

tionships; flipping between micro and macroscopic lenses. But it also means patience as AC programs produce larger numbers for longitudinal studies, and waiting as poorly constructed or unsupported programs fold.

The very idea of circumventing traditional educational training is threatening in some quarters. Alternative certification will definitely have an impact on many areas of education; not the least of which are colleges of education. Does this herald a period of creativity and entrepreneurship at the teacher college level as more districts, college students, and career changers seriously consider training options?

Shifting from debating AC to researching AC could have a healthy effect across all of education.

The age-old debate about whether teaching is a profession, craft, vocation, art, etc. was carefully avoided in this report. It is practically medieval to keep this discourse alive—let it die a natural death and get on with something more useful. Untangling the “can of worms” built into AC nomenclature and chaotic requirements would be an effective first step in establishing an AC line of research. Then beyond conceptual clarification, recent work in the area of educational standards may supply guidance in AC program construction and evaluation. Shifting from debating AC to researching AC could have a healthy effect across all of education. Alternate forms of teacher preparation encourage educators and schools to think outside the envelope when confronted with numerous enrollment crises. The introduction of new approaches to teacher induction and training could become a renaissance for the types of rigorous research needed at all levels of education.

Alternative certification has trouble getting professional respect partly because it has not become part of a national education agenda. Its successes, if they occur, are a matter for local school-district celebration. It looks to almost operate in secret. Alternative certification seems to be an attempt to fashion a modern solution to long neglected but clearly evident problems. It seems rather doubtful if AC could ameliorate the problems that confront so many of our schools; its implementation level remains the classroom. But if we take what has been done with alternative certification over the last decade, could not a parallel program be constructed for alternatively certified school principals? Career changers may be interested at working at other levels of school governance besides the classroom. As Schlechty and Vance noted, “It is time to accept that the quality of teaching personnel is unlikely to substantially improve until the quality of managers improves.”¹⁰⁵

Appendix A: State Alternative Certification Programs¹⁰⁶ (Categories A-H are described on page 216)

	A	B	C	D	E	F	G	H
AL								
AK								
AZ								
AR								
CA								
CO								
CT								
DE								
DC								
FL								
GA								
HI								
ID								
IL								
IN								
IA								
KS								
KY								
LA								
ME								
MD								
MA								
MI								
MN								
MS								
MO								
MT								
NE								
NV								
NH								
NJ								
NM								
NY								
NC								
ND								
OH								
OK								
OR								

	A	B	C	D	E	F	G	H
PA				•				
RI								
SC			•					
SD				•				
TN				•		•	•	
TX	•							
UT			•				•	•
VT			•			•	•	
VA	•							
WA	•	•						•
WV	•					•	•	
WI				•		•		
WY						•		

A longer version of this report was published by the Tomás Rivera Policy Institute in Claremont, California in 1998. The Institute can be reached at 901-621-8897.

- ¹ E. Feistritzer and D. Chester, *Alternative Teacher Certification* (Washington D.C.: National Center for Education Information, 1991).
- ² M. Haberman, "Proposals for Recruiting Minority Teachers: Promising Practices and Attractive Detours," *Journal of Teacher Education* 39 no. 4 (1986): 38-44.
- ³ E. Feistritzer and D. Chester, *Alternative Teacher Certification* (Washington D.C.: National Center for Education Information, 1996), 25.
- ⁴ T. Bliss, "Alternative Certification in Connecticut: Reshaping the Profession," *Peabody Journal of Education* 67 no. 3 (1990): 39-40.
- ⁵ G. Crow, L. Levine and N. Nager, "No More Business as Usual: Career Changers Who Become Teachers," *American Journal of Education* 98 no. 3 (1990): 204.
- ⁶ C. Stevens and M. Dial, "A Qualitative Study of Alternately Certified Teachers," *Education and Urban Society* 26 no. 1 (1993): 65.
- ⁷ D. Kerr, "Teaching Competence and Teacher Education in the United States," *Teachers College Record* 84 no. 3 (1983): 525-552.
- ⁸ F. Lutz and J. Hutton, "Alternative Teacher Certification: Its Policy Implications for Classroom and Personnel Practices," *Educational Evaluation and Policy Analysis* 11 no. 3 (1989): 237-254.
- ⁹ J. Boser and P. Wiley, "An Alternative Teacher Preparation Program: Is the Promise Fulfilled?" *Peabody Journal of Education* 65 no. 2 (1988): 138.
- ¹⁰ E. Feistritzer, "National Overview of Alternative Teacher Certification," *Education and Urban Society* 26 no. 1 (1993): 22.
- ¹¹ L. Huling-Austin, "Research on Learning to Teach: Implications for Teacher Induction and Mentoring Programs," *Journal of Teacher Education* 43 no. 3 (1986): 52, citing others.
- ¹² M. Haberman, "Proposals for Recruiting Minority Teachers: Promising Practices and Attractive Detours," *Journal of Teacher Education* 40 no. 5 (1988): 38.
- ¹³ E. Feistritzer, "National Overview of Alternative Teacher Certification," 24.
- ¹⁴ A. Smith, "The Critical Issue of Excellence and Equity in Competency Testing," *Journal of Teacher Education* 35 no. 2 (1984): 6-9.
- ¹⁵ T. Stoddart, "Los Angeles Unified School District Intern Program: Recruiting and Preparing Teachers for an Urban Context," *Peabody Journal of Education* 67 no. 3 (1990): 84-122.
- ¹⁶ W. Houston, F. Marshall, and T. McDavid, "Problems of Traditional Prepared and Alternately Certified First-Year Teachers," *Education and Urban Society* 26 no. 1 (1993): 3-34.
- ¹⁷ L. Cornett, "Alternative Certification: State Policies in the SREB States," *Peabody Journal of Education* 67 no. 3 (1990): 55-83.
- ¹⁸ Texas Education Agency, *Alternative Teacher Certification in Texas* (Austin: Author, 1990).
- ¹⁹ E. Feistritzer, "National Overview of Alternative Teacher Certification," 18-28.

- 20 Cornett, "Alternative Certification," 73.
- 21 See T. Stoddart, "Los Angeles Unified School District Intern Program," 84-122 for numerous demographic citations.
- 22 Ibid., 95.
- 23 Ibid.
- 24 T. Stoddart, "Who is Prepared to Teach in Urban Schools?" *Education and Urban Society* 26 no. 1 (1993): 29-43.
- 25 T. Stoddart, "Los Angeles Unified School District Intern Program," 100-101.
- 26 Feistritzer and Chester, *Alternative Teacher Certification* (1996); G. Natriello and K. Zumwalt, "New Teachers for Urban Schools? The Contribution of the Provisional Teacher Program in New Jersey," *Education and Urban Society* 26 no. 1 (November 1993): 49-62.
- 27 N. Dessel and G. Mehaffy, "Navy Fast Track Program," speech at the annual conference of the American Association of Colleges of Teacher Education, February 20, 1989; Gursky, 1992; D. Keltner, "Troops to Teachers: Alternative Certification in the Military," *The Educational Forum* 58 no. 2 (1994): 182-186; H. McCree, "Post Career Military as an Alternative Pool of Teacher Candidates," paper presented at the annual meeting of the Mid-South Educational Research Association, New Orleans, LA, 10-12 November, 1993; S. Nyjordet, *Teaching as a New Career for Military Personnel* (Pensacola, Fla: Naval Education and Training Program Management Support, 1991); and T. Taylor, "New to the Ranks: Moving From the Military to Teaching," [online], ERIC Digest: ERIC Documents Reproduction Service, Doc. No. ED370937 (1994).
- 28 E. Dooley and L. Savage, "Encouraging Diversity in Rural Classrooms," paper presented at the National Alternative Certification Conference, 22 April 1994, Washington, D.C.
- 29 M. E. Apodaca, "Improving Foreign Language Teaching Through Innovative Recruitment and Training," in J. F. Lalande, ed., *Shaping the Future of Language Education: FLES, Articulation, and Proficiency*, report of the Central States Conference on the Teaching of Foreign Language, 1988; L. Diaz-Rico, "Recruiting and Retaining Bilingual Teachers: Cooperative School-Community-University Model," *Journal of Educational Issues of Language Minority Students* 14 (1994): 4-17.
- 30 E. Baron and T. Rusnak, *Math and Science Professionals are Becoming Teachers: The Duquesne Model*, paper presented at the summer workshop of the National Association of Teacher Educators, Towson, Md., 5-8 August 1990; L. Cornett, "Serious Shortages of Science and Mathematics Teachers: What SREB States Are Doing," *SREB Regional Spotlight*, 1986.
- 31 K. Gunderson and B. Karage, "Easing the Special Education Teacher Shortage: Are Emergency Permits the Answer?" *Teacher Education Quarterly* 19 no. 3 (1992): 79-80; F. Ikei and B. Hoga, "Respecialization in Special Education (RISE) Program: An Alternative Certification Program," Hawaii State Department of Education, Honolulu, ERIC Document Reproduction Service, Doc. No. ED381944, 1995; B. Ludlow and W. Wienke, "Alternative Certification in Special Education: A Qualitative Study of Two Models," paper presented at the Annual National Conference of the American Council on Rural Special Education, Austin, Tex., 23-26 March 1994; M. Rosenberg and E. Rock, "Alternative Certification in Special Education: Efficacy or a Collaborative, Field-based Teacher Preparation Program," *Teacher Education and Special Education* 17 no. 3 (1994): 141-153; and M. Sullivan, "Stemming Teacher Turnover in Rural Settings: An Alternative Approach," paper presented at the Annual National Conference of the American Council of Rural Special Education, Austin, Tex., 23-26 March 1994.
- 32 S. Kirby, L. Darling-Hammond, and L. Hudson, "Nontraditional Recruits to Mathematics and Science Teaching," *Educational Evaluation and Policy Analysis* 11 no. 3 (1989): 309.
- 33 Ibid., 320.
- 34 Ibid., 321.
- 35 W. Hawley, "The Theory and Practice of Alternative Certification Programs: What Can Be Learned From Teacher Induction Research?" *Action in Teacher Education* 8 (1990): 21.
- 36 Ibid.
- 37 M. McKibbin, "Alternative Certification Programs," *Education Leadership* 46 no. 3 (1988): 32-35.
- 38 E. Guyton, M. Fox, and K. Sisk, "Comparison of Teaching Attitudes, Teaching Efficacy, and Teacher Performance of First year Teachers Prepared by Alternative and Traditional Teacher Education Programs," *Action in Teacher Education* 13 no. 2 (1991): 1-9.
- 39 V. Dill, "Teacher Education in Texas: A New Paradigm," *The Educational Forum* 58 no. 2 (1994): 147-154.
- 40 W. Hawley, "The Theory and Practice of Alternative Certification Programs," 11.
- 41 P. Schlechty and V. Vance, "Recruitment, Selection, and Retention: the Shape of the Teaching Force," *The Elementary School Journal* 83 no. 4 (1983): 469-487.
- 42 G. Adams and M. Dial, "Teacher Survival: A Cox Regression Model," *Education and Urban Society* 26 no. 1 (1993): 97.
- 43 M. McKibbin, "Alternative Certification in California," *Teacher Education Quarterly* 15 no. 3 (1988): 49-59.
- 44 T. Stoddart, "Los Angeles Unified School District Intern Program," 101.
- 45 S. Kirby, L. Darling-Hammond, and L. Hudson, "Nontraditional Recruits," 301-323.
- 46 L. Darling-Hammond, "Teaching and Knowledge: Policy Issues Posed by Alternative Certification for Teachers," *Peabody Journal of Education* 67 no. 3 (1990): 123-54.
- 47 Ibid., 123-54.
- 48 P. Schlechty and V. Vance, "Recruitment, Selection, and Retention," 484.

- 49 E. Feistritzer, "National Overview of Alternative Teacher Certification," 24.
- 50 G. Natriello and K. Zumwalt, "New Teachers for Urban Schools?" 49-62.
- 51 M. Haberman, "Proposals for Recruiting Minority Teachers," (1988), 39.
- 52 M. Haberman, "Proposals for Recruiting Minority Teachers," (1986), 17
- 53 T. Stoddart, "Los Angeles Unified School District Intern Program," 84-122.
- 54 Ibid., 101
- 55 T. Stoddart, "Who is Prepared to Teach in Urban Schools?" 1993, 46-47
- 56 M. Dial and C. Stevens, "The Context of Alternative Teacher Certification," *Education and Urban Society* 26 no. 1 (1993): 14.
- 57 J. Boser and P. Wiley, "An Alternative Teacher Preparation Program," 140.
- 58 W. Hawley, "The Theory and Practice of Alternative Certification Programs," 3-34.
- 59 G. Fenstermacher, "The Place of Alternative Certification in the Education of Teachers," *Peabody Journal of Education* 67 no. 3 (1990): 155-185.
- 60 C. Everston, W. Hawley and M. Zlotnik, "Making a Difference in Educational Quality Through Teacher Education," *Journal of Teacher Education* 36 no. 3 (1985): 2-12; L. Huling-Austin, "Research on Learning to Teach," 173-180.
- 61 T. Stoddart, "Los Angeles Unified School District Intern Program," 117
- 62 W. Hawley, "The Theory and Practice of Alternative Certification Programs," 3-34.
- 63 M. McKibbin, "Alternative Certification in California," 49-59; K. Zumwalt, "Alternate Routes to Teaching: Three Alternative Approaches," *Journal of Teacher Education* 42 no. 2 (1991): 83-92.
- 64 S. Kirby, L. Darling-Hammond, and L. Hudson, "Nontraditional Recruits," 301-323.
- 65 D. Watts, "Alternative Routes to Teacher Certification: A Dangerous Trend," *Action in Teacher Education* 8 no. 2 (1986): 25-29.
- 66 F. Lutz and J. Hutton, "Alternative Teacher Certification," 237-254.
- 67 G. Natriello and K. Zumwalt, "New Teachers for Urban Schools?"
- 68 J. Smith, "The Alternative Route: Flaws in the New Jersey Plan," *Educational Leadership* 49 no. 3 (1991): 32-36.
- 69 W. Hawley, "The Theory and Practice of Alternative Certification Programs," 22; citing G. Natriello, A. Hansen, A. Frisch, and K. Zumwalt, *Characteristics of Entering Teachers in New Jersey*. (Unpublished manuscript. Teachers College, Columbia University, New York, 1990).
- 70 T. Stoddart, "Los Angeles Unified School District Intern Program," 112.
- 71 M. McKibbin, "Alternative Certification Programs," 32-35.
- 72 N. Adelman, J. Bogart, and J. Michie, *An Exploratory Study of Teacher Alternative Certification and Retraining Programs*. (Washington, D.C.: U.S. Department of Education, 1987).
- 73 B. Rodman, "'Alternate Route' Said a Success," *Education Week* 7 (February 24, 1988).
- 74 T. Bliss, "Alternate Certification in Connecticut: Reshaping the Profession," *Peabody Journal of Education* 67 no. 3 (1990): 48; citing J. Love, W. Cooley and J. Trudeau, *First Year Evaluation of Connecticut's Alternative Route to Certification*. (Hampton, N.H.: RC Research Corporation, 1989)
- 75 F. Lutz and J. Hutton, "Alternative Teacher Certification," 251.
- 76 L. Cornett, "Alternative Certification 1990, 55-83; P. Hawk and M. Schmidt, 53-58; and W. Houston, F. Marshall, and T. McDavid, "Problems of Traditional Prepared and Alternatively Certified First-Year Teachers," 78-89.
- 77 F. Lutz and J. Hutton, "Alternative Teacher Certification," 237-254.
- 78 L. Cornett, "Alternative Certification," 1990, 75.
- 79 L. Darling-Hammond, "Teaching and Knowledge" 123-154.
- 80 L. Cornett, "Alternative Certification," 55-83.
- 81 W. Hawley, "The Theory and Practice of Alternative Certification Programs," 17.
- 82 F. Lutz and J. Hutton, "Alternative Teacher Certification," 237-254.
- 83 M. Fetter, "Staffing Up and Dropping Out: Unintended Consequences of High Demand for Teachers," *Education Policy Analysis Archives* [On-line serial] 5 no. 16 (1997): 1-23. Available at <http://olam.ed.asu.edu/epaa/v5n16.html>.
- 84 National Center for Education Statistics, *Projections of Education Statistics to 2006*. (Washington, D.C.: NCES, 1996).
- 85 National Center for Education Statistics, *Digest of Education Statistics*. (Washington, D.C.: NCES, 1996).
- 86 G. Fenstermacher, "The Place of Alternative Certification in the Education of Teachers," 181.
- 87 L. Darling-Hammond, "Teaching and Knowledge," 123-154.
- 88 Ibid., 144.
- 89 M. Haberman, "Proposals for Recruiting Minority Teachers," 38-44.
- 90 M. Fetter, "Where Have All the Teachers Gone?" *Education Policy Analysis Archives* [online serial] 5 no. 2 (1997): 1-14.
- 91 L. Huling-Austin, "Research on Learning to Teach," 51-58 vs. M. Fetter, "Where Have All the Teachers Gone?" 1-14.

- 92 Charles Kerchner, "Shortages and Gluts of Public School Teachers: There Must Be a Policy Problem Here Somewhere," *Public Administration Review* 44 no. 4 (1984): 297. Available: <http://olam.ed.asu.edu/epaa/v5n2.html#1>.
- 93 Ibid.
- 94 Ibid., 297.
- 95 J. Fox, "Restructuring the Teacher Work Force to Attract the Best and the Brightest," *Journal of Education Finance* 10 (1984): 346-353; C. Kerchner, "Shortages and Gluts of Public School Teachers," 292-298; D. Kerr, "Teaching Competence and Teacher Education in the United States," 525-552; B. Oliver and M. McKibbin, "Teacher Trainees: Alternative Credentialing in California," *Journal of Teacher Education* 36 no. 3 (1985): 20-23; P. Schlechty and V. Vance, "Recruitment, Selection, and Retention," 469-487; and W. Weaver, "In Search of Quality: For Talent in Teaching," *Phi Delta Kappa* 61 no. 1 (1979): 29-46.
- 96 C. Book, D. Freeman, and B. Brousseau, "Comparing Academic Backgrounds and Career Aspirations of Education and Non-Education Majors," *Journal of Teacher Education* 36 no. 3 (1985): 27-30; D. Olsen, "The Quality of Prospective Teachers vs. Noneducation Graduates," *Journal of Teacher Education* 36 no. 5 (1985): 56-59; R. Roth, "Alternate and Alternative Certification: Purposes, Assumptions, Implications," *Action in Teacher Education* 8 (1986): 1-6; J. Westerman, "Minimum State Teacher Certification Standards and Their Relationships to Effective Teaching: Implications for Teacher Education," *Action in Teacher Education* 11 (1989): 25-32.
- 97 K. Zeichner and B. Tabachnick, "Are the Effects of University Teacher Education 'Washed Out' by School Experience?" *Journal of Teacher Education* 32 no. 3 (1981): 7-11.
- 98 The conditions of minorities in teaching or in teacher preparation are entire fields of study unto themselves. Numerous researchers have offered dozens of suggestions to rectify teacher hiring disparities. (C. Case, R. Shive, K. Ingebretson, and V. Spiegel, "Minority Teacher Education: Recruitment and Retention," *Journal of Teacher Education* 39 no. 4 (1988). B. deMartinez, "Political and Reform Agendas' Impact on the Supply of Black Teachers," *Journal of Teacher Education* 39 no. 1 (1988). J. Gordon, "Why Students of Color Are Not Entering Teaching: Reflections from Minority Teachers," *Journal of Teacher Education* 45 no. 5 (1994). M. Haberman, "Proposals for Recruiting Minority Teachers." E. Middleton, E. Mason, W. Stilwell and W. Parker, "A Model for Recruitment and Retention of Minority Students in Teacher Preparation Programs," *Journal of Teacher Education* 39 no. 1 (1988). J. Zapata, "Early Identification and Recruitment of Hispanic Teacher Candidates," *Journal of Teacher Education*, 30 no. 1 (1988) to name but a few.)
- 99 M. Haberman, "Proposals for Recruiting Minority Teachers," 38.
- 100 C. Case, R. Shive, K. Ingebretson, and V. Spiegel, "Minority Teacher Education: Recruitment and Retention," *Journal of Teacher Education* 39 no. 4 (1988): 57.
- 101 A. Odden, "Financing Educational Excellence," *Phi Delta Kappan* 65 (January 1984): 314, as cited in J. Fox, "Restructuring the Teacher Work Force," 227
- 102 M. Haberman, "Proposals for Recruiting Minority Teachers," 40.
- 103 E. Middleton, E. Mason, W. Stilwell and W. Parker, "A Model for Recruitment and Retention of Minority Students in Teacher Preparation Programs," *Journal of Teacher Education* 39 no. 1 (1988): 14-18.
- 104 J. Gordon, "Why Students of Color Are Not Entering Teaching: Reflections from Minority Teachers," *Journal of Teacher Education* 45 no. 5 (1994): 346.
- 105 P. Schlechty and V. Vance, "Recruitment, Selection, and Retention," 479.
- 106 E. Feistritzer and D. Chester, *Alternative Teacher Certification*, 1996.

Value-Added Assessment: An Accountability Revolution

J. E. Stone

Value-added assessment is a statistical tool for gauging how much students gain in academic achievement in a given year, i.e., how much "value" has been added to the youngsters by their schooling. By aggregating pupil gains by school, value-added assessment can be used to evaluate schools, regardless of differences among entering students. By aggregating scores by teacher, value-added assessment can be used to identify which teacher's students are learning the most and which teacher's students are learning the least. This provides an objective gauge of teacher effectiveness, replacing traditional modes of identifying good teachers via peer review or paper credentials. This report explores a new and sophisticated version of value-added assessment, developed by William Sanders at the University of Tennessee. The author concludes that value-added assessment is revolutionary because it enables parents, taxpayers, and policymakers to see how well schools are doing without penalizing those with many disadvantaged pupils, and it enables teachers to be evaluated based on the most important factor of all: their results.

Introduction

There is widespread agreement that teacher quality is critical to classroom success. A recent study by William Sanders and June Rivers found substantial differences in the achievement gains exhibited by students who have excellent teachers versus those who have ineffective ones.¹ The differences were large enough to shape decisively the subsequent academic careers of the students in question.

The teachers in the Sanders and Rivers study were all fully licensed, as are most teachers in public schools. Plainly, however, there were important differences in their effectiveness, and these differences were due to a variety of factors including their training, the resources of their school or school system, and their particular teaching methodology. Surprisingly, the impact of these factors on student learning has proven difficult to measure and is not well understood.

For example, there are few clear answers to the question of whether training in pedagogical methods prepares teachers to boost student achievement. The quality of teacher preparation has traditionally been assessed by reviewing inputs: whether the teacher scored well or poorly on entrance and licensure exams, whether the training program was fully accredited, whether the teacher had earned a master's degree, whether the teacher underwent specialized training for a particular teaching assignment, whether the teacher's training included certain courses and field experiences, whether the program was adequately staffed, and whether the institution hosting the program was adequately resourced. All of these indicators are presumed to predict how well a teacher can teach. However, the question of whether "well-prepared" teachers do a superior job of boosting student achievement has not been unambiguously answered by research.

Value-added assessment can be used to appraise fairly and accurately school and system performance regardless of differences among entering students.

Value-added assessment is a statistical tool that can provide an objective answer to questions of teacher effectiveness. Technically, it is a method of education data analysis that summarizes annual gains in student achievement. Applied to the aggregate scores of students taught by a given teacher, however, it becomes an indicator of teacher effectiveness—one based not on traditional input indicators such as training and experience but on demonstrated results, i.e., student learning. It is a measure of educational effectiveness that promises to revolutionize education.

Value-added assessment's use as an impartial and objective gauge of teacher effectiveness is not its only virtue. By comparing students' current achievement to their own past performance and aggregating learning gains by school or school system, value-added assessment can be used to appraise fairly and accurately school and system performance regardless of differences among entering students. In fact, it can be used as an objective performance indicator in judgments about matters ranging from teacher licensure, tenure, and merit pay to the effectiveness of curricular innovations and teacher training programs.²

The newest and most sophisticated version of value-added assessment was developed by Sanders at the University of Tennessee.³ It has been used in Tennessee since the early nineties. A slightly different form of value-added assessment—one that uses a different type of statistical analysis—has been used since 1984 by the Dallas (TX) Independent School District.⁴

In both Tennessee and Texas, value-added assessment is currently used to assess school and teacher performance. Instead of judging school or teacher effectiveness on the basis of measures that may have little relationship to results, school board members, lawmakers, parents, and taxpayers in Dallas and in Tennessee can examine the results for themselves.

Value-Added Assessment in the Volunteer State

The Sanders and Rivers study examined value-added data drawn from the Tennessee Value-Added Assessment System (TVAAS). TVAAS is the heart of Tennessee's educational accountability system, and has been in use since the late 1980s. Since 1995, it has been enlarged to produce value-added teacher effectiveness data for review by principals and other school system personnel.

Tennessee tests all its students annually in grades 3-8 with a customized version of McGraw-Hill's Terra Nova instrument. Tennessee's testing program is called the Tennessee Comprehensive Assessment Program (TCAP) and the results are used to inform students, parents, and teachers about individual pupil achievement. The Tennessee Value-Added Assessment System (TVAAS) produces annual reports of the aggregate student achievement gains produced by each Tennessee teacher, school,

Compared to other methodologies for computing pupil achievement gains, mixed model analysis is more precise and less vulnerable to manipulations that can distort results.

and school system. The reports for school systems are broken down by school and grade for each of the five subjects measured by the TCAP exam (math, science, reading, language, and social studies).

TVAAS reports express achievement gains in scale score points and in the form of comparisons with national, state, and district averages. For example, twenty-five points is a typical gain in student math achievement produced in fourth grade by Tennessee schools. The average gain in math produced in fourth grade by schools nationally is twenty-six points. Therefore, the typical Tennessee school is producing gains in fourth-grade math equal to 96 percent of the national average. In fourth-grade science, by contrast, Tennessee schools are producing gains equal to 115 percent of the national average.

For comparisons on which accountability decisions are made, a three-year rolling average is used to assure statistical stability. For example, Washington County's Boones Creek Middle School produced a three-year average gain (1993-95) of sixty-five scale score points in language arts for grades 5-8. The national average gain in language arts for grades 5-8 is fifty points. Thus, Boones Creek Middle school produced gains equivalent to 130 percent of the national average language arts gains.

The TVAAS reports for individual teachers aggregate the gains earned by all students for which a teacher was responsible and compare them to national, state, and local averages. Instead of stating teacher performance as a percentage of the various benchmark averages, teacher averages are classified as "above the norm," "below the norm," or "not detectably different" from the norm. An example of a TVAAS "Teacher Report" is provided in Appendix I.

The statistical analysis employed in Tennessee's value-added assessment is an advanced form of "analysis of variance" known as Henderson's "mixed model." It is described in "The Tennessee Value Added Assessment System" by Sanders, Saxton,

and Horn and in several other sources.⁵ It produces a "best linear unbiased estimate" of the influence on annual student achievement gains attributable to teachers, schools, and school systems.⁶

Compared to other methodologies for computing pupil achievement gains, mixed model analysis is more precise and less vulnerable to manipulations that can distort results. For example, the hierarchical linear regression analysis used in Dallas, TX can exaggerate the differences attributable to a factor such as funding and correspondingly underestimate the differences attributable to teaching effectiveness if a variable such as per-pupil spending is prematurely entered into the statistical adjustment of student gains.

Preconditions

Value-added assessment is statistically robust but the validity of its results depends on certain preconditions. At a minimum, it requires annual testing of students in all grades with a reliable and valid achievement test. Portfolio assessment and other forms of assessment that lack reliability and objectivity will not suffice. Neither will standardized-achievement tests that have been revised to enhance their marketability to educators at the expense of diminished academic content. No amount of analysis can transform the substance and meaning of fundamentally flawed data. Perhaps this limitation is best expressed in the statistician's time-honored adage "Garbage in, garbage out."

Yet nothing in the use of value-added assessment precludes teachers from also using portfolios or any other form of assessment they favor. Most educators believe that schooling should serve aims beyond those that can be measured by achievement tests, and so they prefer a variety of assessments. Parents and the public are not necessarily opposed to these broader aims, but they do disagree with the vast majority of educators about priorities. Whereas educators typically view measured academic achievement as only one outcome among many, parents, taxpayers, and policymakers view it as the indispensable core of student (and teacher) results. No matter what other benefits good schooling may produce, those who fund schools and enroll their children in them will not be satisfied if the gains shown on objectively measured academic achievement are insufficient. Like an annual audit conducted by an external auditor, value-added assessment is an objective means whereby the public can see whether its priorities are being respected and its hopes fulfilled.

In addition to requiring the annual use of a valid and reliable achievement test, value-added assessment requires that the items used in each test cycle be fresh, non-redundant, and tied to an underlying scale. The forms used at each grade level must include a sufficiently wide range of items such that "ceiling" and "floor" effects are highly unlikely. Also, scores must be reported on a common scale that spans the range of grades for which the test is appropriate.

Empirical studies of value-added assessment have demonstrated that it does remove differences among students and thereby levels the playing field for teachers, schools, and systems.

The purpose of these requirements is to ensure that the effectiveness of teachers, schools, and systems is tracked yearly, measured in understandable terms, and not artificially limited by the assessment process itself. In particular, the use of fresh test items insures that the gains calculated from value-added assessment represent student progress along the full spectrum of curricular objectives and not just improvements in the material sampled by the test.

In order to ensure fair assessments of teachers, Tennessee's value-added assessment reports include only those pupils who have attended school for at least 150 days and are not eligible for special education services. (Special education students are assessed through their own "individual education plans.") For teachers who have taught a given student for less than a full year, only those students who have been the teacher's responsibility for more than seventy-five days are counted. Teachers whose subjects are not covered by the annual achievement examinations (e.g., art and music) are not assessed by value-added indicators.

Advantages

Tennessee's value-added assessment offers six important advantages when compared to other forms of educational accountability.

1. It expresses teacher, school, and school system effectiveness in terms of increases in achievement over previous performance. Each student is compared to his or her own record over a period of several years. By contrast, most present-day education accountability systems assess effectiveness by comparing current student achievement to an average or to an arbitrarily set criterion level. The failure of education accountability systems to consider gains relative to previous achievement can result in misleadingly negative evaluations for educators who are producing substantial but insufficient gains with disadvantaged students or misleadingly positive evaluations of educators who are producing mediocre gains with talented and advantaged students.
2. It excludes from the estimates of teacher, school, and school system effectiveness the influence of preexisting differences among students. These include race, socioeconomic status, previous learning, intelligence, and other factors, known and unknown, that have influenced previous achievement. In contrast to "regression analysis" approaches, Tennessee's "mixed model" approach employs statistical "blocking" to remove the contribution of suspected biasing influences. Blocking has the advantage of removing differences without the necessity of measuring and computing the magnitude of each of the excluded factors.

As counterintuitive as the notion of deleting differences may seem, empirical studies of value-added assessment have demonstrated that it does remove differences among students and thereby levels the playing field for teachers, schools, and systems. Statistical analyses of Tennessee's value-added scores have shown no relationship between annual gains and previous achievement, race, eligibility for free or reduced lunch, or any other of a variety of potentially biasing differences among students.⁷

Although value-added assessment removes preexisting differences, these are not the only factors beyond an educator's control that can influence student gains. Neither mixed-model analysis nor any other means of education gain assessment automatically removes the effects that might result from "exogenous" influences arising during the course of the school year. For example, student illness or a natural disaster might adversely affect a student's achievement, just as improvements in family income or the introduction of better community health care might contribute positively.

The influence of exogenous variables can and must be considered—especially as they impact a given school year—and mixed model I methodology is able to incorporate such considerations. Happily, however, value-added analysis, properly interpreted, minimizes the need to do so. First, data are averaged over a period of years (permitting positive and negative influences to counter-balance each other) and, second, the gains of teachers, schools, and systems can be compared to the gains of other teachers, schools, and systems that have been exposed to the same or similar influences.

3. Mixed model value-added assessment is able to isolate the achievement effects produced by an individual teacher so long as students have had that teacher for at least seventy-five days per semester. As a result, it is possible to assess teaching effectiveness regardless of whether teaching has been undertaken on a departmental basis, a team basis, or a traditional self-contained classroom basis.
4. The influence of a given teacher on student gains is expressed in the form of a "shrunk" or "regressed" estimate, i.e., an estimate that guards against an unfair assessment. In other words, the value-added system takes a very conservative approach to assessing teacher impact and thus ensures that those who are identified as effective or ineffective are deserving of their classification.
5. Value-added assessment using mixed-model methodology makes use of all student scores despite the fact that some students will have missed tests and have incomplete data. By contrast, methodologies such as regression analysis exclude students for whom complete data is lacking and thus typically remove substantial numbers of students when analyses span four or five years. Because poorer-performing students are often most apt to miss tests, the exclusion of such students can substantially inflate achievement gain estimates.

Now schools can produce a balance sheet and report an objective bottom line that is understandable to the interested citizen.

6. As described above, value-added assessment permits comparisons to national average gains of all enrolled students and thus provides an understandable measure of student progress. However, a caveat must be added. Gain scores depict how well students are progressing beyond their previous skills and knowledge but do not show how they stand with respect to an external benchmark of attainment, i.e., a national norm or state criterion referenced standard. For this reason, comparison to national average gains is not a sufficient basis for judging education outcomes. A complete assessment requires consideration of both value-added performance and performance referenced to an external standard. Tennessee's value-added reports, for example, are concerned primarily with average student gains and the comparison of those gains to national averages. But TVAAS reports also include average levels of achievement and appropriate national norms against which these may be judged.

An alternative to Tennessee's reporting system is one in which the annual learning gains produced by a given teacher, school, or system are compared to the gain necessary to bring students to an externally referenced benchmark. Although not currently used by any state, such a report would make it possible to consider both indicators simultaneously. For example, a school system with a substantial number of disadvantaged students might need to produce learning gains equal to 110 percent of the national average annual gains in order to reach national grade level standards by the eighth grade.

Conclusion

Although it employs complex statistics, value-added assessment creates a simple but enormously important change in the educational landscape. It enables parents, taxpayers, and education decision-makers to see for themselves whether schools are working. It does so by greatly simplifying the process of interpreting reports on school effectiveness. Of course, it also provides a means of assessing teacher quality and, potentially, of the programs that train teachers.

Value-added assessment holds the promise to revolutionize education. The public has been flooded with information about school and teacher quality, but making sense of it has required experts; and most of the experts have been educators who work for or with the schools. Now schools can produce a balance sheet and report an objective bottom line that is understandable to the interested citizen. The next step is to free up the system so that resources and students can flow to the most effective schools and the best teachers.

Appendix I: Sample Value-Added Report

The following sample report summarizes the value-added learning gains of the students taught by a single (unnamed) teacher for one year. Under current Tennessee law, it is sent to the teacher, with copies going to the school principal and the system superintendent.

The TVAAS reports that are provided to the public and to the media for school and school system accountability are not shown here.

For each of the five subjects tested by the annual TCAP exam (math, reading, language, social studies, and science), the report shows the average gain earned by students nationally (U.S.A. Norm Gain), the average gain earned by students in the school system (System Gain), the average gain earned by students in Tennessee (State Mean Gain), and the gain earned by this teacher's students in each of the last three years (i.e., 1995, 1996, and 1997). The gains are accompanied by the standard error of measurement associated with each score (in parentheses). For example, the average gain in reading earned by students nationally was 21, by students in Tennessee 20.4, and by students taught by this teacher 24 in 1995., 28.5 in 1996, and 18.7 in 1997. This teacher produced gains greater than the national average in 1995 and 1996 but less than the national average in 1997. A separate row of scores shows the average gain in each subject for the most recent three years. This teacher's students gained 23.7 points on average in reading over the last three years.

In the area below the numerical report, the three-year average gains in each subject are compared with national, state, and systemwide average gains. Differences between the gains earned by this teacher and relevant averages are designated as NDD (not detectably different) if they are separated by less than two standard errors of measurement—a very conservative estimate of difference. Most of the gains produced by this teacher are marked as NDD; the gains she produced are within two standard errors of the average gains for the state and the system.

The bar graph at the bottom of the report displays the relationship of this teacher's gains to the same national, state, and system averages.

1997 TVAAS Teacher Report

Teacher: XXXXXXXXXXXXXXXX (000000000)
 System: XXXXXXXXXXXXXXXX (000)
 School: XXXXXXXXXXXXXXXX (000)
 Grade: 4

** TEACHER COPY **

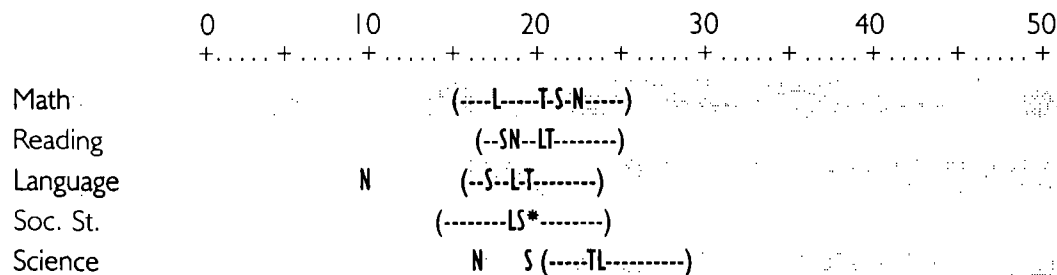
Estimated Mean Gains and (in parentheses) their Standard Errors

	Math	Reading	Language	Soc. St.	Science
USA Norm Gain:	26.0	21.0	11.0	22.0	19.0
State Mean Gain:	25.0	20.4	19.3	21.4	21.4
1995 Teacher Gain:	22.7 (6.8)	24.0 (3.8)	22.3 (4.2)	30.6 (4.9)	28.3 (4.1)
1995 System Gain:	12.3 (2.9)	22.5 (3.4)	19.8 (3.0)	27.9 (3.8)	27.5 (3.1)
1996 Teacher Gain:	33.3 (4.9)	28.5 (4.0)	28.4 (3.8)	18.3 (5.8)	26.8 (4.0)
1996 System Gain:	32.6 (3.0)	28.6 (3.6)	29.5 (3.2)	17.4 (4.0)	28.3 (3.3)
1997 Teacher Gain:	15.6 (5.7)	18.7 (4.9)	17.8 (4.9)	17.9 (5.7)	25.6 (4.5)
1997 System Gain:	16.3 (3.1)	19.1 (3.6)	16.0 (3.2)	18.1 (4.0)	26.1 (3.3)
Teacher 3-Yr-Avg:	23.9 (3.0)	23.7 (2.5)	22.8 (2.3)	22.3 (3.2)	26.9 (2.4)
System 3-Yr-Avg:	20.4 (1.7)	23.4 (2.0)	21.8 (1.8)	21.1 (2.3)	27.3 (1.9)

Teacher 3-Year-Average Gain Comparisons

Teacher vs Norm:	NDD from Norm	NDD from Norm	Above Norm	NDD from Norm	Above Norm
Teacher vs. State:	NDD from Mean	NDD from Mean	NDD from Mean	NDD from Mean	Above Mean
Teacher vs System:	NDD from Mean	NDD from Mean	NDD from Mean	NDD from Mean	NDD from Mean

Teacher 3-Year-Average Gain in Scale Score Units with Approximate 95% Confidence Intervals



Legend: T = Teacher Gain, L = System (LEA) Mean Gain, S = State Mean Gain, N = National Norm Gain.

An asterisk (*) indicates that 2 or more of the above symbols coincide.

The estimated teacher gains presented here are the official TVAAS estimates from the statistical mixed model methodology which protects each teacher from misleading results caused by random occurrences. Each teacher's gain is assumed to be equal to the average gain for the district until the weight of the data pulls the estimate away from the district average.

This year's estimates of previous years' gains may have changed as a result of incorporating the most recent student data.

Appendix 2: References

For the interested reader, there is a fairly extensive literature pertaining to value-added assessment. Although no one has yet written an account of mixed model methodology that is suitable for a general audience, it has been critically examined by a number of scholars and policy experts.

Perhaps more importantly, value-added assessment has been used successfully in Tennessee for nearly ten years, and many educators have learned how to interpret and make use of it. Some schools have used it to identify weaknesses and have, as a result, made phenomenal gains. Other schools—notably rural schools and schools with many disadvantaged students—have been able to show that they are doing a better job of teaching than had been evidenced by indicators such as expenditures and the use of the latest educational practices. On the whole, student achievement in Tennessee has been improving over the years that value-added assessment has been in place.

For a general description see:

J. Archer, "Putting 'Value Added' Data to Good Use," *Education Week*, 5 May 1998.

J. Archer, "Sanders 101," *Education Week*, 5 May 1998.

W.L. Sanders and Horn, S. P, "The Tennessee Value-Added Assessment System (TVAAS) Mixed-model Methodology in Educational Assessment," *Journal of Personnel Evaluation in Education*, 8, no.1 (1994): 299-311.

For an evaluation and policy analysis see:

W. T. Gormley and D.L.Weimer, *Organizational Report Cards* (Cambridge: Harvard University Press, 1999)

For a comprehensive technical review see:

R.D. Bock and R. Wolfe, *Audit and Review of the Tennessee Value-Added Assessment System (TVAAS) Preliminary Report*, 23 January, 1996. (Available from the Tennessee Office of Education Accountability, Comptroller of the Treasury, State Capitol, Nashville, TN 37243-0260)

For sample value-added reports:

Tennessee Department of Education
6th Floor, Andrew Johnson Tower
710 James Robertson Parkway
Nashville TN 37243-0375
Phone (615) 741-2731
<http://www.state.tn.us/education/rptcrd98/rcstudent.htm>

For value-added reports in a consumer-friendly on-line display:

The Tennessean (Nashville)
<http://www.tennessean.com/schools/>

For technical information regarding value-added analysis and its implementation

Dr. William Sanders, Director
 University of Tennessee Value-Added Research and Assessment Center
 225 Morgan Hall, P. O. Box 1071
 Knoxville, TN 37901
 Phone (423) 974-7336
 Fax (423) 974-7448
wsander2@utk.edu

This paper was completed with the support of The Foundation Endowment.

- ¹ W.L. Sanders and J.C. Rivers, *Cumulative and Residual Effects of Teachers on Future Student Academic Achievement*, 1996 (Available from UTVARC, 225 Morgan Hall, P.O. Box 1071, Knoxville, TN 37901-1071).
- ² Although not currently used for this purpose, value-added assessment can provide an independent and objective indication of the quality of teacher training. By aggregating the value-added gains of novice teachers on a program-by-program basis, state policymakers would be able to determine which teacher-training programs are doing the best job of equipping teachers with the skills necessary to improve student achievement. Gains in learning aggregated by teacher-training institution would be a vastly superior indicator of program quality in comparison to indicators now used. Factors such as admission requirements, course offerings, faculty credentials, facilities, and the many other program characteristics now reviewed by state agencies are not demonstrably linked to teacher success in producing student achievement. A value-added indicator of program effectiveness would also be vastly superior to the so called "performance-based" indicators of teacher effectiveness that were recently proposed by the National Council for the Accreditation of Teacher Education. These indicators include items such as credits in educational methods courses, classroom demonstrations, and scores on tests of pedagogical knowledge. For the most part, they are at best weakly related to the primary phenomenon of interest, i.e., objectively measured student achievement.
- ³ W.L. Sanders, A.M. Saxton, and S.P. Horn, "The Tennessee Value-Added Assessment System, A Quantitative, Outcomes-Based Approach to Educational Measurement" in *Grading Teachers, Grading Schools: Is Student Achievement a Valid Evaluation Measure?* ed. Jason Millman (Thousand Oaks, CA: Corwin Press, 1997), 137-162.
- ⁴ W.J. Webster and R.L. Mendro, "The Dallas Value-Added Accountability System" in *Grading Teachers, Grading Schools: Is Student Achievement a Valid Evaluation Measure?* . 81-99.
- ⁵ W.L. Sanders, A.M. Saxton and S.P. Horn, "The Tennessee Value-Added Assessment System, A Quantitative, Outcomes-Based Approach to Educational Measurement;" D.A. Harville, "Extension of the Gauss-Markov Theorem to Include the Estimation of Random Effects," *Annals of Statistics* 4, no.2 (1976): 384-395 and "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems," *Journal of the American Statistical Association* 72 (1977): 320-338; C.R. Henderson, "Analysis of Variance in the Mixed Model Higher Level, Nonhomogenous, and Random Regressions," *Biometrics*, 38 (1982): 623-640; R.A. McLean, W.L. Sanders and W.W. Stroup, "A Unified Approach to Mixed Linear Models," *American Statistician* 45, no.1 (1991): 54-64; H.D. Patterson and R. Thompson, "Recovery of Interblock Information When Block Sizes Are Unequal," *Biometrika* 58 (1971): 545-554; S.W. Raudenbush and A.S. Bryk, "Methodological Advances in Analyzing the Effects of Schools and Classrooms on Student Learning," *Review of Research in Education* 15 (1988): 423-479; W.L. Sanders, *A Multivariate Mixed Model in Applications of Mixed Models in Agriculture and Related Disciplines*, Southern Cooperative Series Bulletin No. 343 (Baton Rouge, Louisiana Agricultural Experiment Station), 138-44.
- ⁶ Of technical significance, value-added estimates of teacher influence are derived from a multi-year "layered" computational model and corrected by a "shrinkage estimate." These two features substantially reduce the possibility of false negative or false positive estimates and ensure that the resulting indicators of achievement gain are as exact as fairness will permit.
- ⁷ University of Tennessee Value-Added Research and Assessment Center, *Graphical Summary of Educational Findings from the Tennessee Value-Added Assessment System (TVAAS)*, 1995. (Available from UTVARC, 225 Morgan Hall, P.O. Box 1071, Knoxville, TN 37901-1071.)

Contributors

Dale Ballou is Associate Professor of Economics at the University of Massachusetts (Amherst).

Dominic J. Brewer, an economist, is Director of RAND's education program.

Chester E. Finn, Jr. is John M. Olin fellow at the Manhattan Institute and president of the Thomas B. Fordham Foundation. He served as Assistant U.S. Secretary of Education from 1985 to 1988.

Dan D. Goldhaber is a labor economist who serves as Research Associate in the Urban Institute's Education Policy Center and as a member of the Alexandria City School Board.

Eugene W. Hickok is Secretary of Education for Pennsylvania. He chairs the Education Leaders Council, a network of state and local education leaders who promote fundamental education reforms.

Marci Kanstoroom is Research Director at the Thomas B. Fordham Foundation and Research Fellow at the Manhattan Institute.

Michael Kwiatkowski is Director of Institutional Research at Citrus College in Glendora, California. He was formerly Director of Educational Research at the Tomás Rivera Policy Institute.

Tyce Palmaffy is a journalist and writer who has worked on the staff of *The New Republic* and *Policy Review: The Journal of American Citizenship*, where he wrote on education issues.

Michael Podgursky is Professor of Economics and Chairman of the Economics Department at the University of Missouri-Columbia.

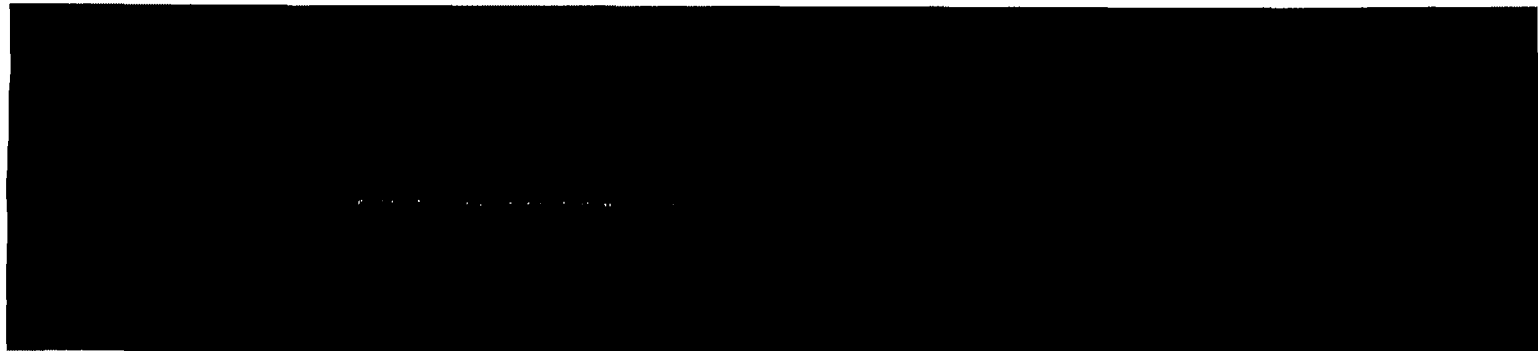
Michael Poliakoff is Pennsylvania's Deputy Secretary of Education for Postsecondary and Higher Education.

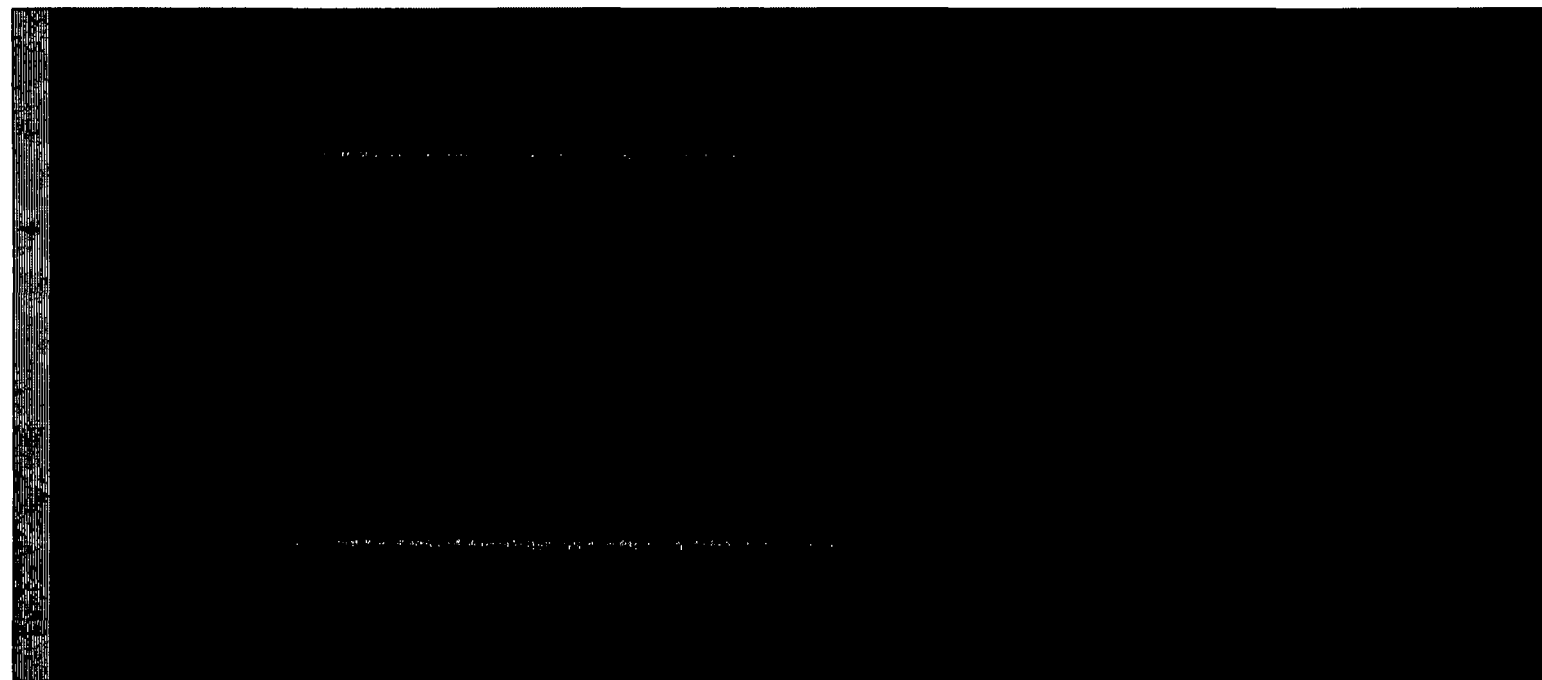
Naomi Schaefer is associate editor of *The Massachusetts News*.

J.E. Stone is an educational psychologist and professor in the College of Education at East Tennessee State University. He also heads the Education Consumers ClearingHouse (<http://education-consumers.com>).

Robert P. Strauss is Professor of Economics and Public Policy at the Heinz School at Carnegie-Mellon University.

Danielle Dunne Wilcox is pursuing a doctorate in education policy and administration at Teachers College, Columbia University.







OUTSIDE THE BOX

The Thomas B. Fordham Foundation

1627 K Street, N.W., Suite 600

Washington, DC 20006

Telephone: (202) 223-5452

Fax: (202) 223-9226

<http://www.edexcellence.net>

To order publications: 1-888-TBF-7474
(single copies are free)