DOCUMENT RESUME

ED 448 151 SP 039 680

AUTHOR	Wayne, Andrew J.
TITLE	Teaching Policy Handbook: Higher Passing Scores on Teacher Licensure Examination. Working Paper.
INSTITUTION	National Partnership for Excellence and Accountability in Teaching, Washington, DC.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	1999-04-00
NOTE	12p.
CONTRACT	RD97124001
AVAILABLE FROM	For full text: http://www.ericsp.org/digests/LicExams.htm.
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Academic Achievement; Educational Quality; Elementary Secondary Education; *Knowledge Base for Teaching; *Teacher Certification; *Teacher Competencies; Teacher Evaluation; *Teaching Skills
IDENTIFIERS	*Teacher Knowledge

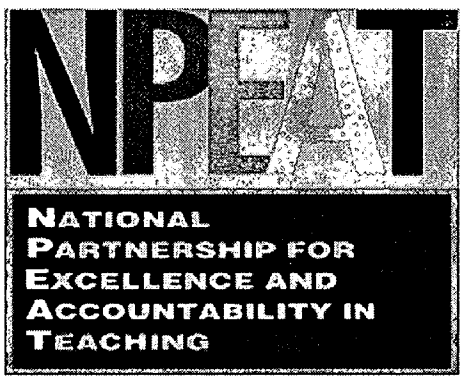
ABSTRACT

This paper provides information about teacher quality proposals that are drawing attention from state and local lawmakers. For each proposal, the paper uses the following framework to organize evidence: definition or explanation of the policy, rationale, trends, evidence, and key issues. The paper demonstrates how the framework would organize evidence about setting higher passing scores on teacher licensure examinations, noting that some people believe that raising the passing score is a means of promoting higher teacher quality. The paper justifies testing related to subject matter, pedagogy, content pedagogy, basic skills, and general knowledge. Courts have not required states to show that students can learn more from teachers with higher scores on licensure examinations. However, a small group of studies of student achievement have assessed the importance of teachers' tested knowledge, and results show that teachers' tested knowledge is an important quality indicator. Raising passing scores will affect the quality of the teaching force, though the overall effect will probably be modest, and this effort will reduce teacher supply. Recent data suggest that disadvantaged students will be either the beneficiaries or those most negatively affected if higher passing scores are implemented. (Contains 33 references.) (SM)

Teaching Policy Handbook

Higher Passing Scores on Teacher Licensure Examinations

Analyst: **Andrew J. Wayne**



April 1999

DO NOT CITE WITHOUT PERMISSION OF THE AUTHORS

The National Partnership for Excellence and Accountability in Teaching (NPEAT) was a voluntary association of 29 national organizations. NPEAT engaged in collaborative, research-based action to achieve teaching excellence and raise student performance.

Primary funding for NPEAT comes from the Office of Educational Research and Improvement, U.S. Department of Education under contract number RD97124001. The positions expressed here do not necessarily reflect the positions or policies of the Office of Educational Research and Improvement or the U.S. Department of Education.

Teaching Policy Handbook

Higher Passing Scores on Teacher Licensure Examinations

Working Paper, October 11, 1999

Analyst: Andrew J. Wayne

Preface

The Handbook project will provide reliable knowledge about teacher quality proposals that are drawing attention from state and local lawmakers. For each proposal, the project uses the following framework to organize evidence:

1. Definition or explanation of the policy
2. Rationale
3. Trends
4. Evidence
5. Key Issues

In addition, individuals who can introduce additional evidence to the analyses will have the opportunity to do so. The working paper that follows demonstrates how the

framework would organize evidence about setting higher passing scores on teacher licensure examinations.

Proposal

Establish higher passing scores on teacher licensure examinations.

Definition or explanation of the policy

Most states require prospective teachers to pass pencil and paper tests as a condition for licensure. States can adjust the difficulty of these tests by raising or lowering the passing score. Raising the passing score is seen by some as a means to promote higher quality teaching.

Rationale

Tests can presumably determine whether prospective teachers possess knowledge that enables better teaching. Advocates for raising cutoff scores on tests taken by teacher candidates assert that this will screen out applicants lacking necessary knowledge and thereby raise teacher effectiveness and better ensure quality teaching for all schoolchildren. Justifications for the five types of tested knowledge appear below.

Subject Matter. Teachers use college-level subject matter knowledge for day-to-day decisions, such as how to design lessons, as well as immediate decisions about how to respond to student questions. Teachers also use subject matter knowledge to design curricula with proper emphases and attention to disciplinary concepts essential to continued student learning. Simply put, one cannot teach what one does not know.

Pedagogy. Decades of academic research have produced a body of knowledge about how to teach. Teachers who possess pedagogical knowledge can apply it in ways that enhance student learning.

Subject Matter or "Content" Pedagogy. Much knowledge about how to teach is specific to the particular subject taught. Teachers need subject-specific pedagogical knowledge to effectively teach a subject.

Basic Skills. A competent teacher must be able to read college-level material, write cogent essays in correct English, and perform at least those mathematical tasks required of high school students. These skills are fundamental to the job of teaching.

General Knowledge. Teachers who possess broad, general knowledge can help students learn by connecting content from various disciplines. Furthermore, teachers with more general knowledge can better respond to students' broad ranging curiosity and interests.

Trends

[Identify states that recently raised or considered raising cutoff scores and related developments.]

Evidence

The courts have not required states to show that students learn more from teachers with higher scores on licensure examinations.[1] Nevertheless, a small handful of sophisticated studies of student achievement have assessed the importance of teachers' tested knowledge. Six studies are longitudinal – involving pre- and post testing – and control for students' socioeconomic background and other likely influences on student achievement.

Collectively, these studies show that teachers' tested knowledge is an important quality indicator. Although the earliest study raised suspicion about the National Teachers Examination, subsequent student achievement studies demonstrated positive relationships between student learning and other measures of teachers' tested knowledge.

The National Teachers Examination. Until recently the National Teachers Examination (NTE) was the most widely used licensure examination. Specialized teachers took a two-hour NTE Area Examination covering both content and teaching methods specific to one of about thirty specialties. Other teachers took the NTE Common Examinations, which consumed 3 hours and 15 minutes and included the following four component tests: (1) general principles of pedagogy and psychological and social foundations of education, (2) written English expression, (3) social studies, literature, and fine arts, and (4) science and mathematics.

The one student achievement study involving teachers' NTE scores was conducted by Anita Summers and Barbara Wolfe (1977, 1975). The authors sampled students in Philadelphia schools in 1971. Among the 627 sixth graders in their elementary school sample, students learned *less* when their teachers scored higher on the NTE Common Examinations. The junior high and high school samples employed teachers' NTE Area Examination scores and yielded indeterminate results.

Summers and Wolfe had controlled for other influences extensively, so their findings reinforced other evidence about the NTE. In studies of relationships among NTE scores, teachers' supervisor ratings, and observer ratings, correlations that did exist were modest, and a few negative relationships arose (Andrews et al., 1980; Ayers and Qualls, 1979; Lovelace and Martin, 1984; Quirk et al., 1973).

Verbal Skills. Three subsequent student achievement studies employed tests of teachers' verbal skills. All three yielded reasonably consistent, positive findings.

Two of the verbal skills studies are reanalyses of old data sets which included teachers' verbal facility scores. Eric Hanushek (1992) reexamined data from an early 1970s welfare reform experiment and found that teacher "word test" scores had affected elementary students' reading score gains – though not their vocabulary gains (see also Murnane and Phillips, 1981).

The second reanalysis employed the Coleman report data -- a very large cross-sectional data set from the mid-1960s. Ronald Ehrenberg and Dominic Brewer (1995) constructed artificial gain scores by subtracting schools' lower grade-level achievement averages from higher grade-level achievement averages. Teachers' scores on a short verbal facility test explained some school-to-school variation in the

gain scores.

The third verbal skills study employed the reading skill component of a basic skills test. Ronald Ferguson (1991, 1998) took advantage of scores available for over half of all Texas teachers tested in statewide teacher testing in 1986 (see also Kain and Singleton, 1996).

Ferguson computed mean teacher test scores for every district. Data on students were also aggregated to the district level. The author computed the difference, for each district, between the mean achievement score of all third graders in 1986 and the mean score of all seventh graders in 1990 -- the same cohort of children if migration is neglected. The analysis revealed that a district was more likely to have a higher difference when the district mean for teachers' test scores was higher. Ferguson later reinforced these findings by showing that the gains of each district's elementary students differed from the gains of its secondary students, depending on the reading proficiency difference between the district's elementary teachers and its secondary teachers.

Two additional well-designed studies appeared recently. The first study took advantage of teachers' responses to a single multiple choice mathematics test item. Rowan, Chiang, and Miller (1997) analyzed nationally representative achievement data from the National Educational Longitudinal Study of 1988. The teacher survey included a single, high school level mathematics test item. The researchers found that students whose teachers answered the item correctly posted larger mathematics gains between eighth and tenth grade.

A second recent study capitalized on Alabama personnel records, some of which contained teachers' ACT College entrance examination test scores. Ronald Ferguson and Helen Ladd (1996) used composite ACT scores. These scores combined English, mathematics, social studies reading, and natural sciences reading components (American College Testing, 1989).

Their initial analysis showed that student reading score gains from third grade to fourth grade were positively related to the average teacher ACT score at the students' school. The relationship was unclear for math score gains. Their second analysis used district level averages from 127 school districts. The authors created artificial gain scores using achievement data from third and fourth graders, and eighth and ninth graders, all for the 1990-91 school year. The average differences in each district between the math scores of younger students and older students were positively related to the average teacher ACT score in the district.

Summary. The evidence shows that tested knowledge is a relevant indicator of teacher quality. The research findings do not map cleanly onto currently used licensure tests, but Tables 1 and 2 show that basic skills tests and subject matter tests are reasonably well supported.

Table 1 tentatively connects the determinate findings involving elementary students to the three types of licensure examinations for elementary teachers. Table 2 depicts the same tool for secondary students and teachers.

The connections drawn do involve some leap-of-faith generalizations. The reading component of the TECAT, for example, represents only one component of a typical basic skills test. Conversely, the connections also leap from general to

particular. The Ferguson and Ladd study that used teachers' ACT composite scores, for example, connects to three columns of Table 2. We cannot know whether Alabama's secondary school students learned more because of teachers' knowledge of the subject taught, general knowledge, or basic skills.

Table 1
Determinate findings on teachers' tested knowledge: Elementary school teachers

Study	*Measure of Teachers' Knowledge*	*Licensure test type*		
		Subjects/ General	Basic Skills	Pedagogy
Summers and Wolfe	NTE Common		-	-
Ferguson	TECAT reading		+	
Hanushek	Word test		+	
Ehrenberg and Brewer	EEOC teacher verbal facility test		+	
Ferguson and Ladd	ACT composite		+	

Table 2
Determinate findings on teachers' tested knowledge: Secondary school teachers

Study	*Measure of Teachers' Knowledge*	*Licensure test type*				
		Subject	Subject-specific pedagogy	General	Basic Skills	Pedagogy
Summers and Wolfe	NTE Area					
Ehrenberg and Brewer	EEOC teacher verbal facility test				+	
Ferguson and Ladd	ACT composite			+	+	
Rowan et al	NELS math item				+	

Key Issues

Effect on Teacher Quality. Raising passing scores will affect the quality of the teaching force, but the overall effect is probably modest. Teachers just above the old passing score will presumably be replaced by teachers just above the new passing score. In addition, student achievement studies show that the overall contribution of teachers to student learning is much larger than teachers' tested knowledge can explain (Goldhaber & Brewer, 1997; Hanushek, 1995; Hedges, et al., 1996; Murnane, 1975, Murnane and Phillips, 1981; Rivkin, et al., 1998; Sanders and Rivers, 1996).

Effect on Teacher Supply. Raising passing scores necessarily reduces supply. Some evidence from the Educational Testing Service (ETS) helps gauge that effect. Gitomer et al. (1999) created a data set using tens of thousands of licensure test records collected from 1994 through 1997. If every state using Praxis I[2] had applied the cutoff score from the state with the highest cutoff score, the pool of test passers would have shrunk by one-third. The same analysis performed for Praxis II[3] shows that the pool of Praxis II passers would have shrunk by one quarter. The authors also described the pool of passers according to its racial makeup. In the Praxis I and Praxis II analyses, minority representation in the pool of test passers fell by almost one third, and by one fifth, respectively.

The ETS estimates of the effects of higher passing scores on teacher supply may overstate the impact. With higher cutoff scores actually in place, low scoring aspirants have strong incentives to acquire the tested knowledge. Nonetheless, substantially raising pass scores on teacher tests will have a big effect on the number of people licensed to teach. The reduction in the pool of teacher candidates will be particularly great among persons who are African American or Latino.

Effect on Disadvantaged Schoolchildren. In the past, low-income and minority students were the most likely to have teachers with low tested knowledge (Ehrenberg & Brewer, 1995; Summers and Wolfe, 1976). Those students therefore had the most to gain if higher pass scores were to result in the recruitment and retention of teachers with teachers who did well on such tests. However, low income and minority students have the most to lose from licensure test cutoffs if those individuals who then teach them have no preparation to teach or are unlicensed and have even lower verbal ability and content knowledge.

Recent tabulations indicate that disadvantaged schoolchildren will be either the beneficiaries or those most negatively affected if higher passing scores are implemented. Low-scoring teachers remain concentrated with both minority and low-income schoolchildren (Ferguson, 1998; Kain and Singleton, 1996; Rowan et al., 1997). In addition, matched records from the Educational Testing Service show that teachers who pass the Praxis II examination for special education have markedly lower SAT scores than teachers who pass other Praxis II examinations (Gitomer et al., 1999).

Concluding Observations

It seems clear that states will continue to increase the passing scores on various tests of teacher knowledge. Of the tests now in use, we have evidence that subject matter and verbal skills TESTS are related to teacher effectiveness. This does not mean that other tests do not capture aspects of effective teaching, only that we cannot say from existing research that this is so.

While students whose teachers have high verbal skills and know their subject matter are more likely to achieve at somewhat higher levels than students whose teachers do less well on relevant tests, EXISTING EVIDENCE IS NOT SUFFICIENT to determine what the gain in student achievement would be as a result of a given increase in the passing score.

It is possible that the quality of the teaching corps would actually decline in hard-to-staff schools as passing scores are raised above current levels. Districts, confronted with fewer licensed applicants may persuade states to issue emergency permits. Such declines could be avoided I passing scores

increases were complemented with salary increases or increases in other incentives to teach.

The fact that pencil and paper tests of knowledge are weak, if positive in some cases, indicators, has led an increasing number of people wanting to raise the quality to urge that prospective candidates be required to. Such performance assessment would include tests of knowledge but the overall score of a prospective teacher would depend on several measures, including assessments of actual classroom teaching performance.

References

Alexander, K. L., Entwisle, D. R., & Thompson, M. S. (1987). School performance, status relations, and the structure of sentiment: Bringing the teacher back in. American Sociology Review, 52, 665-682.

American College Testing, Inc. (1989). The enhanced ACT assessment: Using concordance tables. Iowa City, Iowa: Author.

Andrews, J. W., Blackmon, C. R., & Mackey, J. A. (1980). Preservice performance and the National Teacher Examinations. Phi Delta Kappan, 61(5), 358-359.

Ayers, J. B., & Qualls, G. S. (1979). Concurrent and predictive validity of the National Teacher Examinations. Journal of Educational Research, 73(2), 86-92.

Carlos, L., & Howell, P. (1999). Class size reduction in California, 1996-98: Early findings signal promise and concerns. Palo Alto, CA: EdSource, Inc.

Educational Testing Service. (1999a). Setting passing scores: The need for justifiable procedures. Princeton, New Jersey: Author, Teaching and Learning Division.

Educational Testing Service. (1999b). Significant decisions in testing litigation. Princeton, New Jersey: Author, Teaching and Learning Division.

Educational Testing Service. (1999c). Validity for licensing tests. Princeton, New Jersey: Author, Teaching and Learning Division.

Ehrenberg, R. G., & Brewer, D. J. (1995). Did teachers' verbal ability and race matter in the 1960s?: *Coleman* revisited. Economics of Education Review, 14(1), 1-21.

Ehrenberg, R. G., Goldhaber, D. D., & Brewer, D. J. (1995). Do teachers' race, gender, and ethnicity matter?: Evidence from the National Educational Longitudinal Study of 1988. Industrial and Labor Relations Review, 48(3), 547-561.

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

Ferguson, R. F. (1998). Can schools narrow the black-white test score gap? In C. Jencks, & M. Phillips (Eds.), The black-white test score gap (pp. 318-374). Washington, DC: Brookings Institution.

Ferguson, R. F., & Ladd, H. F. (1996). How and why money matters: an analysis of Alabama schools. H. F. Ladd (Ed.), Holding schools accountable: performance-based reform in education (pp. 265-298). Washington, D.C.: Brookings Institution.

Gitomer, D. H., Latham, A. S., & Ziomek, R. (1999). The academic quality of prospective teachers: the impact of admissions and licensure testing. Princeton, New Jersey: Educational Testing

Service.

Goldhaber, D. D., & Brewer, D. J. (1997). Evaluating the effect of teacher degree level on educational performance. In W. J. Fowler (Ed.), Developments in School Finance, 1996 (pp. 197-210). Washington, D.C.: National Center for Education Statistics, U.S. Department of Education.

Goldhaber, D. D., & Brewer, D. J. (1999). Teacher licensing and student achievement. In M. Kanstoroom, & C. E. Finn (Eds.), Better teachers, better schools (pp. 83-102). Washington, D.C.: Thomas B. Fordham Foundation.

Hanushek, E. A. (1992). The trade-off between child quantity and quality. Journal of Political Economy, 100(1), 85-117.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (1999). Do higher salaries buy better teachers? National Bureau of Economic Research Working Paper 7802.

Henke, R. R., Choy, S. P., Chen, X., Geis, S., & Alt, M. N. (1997). America's teachers: Profile of a profession, 1993-94. Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Kain, J. F., & Singleton, K. (1996). Equality of educational opportunity revisited. New England Economic Review, 87-111.

Kennedy, M. M. (1999). The problem of evidence in teacher education. In R. A. Roth (Ed.), The role of the university in the preparation of teachers (pp. 87-107). Philadelphia: Falmer.

Lovelace, T., & Martin, C. E. (1984). The revised National Teacher Examinations as a predictor of teachers' performance in public school classrooms. Lafayette, LA: University of Southwestern Louisiana.

Mitchell, R., & Barth, P. (1999). How teacher licensing tests fall short. Washington, D.C.: The Education Trust.

Murnane, R. J. (1975). The impact of school resources on the learning of inner city children. Cambridge, MA: Ballinger.

Murnane, R. J., & Phillips, B. R. (1981). What do effective teachers of inner-city children have in common? Social Science Research, 10, 83-100.

Perkins, L. M. (1989). The history of blacks in teaching: growth and decline within the profession. In D. Warren (Ed.), American teachers: histories of a profession at work (pp. 344-369). New York: MacMillan.

Quirtk, T. J., Witten, B. J., & Weinberg, S. F. (1973). Review of studies of concurrent and predictive validity of the National Teacher Examinations. Review of Educational Research, 43, 89-114.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (1998). Teachers, schools, and academic achievement. National Bureau of Economic Research Working Paper Number 6691.

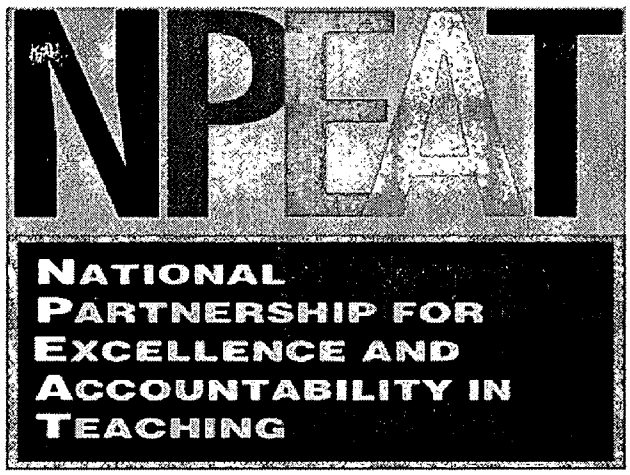
Rowan, B., Chiang, F., & Miller, R. J. (1997). Using research on employees' performance to study the effects of teachers on students' achievement. Sociology of Education, 70(October), 256-284.

Sanders, W. L., & Rivers, J. C. (1996). Cumulative and residual effects of teachers on future academic achievement. University of Tennessee Value-Added Research and Assessment Center.

Summers, A. A., & Wolfe, B. L. (1975). Equality of educational opportunity quantified: a production function approach. Philadelphia Fed Research Papers.

Summers, A. A., & Wolfe, B. L. (1976). Intradistrict distribution of school inputs to the disadvantaged: Evidence for the courts. Journal of Human Resources, 11, 328-342.

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



University of Maryland, College Park
College of Education
College Park, MD 20742

Phone: (301) 405-2341 ♦ Fax: (301) 405-3573

Website: **WWW.NPEAT.ORG**

[1] Defendants in employment discrimination cases must provide evidence that employment criteria are job-related. For teacher licensure examinations, documentation from a proper test development process suffices. At the Educational Testing Service, for instance, developers analyze job requirements by surveying practicing educators. These analyses inform the efforts of committees of experienced teachers and teacher-educators who develop test specifications. Finally, committee members and other teachers collaborate with subject matter and test development experts to specify actual questions. The courts also require information about how passing scores are set. For more information, see Educational Testing Service (1999a, 1999b, 1999c).

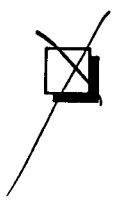
[2] Praxis I is a basic skills test developed by the Educational Testing Service and currently used by about twenty states.

[3] Praxis II is a licensure test developed by the Education Testing Service. Praxis II examinations cover content and pedagogical knowledge in a specific subject field.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

REPRODUCTION BASIS



This document is covered by a signed "Reproduction Release (Blanket) form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").