The Problem of Improving Teacher Quality While Balancing Supply and Demand

MARY M. KENNEDY

INTRODUCTION

At the same time improvements in education contribute to improvements in the social and economic well-being of the country, these social and economic improvements place further demands on the education system. Consequently, the history of education is in part a history of continuing efforts to improve both the quantity and quality of education provided to American youth. The reform effort we face today could be interpreted as one further step in this continuing effort.

Yet, while similar to previous reform efforts in its broad purposes, this reform also carries its own unique features, and these need to be taken into account. The problem of supply and demand of qualified teachers, for instance, is one that regularly confronts U.S. policy makers, but it takes on a special cast in the current reform in two different ways. First, the concerns about supply that we face today are not a function of an increase in the number of students we want to serve, as has often been the case in the past, but instead are a function of changes in the attractiveness of a teaching career relative to other careers. Moreover, problems of supply are uneven. In some regions and some subject areas, we have an oversupply, while in other regions and other subjects, we face an undersupply. Second, the problem of teacher quality is not one of increasing the required credential, as has often been the case in the past. Virtually all teachers now hold a bachelor’s degree, and a great majority hold a master’s degree as well. Yet we are still not satisfied with the quality of classroom teaching practices.
Before examining policy strategies of the current reform, I want to examine each of these unique problems.

**The Contemporary Version of Teacher Supply and Demand**

With the exception of the past two or three decades, education has almost always faced a problem of teacher shortages, in large part because the population of students needing an education was continually growing. Shortages in the past have largely been filled by women and minorities who had few other career opportunities (Sedlak and Schlossman, 1986). However, the past two decades have seen an increase in the career opportunities for these populations that have traditionally filled teaching positions. Thus, the supply problem we now face derives more from competition from other professions than it does from an increase in the actual number of teachers needed. This fact has led policy makers and scholars to attend more than they have in the past to such issues as the attractiveness of the profession as a whole, the extrinsic and intrinsic rewards of teaching, and the quality of teachers’ working conditions.

At the same time, there has been an increase in the amount and quality of data available to assist in policy formulation. Longitudinal data bases have increased our understanding of the career paths of teachers, giving us more information about the length of time teachers tend to stay in teaching positions and more information about the source of new teachers. We now know, for instance, that many people who are prepared to teach never enter the profession, that many of those who do teach leave the profession after only a few years, and that only about 25 to 30 percent of leavers are likely to return after a career interruption. These returning teachers are often referred to as a reserve pool of teacher supply that augments the pool of new college graduates available for teaching. And we know that attrition rates are higher in chemistry and physics than in other subject areas (Murnane et al., 1988, 1989). Finally, we know that the reasons for leaving are often not the reasons presumed by policy makers. For instance, even though young women are more likely to leave than other groups, they are not necessarily leaving to rear children, and they are not necessarily leaving schools that are considered to present more difficult teaching assignments (Heyns, 1988). Most important, we now know that the most important predictor of teacher demand is teacher attrition (Haggstrom et al., 1988), and that the most unknown factor in predicting teacher supply is the size and character of the reserve pool (Gilford and Tenenbaum, 1990).

Since schools have a fixed clientele and must provide services to all those who are eligible for them, they must necessarily fill every classroom position with someone. Before teacher unions began negotiating class size, school districts often adjusted their services by expanding class sizes. As
this strategy becomes less feasible, districts have resorted to other options, in particular assigning teachers to subjects they are not certified to teach or hiring new teachers who do not have the formal credentials required. States tacitly approve of these strategies by granting emergency credentials for these teachers. Thus, even though most states have steadily increased their requirements for teaching credentials, they have, at the same time, continually permitted teachers who do not have the requisite credentials to enter the profession through the back door, so to speak. These practices have become criticized recently, both because they threaten the quality of classroom practices and because they diminish the status of the profession as a whole.

Finally, much of the evidence currently available regarding teacher supply and demand suggests that the issue is not a universal one, but instead the ratio of supply to demand varies by school, subject matter, and grade level. That is, urban schools have more difficulty filling positions than suburban schools do, secondary schools have more difficulty filling physics positions than math positions, and so forth. That such patterns exist suggests that perhaps projections of teacher supply and demand would be more fruitful if they included particular school and job characteristics in their models.

**The Contemporary Version of High-Quality Classroom Practice**

Contemporary reform demands also differ from past demands in that they are based more heavily on empirical examinations of teaching and the outcomes of teaching than they have been in the past. Findings of recent research have contributed heavily to contemporary dissatisfaction with classroom practice and have seriously complicated the problem of maintaining an adequate supply of teachers. Let us review some of the evidence.

First, national assessments in virtually every subject indicate that, although U.S. students can perform basic skills pretty well, they are not doing well on thinking and reasoning. U.S. students can compute but they cannot reason through complex mathematical problems (McKnight et al., 1987). They can write complete and correct sentences but cannot prepare arguments (Applebee et al., 1990). They cannot reason through scientific problems very well (International Association for the Evaluation of Educational Achievement, 1988). Moreover, in international comparisons, U.S. students are falling behind not only students in European countries, but also those in many Third World countries as well, particularly in those areas that require higher-order thinking, problem solving, or conceptual work.

The second research finding gives us a clue as to why we have the first: textbooks in this country typically give no attention to big ideas, offer no analysis, and pose no challenging questions. Instead, they provide a tre-
mendous array of information or “factlets” and pose questions that require nothing more than the ability to recite back the same empty list of facts. In fact, our textbooks often don’t even provide much in the way of organization or coherence for these facts (Tyson and Woodward, 1989). So whatever real understanding students get about the subjects, whatever intellectual challenge they get, must come from their teachers.

The third finding from research is that most teachers teach content only for exposure, not for understanding (Porter, 1989). That is, their aim is not to ensure that students really understand the concepts they present, but rather only that they have been exposed to them. And the fourth finding is that teachers tend to avoid thought-provoking work and activities and stick to predictable routines. Why? Because students are easier to manage, and student outcomes easier to control, when the tasks are routine (Doyle, 1983, 1986; Doyle and Carter, 1984). So if we were to describe our current K-12 education system, we would have to say it provides very little intellectually stimulating work for students and that it produces students who are not capable of intellectual work. These problems are not, of course, all due to bad teaching. I have already pointed out that U.S. textbooks often don’t provide intellectually defensible material, and I should add now that many features of school organization and school policy can yield these results as well. Still, these findings about what happens in U.S. classrooms help explain the findings from national and international assessments, and they demonstrate why policy makers care about the quality of classroom teaching practices.

Yet classroom practice is highly influenced by teachers’ a priori ideas about what counts as good teaching. Teachers acquire seemingly indelible imprints of teaching from their own experiences as students, and these imprints are tremendously difficult to shake. The dominant impulse in any new teacher is to imitate the behaviors of his or her own former teachers. And this impulse will remain strong unless teachers are offered an equally strong and compelling alternative approach to teaching and are provided with a great deal of help learning to adopt this new approach.

And this is the fifth relevant finding from research: people who teach are highly likely to teach in the way they themselves were taught (Haberman, 1985; Lortie, 1975; Nemser, 1983). All of us, whether we choose to enter teaching or not, learn about teaching throughout our lives. From kindergarten through twelfth grade, we observe teachers. Those of us who go on to college observe even more teachers, and these teachers are not necessarily any better or different that those we observed when we were younger (McDiarmid, 1990). By the time we receive our bachelor’s degree, we have observed teachers for over 3,000 days. For those of us who choose to enter teaching, the pervasiveness of this “apprenticeship of observation,” both across grade levels and across subject areas, coupled with the sheer volume of time spent
observing, yields a deeply entrenched and tacit set of beliefs about what can and should happen in schools: about the nature of intellectual work and the nature of school subjects, about the teachers’ role in facilitating learning, and about the pedagogical implications of student diversity.

So, for instance, if your elementary school teacher presented mathematics to you as a set of procedural rules with no substantive rationale, you are likely to think that this is what mathematics is and that this is how mathematics should be studied. And you are likely to teach it this way. If you studied writing as a set of grammar rules rather than as a way to organize your thoughts and to communicate ideas to others, then this is what you will think writing is and this is how you will think it should be taught and learned.

Moreover, these beliefs are highly likely to be accompanied by an emotional commitment. People who choose teaching as a career not only have ingrained beliefs about what the enterprise is all about, but also have chosen it as a career precisely because this is what it is all about. They are not committed to the enterprise in the abstract; they are committed to it as they understand it.

POLICY EFFORTS TO IMPROVE THE SUPPLY OF QUALITY TEACHERS

The special problem facing contemporary policy makers is that of, at one and the same time, increasing the supply of teachers available and improving the quality of their classroom teaching practices. Certain features of the teaching population bear notice when considering either supply and demand or quality. One is that the teaching force in the public sector consists of some 2.3 million teachers. Any efforts to improve the quality of this large teaching force will be labor-intensive and therefore expensive. Second, some 15 percent of U.S. adults now have at least a baccalaureate degree, as do almost all teachers. And some 20 percent of all baccalaureate degree holders received their degrees in education, even though less than half of those who received such degrees are currently teaching (National Center for Education Statistics, 1989). Moreover, the number of college-educated adults is rising, so that the fraction is much higher among younger adults who are likely to be enrolling children in schools in the next decade or two.

The size of the college-educated population has two implications for policy makers concerned about balancing supply and demand and improving quality. First, as consumers of education, these segments of population-bachelor’s degrees in general or education degrees in particular-form an increasing fraction of education’s clientele. And clients of education have more direct knowledge of teachers’ practices than clients of other
professions have. Parents see their children’s homework and often visit
classrooms and they form judgments about quality from these observations.
To the extent that their own education levels lead them to hold greater
expectations for quality in teaching, or greater demands for teaching of the
subjects they studied in college, we can expect public demands for improve-
ments in teacher quality to continue or to increase over time, rather than to
subside. Indeed, we expect that questions about the quality of teaching
will necessarily continue to exist, given the educational levels of parents
and given their direct knowledge of their children’s teachers.

A second implication of these facts is that this entire population consti-
tutes a potential reserve pool that could be recruited into teaching. Indeed,
the current teaching force represents only 43 percent of the those who re-
ceived degrees in education, and only 10 percent of those who have re-
ceived degrees in all fields. If one assumes that even those who did not
receive degrees in education could still teach, the reserve pool expands to
include over 25 million adults.

Since actual teaching practices cannot be efficiently monitored, policy
makers tend to work toward improving the supply of other aspects of teach-
ers or teaching that are presumed to lead to improvements in practice. For
purposes of this discussion I would like to distinguish five dimensions of
quality that form the basis of most policy activities. The first two dimen-
sions, quality of credential and tested ability, refer to characteristics of
individual teachers: policy makers want to ensure that those who enter
teaching hold credentials that are appropriate to their task, that they have
acquired the subject matter knowledge and pedagogical training needed, or
that they have some minimal level of intellectual ability deemed appropriate
for teaching. The third dimension, quality of demographic representation,
refers to the mix of teachers working in individual schools: policy makers
want a population of teachers that roughly matches the racial and ethnic
backgrounds represented in the student body. And as the composition of the
student body changes, the demand for different types of teachers will also
change. The fourth and fifth dimensions refer to the character of practices
that occur within individual schools. The fourth dimension, quality of pro-
fessionalism in schools, refers to the extent to which teachers are given real
responsibility for their work and the extent to which they are able to make
sound professional decisions in their work.\textsuperscript{1} Many policy makers want to
change the character of teaching practice so that teachers have the kind of
independent responsibility for managing their work that other professionals
have, believing either that sound educational practices cannot be dictated
from afar or that teaching will be a more attractive line of work if teachers
have more professional autonomy. Finally, the fifth dimension is the qual-
ity of classroom teaching practice: even if policy makers find ways to
bring people into teaching who have the proper credentials, the right demo-
graphic characteristics, and higher tested ability, they need assurance that the actual classroom teaching practices of these teachers matches some level of acceptable practice.

Although nearly all policy makers recognize nearly all of these dimensions of quality, many assume the fifth dimension will be solved automatically by solving one or a combination of the first four. Indeed, much of the current policy debate in the area of teacher supply and demand has to do with which of the first four dimensions of quality is most likely to ensure the fifth. This is not to say, however, that the fifth dimension is the only one that matters. Arguments have been made for the independent value of most of the other dimensions, even if they do not necessarily yield improvements in classroom practice. The ethnic mix of teachers, for instance, may have positive benefits on both student motivation and learning that will not be easily discerned from observations of teachers’ classroom practices, and changes in the professionalism of teaching may alter teachers’ regard for their work in ways that influence both their retention in the job and the tacit messages they communicate to students about the importance of school and learning to a variety of life’s endeavors. The problem for policy makers is one of sorting out these several dimensions and recognizing each as a separate issue that warrants its own attention.

In this paper, I argue that these several dimensions of quality are not the same, and that improving quality on any one dimension will not necessarily improve quality on any of the others. The argument has two implications for policy. First, we cannot assume that we can improve classroom practice, the dimension of most importance to most policy makers, simply by improving any of the other dimensions. And second, to the extent that policy makers wish to alter any of the other dimensions of quality, they need to address them separately.

The remainder of this paper is divided into five sections, one addressing each of these five dimensions of quality. In each main section, I address three questions: how important this dimension of quality is to contemporary education policy, what policy activities are currently under way or under consideration to increase supply and quality within this dimension, and what are the measurement and data collection implications of these issues and activities for generating useful policy information about supply and demand within this dimension of quality.

**SUPPLY OF QUALITY CREDENTIALS**

Much of the contemporary work on measuring teacher supply and demand defines quality in terms of the teachers’ credentials. That is, a teacher is qualified to teach if he or she is certified in the area in which he or she is teaching. So if, for instance, a history major is certified to teach history
and political science, but not sociology, that teacher is considered *qualified* when teaching history or political science but not when teaching sociology. If such a teacher is found teaching all three of these subjects—a likely event in a small high school—it is difficult to know whether to count the teacher as qualified or not. The confusing array of assignments, misassignments, and joint assignments, combined with the variety of combinations of certification offered by different states and taken by different teachers, make it extremely difficult to define the supply or the demand of credential-qualified teachers.

**How Important are Credentials?**

One argument for emphasizing teaching credentials has to do with their symbolic function. If teachers are to maintain a professional standing, there should be some indication that they are entitled to teach, and formal credentials can play an important role in entitling teachers as professionals. Yet the empirical evidence of a relationship between formal credentials and student achievement is remarkably weak. Several aspects of teachers’ formal education have been presumed important indicators of teacher quality. Three popular measures are: (a) whether the teacher is formally credentialed to teach the subject or grade being taught, (b) whether the teacher majored in the subject being taught, and (c) whether the teacher has taken courses beyond the bachelor’s degree.

Researchers who have tried to connect these measures of teachers’ education and certification to gains in student achievement have been unable to find any clear or meaningful relationship (Gilford and Tenenbaum, 1990; Hanushek, 1986). Particularly discouraging is research on teachers’ subject matter majors, which has found very little relationship between subject matter major and teaching ability (e.g., Begle and Geeslin, 1972). Even college majors who have reasonably high grade points in their majors are often unable to explain important concepts from their disciplines or to illustrate them accurately (see, e.g., Ball, 1990). The lack of convincing evidence has motivated Richard Murnane and others (Murnane et al., 1991) to argue for abandoning state program requirements that guide teachers’ credentials, so that the field can experiment with a wider array of ideas for preparing teachers to teach. If there is no evidence that these program requirements yield better teachers, these authors argue, there is no reason to constrain teacher education programs.

On the other hand, the lack of association between credentials and student achievement gains may be an artifact of the wide range of educational backgrounds that can lead to a credential. Since each state defines its own credentialing system, the possession of a credential can mean that a teacher has either a bachelor’s or a master’s degree, depending on the state; that a
teacher has taken an exam or not, depending on the state; that a teacher has participated in an induction program or not; that the teacher’s classroom practices have been assessed or not; that the teacher has majored in an academic subject or not; or that a teacher has participated in a university-based preservice program or an alternative route. Given this variation, it should not be surprising that the presence of a credential per se does not correlate strongly with measures of student achievement growth.

**Policy Activity in This Area**

Typically, a formal credential means that a teacher has received both the requisite courses in teacher education and the requisite courses from the arts and sciences departments. Some colleges and universities—notably those that are members of Project 30—are working on courses that combine these two sides of teacher knowledge into integrated courses. However, the structure of most colleges and universities is such that these two sides of teachers’ formal course work exist in a zero-sum relationship, such that additional courses in pedagogy are perceived to mean fewer courses in the subject areas. This means that, when policy makers want to improve the quality of the initial credential, they must choose which kind of courses they want to increase and which they will decrease.

The current trend is to reduce the emphasis on pedagogy courses and to increase the emphasis on subject matter courses, although policy makers and others are far from agreeing on what the exact nature of these subject matter courses should be, particularly at the elementary school level. Three states—Virginia, Texas, and New Jersey—have recently made such changes in their teacher education requirements. In addition, the Southern Regional Educational Board has recommended that Southern states try to reduce the number of credits teacher candidates take in their professional sequence.

These policies, however, manipulate only the supply of new graduates who might enter teaching. They do not address the large reserve pool of teachers, from whom many new hires are drawn. I should mention here, as an aside, that most states also have a mechanism for recognizing credentials from other states. That is, if a teacher moves from state A to state B, the teacher can apply to the state to have his or her original credential recognized. Though many states have such procedures in place, the criteria used to accept credentials from other states are often not synchronized with current policies for new graduates. In addition, states rarely keep track of the number of applicants for out-of-state credential recognition, nor do they know how many such applicants have been turned away. These immigrant teachers themselves could constitute a substantial reserve pool of teachers if their credentials were accepted. Moreover, despite policy makers’ espoused
interest in strengthening teachers’ credentials, few states actually monitor teaching assignments. They often do not know how many teachers, or which teachers, actually have the required credentials to teach in their assigned area. The nature of education is such that schools have a constant supply of students that must be taught; since they cannot turn away students when they have teaching vacancies, they always fill all their vacancies. Many states give emergency credentials to teachers who are hired to work in areas for which they do not have formal credentials, and some, such as Pennsylvania, maintain records on the number and location of such emergency-credentialed teachers. But advocates of improved credentialing argue that, as long as states permit easy entrance through the back door, improved requirements for entrance through the front door are virtually useless.

A related policy activity, often designed to influence both supply and quality, is the alternative route into teaching. Thirty-three states now have provisions for some form of alternative route to certification (Feistritzer, 1990). What this means is that the state drops its normal requirements for obtaining a teaching credential and permits certain individuals to enter teaching through some alternative route. Although it may seem counterintuitive to argue that providing shortcuts into the profession would enhance teacher’s credentials, states offering alternative routes generally hope to attract into teaching people who have stronger subject matter knowledge and stronger liberal arts educations. Their reasoning is that such people were, as college students, either deterred from teaching by its curriculum requirements, or attracted to higher salaries in other career lines. In either case, they may have second thoughts later in life. So the state offers a way to become certified without having to go back to college and take a lot of teacher education courses (Adelman, 1986; Carey et al., 1988).

Many alternative routes are based on the same assumption as those policies that limit the number of courses candidates can take in education, namely that the quality of teaching will improve if teachers have more courses in academic subjects. Others are based on the assumption that such ad hoc training programs are an improvement over the emergency credentialing approach to filling new vacancies. In New Jersey, for instance, as more teachers are hired through alternative certification, the number of teachers hired on emergency credentials has dropped to nearly zero.

We are starting to get some evidence now about who enters teaching through alternative routes. Many of the people who acquire certification through these sanctioned loopholes have been teaching already in private schools, in which certification is not needed, and are now seeking more lucrative positions in public schools. Others are new college graduates with roughly comparable grades and academic backgrounds as graduates who are already certified. Even those who move into teaching from other career
lines are moving from lower-status technical and clerical jobs more often than from higher-status business and science jobs (see Darling-Hammond et al., 1989, for a summary of 64 such programs).

Finally, one of the most controversial recent policy moves with respect to formal credentials was made by the National Board of Professional Teaching Standards (NBPTS), which decided not to require state credentials for teachers who wish to take the NBPTS and become board-certified as master teachers. The board’s reasoning was based in part on the lack of evidence that such initial credentials were relevant to good teaching and in part on the recognition that their task was to directly assess the quality of applicants anyway, so that they did not, in principle, need any other indicators of quality to guide their decisions. The decision has been criticized by others who argue that credentials provide an important symbolic function in the formation of a profession and that they are important for that reason, even if empirical evidence of their value is missing.

At the same time these various policy initiatives are reducing the requirements for formal course work in pedagogy, the Holmes Group, a consortium of deans of research universities, has been pressing for a more demanding credential for teachers. They argue that the road to quality in teaching is through a master’s degree, for this degree enables candidates to obtain the volume of subject matter they need and to still obtain the volume of pedagogical knowledge they need. In addition, they argue that schools need to be restructured to differentiate between teaching staff who are highly qualified and those who are really intending to teach for a few years and then move on to other careers (The Holmes Group, 1986). Their efforts have been criticized by some educators on the grounds that there is little demand in school districts for the more expensive master’s degree candidates and that there is little cost-benefit advantage for the teacher candidates, given the current salary structure of schools, to invest in a master’s degree in teaching (e.g., Hawley, 1986). In fact, however, close to half of all teachers now hold a master’s degree anyway, obtaining it within their first few years of teaching.

In sum, then, states have been quite active in manipulating the educational requirements associated with teachers’ initial credentials, and this activity has been geared largely toward improving the subject matter side of teachers’ undergraduate course work. At the same time, they have initiated quite different policies regarding the reserve pool of teachers. Most states readily provide emergency credentials for misassignments, few monitor the appropriateness of teachers’ assignments, and many promote alternative routes into teaching as a means of increasing the number of teachers who have specifically not obtained formal credentials.
Implications for Data Collection

We already have a relatively strong empirical basis for doubting the significance of formal credentials per se to teaching practice. My first recommendation is that we do not pursue this line of inquiry any further. Coding teachers on the subject they majored in, or on whether they have the correct credential for their teaching area, provides too little knowledge of the actual courses teachers took and of what they actually learned in these courses. Even transcript studies are unlikely to reveal the detail we need about teachers’ formal education, for courses with similar titles can be remarkably different from one another. If researchers wish to pursue further the relationship between credentials or formal course work and teaching practices, they need to invest in more carefully designed studies, studies that follow candidates over time and that examine the actual content they are taught rather than the number of credits they have taken.

Still, there are symbolic reasons for caring about teachers’ credentials. Given the increasing level of education among parents, the demand for highly educated teachers, and for teachers who have majored in the subjects they teach, will increase. Two data collection strategies would be helpful here. First, states should be encouraged to improve the quality of data they maintain on the subject matter certifications of all their teachers, and in particular the certifications of those teachers who are teaching with emergency or provisional credentials and the certifications of those who have applied to have out-of-state certificates recognized. It would be useful to know, for instance, how many out-of-state credentials were rejected and why, and how many secondary teachers are teaching subjects for which they were certified to teach. And since many secondary teachers teach more than one subject, it would be useful to know how many different subjects such teachers are currently teaching and the proportion of their teaching load that is in the subject(s) for which they are certified and how many are misassigned. More detailed data on the relationship between teaching credentials and teaching assignments can help states determine the need for, for instance, targeted alternative route programs to fill particular shortage areas, and they could help them better determine the relationship between initial credentials and likely teaching assignments.

Second, the range of educational packages that can lead to a credential is worth documenting in itself. One way that national data bases might facilitate state policy makers would be by documenting the variety of educational backgrounds that certified teachers have, so that state policy makers can gain a better understanding of their own requirements in relation to others. That is, what college curricula are required to become certified to teach secondary math, for instance, in each of the states, and how different are these requirements from state to state? If these data were available,
together with data on the number of teachers teaching out of field and the
number of emergency credentialed teachers teaching in each area, we may
begin to get a handle on the relationship between college course-taking
requirements, actual teaching assignments, and teaching vacancies.

**SUPPLY OF TEACHERS WITH HIGH TESTED ABILITY**

There is evidence that those who enter teacher education, generally
speaking, score lower on tests of academic achievement than those who
enter other career tracks. Moreover, there is evidence that, among those
who become certified, those with lower scores are more likely to take teach-
ing positions. And finally, there is evidence that those who continue in the
profession tend to have lower academic achievement scores than those who
leave after a few years. Throughout the pipeline, then, we have created a
system that systematically favors the lower scoring of all possible teaching
candidates (Kerr, 1983; Murnane et al., 1989; Schlechty and Vance, 1983;
Robertson et al. 1983).

Tested ability is the dimension of quality that most policy makers as-
sume is related to classroom teaching practice, and consequently it has
received more attention from policy makers than any of the other dimen-
sions of quality. They fear that, by permitting low-achieving adults to
teach, we are creating a bad situation. The statistics are reminiscent of the
old George Bernard Shaw line, “Those who can, do; those who can’t, teach.”

**How Important is Tested Ability?**

The seriousness of the tested ability problem is hard to estimate, for to
some extent, we do decide, as a society, how to allocate our talent. It is also
complicated by the sheer size of the teaching population. Our society now
employs about 10 percent of its college-educated women and 4 percent of
college-educated men as teachers; no other profession draws such a large
proportion of educated adults (Lanier, 1986). One implication is that if we
were so successful in solving this problem that all, or even most, of our best
and brightest went into teaching, we would soon be worrying about a simi-
lar problem in business, law, medicine, and several other fields as well.
The size of the teaching force is simply too large to expect it to be filled
entirely from the uppermost ranges of tested ability.

To gauge how serious the tested ability problem is in education, we
would need to decide the level and range of tested abilities we are willing to
tolerate, given the tradeoffs. We would need to be honest about how many
high-scoring people we will want to bring into teaching when moving them
into teaching means moving them out of business, computer science, medi-
cine, law, and so forth. How many high-scoring teachers do we need, and how low we are willing to permit the bottom end of the achievement score distribution to go?

Even aside from the problem of allocating talent among different occupations, we face a difficult problem in arguing that tested ability is relevant to the tasks of teaching. Tests measure a particular kind of intellectual ability, one that some have argued is far too narrow. The achievement tests that are typically used to assess teachers really measure only their ability to choose the right answer when several possible answers are provided for them. Obtaining a high score on such a test does not necessarily mean a teacher could solve such problems if the right answer were not available there for them to recognize. Neither does it mean they could solve more complicated or open-ended problems for which no answer is clearly best. Nor does it mean they can explain to someone else how to solve such problems, as teachers must do. And, finally, it doesn’t mean they can create problems for someone else to learn from, as teachers must do. Consequently, we don’t know whether moderate-scoring adults actually teach less well than higher-scoring adults, nor do we know whether moderate-scoring adults have a less adequate understanding of the material they would actually teach in elementary or secondary classrooms.

I don’t mean to suggest that tests don’t measure anything useful. Obviously they do. We know that differences in test scores are relevant to a number of academic pursuits, and there is some evidence that scores on verbal ability are relevant to teaching. Nor do I mean to suggest that we can get away with a population of teachers who come from the bottom of the test score distribution. What I am saying, though, is that we don’t know how low a score has to be before it makes a difference to teaching.

The only aspect of tested ability that has been found to relate to teaching is general verbal ability. Several researchers have found that tests of verbal ability are associated with teaching success. It seems reasonable to suppose that general verbal ability would be relevant to teaching. After all, teaching is largely an art of communication. However, even this finding has been contested, and the evidence is far from clear.

So it is hard to say how serious the tested ability problem really is. It is conceivable that, if we took into account the other occupations that need to draw high-scoring individuals, and if we knew more about the relationship between test scores and teaching performance, we would conclude that the problem is not so grave as we have been supposing. It is also conceivable that tests that were structured to target more specifically on teaching valid content, rather than recognizing it when it is presented, would demonstrate a stronger relationship to classroom practice.
Policy Activity in This Area

The most popular approach to enhancing the tested ability of teachers is to build some form of test into state requirements for certification. Virtually every state now includes some form of teacher assessment in its certification requirements (Council of Chief State School Officers, 1988). These assessments vary considerably from state to state but, regardless of their particular features, most are intended to improve the tested ability of the teaching population.

Some of these assessments may accomplish that goal, but most probably won’t. In order for a state assessment to really raise the tested ability of teachers entering the work force, it would have to satisfy three conditions. First, the test would have to include the kind of intellectually rigorous content we think teachers need to be able to handle. If the content of these tests rests mainly on the ability to recall simple facts, as much of the teacher assessment content does, the test will not raise anyone’s confidence in the tested ability of teachers. Second, the test would have to be designed so that a lot of people fail it. If no one fails the test, then it can’t, by definition, be raising the tested ability of new teachers. Yet many state teacher assessments are designed to ensure very low fail rates. And finally, the test would have to pass legal criteria for job relevance, for as soon as very many people fail an entrance test, the state will face a lawsuit challenging the relevance of both the content and the cutoff score to the tasks of teaching. So far, no state has been able to demonstrate job relevance. And to avoid these lawsuits, most states design their assessments so that very few teachers will fail. That is, they adopt relatively low cutoff scores and they provide ample opportunities for those who fail to try again. Consequently, no state assessment system really contributes to solving the tested ability problem. In the end, these programs serve more of a symbolic than a real function.

Even if such testing programs could be shown to improve the tested ability of new teachers, they would make only a small impact on this dimension of quality, since most of them affect only new college graduates. The existing population of teachers are usually exempted from such testing requirements—and so too presumably are those certified teachers in the reserve pool.

A second popular strategy for improving tested ability is the alternative route. Just as policy makers believe that college students with strong interest in the liberal arts might be repelled from vocationally oriented teacher education programs, many also believe that students with higher tested ability tend to avoid the presumably less rigorous teacher education curriculum. However, the evidence suggests that the new recruits who enter teaching
through alternative routes do not differ substantially from conventional teacher education graduates in their tested ability.\(^6\)

One surprising finding, however, is that alternative routes are attracting a more diverse population into teaching. Some alternative routes, for instance, are bringing more males and minorities into teaching than traditional teacher education programs do (Darling-Hammond et al., 1989). So even if this solution doesn’t work well for improving tested ability, it may help improve demographic representation. This is one of the reasons why I emphasize the need to think about each dimension of quality separately.

My main reason for insisting on thinking about these dimensions separately, though, is that policy makers often assume that, if they improve tested ability, they also automatically improve classroom teaching practice. There are several reasons for doubting this assumption. First, evidence for a relationship between tested ability and teaching practice is remarkably scarce. Second, the dominant method for attracting higher-scoring people into teaching (i.e., alternative route programs) is rarely designed to challenge the conventional wisdom that teachers arrive with; instead they are designed to capitalize on this “wisdom.” That is, these programs try to recruit teachers who are smarter than traditional teacher education graduates and give them only brief pedagogical training, presumably on the assumption that these smart future teachers can use their native intelligence to figure out how to teach. One result of this is that alternative route candidates might be even more likely than fully prepared teachers to emulate the practices they observed as children and less likely to develop new approaches to teaching. They receive virtually no preparation that would lead them to rethink the assumptions they bring with them. So even if alternative routes do raise the tested ability of teachers, or bring more minority teachers into teaching, they are not likely to improve teaching practice significantly.

Finally, high tested ability may actually impede a teachers’ ability to teach. Have you ever heard someone say: “He is so smart, I don’t understand a word he says.” Surely that is not the kind of high-ability person we want in our classrooms. Knowing a lot, or being smart, is not enough for teaching, for the main job of the teacher is to get someone else to know a lot and to be smart.

Two policy efforts are currently under way to devise better tests for teachers. One is sponsored by the National Board of Professional Teaching Standards (NBPTS), an organization that eventually hopes to provide experienced, high-quality teachers with a special credential that certifies them as experts in their area. The other is that of the Educational Testing Service, which is developing a new version of its widely used National Teacher Examination (NTE) in the hope of providing a more valid assessment of beginning teachers. Both organizations are working to move beyond the multiple-choice format into items that require teachers to demonstrate how
their knowledge of subject matter and of pedagogy actually influences their teaching decisions and teaching practices.

The research we have been doing at the National Center for Research on Teacher Education (NCRTE) offers some potentially useful new directions for learning about teachers’ and prospective teachers’ subject matter knowledge. We have relied heavily on interviews in the belief that teacher knowledge must be sufficiently explicit (i.e., not tacit), that they can articulate it. This is, after all, their task as teachers. So we ask them questions such as these:

- Suppose a student asks you whether to use *is* or *are* in the following sentence: “None of the books *---* in the library.”
  How would you respond and why?

- Imagine that you are teaching division with fractions. One of your students raises his hand and says, “I am very confused. This doesn’t make sense. The answer I am getting is greater than the numbers I start with! What am I doing wrong?” And he hands you one of the examples he has done:
  \[ 1 \ 3/4 \div \ 1/2 = 3 \ \frac{1}{2} \]
  What is going on here?

These questions have proved to be very difficult for our respondents to answer. English majors have almost as much trouble as elementary teacher candidates articulating the principles that should be used to select a verb for the first problem, and math majors have nearly as much difficulty explaining what is going on in the second problem. In fact, even math majors who have high grade point averages often provide erroneous story problems to illustrate this equation. Many of them generate story problems that divide by 2, rather than by 1/2.’ This should not be surprising; the content that these two problems address is not college-level content. Verb usage is usually addressed in high school and division with fractions is usually addressed in middle school. And most teacher candidates have not thought about these content issues since they first learned them themselves, when they were students in high school or middle school.

The significance of questions of this sort is more apparent when these questions are contrasted with state assessments that focus on the content teachers will teach. Although many states employ some form of subject matter knowledge tests, these tests do not assess teachers’ ability to explain subject matter, to generate illustrations of it, or to generate problems that will help students learn it. Instead, they tend to present teachers with multiple-choice questions that require no more than recall and that focus on less important aspects of the content. A question that asks, for instance, “Which of the following five authors wrote *Moby Dick*?” will not tell us much about subject matter knowledge that is relevant to teaching.
Implications for Data Collection

The question of the relationship between test scores and teaching performance is clearly one that needs further examination. But rather than continuing to conduct correlational studies that rely on existing tests, I would recommend that the greatest research effort be placed in the area of test development. At present, we really have no efficient, standardized way to test the pedagogical subject matter knowledge that is relevant to teaching. Interviewing teachers in the way my colleagues and I did at Michigan State would be far too costly a strategy to implement on a broad scale. The examples I describe, though, do provide some alternative ways of thinking about tested ability and about the kinds of questions we need to be asking teachers.

My second recommendation, though, is that this test development activity should not be undertaken by the National Center for Education Statistics (NCES), but instead should continue under the auspices of the NBPTS and the ETS, organizations whose primary reason for being is to develop such assessments, or through research centers such as the new Center for Research on Educational Accountability and Teacher Evaluation. This is not to say that there is no role for the NCES, however, for the validity of these alternative assessments still needs to be assessed. One useful contribution that could be made by the NCES would be to attach some items such as those devised by my research center, or those devised by ETS or the NBPTS, to its future National Assessment of Educational Progress (NAEP) studies, so that teachers’ responses to these items could be correlated to the skills and abilities demonstrated in students.

There is probably also little value in maintaining, at a national level, records on the scores of teachers who take state-specific tests, for the content and character of these tests vary widely from state to state. And, since only those teachers who pass the test are permitted to teach, the data would not be very useful for testing relationships between test scores and teacher performance. Moreover, since most state testing systems are relatively new, data will not be available for most practicing teachers. However, within particular states, there might be some value in tracking the relationship between test scores and teaching careers. For instance, are teachers in certain subject areas more likely to pass or to fail these tests, or are teachers who pass with high scores more likely to be hired in certain locations?

SUPPLY OF DEMOGRAPHICALLY REPRESENTATIVE TEACHERS

Whereas the first two dimensions of quality refer to characteristics of individual teachers, this third dimension refers to the characteristics of groups
of teachers. By now the demographic projections for the next several decades are well known. The student population is changing dramatically. In the near future, Hispanics will replace blacks as the dominant minority in this country; the total minority population will become a substantial portion of the total population; and in some states, Caucasians will become a minority group.

Yet, despite these changes in the student population, our teaching population is still largely white, suburban, middle class, and female. Our teaching force no longer represents the population at large and will represent it even less in the future unless we work actively to change it. By the mid-1980s, whites represented only 71 percent of the student body but 90 percent of the teaching force. Moreover, minority enrollment in teacher education programs decreased from 17 percent in 1980-81 to 10 percent in 1984-85. Several authors have argued that we need to recruit into teaching people who better represent the students being taught.

**How Important is Demographic Representation?**

Is demographic mix really a valid measure of quality? How serious a problem is it, really, if we have mainly white, working-class suburban women teaching high-income students, low-income students, urban students, rural students, whites, blacks, Hispanics, and a host of immigrants? After all, we have allowed, for decades, both girls and boys to be taught mainly by women. One could argue that it shouldn’t matter whether the teacher and student come from culturally similar or dissimilar backgrounds provided that the teacher can actively engage students with important content. In principle, you shouldn’t need to be the same sort of person as your students to have an impact on them.

Furthermore, as a practical matter, we can never completely match students and teachers by their demographic characteristics. Any given student is likely to encounter some 30-35 teachers by the time he or she has finished 13 years of school. Simply from the standpoint of probabilities, many, if not most, of these teachers will be demographically different from the student even if the overall population of teachers is perfectly representative of the student population. So even if we successfully meet all our goals regarding demographic representation in the teaching work force, we would still have numerous individual teaching situations in which teacher and student come from different demographic backgrounds.

Given all of this, one could argue that, if our main concern is that minority students be able to learn school material, we should worry as much or more about measuring the dimension of quality of classroom practice as we do about measuring the demography of the work force, for no matter
how representative the total population of teachers is, individual teachers will still need to serve students from a variety of backgrounds.  

Yet students are aware of the full population of teachers in their schools, not just the teachers they encounter in their own classrooms. And this population as a whole does much more than literally teach content. They also personify content. They stand as models for what it is like to be an educated person, to be a member of the community of scientists, writers, mathematicians, or political scientists. They also serve as ex-officio parents, guides, and mentors to young people. And if we want students to believe that they themselves might one day be scientists, writers, or mathematicians or that they might be mentors, guides, and educated people, then we need them to see diverse examples of such people, including at least one who looks like they, the students, look. Moreover, diversity among teachers may increase both the students’ and the teachers’ knowledge and understanding of different cultural groups, thereby enhancing the abilities of all involved to interact with different cultural groups. For students, such diversity among teachers may improve their future chances of interacting successfully with the different groups they will encounter in their own work. For teachers, such diversity may enhance their ability to interact successfully with their own students, thereby improving the quality of their classroom practices.  

This suggests, then, that even if better representation does not always, or even often, yield matched teachers and students, it is still important that we solve the representation problem, for students need role models.  

Policy Activity in this Area  

Demographic quality can be thought of as a “pipeline” problem. To get more minorities into teaching, policy makers want to help more minorities get into college in the first place. One popular solution for increasing minority enrollments in college is to offer financial assistance to minorities and low-income students. Financial assistance alone, of course, cannot ensure that any of its beneficiaries will actually enter teaching, for just as new employment opportunities have opened up for women, so have they opened up for minorities. But it can increase the pool of graduates available who might consider teaching as a career.  

Another idea, currently being considered by Congress, is to revive some form of the Teacher Corps program. The original Teacher Corps, initiated in the 1960s, provided financial assistance during college to encourage students to teach in low-income areas for some period of time after graduation.’ The idea is analogous to the Peace Corps: you don’t assume these teachers will remain in these schools forever, but you give them an incentive to work there at least for a short time. Some Teacher Corps programs
funded students while they were in college and encouraged them to go into teaching; some solicited liberal arts graduates and gave them graduate preparation in teaching. All sought candidates who came from low-income neighborhoods in the hope that they would return to those neighborhoods. Whether a new Teacher Corps program could accomplish this outcome, given the new job opportunities for minorities, is not clear.

Interestingly, very few policies move further back in the pipeline and strive to improve the quality of high school programs for these students. Yet if high school minority students aren’t receiving a good education in the first place, financial aid will be inadequate, for the most it can do is help them get into less prestigious colleges. It won’t make them eligible, or capable of responding to, more demanding colleges and universities. Many colleges and universities now have special remedial programs to assist their less qualified students, but their success rates often are not very good. If we really want to ensure a larger pool of qualified college students, one strategy we should attend more to is increasing the number of minority high school graduates who are qualified for college.

A significant weakness in all of these strategies for altering the demography of the teaching population is that they attend largely to new college graduates. To the extent that new teachers are actually hired from a reserve pool rather than from a pool of new graduates, these policies will not influence the demographic mix of teachers for many years. Indeed, the fact that the current teaching force is so short of minority teachers suggests that, if we were to redress this imbalance, we would need to be preparing far more than a proportionately representative group of new teachers. And, given the size of the teaching force and the continually changing student body, the problem of altering the demographic representation of teachers within schools is doubly challenging.

Implications for Data Collection

Of the five dimensions of quality that I have defined here, the demographic dimension is easiest to document. My first recommendation, therefore, is that any system for measuring supply and demand for teachers should include demographic variables as indicators of quality. Most universities already keep records of the gender, ethnicity, and language origin of their students, including those who are prospective teachers, and therefore could report the demographic characteristics of graduates who are credentialed to teach. Many school districts also keep such records both on their students and on their practicing teachers, and so could report on the makeup of their current teaching force and the similarities between their teaching populations and their student populations. And many NCES databases include data on demographic characteristics. These data can give us a good sense for the
match between the demographic characteristics of prospective and practicing teachers and characteristics of the population of students we project will fill our schools in the next several years.

Of course, demographic data can be made more complicated rather quickly if we also want to know how these different groups are dispersed across subject matter areas within the secondary level. Whether such fine-grained distinctions matter is not clear. We have known for years that most math and science teachers are males, and that female students tend to move away from these courses. However, we also have evidence to suggest that the reason for girls avoiding science courses may not be a lack of female role models in these particular content areas, but instead may be due to the teaching practices employed by these male teachers. If that is the case, it is difficult to know whether we really need representativeness within each field of study or whether, instead, we might be satisfied with representativeness among teachers as a whole, provided that we also address the improvement of teaching practices within particular subject areas.

My second recommendation is that predictions of the number of potential minority college graduates should become part of NCES’s routine projections. Because the supply of ethnically diverse teachers is a pipeline problem, policy makers need access to long-term projections in this area. To project the possible supply of minorities entering teaching, we will need first to estimate the size of the potential hiring pool, that is, the number of minority college graduates who have at least majored in teachable subjects, even if they did not obtain teaching credentials. Better still would be to step back yet another level and estimate the number of college-bound minorities as they graduate from high school. Although high school transcripts and Scholastic Aptitude Test scores are rarely used to project potential future teachers, data on the academic status of minorities while in high school could prove to be quite useful in estimating the total pool of minorities who are eligible for college and who therefore could become teachers if the appropriate incentives were provided. I am not sure what would be involved in gathering such data on minority students, but I think that some efforts in this area would be useful both in making projections and in helping policy makers recognize that the progress of minorities through college and into teaching does not begin once they are in college, but instead years earlier.

My third recommendation moves to the demand side of the equation. I think we should measure demand at two levels: the individual teacher level and the school composition level. At the individual teacher level, it would not be too difficult to portray the total number of students served from each racial/ethnic category and the total number of teachers from each group. More difficult, though, would be an index of the dispersion of these teachers that would indicate the relative overlap between teaching populations
and student populations. I do not propose a quota-system for hiring teachers, but I do suggest a system for monitoring schools to determine how many have attained various levels of demographic mix and how well their mixes of teachers match their mixes of students. Such a system would enable us to identify and locate those schools that are most or least successful in matching their teacher demography with their student demography and thereby providing role models for their students.

SUPPLY AS A FUNCTION OF THE QUALITY OF PROFESSIONAL LIVES OF TEACHERS

The fourth dimension of quality that is of interest to policy makers today is the quality of professional lives of teachers. After a period in the 1960s and 1970s when curricula were designed to be "teacher proof," and a period in the 1980s when policies were designed to dictate the teaching behaviors in classrooms, policy makers are now turning to the belief that teachers should have the same professional responsibilities, and associated autonomy, that professionals in other areas have. Arguments for "empowering" teachers stem in part from a recognition that teaching is too complicated and context-specific to regulate centrally, and in part from a recognition that such central regulation may be an important contributor to teacher attrition.

At the same time, there are those who would argue that the quality of teaching and of teachers is so low that we cannot possibly entrust them with more autonomy and responsibility, for they are not prepared to handle these responsibilities well. Thus, even though there is a tremendous amount of rhetoric within this reform about empowering teachers and converting teaching to a True Profession, there is, at the same time, some countervailing pressure against this rhetoric.

How Important is Quality of Professional Life?

Arguments for improving the quality of teachers' professional lives are numerous. Some people advocate professionalism simply because teachers deserve to be treated better. Compared with other college graduates, for instance, teachers lack such basic amenities as private office space, telephones, and access to photocopying machines. Others advocate professionalism as a means to stem attrition by making the job more attractive. They extend the argument from telephones and private space to professional control over work and the work environment. Finally, some advocate professionalism as an avenue toward improved quality of classroom practice. These people believe that teachers are more qualified than central administrators to make decisions that will improve classroom practice, and that teachers
should learn to examine and improve their own practices. The arguments are so diverse that it is difficult to assess the real potential of professionalism for improving either teaching quality or teacher supply. In addition, the arguments are still mainly at the stage of speculation, with few data available to test the merits of any of these propositions. At a minimum, we can separate their potential for reducing attrition from their potential for improving the quality of classroom practice.

With respect to professionalization as a means of improving the attractiveness of teaching as a career, we need to be more clear about what aspects of teaching make the job attractive in the first place, and what aspects of it are most responsible for decisions to leave it. Certain patterns of attrition could be interpreted as indicating that the quality of work life is not a major reason for leaving. For instance, young single women are more likely to leave teaching than are older women or men (Heyns, 1988; Murnane et al., 1988), a pattern that suggests the reasons for attrition may differ across subgroups of teachers. Moreover, leavers are not leaving those schools that might be considered “difficult” schools.

In fact, the pattern of entering and leaving teaching suggests that there is a subgroup of people are simply dabbling with teaching, who perhaps never really considered it as a long-term career option. Of those who were prepared to teach, 25 percent never entered the profession in the first place, a finding I take to mean that they were probably not seriously interested in teaching to begin with, but perhaps picked up a credential as a fallback career option. And those who take teaching jobs but leave them relatively soon tend to come from higher socioeconomic backgrounds, to have higher test scores, and to be teaching in private or suburban schools rather than urban schools. They are not leaving because they are less capable, nor because they have difficult teaching assignments. Perhaps they might not have taught at all if they couldn’t find these more comfortable teaching assignments to start with. I suspect that, for this group, teaching is more of an avocation than a career.

This casual approach to teaching may account for the unevenness of supply across schools and teaching positions. The supply of teachers available to a particular school, for instance, depends, among other things, on how far the school is from a teacher’s college,” the proportion of the student body which is nonwhite (and therefore more difficult to teach for the typical white teacher), and the safety of the school. Within schools, the supply available for particular teaching positions depends in part on the number of preparations required per day, whether the position is itinerant or includes a home room, and the proportion of classes taught outside the primary subject. These are aspects of the quality of work life may be more relevant to dabblers than aspects of professionalism that entail greater commitment to the task. If professionalizing the job requires more commitment
from teachers, it is possible that this population of dabblers may be discour-
geraged, rather than encouraged, to take teaching positions even for short peri-
ods of time.

With respect to professionalization as a means of improving the quality of
classroom practice, we also have reason to be skeptical. If it is true that
teachers teach as they were taught, then there is no particular reason to
believe that, left to their own devices, teachers would be very likely to
make radical changes in their own practices. Even if there were ample
opportunities for teachers to observe one another and to discuss and com-
pare their ideas about teaching, they may be more likely to reinforce tradi-
tional methods than to devise new ones. Houston’s (1988) description of
restructuring efforts in two schools illustrates this point: one group of teachers
tried to devise new structures for educating their children but were over-
whelmed with the difficulty of achieving the goals they had set for them-
selves, largely because their students were so diverse. They were unable to
teach the students in whole groups, did not have the time to individualize
their instruction, and did not have the wherewithall to devise new classroom
structures that could accommodate their students. Eventually, they limited
enrollments in this special program so that their student body was more
homogeneous, thereby defeating at least one important goal of restructuring,
that of improving education for the difficult-to-serve students.

One aspect of professionalism argues not merely to give teachers more
responsibility for their own practices, but to convert teachers to researchers
or self-examiners who can engage in critical analyses of their own prac-
tices, test new ideas, and improve their practices based on these examina-
tions. In this context, one often hears the term reflective practitioner. Whether
teachers can be rendered capable of such self-examination remains to be
seen, particularly given the current workloads and time constraints most of
them face. The only reflective practitioners I have met are working in
highly unusual situations, characterized in part by severely restricted teach-
ing loads and in part by extensive access to university colleagues. Unless
policy makers are willing to provide the finances necessary to make such
conditions commonplace, the concept of professionalism may never yield
more than rhetoric.

**Policy Activity Promoting Quality Professional Life**

Most policies enacted in the last decade have increased, rather than
decreased, regulation of teachers and teaching (Darling-Hammond, 1988;
Darling-Hammond and Berry, 1988; McDonnell, 1989). Ideas promoting
professionalism that have generated the most attention are school-based
management and restructured schools, career ladders, merit pay, and peer
evaluation.
School-based management programs are intended to release schools from strong district regulation and to encourage teachers to work together within each school to decide such matters as what areas they want to improve, what curricula they will use, how they want to allocate themselves and their students to classrooms, how they will coordinate their activities, and how they will evaluate their progress. While the idea is popular, and many districts are implementing some sort of program that presumably gives schools more autonomy, the degree to which these programs make a real difference in the quality of teachers' professional lives is highly variable. In some districts, teachers may simply be asked to fill out numerous forms listing their evaluation data, objectives, and plans. In others, teachers may be given resources such as released time or consultant services to help them plan and may be given responsibilities for their own budgets.

A related idea, restructured schools, is the central focus of the Coalition for Essential Schools. The goal of the coalition is not teacher empowerment per se, but instead improved teaching and learning. However, in the process of establishing these improvements, teachers become heavily involved in deciding how their respective schools will be organized, what it will be trying to accomplish for students, and what the teachers themselves will be doing to achieve these goals.

Programs that establish merit pay, career ladders, and/or peer evaluation focus more on individual teachers than on teams of teachers in schools. They are intended to recognize differences in teacher quality and to reward those differences in some formal way. In the case of peer evaluation, they are also intended to give teachers more voice in judging their professional colleagues. Like school-wide programs, these programs exist in only a few states or school districts, and have not existed long enough to have a noticeable impact on teachers. Or, if they do persist, it is often because they have been altered to the point that they are not so radical a change as reformers might have hoped for.

Moreover, many of these empowerment activities are extremely difficult and time-consuming and are not necessarily embraced by teachers. Teamwork requires a tremendous commitment of time, time which many teachers do not have. In addition, plans that require teachers to observe and comment on one another’s classroom practices fly in the face of the culture of teaching, in which teachers are very careful never to comment on one another’s work (Little, 1988). Similarly, ideas that differentiate among teachers, either through staffing arrangements or through career ladders or merit pay, are generally distasteful to teachers. The evidence to date does not suggest that teachers perceive clear benefits from efforts to professionalize their work and that, moreover, these efforts increase their burdens.

Ironically, although the aim of these policies is to increase teachers’ professional authority, most of these new policies are not originated by
teachers themselves but instead by state or school district administrators. Thus, even though their rhetoric is that of teacher empowerment, the strategy continues to be that of establishing new rules for teachers to comply with. To the extent that these policies are viewed by teachers as burdensome new requirements, rather than opportunities, they are likely not to have their intended effect (see, e.g., Johnson, 1990). The Coalition for Essential Schools offers an alternative to the top-down strategy for school reform, as do some of the district-level devices that have been designed through negotiation between the district administration and the teachers' union.

There are at least two reasons, then, to be pessimistic about these proposals. First, the vision of professionalism espoused by researchers and policy makers is somewhat at odds with teachers' own vision of their work. Whereas the advocates of professionalism want to recognize the differential abilities of teachers and want teachers to strive to improve their own practices by engaging in research, making their own decisions about what to teach and how, engaging in critical reflection about their own practices and interacting with one another about all of these issues, teachers themselves prefer not to distinguish among themselves, not to discuss their practices with one another, not to examine one another's practices, and often not to make difficult curriculum decisions. Second, even if we were to achieve this radical alteration in the character and quality of the teaching profession, we have no evidence that such changes will in fact improve the quality of classroom practices. The notion that teachers will somehow, simply by virtue of having the opportunity to make their own decisions, discover better practices, is weak at best, and no clear model has yet been articulated regarding where the better teaching practices are to come from when the existing population of teachers are committed to practices that derive from their own childhood experiences.

**Implications for Data Collection**

The problem with policies aiming for professionalism is that they are tremendously various. If we begin gathering data on the presence or absence of a particular idea, such as school-based management or merit pay, and then look for evidence that such programs have altered teacher attrition or improved classroom practices, we may easily find the same result we have found when studying the role of credentials. And the finding would be similarly difficult to interpret, for we cannot know whether such programs could make a difference, if done in such a way that teachers really feel more professional responsibility. Perhaps, instead of documenting the presence or absence of particular policy practices, NCES might contribute more to policy by documenting teachers' perceptions of their professional control
over different aspects of their work, and see whether these perceptions are related to, say, attrition. These data might shed more light on what teachers really want in their jobs.

One thing that makes teacher attrition difficult to measure is that we often don’t know, when a teacher leaves a particular position, whether that teacher is leaving the field or is leaving that particular position. Certainly improvements could be made in our knowledge of teachers who leave particular teaching jobs. I recommend that such data should not be limited to teachers’ perceptions of their job satisfaction but include data on the job itself. For instance, how far it is from the teacher’s home town or alma mater, what fraction of the teacher’s courses are in his or her subject matter major, how many different preparations are required of the teacher each day, and how many students the teacher comes into contact with weekly. Such data, combined with data on the teachers’ perception of job-related working conditions, responsibilities, and constraints, could be used to begin developing a predictive model of teacher attrition.

SUPPLY OF HIGH-QUALITY CLASSROOM TEACHING PRACTICE

Virtually all of the blue ribbon commissions that have studied education in the last decade have argued that we need a new and better kind of teaching: one that challenges students more than current methods do, that expects more of students, that demands higher-order thinking from them, that prepares them for the work place of tomorrow. More than previous generations, today’s students must learn to work collaboratively in teams, to solve problems, to be flexible and adaptable. Yet traditional teaching practices encourage students to work in isolation and compete with one another, to learn discrete facts and skills rather than to solve complex problems, and to follow fixed routines rather than experiment with novel tasks. Preparing students for tomorrow’s work place requires a different kind of teaching, so policy makers face the task of finding ways to improve classroom teaching practices.

But improving the quality of classroom teaching entails more than merely a change in substantive orientation. I mentioned earlier that the student body itself is changing rapidly, and that even a good solution to the representation problem would not yield a complete match between teachers and students. This means not only that teachers must learn to teach different content, but also that they must learn to teach a different kind of student, one that has traditionally been alienated from schools and from academic subjects, and one that differs substantially, in most cases, from those the teacher grew up with. Schools have always served diverse groups of students and have never been able to serve them all equally well. As a society,
we have vacillated in the degree to which we tolerate this problem. But as our population changes so that a larger fraction of students represent those that schools have been unable to serve, the problem of school failure becomes much more severe.

How Important is the Quality of Classroom Teaching Practice?

I have already discussed the importance of classroom teaching practice. My main argument here is that the importance of this dimension of quality stems from a recognition that the content being taught to students is at a far lower level than we want to see. If our educational goals were limited to basic skills and the memorization of specific facts, we might not be worried about the quality of classroom practice. But we have, as a nation, set much higher goals for education than we are now able to achieve.

The quality of classroom practice problem boils down to this: If we know that teachers are highly likely to teach as they were taught, and if we are not satisfied with the way they were taught, how, then, can we help them develop different teaching strategies? And how can we create schools and policies that support their use of these alternative strategies?

Of the five quality issues I have described, the improvement-of-practice problem is the most serious. We are caught in a vicious cycle of mediocre practice modeled after mediocre practice, of trivialized knowledge begetting more trivialized knowledge. Unless we find a way out of this cycle, we will continue to recreate generations of teachers who in turn recreate generations of students whom we already know are not prepared for the kind of information/technology society we are entering.

The heart of the problem is that if teachers and prospective teachers are driven to emulate the teachers they observed when they themselves were students, they will not quickly adopt new ideas. In our work at Michigan State, for instance, we have found that undergraduates who plan to enter teaching almost universally hold a limited view of their role as teacher, thinking that teaching entails little more than telling students what they know and assessing students’ recall of that knowledge (see, e.g., Ball, 1988b; Feiman-Nemser et al., in press). When asked what they might do if their students did not understand a particular concept, their sole recourse was to “go over it again.” They simply could not envision any alternative teaching strategies. They do not even imagine that the telling itself must be done with care. Their tacit view of the teacher’s role limits their ability to envision, let alone enact, any more demanding forms of teaching. Having witnessed teaching throughout their lives, and having been successful students when their teachers taught mainly by telling students and then testing students, teachers cannot imagine the effort that is implied by the current demands for improvement.
More than any other aspect of instruction, the role teachers adopt in their classrooms is critical to the improvement of practice. Moreover, helping novices conceive of a new role may be the most difficult challenge facing teacher educators. Teachers cannot follow procedural manuals defining a list of new activities; they must adopt a role and make it their own. A role is an abstract concept to grasp, and the roles envisioned by education reformers must be adopted cognitively, affectively, and kinesthetically by teachers. A teacher’s ability to adopt a new, more demanding role depends not only on his or her understanding of that role, but also on emotional acceptance of and commitment to that role.

**Policy Activity in This Area**

The traditional strategy of inservice teacher education continues to be a popular one for improving the quality of classroom teaching practices. Although most current efforts at inservice teacher education fall remarkably short of having any real impact, inservice programs are still one of the few policy options that can be expected to alter the practices of existing teachers or alter the practices of those hired from the reserve pool.

Moreover, some two decades of research into strategies for influencing teaching practices have given us a reasonable basis for designing inservice programs. Research on teaching meaningful content, for instance, suggests that the key to teaching lies in subject-specific pedagogy rather than generic pedagogy, and research on promoting changes in teaching practice suggests that changes are more likely to occur when they are introduced at the school building level, rather than to individual teachers working in isolation, and when the desired changes are promoted over a long period of time with ongoing classroom support.

Yet all of these features are absent from most state and locally supported inservice programs. Interestingly, one of the strongest inservice systems now in existence, the National Writing Project, is sponsored almost exclusively by teachers rather than by policy makers. Three weaknesses are particularly salient to state- and district-supported inservice programs. First, most inservice activities are not really full-blown programs, but instead are brief workshops. School districts typically build into their budgets one or two workshops each year, with each workshop consisting of a half-day to a two-day session on a particular issue. Given the depth of teachers’ commitments to their current practices, it is highly unlikely that such brief exposures to new ideas will substantially alter their classroom practices. Second, the content of these workshops often changes from one to the next, so that teachers receive inconsistent messages regarding what counts as good practice. One may promote assertive discipline while another promotes teaching higher-order skills and a third promotes increases in time on task.
Because these workshops are independent of one another, teachers not only receive mixed messages regarding the relative importance of these ideas, but also receive no assistance in rectifying philosophical discrepancies among the different ideas. Finally, very few inservice programs provide follow-up assistance in the classroom to help teachers implement the ideas they have learned about. Consequently, teachers are left with only the general idea and with no specific means for implementing these ideas.

Some teacher educators are also investigating ways to help teacher candidates learn more from teacher education than they have tended to learn from it in the past. We now know more than we did even 10 years ago about what is involved in teaching higher-order thinking, and we know more about the nature of the knowledge teachers need.” Teacher educators are now experimenting with ways of getting more out of their candidates’ formal classroom learning time (e.g., Schramm et al., 1988; Feiman-Nemser et al., in press; Ball, 1988a). One way they do this is through assignments that force their candidates to be more analytic about what they see in classrooms—to raise questions about what is being taught and what is being learned and to generate hypotheses about better ways to do it. Teacher educators are also experimenting with ways of helping candidate teachers develop alternative teaching strategies.

A popular new strategy designed to improve the classroom practices of new teachers is the inclusion of on-the-job assessments in hiring policies, such that teachers are placed on probation for the first year or two and their practices are assessed during this probationary period. The most visible systems, those in Virginia and Florida, focus on discrete teaching skills that are correlated with standardized test scores in early elementary grades, in which the content is heavily oriented toward basic skills. The research on which these assessments are based, often referred to as process-product research, has relied on measures of student outcomes that are themselves limited: standardized achievement tests that tend to measure students’ acquisition of basic knowledge and skills. Moreover, this research has been conducted for the most part in the early elementary grades, in which students are learning largely basic skills.

Since the original research addressed the way teaching is currently done, most of its findings about pedagogical practices are valid only for the teaching of routine skills, not for teaching higher-level thinking, nor for adequately representing subject matter to diverse learners. Thus, the teaching behaviors identified through this research are not necessarily those that will yield the kind of improvements in practice that we now seek. Indeed, some reformers argue that these are the very skills that hinder higher-order reasoning in the classroom. Perhaps the most remarkable evidence of the limitation of this basic skills orientation appears in a response by Barak Rosenshine (1986) to a high school history lecture given by then-Secretary
of Education William Bennett. Rosenshine confesses that he doesn’t know what to make of the secretary’s lesson because research on teaching has focused on teaching skills, not on teaching content. As Rosenshine tries to evaluate the secretary’s lesson, using his own knowledge of teaching skills, he assumes that the goal of the lesson is to get students to recite the knowledge as opposed to getting them to understand it, interpret it, or use it to interpret current events. These performance assessment systems, then, are not likely to yield the kind of improvements in classroom practice that most reformers seek.

Another popular strategy for improving the classroom practices of beginning teachers is the induction program. Thirteen states now require some form of induction program beyond student teaching. In most cases, a state-level requirement for induction means that school districts must provide first-year teachers with some sort of guidance or assistance as they struggle through their first year or two of learning to teach (see Schlechty, 1985, for a discussion of induction rationales). Often this assistance appears in the form of an experienced teacher who serves as a mentor to the new teacher. There are numerous variations on this theme. Mentors can work independently in one-on-one relationships or they can provide a standardized orientation to all newcomers. Districts may release their mentors from full-time teaching while they serve as mentors, or they may require them to help novices in the cracks between classes. Given the range of strategies for providing assistance, it would be reasonable to expect these programs also to differ in the benefits that novices gain from them.

Gauging the benefits of induction programs forces us to face the difference between efforts to increase the supply of teachers and efforts to improve the quality of teaching practices. One strong argument for induction programs is that they may decrease the attrition rates among novice teachers. Yet, like alternative route programs, induction programs will probably not improve the quality of classroom practice. They may help new teachers learn what the current population of teachers already knows, and may help them over the difficult adjustments into teaching, but they will not help them learn new or different approaches to teaching. If anything, induction programs further reinforce the same kind of teaching we already have—the kind that emphasizes facts and skills rather than reasoning and analysis and that encourages student passivity and compliance rather than active engagement with subject matter. This is the kind of teaching that has led to the assessment results we are unhappy about.

Yet another proposal often put forward to improve classroom teaching practices is the professional development school, or PDS (The Holmes Group, 1986; Levine, 1988: Lieberman, 1989). There was a time when most colleges of education operated lab schools—schools located on or near university campuses where student teachers received their practical experience. It
might be easy to think that professional development schools are a newer version of the same thing, and in many respects they are. But in one important respect they are not. The lab school was operated by the university, not by the community, and its students tended to be faculty offspring. Lab schools prepared teachers in a sort of ivory tower environment—no poverty, no uneducated parents, few children from culturally different backgrounds, and often plenty of resources.

Professional development schools, in contrast, are genuine collaborations between the university and the school. They serve the children who attend public schools, not those who attend private schools. Moreover, their reason for being is not just to provide a context for new teachers to learn to teach, but to allow all teachers to learn more about teaching. In that sense, they are truly laboratories.

Not many full-blown professional development schools exist now, so I can describe here only their idealized features. Modeled after the teaching hospital, one feature of the PDS is that university faculty would regularly teach there, and that teachers from these schools would regularly teach at the university. Another feature is that virtually everyone associated with a PDS—faculty, teachers, and student teachers—would be experimenting with new ways to teach children. And finally, the staffing patterns and the physical layout of PDSs would look different from regular schools. They would contain private places for teachers to plan and design new strategies and materials, conference rooms where teachers could work together on new ideas, and observation rooms connected to classrooms so that teachers, teacher candidates, and other visitors could observe these teaching practices without disrupting the activity itself.

Getting a sense of the layout and the kind of work that occurs in these places should indicate some of the staffing changes that would have to occur as well, for we cannot have genuine experimentation in the typical egg-carton school organization, where each teacher is tied to one or another group of children almost every hour of every day. Yet someone must be with these children, and the children must be learning throughout the day. To get a professional development school going, then, districts would need either to double the existing staff and be very creative in grouping students and teachers, or to develop a new cadre of teaching assistants who would free teachers to do the kind of experimentation and development that is needed.

The PDS proposal is clearly designed to contribute to the improvement of practice. It does nothing for credentials, demographic representation, or tested ability. Moreover, PDSs are intended to improve the practices of the existing teaching force as well as to provide a better preparation for new teachers. But while there is a tremendous amount of activity in this area, we are hardly at a point at which we can say that this strategy has clear
potential. Moreover, most proposals for PDSs, even though they recognize the need to improve the practices of existing teachers, still take their main task as that of preparing new teachers. Existing practices are only improved in PDS schools themselves, not in the rest of the schools.

Implications for Data Collection

When it comes to data collection, the quality of classroom practice is the most muddy. Though nearly everyone believes they can distinguish good teaching from bad, we still face tremendous conceptual problems in defining ordinary teaching, bad teaching, and good teaching and even more problems in trying to measure these things. Add to all of this the problem of measuring changes in classroom practices and we have a real soup.

Gathering data on practice itself offers us both the advantage of directly measuring the thing we are ultimately interested in, and the disadvantage that we have no clear methods for doing this. Performance assessments such as those described above, even if altered to capture more important aspects of teaching, still are susceptible to corruption if teachers learn what they are supposed to do while they are being observed. The work of the Educational Testing Service on a new version of the NTE, and the work of the National Board of Professional Teaching Standards on a certification assessment for teachers, will contribute to our understanding of teacher knowledge and perhaps to our understanding of their pedagogical subject matter knowledge. Both are intended to yield more valid and meaningful measures of teacher quality. And, although neither of these new devices will be ready in the near future, they will probably be ready before we could come up with any other alternative independently.

What these measures will not yield is evidence of teachers’ inclinations to regularly engage in the practices that are advocated by the assessors. The more we recognize the importance of teachers’ commitments to different teaching ideals, the more imperative it becomes to find ways to measure these commitments and to incorporate them in our estimates of the number of teachers available whose classroom practices approach those we are aiming toward. The research being done at the National Center for Research on Teacher Education includes a number of questions designed to tap teachers’ beliefs and values as they are expressed in the context of particular kinds of teaching situations. Recognizing the importance of teachers’ beliefs about what should be taught and about their role in facilitating learning, we ask several questions designed to learn more about these aspects of their teaching. Such questions, if added to, say, an NAEP assessment, might yield important findings regarding the relationship between these teacher characteristics and student outcomes in higher-order areas. Consider, for instance, the finding from the NAEP assessment that students can manage basic writ-
ing skills, but cannot develop coherent arguments. A likely reason for this outcome may be that teachers of writing emphasize basic skills more than coherence. Here is an item we asked teachers in our study that reveals their tendencies to emphasize different aspects of writing:

I’d like you to imagine that your third grade students are writing stories. Jessie, one of your students hands you the following story.

One day my friend Mary asked me. Do you want to have a picnic? When we got there we started playing. At the picnic people said. Where’s your puppy? He is at home? We went home happy. My mother said. I’m glad you had a picnic.

What do you think of Jessie’s story?
How would you respond to Jessie? Why?
What grade would you give this paper? Why?

Responses to this question have been particularly telling with respect to teachers’ assumptions about their role in helping students learn to write. Teachers can choose to focus on the content and coherence of the story or on its technical correctness. Even when focusing on its technical aspects, they may either identify one or two main problems or can enumerate every problem they see. What Jessie has done here is invent a way to indicate quoted material. Not knowing the convention of using quotation marks, Jessie indicates a new speaker by inserting a period. In fact, three of the errors in this story derive from this one writing decision, and three others derive from an inability to spell the word picnic. So a teacher who focuses on technical correctness may count these as two errors, while another may count them as six errors. The former is diagnosing the student’s writing and concentrating instruction on general skills, while the latter is not really teaching but instead only counting errors.

When responding to student work, teachers have enormous latitude in deciding what role they want to play in helping their students learn. Questions such as these may prove to be highly useful indicators of teaching practice and may be more efficient than direct observation. If such questions were embedded into NAEP surveys, they could improve our understanding of the relationship between teachers’ fundamental orientations toward teaching and the kinds of outcomes they promote in their students.

**SUMMARY**

I have tried in this paper to distinguish among several different dimensions of quality, to review the kinds of policy alternatives that are generally under consideration with respect to each of these dimensions, and to consider data collection strategies that might be relevant to each. The dimen-
sions of quality that I reviewed here differ in their attention to individual teachers and groups of teachers and in their attention to characteristics of teachers versus characteristics of teaching.

These dimensions also differ with respect to how easy they are to study. The dimensions of credentials and demographic characteristics are easier to define; consequently, supply and demand within these dimensions is easier to measure. The dimension of tested ability is somewhat more difficult to define, and consequently more difficult to measure as well. And in the absence of a handy measure of tested ability, it will be difficult to gauge supply and demand within this dimension of quality. The dimensions focusing on teaching behaviors—the professionalism of their work or their classroom teaching practices—are the most difficult to define, the most difficult to influence, and the most difficult to measure. Although we all believe we can recognize professional attitudes and good classroom teaching when we see it, we do not have ready indicators of it. Moreover, to the extent that we want data that can indicate whether teaching practice is improving, especially data that indicate whether we have an adequate supply of teachers who can promote substantive reasoning in their classrooms, we need ways of recognizing and documenting progress toward practices that are not well understood and are only occasionally available for examination. Change in practice will prove difficult not only for teachers, but also for researchers and policy makers, for all of us are more at home with teaching as we knew it when we were young.

I have also argued that solving any one of these aspects of quality would not necessarily solve any of the others and that, in particular, improving the supply of teachers who meet any of the first four dimensions of quality may not lead to improvements in the supply of teachers whose actual classroom teaching practice is better than current practice is. Improving the match between teachers’ credentials and their teaching assignments won’t necessarily lead to improved practice; getting a more demographically representative population of teachers won’t do it; increasing tested ability won’t; and neither, necessarily, will efforts to professionalize teaching. But it is also true that improving classroom teaching practices will not necessarily improve credentials or demographic representation or tested ability or teacher professionalism. So to the extent that we want all of these things, perhaps each for a different reason, each of these dimensions of quality must be considered in any estimates or projections of supply and demand.

My recommendations for improving data on supply and demand within these dimensions of quality are as follows:

With respect to credentials, states should be encouraged to maintain more and better data on the number of teachers who apply to have out-of-state credentials recognized and the number accepted and rejected, on the
number of emergency-certified teachers in each teaching category, and on the actual teaching assignments of teachers with different credentials. In addition, NCES could facilitate state policy by documenting the variety of college curricula that lead to credentials in different states. Such data would enable state policy makers to examine their own policies in relation to those of other states.

With respect to tested ability, existing state tests are too various to warrant cross-state data collection systems. And since two major efforts are currently under way to improve the quality of national-level data on teachers’ tested ability, it would not make sense for NCES to do so. However, NCES could play an important role in testing the validity of some of the alternative assessments currently being developed by attaching portions of them to future NAEP studies, so that teachers’ responses to these items could be correlated with the different kinds of skills and abilities demonstrated in students.

With respect to demographic quality, NCES should capitalize on university data bases to predict the number of potential minority college graduates and, if possible, should step back yet another level and use secondary school data to estimate the number of college-bound minorities as they graduate from high school. When measuring demand for minority teachers, NCES should devise a metric that enables them to document teacher demography as a variable of school composition. An index of the demographic mix of teachers within schools could indicate the relative overlap between teaching populations and student populations. Such a system would enable us define the supply of schools available that offer students a locally representative teaching force and to identify and locate those schools that are most or least successful in matching their teacher demography with their student demography.

With respect to professionalism in teaching, the programs designed to promote professionalism are so various that documentation of programs per se will probably not yield much. Instead, NCES might try to develop better models for predicting teacher attrition-models that incorporate relevant features of teaching jobs and that incorporate teachers’ perceptions of control over various aspects of their work.

With respect to classroom teaching practices, the nature of programs designed to promote improvements in classroom practices are extremely various, even though many of them share the same names. Moreover, direct measures of classroom teaching practices would be too costly to implement on a wide scale. Instead of documenting the presence or absence of such programs, NCES might better facilitate policy making by experimenting with measures of teachers’ beliefs about what should be taught and about their role in facilitating learning. Such experimentation could be done by embedding context-specific questions about teachers’ beliefs and values into NAEP or other NCES surveys. Such questions could improve our under-
standing of teachers’ fundamental orientations toward teaching and the kinds of outcomes these orientations promote in students.

I have also argued that two particular features of the teaching force make policy efforts especially perplexing. One of these features is the sheer size of the teaching force. Numbering over 2 million, teachers represent the largest fraction of college graduates in this country—yet they represent less than half of all college graduates who have been prepared to teach. These facts have implications for all dimensions of quality, and for virtually any policy effort to influence quality and supply. Size has implications for credentialing, especially in light of the large number of people who obtain credentials but do not use them or use them only for a year or two. The data suggest that policy makers need to increase the supply of credentialed teachers at substantially greater rates than are actually needed if they want to ensure that credentialed teachers will actually be available to schools that need them. That is, projections of credential supply need to take into account the fact that a sizeable fraction of people with credentials never enter teaching or will teach for only a year or two.

Size also has implications for altering the demographic composition of the work force. If policy makers tried to encourage an equal proportion of each racial or ethnic group of college students to enter teaching, they would find that they still did not have enough new minority teachers to balance the proportion of minority students in schools. This result would occur in part because the production of new teachers will not alter the current teaching population, which is substantially out of balance with the racial and ethnic composition of the student population, and in part because the student population is changing so rapidly toward nonwhite students that even the most concerted minority recruiting programs are likely to lag considerably behind the student body changes.

The size of the teaching force also has implications for tested ability, for schools must compete with other industries for employees and cannot expect to take a substantial fraction of adults with high tested ability from other professions and businesses. Finally, the size of the teaching force has implications for virtually any efforts to alter teaching behaviors and norms, either through professionalization or through efforts to improve classroom practices, for the kinds of changes that are being sought. They would necessarily entail substantial labor costs, both for teachers, who need more time to develop new teaching strategies, and for consultants and other support services that would be needed to help teachers make these changes.

The second aspect of teaching that makes policy efforts more perplexing is that many of these issues of teacher quality and supply are differentially salient across different types of schools, so that nationally aggregated data will not capture the policy-relevant issues. Problems of supply are uneven across schools and types of teaching positions, so that data on sup-
ply and demand should be defined in terms of the particular school contexts, rather than through national averages. This admonition is relevant to all dimensions of quality and to all methods of measuring supply and demand within these dimensions. Schools differ, for instance, in their reliance on emergency credentials, so that data on the supply and demand of credentialed teachers will be more useful when they take schools into account. Problems of demography are also relevant at the school level, where discrepancies between populations of teachers and of students are most visible.

Tested ability is also likely to vary by schools. If brighter teachers tend to marry brighter spouses, they will live in communities that have jobs for their spouses. That is, we might find tested ability higher in communities that employ more highly educated people in general, rather than in communities in which the dominant employment is unskilled or skilled labor. Attrition is also uneven across types of schools and types of teaching positions within schools. Therefore, rather than measuring national variations, data bases should be constructed to indicate numbers and types of schools that exhibit different mixes of teachers and different mixes of teaching assignments. For all of these reasons, then, it would behoove the NCES to consider methods of measuring both quality and supply and demand through school units rather than through national aggregates of individual teachers.

REFERENCES

Adleman, Nancy

American Association of Colleges for Teacher Education

American Federation of Teachers

Applebee, A.N., J.A. Langer, I.V.S. Mullis, and L.B. Jenkins

Ball, Deborah L.

Begle, E.G., and W. Geeslin

Carey, Neil, B.S. Mittman, and L. Darling-Hammond

Carnegie Forum on Education and the Economy

Cohen, D.K., and R.J. Murnane
1985 The merits of merit pay. The Public Interest 80(Summer):3-30.

Council of Chief State School Officers

Darling-Hammond, Linda

Darling-Hammond, L., and B. Berry

Darling-Hammond, Linda, Lisa Hudson, and Sheila N. Kirby

Doyle, Walter

Doyle, Walter, and K. Carter

Feiman-Nemser, S., G.W. McDiarmid, S. Melnick, and M. Parker

Feistritzer, Emily

Gallagher, James J.

Gilford, Dorothy M., and Ellen Tenenbaum, eds.

Haberman, Martin

Haggstrom, G.W., L. Darling-Hammond, and D.W. Grissmer

Haney, Walter, G. Madaus, and A. Kreitzer
1987 Charms talismanic: Testing teachers for the improvement of American educa-

Hanushek, Eric


Hawley, Willis D.


Heyns, Barbara


The Holmes Group


Homer, B., and J. Sammons


Houston, Holly


International Association for the Evaluation of Educational Achievement


Johnson, Susan M.


Kennedy, M., ed.

1989 *Competing Visions of Teacher Knowledge.* East Lansing, Michigan: National Center for Research on Teacher Education.

Kerr, Donna H.


Lanier, Judith


Levine, Marsha, ed.


Lieberman, Ann, ed.


Little, Judith W.


Lortie, Daniel


McDiarmid, G. Williamson

McDiarmid, G. Williamson, D.L. Ball, and C.W. Anderson  

McDonnell, Lorraine M.  

McDonough, M.W., Jr., and W.C. Wolfe, Jr.  


Murnane, Richard J., J.D. Singer, and J.B. Willett  

Murnane, R.J., J.D. Singer, J.B. Willett, J.J. Kemple, and R.J. Olsen  

National Center for Education Statistics  

National Commission on Excellence in Education  

National Research Council  

Nemser, Sharon Feiman  

Porter, Andrew C.  

Robertson, S.D., T.Z. Keith, and E.B. Page  

Rosenshine, Barak  

Schlechty, Phillip C.  

Schlechty, Phillip C., and V.S. Vance  

Schramm, Pam, S. Wilcox, P. Lanier, and G. Lappan  
Sedlak, Michael, and Steven Schlossman

Shulman, L.S.

Sloddart, Trish

Tobin, Kenneth
1987 Forces which shape the implemented curriculum in high school science and mathematics. Teaching and Teacher Education 3:287-298.

Tobin, Kenneth, and James J. Gallagher

Tyson, H., and A. Woodward

Wilson, S., L.S. Shulman, and A.E. Richert

Zumwalt, Karen, G. Natriello, A. Hansen, and A. Frisch

NOTES

1. I am indebted to James Stedman for suggesting this dimension of quality as relevant to issues of supply and demand.

2. A recent report from the American Federation of Teachers (1985) and the Council of Basic Education indicates that, even though virtually all states have requirements about who can teach in various areas, almost none has any methods for checking on local compliance with these rules.

3. Some unpublished work done by the National Center for Research on Teacher Education is relevant to this point. We found, for instance, that one of the staples of teacher education, the educational psychology course, can look remarkably different in different programs. We have also found that, when states change their curriculum requirements for undergraduate teacher education, programs are likely to simply rename their courses rather than actually change the content to match new requirements.

4. Hanushek’s work tends to support the idea that verbal ability is relevant to teaching; however, other reviewers have questioned this finding. See, for instance, Haney et al. (1987).

5. For a history of the legal precedents in this area, see McDonough and Wolfe (1988). For an example of a report questioning the validity of such assessments, see Horner and Sammons (1987).

6. Few studies actually examine the premises of these programs; most look instead at the number and biographical backgrounds of recruits. Two studies sponsored by the National Center for Research on Teacher Education, however, did examine the test scores and grade point averages (GPAs) of alternative route candidates. One is Zumwalt et al. (1987). The other is Stoddart (1987). Stoddart found that grade point averages, for instance, were rela-
TEACHER QUALITY AND BALANCING SUPPLY AND DEMAND

tively high among alternative route English teachers but were relatively low among alternative route mathematics teachers. Based on data from the American Association of Colleges for Teacher Education, I would guess that these English teachers have grade point averages a bit above those of traditional teacher education graduates, whereas the math teachers have grade point averages a bit below those of traditional teacher education graduates. See American Association of Colleges for Teacher Education (1987a).

7. A problem using 2 as the divisor, for instance, might be: “My roommate and I have 1 3/4 pizzas and we want to divide them equally between us. How much pizza can we each have?” A problem that uses 1/2 as a divisor would be: “I have 1 3/4 yards of cloth to make tutus for the ballerinas. Each tutu requires 1/2 yard of cloth. How many tutus can I make?”

8. The American Association of Colleges for Teacher Education (1987a) has documented these trends and has developed a strong set of policy initiatives (1987b) related to it.

9. For purposes of this discussion, I am describing a few general ideas. For a more detailed discussion of a variety of other approaches to the representation problem, see the American Association of Colleges for Teacher Education (1987b).

10. For instance, Tobin and Gallagher (Gallagher, 1989; Tobin, 1987) have shown that science teachers often do not include all students in their lessons and call on only a few students whom they believe to be most capable. They have also argued that teacher’s beliefs about their role in promoting learning can severely limit their practices.

11. There is considerable evidence that college students who intend to teach enroll in colleges within roughly 50 miles from their home town, and that they intend to return to this same region when they take teaching positions. Since most teacher’s colleges, by design, are scattered through the less populous regions of states, this pattern means that few teacher candidates are drawn from, and intend to return to, urban areas.

12. See, for instance, Cohen and Murnane (1985). The districts they examined, which had been operating merit pay systems for at least six years, had altered the concept in one or more important ways: they made the system voluntary rather than universal; they made the size of financial rewards small or distributed rewards widely; they defined merit in terms of extra work rather than better work; or they allowed teachers to define their own criteria for effectiveness. These authors argue that merit pay will probably not have an impact on classroom teaching practices and may not be feasible in those districts that are most in need of improvement.

13. Aspects of this argument can be found in a wide range of places. For instance, the Carnegie Forum on Education and the Economy (1986) specifically addresses the problem of teacher quality; the National Commission on Excellence in Education (1983) addresses mainly the quality and rigor of U.S. secondary education; the National Research Council (1989) addresses the issue of mathematics education in particular; and The Holmes Group (1987) addresses both teacher education and the attractiveness of the workplace to teachers.

14. Much of this development has occurred through the work of Lee Shulman and his colleagues at Stanford University and the work of the National Center for Research on Teacher Education at Michigan State University. And some professional associations, such as the National Council for Teachers of Mathematics and the National Council of Teachers of English, are now incorporating these new ideas into their teaching standards. Some illustrative work from Stanford include Shulman (1987) and Wilson et al. (1987). From the NCRTE are Kennedy (1989); McDiarmid et al. (1989); and Bali (1990).

15. A good book on some of the dilemmas of professionalizing teaching is Lieberman (1988). Several of the chapters, both empirical works and general analyses, are relevant to this point.

16. The story is taken from a real teaching situation, and when Jessie’s teacher asked Jessie why the periods were put into these sentences, Jessie said that they indicated a pause, and that when a new speaker was mentioned in a story, there was always a pause to indicate this new speaker.