Teacher Inequality: New Evidence on Disparities in Teachers' Academic Skills

Andrew J. Wayne
SRI International


Abstract
When discussing the teacher quality gap, policy makers have tended to focus on teacher certification, degrees, and experience. These indicators have become key benchmarks for progress toward equality of educational opportunity, in part for lack of additional teacher quality indicators. This article turns attention to teachers' academic skills. National data on teachers' entrance examination scores and college selectivity reveal substantial disparities by school poverty level. The findings commend attention to the gap in academic skills in the formulation of future policy and research on the teacher quality gap.
The teacher quality gap has received attention from a broad spectrum of policy makers (see, e.g., American Council on Education, 1999; National Association of State Boards of Education, 1998; National Commission on Teaching and America's Future, 1996). Even federal legislators have recently proposed ways to close the gap in teacher qualifications between low-income and affluent children (see Stedman, 1999; Wayne, 2000, 6-7). This article reports on the results of an examination of national data on disparities in teachers' academic skills. Analysts have thoroughly documented disparities in teachers' experience, certification, and degrees, but very few studies assess differences using academic skills indicators such as teachers' college ratings or entrance examination scores. (Note 1)

Opportunities to examine disparities in teachers' academic skills at the national level exist because of two national survey efforts, overseen by the National Center for Education Statistics. First, data from the Schools and Staffing Survey portray the public school teaching force as a whole. Second, data from the Baccalaureate and Beyond Longitudinal Study portray the contribution of a cohort of college graduates. This article presents analyses of both data sets. For organization, the article begins with a discussion of relevant theory. It then dedicates one section to each of the two national data sources. Its conclusion briefly discusses implications for policy making and for future research.

Theory

The basic premise undergirding this investigation is that students with lower quality teachers experience a disadvantage. Put into theoretical terms, student learning is a function of teacher quality. Theorists have conceptualized teacher quality as a set of specific knowledge areas and skills (see e.g., Shulman, 1987). But in order assess teacher quality disparities at the national level, one must rely on quality indicators that can be readily measured by teacher questionnaires, administered via sample surveys. Measures such as teachers' experience, certification, and degrees are taken to be indicators of the theorized components of teacher quality, such as pedagogical knowledge and content knowledge. Analysts have thoroughly documented disparities using questionnaire data containing these indicators (see Henke et al., 1997; Ingersoll, 1996; Lewis et al, 1999).

Less attention has been paid to differences in what Mayer, Mullens, and Moore (2000) refer to as indicators of teachers' academic skills, such as teachers' college ratings or entrance examination scores. The case for such attention is clear: studies of student achievement gains confirm—more resolutely than for many other indicators of teacher quality—that students learn more from teachers with better academic skills (for reviews see Hanushek, 1997; Mayer, Mullens, and Moore, 2000, 5-7). But it is important to point out that the relationship between academic skills indicators and the conceptualized elements of teacher quality has not received attention from theorists. Why might students learn more from teachers with better academic skills? Teachers who read faster may acquire new content knowledge more quickly. Teachers with greater verbal facility may spend their preparation time more focused on lesson design than on deciding what exactly to say. Maybe teachers with better college entrance examination scores learned more in college, or maybe entrance examination scores and college selectivity correlate with the quality of teachers' precollegiate education—a much longer educational experience than undergraduate education.
The development of a formalized theory of teachers' academic skills is beyond the scope of this article. But such a body of theory will clearly become necessary in the future as researchers and policy makers think about policies that might remedy disparities in academic skills.

**Evidence from the Schools and Staffing Survey**

**Description of the Data**

The 1993-94 Schools and Staffing Survey (Note 2) (SASS) encompasses several distinct surveys whose findings are readily linked. This investigation linked teacher quality indicators from SASS's teacher survey, the Public School Teacher Questionnaire, to school poverty levels obtained via the SASS Public School Questionnaire.

The teacher survey is known for its detailed questions about teacher degrees and certification, but it also contained an indirect measure of academic skills. The SASS asked teachers to identify their undergraduate institutions. Peterson's Guides (1995) rates institutions on its Entrance Difficulty Index, a reasonable albeit rough proxy for academic skills. (Note 3) Ratings include 'most difficult,' 'very difficult,' 'moderately difficult,' 'minimally difficult,' and 'noncompetitive.'

To properly consider the relationship between these ratings and academic skills, one must ask exactly how Peterson's assigns ratings. In truth, Peterson's asks institutions to rate themselves. To guide responses, instruction booklets that accompany the Peterson's surveys specify thresholds on three optional criteria: (1) entering students' high school class rank, (2) entering students' college entrance examination scores, and (3) the percentage of applicants accepted. The ratings probably do not predict academic skills perfectly, since applicants' decisions depend on other factors as well (e.g., wealth).

Another SASS measure that requires some discussion reveals the prevalence of poverty at each teacher's school. The SASS collects, from each sampled school, the National School Lunch Program participation rate. This measure comes with two notable flaws. First, although the income eligibility threshold varies for different family sizes, it does not account for geographic differences in the cost-of-living. Therefore lunch program eligible students in rural areas, for instance, may be better off than eligible students in central cities.

A second problem with this poverty metric is that elementary school students exhibit higher lunch program participation rates than secondary school students—40 percent and 28 percent respectively according to the 1993-94 SASS (Henke et al. 1997, 16). As a result, apparent teacher quality disparities by school poverty level may actually represent differences between elementary and secondary teachers. However, analyses of the disparities in teacher academic skills that disaggregated elementary and secondary teachers revealed disparities substantially like those found for all teachers.

The final source of uncertainty worth mentioning is nonresponse. Missing records, records without school poverty information, and records that could not be associated with a Peterson's rating together conspired to reduce the usable sample by 32 percent, to 37,874 teachers. On the one hand, the availability of so many observations ensures that
sampling errors need not even be mentioned in the figures presented below; all standard errors were less than two percentage points. On the other hand, the assumption that respondents and nonrespondents do not differ systematically is somewhat risky. Thus the true gap in academic skills may be somewhat larger or smaller than depicted.

**Results**

Before presenting results, two analytic issues must be addressed. The first involves what specifically to compare. Because SASS observes poverty at the school level—not the classroom level—all findings about disparities in teacher quality really denote differences between teachers at low-poverty schools and teachers at high-poverty schools. In other words, the selection of comparison categories is a matter of dividing up schools, not teachers. Analysts divide up the schools using lunch program participation rates; the rates act as category boundaries. For example, Ingersoll (1996) divided schools into 'low-poverty,' 'medium-poverty,' and 'high-poverty' according to whether lunch program participation was less than 15 percent, between 15 and 50 percent, or 50 percent or more.

To best meet the needs of policy maker audiences, the present investigation examines disparities using three separate category schemes. The first designates all schools as either low-poverty or high-poverty. It divides them according to whether their lunch program participation rates are above or below 28 percent. That threshold, determined via an analysis of SASS data on schools' enrollments, divides schools such that the low- and high-poverty categories each enroll one half of all U.S. public school students.

Similar computations yielded the remaining two category schemes: one that divided U.S. students into fourths, and another that divided them into eighths. (Note 4) The divisions remain divisions of schools, so the text refers to the individual categories as school poverty quartiles and school poverty octiles.

A second analytic note that must preface the findings involves the teacher quality variable. Even skeptics of the student achievement research would admit that students' opportunities are diminished when their teachers' academic skills fall below some minimum threshold. Therefore the analysis collapses the quality variable to focus on the percentage of teachers from institutions Peterson's rated either 'minimally difficult' or 'noncompetitive.' It subsumes such teachers under the new analytic label, 'less selective.' (Note 5) For perspective, in 1993, these institutions conferred only about one fifth of all bachelor's degrees, according to a weighted tabulation based on the Baccalaureate and Beyond Longitudinal Study.

Having established the comparison categories and the quality metric, the discussion can now turn to results. Figure 1 applies the three category schemes to the SASS teachers to investigate disparities in the percentages from less selective institutions. All three category schemes show disadvantages for higher poverty schools. The juxtaposition of the three comparisons into a single graphic shows that the comparison of halves hides some important variation, evident in the comparison in quartiles. But the further breakdown into octiles does little. An additional breakdown into sixteenths (not depicted) was also not fruitful. The quartile comparison thus properly summarizes the disparities.
Figure 1. The percentages of teachers from less selective institutions, compared using three school poverty categorizations: halves, quartiles, and octiles. Weighted tabulations from the 1993-94 Schools and Staffing Survey.

Evidence from the Baccalaureate and Beyond Longitudinal Study

Description of the Data

The Baccalaureate and Beyond Longitudinal Study (Note 6) (B&B) differs markedly from the SASS. Rather than portray only teachers, B&B allows a unique look at the qualities of graduates from the college class of 1993 who had entered teaching by 1997.

The full B&B sample actually includes many nonteachers as well; it is representative of the entire college class of 1993. But B&B followed the sample members over time, and the 1994 and 1997 interviews determined which graduates had become teachers.
B&B’s teacher academic skills measure and school poverty measure are almost identical to those described above for the SASS. Respondents' undergraduate institutions were linked to Peterson's ratings to yield a measure of academic skills. And the school poverty measure is again the National School Lunch Program participation rate for each teacher's school. (Note 7) B&B does contain one additional academic skills measure, however. B&B’s authors drew records from several sources to determine each graduate's college entrance examination scores. (Note 8)

Two final, important characteristics of B&B are its response rate and the number of observations. A remarkably high 90 percent of the original B&B sample completed the 1997 interview. However, many interviews failed to obtain the quality and school poverty information identified above. Thus, although the 1997 interview located some 967 respondents who reported having entered public school teaching jobs, (Note 9) school poverty information was obtained for only 646 of them. And among those with school poverty information, 630 had Peterson's ratings, and only 530 had entrance examination scores. These problems limited the conclusions substantially, as discussed below.

Results

As in the analysis of SASS, some preliminary remarks are needed regarding the school poverty categories and the collapse of the teacher quality variables. The school poverty thresholds used in the B&B analysis were exactly those identified above in the SASS analysis. Unless the distribution of students across schools of different poverty levels changed substantially between the SASS measurement in 1993-94 and the B&B interviews in 1997, the SASS-derived thresholds provide a sufficient approximation.

Like the analysis of SASS, the B&B analysis presented here collapses the quality variables to focus on the proportions of teachers falling below what might be considered a minimum threshold. The label 'less selective' thus retains the meaning established earlier. For the other quality variable, entrance examination scores, the analysis focuses on the percentages of teachers with bottom quartile college entrance examination scores, where bottom quartile is defined in reference to the examination score distribution of all class of 1993 graduates.

Turning to the results, the weaknesses of the B&B data set constrained the conclusions such that, ultimately, replicating Figure 1 for B&B was not possible. Besides nonresponse, the principal barrier was the very limited number of observations; estimates for school poverty quartiles and octiles were simply not reliable, yielding standard errors as high as nine percentage points. (Note 10) Figures 2 and 3 therefore each compare teacher quality in the low-poverty and high-poverty halves only.

Substantial disparities are evident in these comparisons. In Figure 2, the proportions of teachers from less selective institutions were 22 percent at low-poverty and 37 percent at high-poverty schools. In Figure 3, the proportions with bottom quartile entrance examination scores were less disparate: 26 percent and 34 percent, respectively. This disparity was significant only at the .15 level, while all other differences were significant at the .05 level.
In interpreting the B&B results, readers must remember that B&B does not portray disparities among the entire teaching force. It portrays disparities in the flows of teachers—particularly the flow from undergraduate institutions, within four years of degree receipt. Other flows also play a role in determining the quality of low-income students’ teachers, such as teacher attrition and mobility. In some state-level analyses researchers have found that such departures are especially prevalent among those with relatively high academic skills (see e.g., Murnane, Singer, and Willett, 1989).

**Conclusion**

**Implications for policy making**

When discussing the teacher quality gap, policy makers have tended to focus on certification, degrees, and experience. These indicators have been the benchmarks for progress toward equality of educational opportunity, in part for lack of other indicators. This article shows clearly that policy makers need to consider teachers’ academic skills. According to the evidence presented here, an academic skills gap exists, and it is quite

---

**Figure 2.** The percentages of teachers from less selective institutions, compared using school poverty halves. Weighted tabulations from the 1993 Baccalaureate and Beyond Longitudinal Study, Second Follow-up.

**Figure 3.** The percentages of teachers with bottom quartile entrance examination scores, compared using three school poverty halves. Weighted tabulations from the 1993 Baccalaureate and Beyond Longitudinal Study, Second Follow-up.
large. The proportions of teachers who graduated from institutions rated either 'minimally difficult' or 'noncompetitive' were 21 percent and 39 percent in low- and high-poverty schools, respectively. Closing this gap would require the replacement or upgrade of about one sixth (18 percent) of the teaching force at high-poverty schools.

Disparities of similar size have appeared in the literature for only two other teacher quality indicators. One is graduate degree holding, which is of debatable importance (see Ballou and Podgursky, 1999, 2000; Darling-Hammond, 1999). The other is in-field degree holding, which, though probably important, cannot benchmark those teachers who lack subject-specific assignments (e.g., general elementary teachers). Thus academic skills indicators are a relatively powerful tool for understanding how far away the nation is from providing equally qualified teachers to schoolchildren from different income groups.

Consideration of policy options that might affect academic skills disparities would require clearer theories about the role of teachers' academic skills. Some policy options discussed in the literature thus far relate to school finance (Figlio and Reuben, 1999), school choice (Hoxby, 2000), teacher licensure examinations (Ferguson, 1998; Gitomer, Latham, and Zimeo, 1999), and the use of academic skills indicators in determining eligibility for teacher scholarships and loan forgiveness (Wayne, 2000).

Implications for future research

Data quality limited the certainty of the findings; due to nonresponse and sampling error, the true gap in academic skills may be somewhat larger or smaller than depicted. But data quality sufficed insofar as it showed (1) that a gap exists and (2) that the gap is substantial. Those findings have a clear implication for future research: efforts are needed to bring data on academic skills up to par with data for other quality indicators.

In the short-term, some additional investment to incorporate academic skills indicators into ongoing data collection efforts seems worthwhile. The cost may be substantial. Most quality indicators can be reliably measured via one or two items on a pencil-and-paper questionnaire for teachers (e.g., experience). The same is not true for academic skills measures, as additional effort is required to match entrance examination score records or to code institutional identifiers. Presumably costs and competing priorities explain why recent federally funded teacher surveys (e.g., Lewis et al., 1999) have not measured teachers' academic skills.

Over the long-term, as academic skills indicators receive greater attention, the need to improve the indicators will become obvious. College selectivity ratings offer a fairly rough proxy. And if researchers could administer standardized tests to teachers, it is not clear that they would choose to administer college entrance examinations. Thus, although the disparities reported here are very real, serious thought will be required about what measures could better represent academic skills.

References


About the Author

Andrew Wayne
Center for Education Policy
SRI International
Arlington, VA 22209

Email: andrew.wayne@sri.com

Andrew Wayne works at SRI International in the Center for Education Policy, where his work addresses teacher quality, educational technology, and standards-based reform. Before coming to SRI, he served as a policy analyst for two national efforts to improve teacher quality. He holds degrees in social policy, physics, and education, and has worked as a teacher of science and computers. He can be reached at (703) 247-8491 or by email (above). Correspondence may be sent to the author at SRI International, 1611 North Kent Street, Arlington, VA 22209.

Notes

Some of the findings presented in this article appear in an introductory chapter of the author's doctoral dissertation. Although the responsibility for errors belongs solely to the author, I would like to thank Daniel Goldhaber and Willis Hawley for very helpful comments on an earlier draft. I would also like to thank David Figlio for generously
sharing the data he and his colleagues entered on college selectivity.

1. Disparities were evident in an analysis of very old national data, from 1966 (Ehrenberg & Brewer, 1995), and in recent data from the states of New York and Texas (Lankford, Loeb, and Wyckoff, 2002; Ferguson, 1998; Kain and Singleton, 1996).

2. A variety of technical and methodological reports on the Schools and Staffing Survey are available online at http://www.nces.ed.gov/surveys/sass.

3. Other researchers have completed the requisite data entry to link each institution to its rating in the 1995 edition of Peterson's. The use of the 1995 ratings is defensible if, as Hoxby (2000) has claimed, ratings are sufficiently stable over time. To better reflect teacher characteristics, each teacher would need to be linked to the ratings issued approximately four years before his or her college graduation.

4. The lunch program participation thresholds for all three category schemes were computed via weighted tabulations of the data yielded by the SASS Public School Questionnaire. The threshold that divided them into halves was 28.37 percent. The exact thresholds that divided U.S. students into quarters were as follows: 12.59, 28.37, and 51.80 percent. And the seven thresholds that divided them into eighths were as follows: 6.35, 12.59, 19.67, 28.37, 38.58, 51.80, and 71.80 percent.

5. Richard Ingersoll (1996, 4-5) used a very similar rationale to justify his particular construction of out-of-field teaching. He opted to treat teachers as in-field even if they held only a minor related to the subject taught, and even if that minor was in a subject-related education field (e.g., mathematics education).

6. A variety of technical and methodological reports on the Baccalaureate and Beyond Longitudinal Study are available online at http://www.nces.ed.gov/surveys/B&B.

7. The architects of B&B foresaw that knowledge about the schools at which the B&B teachers taught could be useful. Therefore respondents who taught were asked to identify the school at which they taught, and responses were coded so that basic school characteristics could be obtained via the Common Core of Data—another data set created by the National Center for Education Statistics.

8. Sources included records from the Educational Testing Service and higher education institutions' records of sample members' SAT and ACT scores. See Henke et al. (2000, 83).

9. This total does not include teachers who first taught at private schools. It also excludes teachers who had taught before graduation or had received their certification more than one year before graduation. This latter group is often excluded in analyses of the B&B cohort's contribution to the teaching force. See Henke et al. (2000, 9).

10. Further disaggregation of teachers into school poverty quartiles and octiles yielded estimates with standard errors as high as nine percentage points. Some interesting spikes occurred in the percentages of teachers with low academic skills in the fifth octile, and to a lesser extent in the third quartile. But given that the spikes may be largely an artifact of sampling error or nonresponse bias, it was judged that the best summary would be the comparisons that focus on the low-poverty and high-poverty halves.