For example, Lampert and Graziani (2009) describe the study and improvement of teaching in a language school in Italy as focusing on the development of “instructional activities” that are used to help students achieve particular learning goals. These highly structured activities get refined as instructors use them and information is collected on their effectiveness. Ball and colleagues (2009) describe their work in developing and continuously improving a curriculum for a mathematics methods course for elementary-preservice teachers. The curriculum consists of lesson plans, electronic slides, course assignments, assessments, and summary course guides. Work on these materials focuses attention on the teaching involved in these courses rather than on the instructors who teach them.

A question that begs to be answered at this point is, “What about the development of commercial curricula? Is this not the development of instructional products and does it not fit within the third approach?” Our answer is, “Yes, in principle, the writing of curricula could yield instructional products of the kind we have in mind. But, in practice, it rarely does.” The reason is that curriculum development in the United States is usually conceived as an enterprise that produces a final product rather than a product that invites continual testing and revising.1 Pilot tests are often run on new versions of textbooks, and feedback received on current versions might be used to inform future versions, but the process assumes that a final version will be ready for publication at some point and that this version will be used, as is, for as long as...
it is commercially viable. As we describe in the following section, the instructional products that we have in mind are subjected to continuing implementation studies and are never seen as final versions.

**Fore grounding Two Products: Lesson Plans and Assessments**

Building and using annotated lesson plans and common assessments. Annotated lesson plans contain knowledge of two kinds—what to do and why/how to do it that way. “What to do” offers prescriptions that teachers can implement to help students achieve the specified learning goals; “why/how to do it that way” provides a rationale or local theory for why the prescription might work along with information that teachers will likely need to understand and implement the plan as described.

More precisely, the annotated lesson plans we have in mind contain the following features. First, the learning goals for the lesson are stated as explicitly and completely as possible. The more explicit and precise the learning goals, the more clearly they guide the selection and implementation of instructional activities and the easier it is to assess whether the activities are helping students achieve the goals. Second, the rationales for key instructional moves are presented so teachers understand the reasons for the instructional decisions and can adapt them to local settings without changing the core aims of the lesson. Third, the learning goals, rationales, and instructional activities are described in enough detail that teachers can implement them as intended. Fourth, students’ likely responses to instructional tasks and questions are predicted to allow teachers to plan how to use students’ thinking during the lesson. Suggestions for the teacher are provided. Finally, information is presented to help teachers implement the lesson. This information moves beyond rationales and includes things like background information on the key concepts, helpful hints to prevent common difficulties for the teacher or the students, and markers in the lesson where particular kinds of explanations will be especially beneficial.

It is not hard to see how each feature of a lesson plan can improve with each implementation. Information can be gathered on the completeness and clarity of the learning goals and of the rationales, on the effectiveness of the instructional activities, on students’ responses, and on the aspects of the lessons for which instructors need more assistance. Those who revise lessons can take advantage of this information to elaborate and refine each feature. This yields continuously improving lessons and increasingly useful knowledge that can guide teachers’ implementations of them.

Information on the effectiveness of lessons is most useful if implementations in different classrooms can be compared. This allows teachers and researchers to judge whether different outcomes are due to different contextual variables, implementation of different lessons, or different implementations of the same lesson. Common assessments enable these kinds of comparisons. Indeed, without common assessments, keyed to common learning goals, there is no way to relate teaching to learning and sort out the lessons and implementations that are most effective (Bryk, Gomez, & Grunow, 2010).

Providing this array of information on lesson effectiveness requires assessments to measure the nature of lesson implementation as well as the nature of student outcomes (Stigler, 2010). Measuring implementation not only informs teachers and researchers about potential causes for different student outcomes but also provides information that can help identify difficulties with lessons and suggest possible fixes. Some of these fixes might improve the instructional activities for students, and some might provide information for teachers that would help them implement the lesson more effectively.

In addition to building increasingly effective lessons through cycles of testing and refining, teaching will improve only if the lessons are actually implemented in classrooms that share the same learning goals. This requires, at the least, an educational system (e.g., even just a school district) that shares the same learning goals for students. Shared learning goals create a demand for lessons that help students achieve these learning goals most effectively. As Elmore (1996) notes, it is the absence of demand for good ideas that prevents widespread improvement in the system. If teachers hold the same learning goals for students, this demand rises in a natural way. Driven by the demand for the best lessons for specific learning goals, one can imagine that teachers (and researchers) are motivated to study, test, improve, and share lessons that show increasingly good results. Teachers who wish to help students achieve the learning goals more effectively are naturally inclined to replace their lessons with ones that have been shown to work better. In this way, Al Shanker’s (1997) dream could become a reality: The wisdom of the profession becomes its standard practice.

**Teachers’ skills as a by-product of testing and improving instructional products.** Improvement of teaching comes not only from the continuously improving instructional products but also from improving the skills of teachers to use these products. Earlier, we noted that the second approach to improving teaching—improving the qualifications of teachers—assumes that improved teaching methods will result from improved knowledge and skills of individual teachers. We suggested that this relationship could be reversed: Improved skills for teaching can result from studying, testing, and revising instructional methods (as specified in the instructional products). We now elaborate this hypothesis.

As teachers (and researchers) work together to study and improve annotated lesson plans from which they will teach, they have the opportunity to acquire new knowledge of content and pedagogy (Huang & Bao, 2006; Lampert & Graziani, 2009; Lewis & Tsuchida, 1998; Ma, 1999). An important feature of this knowledge is that it is developed in the exact context in which it will be used. By studying particular
lessons, teachers can acquire (and are likely to feel some urgency to acquire) knowledge and skills that help them implement just those lessons. There is no distance between the context in which teachers’ knowledge and skills for teaching are developed and the context in which they are used. The transfer problem is greatly reduced.

Based on the reports of teachers engaged in studying and improving annotated lesson plans—one form of which is called “lesson study”—this work also results in deeper understandings of teaching, in general. Japanese teachers comment about the general benefit of lesson study to their teaching of all lessons (Lewis & Tsuchida, 1998). Capturing the sentiments of many, one Japanese teacher said, “Through these experiences, I believe that my teaching method has improved” (Fernandez, 2002, p. 395). We interpret these comments to mean that a more general payoff for teachers includes the opportunity to induce instructional routines or core practices for teaching. We refer here to the instructional routines described earlier that, within the second approach for improving teaching, are targeted for rehearsal, sometimes apart from particular lessons in which they play a central role.

Consider the example of leading a class discussion, an instructional routine that could appear in multiple lessons. In our view, the routine would best be learned by enacting it as a part of lessons designed to support productive class discussions. Such lessons would include tasks that lend themselves to rich discussions, questions that help begin and sustain such discussions, predicted student contributions to the discussions, clearly stated aims for such discussions related to the lesson learning goals, and so on. Learning the routine of leading class discussions would include learning how the routine fits within the surrounding features and how these features support its success. Applying the routine to new lessons would require detecting, or introducing, similar features so the routine would work as well in the new lessons as in the old.

Our argument for keeping instructional routines linked with surrounding, facilitative features is based on the claim that teaching is a system of interacting elements (Stigler & Hiebert, 1999). Studying and improving lessons does not leave the linking or integration of individual features to chance; it requires that the essential features work together to create lessons that are the best teachers know for helping students achieve particular learning goals. In this way, generalized instructional routines are not isolated for practice and then introduced into real classrooms but, rather, are always part of lessons in which they link with surrounding features that, together, help students achieve the learning goals.

In our view, these general pedagogical practices are better learned as they arise repeatedly in specific lessons because we believe the two subproblems related to transfer, and voiced earlier, are resolved naturally. First, which are the right instructional routines? The instructional routines that emerge as more general when studying and improving individual lessons do so because they are shown to be effective in many different lessons. This provides the best kind of evidence for their selection.

The second subproblem gets to the heart of the transfer issue: How can novice teachers learn to transfer skills acquired in one setting to another? In the third approach, the conditions shift so that the challenge changes from transferring an isolated instructional routine from the laboratory to the classroom, to repeated application of an instructional routine in lessons designed to support such a routine. We conjecture that when instructional routines are learned in the context of studying, improving, and implementing particular lessons, they are learned as skills or routines embedded in lessons. Then, as the routine is found to appear in many lessons, novice teachers can induce that the routine is a more general skill that might play a useful role in lessons not yet taught. In fact, we expect that novice teachers could introduce, appropriately, the routine in lessons where it did not (yet) appear because they would have already acquired an understanding of how it plays out in lessons to help students achieve the lesson’s learning goals, how it interacts with other features of the lesson to work effectively, and so on.

What Is the Evidence?

Whether it is more productive to focus on teachers or teaching has not been addressed in anything like an empirical test, so the evidence regarding the effectiveness of one approach or the other is indirect, at best. Still, we believe several findings are relevant and justify seriously considering the third approach that has frequently been ignored—improving teaching by improving the methods of teaching.

Persistence of Teaching Methods Over the Past Century

The most common U.S. approach to improving teaching—improving teachers—has had little effect on changes in classroom practice. The persistence of teaching methods is well documented (Cuban, 1993; Fey, 1979; Hoetker & Ahlbrand, 1969; Jacobs et al., 2006). The core of teaching—the way teachers and students interact about content—has remained the same for a century or more. If recruiting better teachers or preparing them more effectively is the answer, it surely would have had an effect by now. Branson and Grow (1987) argue that the model of improving individual teachers has reached maturity.

All obvious gains were made long ago. Without a fundamental change in the operating concept, no additional improvement is likely. Trying to increase the quality of the education system by somehow improving the quality of teachers without, at the same time, implementing an improved model of operations is an approach doomed to failure. (p. 422)
Focusing on the development and refinement of instructional products is, in our view, an “improved model of operations.”

**Changing Teachers in Hopes of Improving Teaching**

Is there any direct evidence that changing teachers improves teaching? It is clear that teachers make a big difference in how well students learn (Konstantopoulos & Chung, 2011; Nye, Konstantopoulos, & Hedges, 2004; Sanders & Rivers, 1996; Sanders, Saxton, & Horn, 1997). But these findings are not conclusive with respect to the teachers versus teaching issue. The findings could be interpreted to mean that effective teachers share similar characteristics and these characteristics should be identified and used to hire or train other teachers. Or, the findings could be interpreted to mean that some teachers are more effective than others because they use better methods, and it is the methods that should be captured in instructional products so other teachers could use, test, and refine them. In other words, these data do not say whether it is the individuals or the instructional products they use that make the difference.

One large case study addresses the teacher versus teaching question in an unintended but revealing way (Silva, 2008). Chattanooga, Tennessee, was concerned about the achievement of students in eight of its inner city schools. In 2001, the Benwood Foundation donated US$5 million to improve the schools. Over the next 5 years, the achievement of students steadily increased until, in 2006, their performance levels surpassed the district average. The dramatic improvement was noticed and applauded in the public media and the halls of Congress. How did the schools improve? The most popular explanation and the one most widely reported was the dramatic release of all teachers in these schools in 2001. Changing the teachers was part of the intended solution. But, in fact, most of the teachers were hired back. The underreported part of the story is that when the teachers returned, they were asked what they needed to teach more effectively. They answered, in part, that they needed more opportunities to work together and prepare more effective lessons. The money was used to help support teachers’ requests. In the end, it is reasonable to interpret the cause for the steady increases in learning as improvements in teaching rather than as changes in teachers.

**Findings from Large-Scale School Reforms**

Many school districts have undertaken reforms to improve students’ learning albeit in a less dramatic fashion than the Chattanooga case just recited. Correnti and Rowan (2007) and Rowan and Miller (2007) studied the implementation and outcomes of three different models for school improvement: Success for All (SFA), America’s Choice (AC), and Accelerated Schools (AS). The first two are curriculum-based models that provide schools with instructional materials and detailed descriptions of the instructional practices that should be used to implement them. The third aims to change the culture of schools by supporting teachers and leaders to collaborate in developing their own improvement strategies. Results show that the first two models, SFA and AC, yielded greater changes in instructional practices and significantly higher student achievement, especially in areas targeted by the instructional products and practices. Correnti and Rowan (2007) and Rowan and Miller (2007) conclude that changes in instructional practices—in teaching methods—were largely responsible for the gains in student achievement. Although the data are only suggestive, they do indicate that well-designed instructional products that are used as intended can improve students’ learning across large numbers of classrooms.

**Lessons from Japan and China**

The system of lesson study in Japan is seen by many as largely responsible for the high-quality teaching in Grades 1 to 8 Japanese classrooms (Lewis & Tsuchida, 1997; Shimahara & Sakai, 1995; Stigler & Hiebert, 1999). If this is true, it provides considerable support for the approach to improving teaching that we describe because lesson study focuses directly on developing and refining annotated lesson plans (Lewis, 2002; Wang-Iverson & Yoshida, 2005) using common assessments to gather relevant information (Stigler, 2010). In other words, by focusing directly on the continuous improvement of the instructional products we described earlier, teaching has become of sufficiently high quality to attract international attention. In addition, those who are familiar with lesson study note that by engaging in studying and improving lessons, teachers acquire more general skills, which they apply throughout their teaching (Lewis, 2002; Stigler & Hiebert, 1999; Wang-Iverson & Yoshida, 2005). That is, teachers do induce core practices or instructional routines by working directly on creating and improving individual lessons.

In China, school teachers have, for years, worked to improve their teaching by collaboratively studying teaching materials, planning lessons, and observing each other teach (Ma, 1999; Paine, 1990, 1992). With the release of the National Mathematics Curriculum Standards (NMCS) in 2001, more systematic strategies for improving teaching have emerged. Chief among these is the “Keli” model (Huang & Bao, 2006). Keli, or “Exemplary Lesson Development,” engages teachers and experts (often university researchers) in developing a lesson that illustrates the innovative teaching methods prescribed by the NMCS. In addition to producing a model lesson that other teachers can use, teachers who work on these lessons by developing or adapting them are encouraged to reflect on the gap between their conventional methods and
the new methods described in the lesson. In this way, teachers acquire new methods of teaching in the context of particular lessons rather than being taught these methods separately and then asked to apply them. The process has shown sufficient success that it is now promoted by the Chinese government.

Why Has the United States Continued to Focus on Teachers Rather Than on Teaching? Deeply Ingrained in U.S. Culture

Stigler and Hiebert (1999) argue that teaching is a cultural activity. Teachers learn to teach by growing up in the culture, by observing their teachers teach. Similarly, teacher education is a cultural activity. Its origin as training teachers in generic skills has become the accepted practice, and the assumption that teacher quality is synonymous with teaching quality is more deeply ingrained in U.S. culture than one might imagine. Several examples reveal its pervasiveness.

In a highly publicized government report titled “Before It’s Too Late: A Report to the Nation from the National Commission on Mathematics and Science Teaching for the 21st Century,” the national blue-ribbon panel argued that improving the quality of U.S. classroom instruction is the key to improving students’ learning (National Commission on Mathematics and Science Teaching for the 21st Century, 2000). The recommendations, however, dealt almost entirely with improving the characteristics, training, and resources of teachers. Nothing was said about how one might improve teaching.

In a similar example confusing teaching with teachers, Education Week’s 2008 annual report on the quality of American schools was titled “Tapping Into Teaching.” One of the lead articles was “Taking Teaching Quality Seriously.” With these banners, and an introduction that claimed good teaching is a key to good learning, one might expect an analysis of how schools can improve the quality of teaching in their classrooms. Instead, the analysis focused only on improving the characteristics of teachers: improve training opportunities, recruit more qualified individuals, retain the best qualified people, and so on. The belief that good teaching is achieved by getting the right people into the classroom is so ingrained in the U.S. educational culture that teaching and teachers are treated interchangeably.

Because little distinction is made between teachers and teaching, many believe that the right teachers can magically improve students’ learning and that the wrong teachers are a major reason for perceived failures of U.S. education. Malcolm Gladwell’s (2008) analysis in The New Yorker likens finding the right teachers to finding the right professional quarterbacks. Assuming that hidden talent is the key, Gladwell buys into and reinforces the belief that good teaching is tied to innate characteristics of the individual. If society could only identify the talented teachers and entice them into the classroom, they would teach using better methods and produce higher student achievement.

The opposite, and darker, side of the coin is the growing assumption that teaching would improve if the poor teachers could be exposed and weeded out. Joel Klein, chancellor of the New York City public school system; Michelle Rhee, former Superintendent of Washington, DC, schools; and 14 other superintendents published a “manifesto” in the Washington Post (Klein et al., 2010) claiming that the obstacles to removing incompetent teachers made their job of improving the education for their district’s children very difficult. One could just as easily believe that it is the poor teaching methods that are to blame and it is these poor methods that must be weeded out and replaced by better ones. But that is not the direction U.S. education has taken.

Based on Widely Shared Beliefs

The cultural tenacity of the conventional approach to improving teaching in spite of a suspect logic and little empirical evidence is reinforced by deep-seated beliefs that align with this approach. We can identify several.

Teaching is inherently uncertain. A common belief about teaching is that teaching is so complex and requires so many spontaneous decisions that it is inherently uncertain (Sawyer, 2004; Schön, 1988; van Manen, 1995). This belief calls into question the goal of creating instructional products that could be implemented by different teachers in different classrooms with predictable results. The belief also challenges the second (updated) approach to improving teachers’ generic skills. As Ball and Forzani (2009) acknowledge, “A second problem in identifying the core tasks of teaching rests with the dominant contemporary view of teaching as highly improvisational and wholly context dependent” (p. 503).

The consequences of believing that teaching is uncertain gain strength when combined with the negative reaction to previous attempts to “script” teaching (Sawyer, 2004; Schön, 1988). Scripting teaching has become interpreted as taking away teachers’ intellectual and pedagogical capabilities and turning them into robots. It has become an incendiary phrase associated with de-professionalizing and de-skilling teachers. In contrast, we believe, with others, that asking teachers to design, study, and improve detailed annotated lesson plans that can be shared with other teachers and taught as scripts (appropriately interpreted) is one way of professionalizing teaching (Hiebert, Gallimore, & Stigler, 2002; Morris & Hiebert, 2011; Stigler & Hiebert, 1999). It places teachers in the position of working in the domain they know best—teaching students content—and it allows them to take ownership of improving their profession’s performance.

Teaching is not researchable. A corollary to the belief that teaching is inherently uncertain and must be improvised for each setting is a belief that teaching defies systematic analysis. Teaching is too complex and too dependent on local
conditions, says this belief, to ever succumb to research efforts aimed toward, for example, creating empirically based instructional products that could guide improved teaching across multiple classrooms. A manifestation of this belief often casts teaching as an art rather than a science (Eisner, 1979; Gage, 1978).

Although “research on teaching still lacks powerful ways of parsing teaching that provide us with the analytic tools to describe, analyze, and improve teaching” (Grossman & McDonald, 2008, p. 185), our proposal assumes that this is not because teaching defies systematic analysis but rather because of the conditions under which research on teaching has been conducted. An essential aspect of conducting research on any phenomenon is using consistent assessments to measure a common enactment of the phenomenon. Returning to Cohen’s (2010) historical analysis, the absence of a common curriculum and common assessments in the United States has made it nearly impossible to measure teaching. Teaching is designed to teach something, and if this something is not standard, then it is difficult to measure teaching in any kind of standard way. We believe that this measurement problem, more than any inherent quality of teaching itself, has contributed to the current limitations of research findings.

The measurement problem also helps to explain the persistent focus on teachers rather than on teaching. Teacher characteristics are relatively easy to measure. In fact, the more generic the characteristic, the easier it is to measure. For example, it is easy to measure the number of subject matter courses that teachers have taken during their preparation. It is likely that ease of measurement has contributed to the belief that it is more productive to build improvement strategies on improving the characteristics of teachers than on improving the quality of teaching.

Teachers are victims of the “fundamental attribution error.” As described by Kennedy (2010), the fundamental attribution error, a well-known construct in social psychology, overestimates the influence of enduring personal characteristics on behavior and underestimates the influence of situational factors. In teaching, this can mean that if a teacher misses a chance to point out a critical relationship between poverty and unrest leading to a country’s civil war, or misunderstands a student’s response and offers a confusing reply, or makes a mathematical error when factoring a quadratic expression, educators are quick to attribute these deficiencies to enduring teacher characteristics (e.g., poor pedagogical or content knowledge) rather than to situational factors, such as the lack of planning time, pressure to cover material quickly, or momentary distractions. This attribution error reinforces the belief that teachers’ characteristics determine their classroom behavior and, consequently, the preferred strategy for improving teaching is improving these characteristics.

Final Thought: Not an Easier Problem, But a More Productive One

In our view, the beliefs that might explain why educators in the United States have focused on teachers and often avoided trying to improve teaching by improving teaching directly do not justify a continuation of this practice. They simply indicate that creating a system in which teachers and researchers work directly on improving the methods of teaching will be challenging.

We do not believe that making progress in improving teaching by building and continuously refining instructional products is easier than training teachers to acquire skills and knowledge to use in the classroom. Changing cultures presents enormous challenges, and this alternative would require a change in culture. But, we do believe that working directly on improving teaching is a more productive approach because as parts of this problem get solved, teaching necessarily improves.

At least, we believe that the second approach—training teachers to acquire the core practices of teaching—would benefit significantly from placing the training directly into the context of studying and improving instructional products. The easiest way to do this would be to situate teacher education in the context of learning to teach a particular curriculum designed to help students achieve learning goals valued by the stakeholders. Training would involve testing and refining the lesson plans specified by the curriculum, highlighting those instructional routines that appeared in multiple lessons. No longer needing to train teachers to acquire these routines disembodied from their use, this approach could now situate the practice of these routines directly within the lessons that already contain them.

In the end, we believe there is compelling logic and suggestive empirical evidence for bringing the third approach into much greater prominence in the United States as educators search for ways to improve the core of education—the classroom interactions between teachers and students around content.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Preparation of this article was supported, in part, by the National Science Foundation, Grant 0083429 to the Mid-Atlantic Center for Teaching and Learning Mathematics. The opinions expressed in the article are those of the authors and not necessarily those of the National Science Foundation.
Note

1. Commercially produced curricula can, of course, be treated by local teacher/researcher groups as subject to continuous testing and improvement. That is exactly what is reported by the El Barrio–Hunter College PDS Partnership Writing Collective (2009).

2. Recent events indicate that the culture might be changing in ways that would support the direct study and improvement of teaching (e.g., Donovan, Wigdor, & Snow, 2003). Boosting the possibility of national efforts to study teaching directly, a bipartisan group of 75 leaders in education, business, and government recently called for the development of a common school curriculum that all states could adopt (the Common Core State Standards; Dillon, 2011; Gewertz, 2011a). Because the absence of a common curriculum has undermined efforts to measure and improve effective teaching across teachers and classrooms, adoption of a common curriculum would remove this significant obstacle. The level of detail will make a big difference (the more detail the better), but the calls signal the beginnings of important changes in the U.S. educational culture. Of course, there already is pushback against the call for a common curriculum (Gewertz, 2011b).

References


**About the Authors**

James Hiebert is the Robert J. Barkley professor of education at the University of Delaware, where he teaches in programs of teacher preparation, professional development, and doctoral studies. His research interests focus on mathematics teaching and learning in classrooms, and the mathematics preparation of teachers.

Anne K. Morris is professor of education at the University of Delaware, where she teaches mathematics content courses for prospective teachers and doctoral courses in mathematics education. Her research interests focus on prospective teachers’ skills for studying and improving teaching, and on K-16 students’ understanding of mathematical proof.