Appendix D

NATIONAL BOARD FOR PROFESSIONAL TEACHING STANDARDS CERTIFICATION PROGRAM

The National Board for Professional Teaching Standards (NBPTS) was established to develop and administer a voluntary system of advanced professional certification to recognize highly accomplished K–12 teachers. The standards guiding task development were developed by committees of teachers and scholars. To obtain the NBPTS certificate, teachers prepare an extensive portfolio demonstrating their preparation, classroom work, teaching strategies, instructional goals and results, students’ work, and professional activities. In addition, they participate in one day of performance activities at a regional assessment center. The process takes up to one school year to complete.

NBPTS certification offers benefits to teachers, school districts, and teacher training institutions. Teachers have an opportunity to reflect on and perhaps improve their teaching skills and professional life. School districts have an independent standard against which to measure the ability of their experienced teachers, and the process clarifies for teacher training institutions what accomplished teachers should know and be able to do. Ultimately, the most important beneficiaries of improved teaching practices are students.

DESCRIPTION AND PURPOSE

NBPTS, a nonprofit organization, was founded and initially funded by the Carnegie Corporation of New York to provide avenues for teachers to demonstrate their professional achievement. Establishment of NBPTS in 1987 fulfilled a major recommendation of A Nation Prepared: Teachers for the 21st Century, a report issued by the
Carnegie Task Force on Teaching as a Profession (1986). NBPTS hopes to improve the public’s perception of teachers, enhance teachers’ own view of their profession, and thereby attract and retain high-quality teachers.

A majority of the sixty-three NBPTS members are classroom teachers actively engaged in instruction, and many others are teacher educators or leaders of teacher professional associations. Other members include public officials, board of education members, administrators, presidents and faculty of higher education institutions, parents, minority student rights advocates, and business leaders. A majority of the public and other educator members are elected or appointed public officials. Reaching consensus on the standards among such a broad stakeholder group has been a long process, as has developing, testing, and administering assessment tasks. To date, these activities have cost tens of millions of dollars. Although NBTPS planned on high costs, especially for developing and testing initial assessments, it aims to make development and administration of the remaining assessments more efficient and economical.

NBPTS has a threefold mission: “To establish high and rigorous standards for what teachers should know and be able to do, to certify teachers who meet those standards, and to advance other education reforms for the purpose of improving student learning in American schools” (NBPTS, 1989, p. 1). Its core activity is an assessment system organized around subjects and age levels. NBPTS will combine four age groupings and fourteen subjects to produce its range of certificates (see Table D.1), which will total thirty or more (some subjects cover two age groups combined). In 1995–1996, NBPTS certification was available in two categories: early childhood/generalist (EC/G) and middle childhood/generalist (MC/G).

**RELATIONSHIP TO OTHER PROGRAMS**

NBPTS certification complements but does not replace state licensing. State licensing systems set compulsory minimum standards for novice teachers; NBPTS certification creates voluntary standards for accomplished teachers. Similarly, NBPTS standards should build upon but not substitute for requirements for preservice training.
Table D.1
Current and Planned NBPTS Certificates

<table>
<thead>
<tr>
<th>Levels</th>
<th>Subject Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early childhood (ages 3–8), middle childhood (ages 7–12), and early adolescence (ages 11–15)</td>
<td>Generalist</td>
</tr>
<tr>
<td>Middle childhood (ages 7–12), early adolescence (ages 11–15), and adolescence and young adulthood (ages 14–18+)</td>
<td>English language arts, mathematics, science, social studies/history</td>
</tr>
<tr>
<td>Early-middle childhood (ages 3–12), and early adolescence-young adulthood (ages 11–18+)</td>
<td>Foreign language, art, music, exceptional needs/generalist, English as a new language, physical education, library/media, guidance counseling</td>
</tr>
<tr>
<td>Early adolescence-young adulthood (ages 11–18+)</td>
<td>Health, vocational education</td>
</tr>
</tbody>
</table>

Teacher training institutions develop curriculum to comply with state laws; NBPTS standards establish a set of profession-endorsed guidelines for best practices that schools can use in improving in-service training (e.g., when developing curriculum to organize the continuing education of teachers).

The teaching profession is attempting to link accreditation, licensure, and advance certification with the goal of ensuring that all students are taught by competent, professional teachers (Rahn, 1995). Two national organizations (the National Council for Accreditation of Teacher Education [NCATE] and the Interstate New Teacher Assessment and Support Consortium [INTASC], a project of the Council of Chief State School Officers), the two teacher unions (the National Education Association and the American Federation of Teachers), and others are working with NBPTS to improve the profession of teaching for both teachers and students. They envision a linked system of preservice preparation, extended clinical training, and continuing professional development in which National Board certification plays an important role. Although much of this linkage is still under development, both NCATE and INTASC are taking steps to see that their own components of this process are aligned with the work of NBPTS. Similarly, input from relevant professional commit-
tees and other stakeholders is actively sought in every phase of implementing the NBPTS system—from the composition of the board of directors, standards committees, and field test network that operated from 1993 to 1995, to the broad-based review of documents and test packages.

IMPLEMENTATION AND ADMINISTRATION

Implementation of the teacher assessment system builds on a strong base of educational research and broad support among established educational organizations and stakeholders. The first task of NBPTS was “to identify the knowledge, skills, and dispositions that describe accomplished teaching and to convert those attributes into high and rigorous standards upon which to base the National Board Certification system” (NBPTS, Executive Committee, 1989, p. 2). NBPTS staff reviewed the relevant literature on these issues and the standard-setting work of other occupations and obtained comments from leaders in the education community before issuing Toward High and Rigorous Standards for the Teaching Profession (NBPTS, 1989).

This 1989 document established the philosophical underpinnings for the program, including the prerequisites for applying (three years of teaching and, at a minimum, a baccalaureate degree); five propositions that set forth broad principles to guide the development of standards; and assessment activity development guidelines. NBPTS established a comprehensive organizational structure and process to develop the assessment system, with teachers playing a major role in almost every area.

Through a national competitive process, contractors were selected for several development and implementation tasks. Assessment development laboratories (ADLs) developed and tested the assessment tasks. A field test network (FTN) was engaged in the early years to provide candidates, administrators, scorers, and evaluations of the methods and systems. The operation of the system was assigned to another contractor. A technical analysis group (TAG) provided research support to the other contractors (e.g., a literature review of assessment methods in other professions; development of the sampling frame of teachers for the field-test trials) and analysis of the work of the standards committees, ADLs, and other contractors. In
the past year, most of these functions have been consolidated in the hands of a single contractor, Educational Testing Service.

The development process was implemented gradually. Initially, two standards committees were appointed; then additional committees were added each year. Once the initial standards committees began work, an ADL was appointed to work with each of the committees. A year and a half later, a request for proposal (RFP) for six additional ADLs was issued. This strategy of gradual growth allowed NBPTS to learn from early experiences and adjust the process.

Because the process of establishing the assessment system may provide a useful model for establishing a system for recognizing accomplished vocational teachers, it is described in considerable detail here. A vocational education standards committee has been appointed and has issued a set of draft standards that have now been through the NBPTS public comment and critique process.

**Developing Standards**

Standards committees are composed primarily of teachers but also include researchers and others involved in the field of interest. The first four committees were appointed in 1990 for early adolescence/English language arts (EA/ELA), early adolescence/generalist (EA/G), adolescence and young adulthood/mathematics, and early adolescence through young adulthood/art. By 1995, seventeen committees had been established to set standards in twenty-one of the more than thirty certification fields (Bradley, 1995a); by 1997, initial development had been completed and standards were released for public comment in these twenty-one areas.

Committee members are selected through a process that gathers nominations from a broad group of professional organizations, NBPTS members, and NBPTS staff. For the EA/G committee, for example, over 130 nominations were reviewed to find thirteen members who provided a balance of gender, location, and profession (Hattie, Sackett, and Millman, 1994). Members of the relevant ADL participate in standards committee meetings. Key professional organizations are also invited to appoint liaisons who attend all committee meetings, e.g., the International Reading Association provided a liaison to the EA/G committee.
These committees develop draft standards for the knowledge and skills that teachers should have to achieve NBPTS certification. Draft standards are reviewed by NBPTS, its certification standards working group (CSWG), field site members, the ADL, and a broad spectrum of stakeholders. Reviewers were asked to rate the standards for clarity and for their relevance to highly accomplished teaching, among other factors. One issue that arose in the review of the initial EA/G standards was the difficulty of determining appropriate subject matter for generalist teachers. The standards committee makes recommendations to the board, which has final authority in making decisions. Almost thirty months after their first meeting, NBPTS and its CSWG approved the standards for the 1993–1994 EA/G field test. Revised standards were approved the following year for the 1994–1995 administration. While standard setting proved to be a lengthy process, a survey of the reviewers indicated widespread approval for the standards’ validity among both teachers and nonteachers.¹

Assessment Development Laboratories

ADLs worked with one or more standards committees to develop and produce an assessment package. ADLs were selected through a competitive merit review RFP process (NBPTS, 1990a and 1990b). Guidelines for the ADLs stated that all assessments had to be professionally credible, publicly acceptable, legally defensible, administratively feasible, and economically affordable. Potential contractors were directed to include multiple forms of assessment and to consider how student learning as a measure of teacher effectiveness might be demonstrated.

The following six assessment methods were specified for exploration: (1) a portfolio of classroom teaching accomplishments that includes evidence of the teacher’s participation in a learning community, samples of student work, and artifacts produced by the teacher; (2) observations of the teacher in his or her classroom; (3) structured interviews based in part on the portfolio; (4) exercises typical of the

¹About 87 percent of teachers and 88 percent of nonteachers responded “agree” or “strongly agree” that “each of the 11 standards describes a critical aspect of highly accomplished teaching practice within this field” (Hattie, Sackett, and Millman, 1994, p. 34).
teacher’s work, e.g., viewing videotape of a teaching situation and grading samples of students’ work resulting from that situation; (5) simulations that are “contextual assessments,” e.g., suggesting more effective strategies after viewing a videotape of the teacher’s performance; and (6) written tests of subject matter knowledge and pedagogy. The following components were also to be included in the assessment procedures: (1) documentation of the teacher’s practice and thinking through videotapes, student work, other artifacts, and commentaries; and (2) assessment of the candidate’s subject matter knowledge, knowledge of pedagogy, and knowledge of child development for the specific age group.

The first ADL contract was awarded in 1990 to the University of Pittsburgh School of Education and the Connecticut Department of Education to develop assessments for the EA/ELA certificate. The second contract was awarded to the Performance Assessment Laboratory at the University of Georgia for the EA/G certificate.

Both labs devised roughly similar assessment activities that could be closely integrated with class lessons, such as developing appropriate applications for a new classroom resource, recording actual classroom plans and activities and analyzing what occurred, or analyzing and evaluating samples of student writing. Emphasis in the assessments is on giving teachers an opportunity to show what they know in an authentic context, rather than on pinpointing what they do not know. For both certificates, teachers complete activities at the school site, collecting documentation in a portfolio, and perform additional activities at the testing center. Tables D.2 and D.3 compare the skills and activities targeted by each lab at, respectively, the school site and the assessment center (Bradley, 1994).

The first field test, which functioned like a real assessment, revealed strengths and weaknesses in the assessment center model. More than one-third (eighty-one out of 289) EA/G candidates participating in the 1993–1994 field test were certified (the certificates were awarded in January 1995). Successful EA/ELA candidates were certified that summer. The two certificates were offered again in 1994–1995 to fee-paying candidates, and about 200 candidates participated. At a minimum, fourteen assessors were required to score each candidate’s work (two scores for each of seven exercises) during this first year. Problems with the scoring procedures resulted in a
Table D.2

School Site Skills and Activities Targeted by Each Lab

<table>
<thead>
<tr>
<th>Early Adolescence/Generalist</th>
<th>Early Adolescence/English Language Arts</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Professional development and service: submit vita; write accounts of 1) an impact of professional development on practice and 2) professional service activity; obtain letters of support from colleagues.</td>
<td>1. Professional background: submit resume; write one- to two-page description of participation in a learning community.</td>
</tr>
<tr>
<td>2. Teaching and learning: write narrative describing a selected class over a period of time; describe the progress of three students, reflecting different learning characteristics; videotape class activities; provide samples of student work and teaching practices.</td>
<td>2. Teaching and learning: describe and analyze the writing of three students, including the influence of instruction (submit with five to eight samples of the students’ writing).</td>
</tr>
<tr>
<td>3. Lesson analysis: select unedited 30–45 minute videotape from a class and write account of the teaching and learning that occurred, highlighting five to seven particularly important points.</td>
<td>3. Interpretive discussion: videotape 15 to 20 minutes of a class discussing a piece of literature; write an evaluation of the discussion.</td>
</tr>
<tr>
<td>4. Planning and teaching: write eight-page commentary describing planning and instruction over a three-week period, using an integrated curriculum that demonstrates cultural awareness; include videotape of one class session.</td>
<td></td>
</tr>
</tbody>
</table>

costly redesign of the scoring system and delays in announcing the results (Bradley, 1995a). This particular certification was revised and then offered again in 1996–1997. Reducing the complexity of the scoring process was critical. Modifications focused on making the assessments less burdensome to scorers and candidates. According to James R. Smith, NBPTS senior vice president, the initial portfolios, for example, asked for more material than was necessary and needed to be more focused (Bradley, 1995a).
<table>
<thead>
<tr>
<th>Early Adolescence/Generalist</th>
<th>Early Adolescence/English Language Arts</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Instructional resources: write analysis of the potential of SimCity for teaching social studies, math, history, and science. (SimCity is a computer simulation supplied to the candidate at the school site.)</td>
<td>1. Group discussion: with other candidates, develop a curriculum unit on personal relationships. Unit materials are selected from eight novels provided previously at school site. Discussion is videotaped.</td>
</tr>
<tr>
<td>2. Instructional analysis: write analysis of videotape and materials from a mathematics instructor, including suggestions for more effective strategies and extension of the topic to the arts.</td>
<td>2. Instructional analysis: analyze videotape of teacher-led discussion, including suggestions for improving instruction; show knowledge of young adolescent learning, and demonstrate cultural awareness and understanding of discussion dynamics.</td>
</tr>
<tr>
<td>3. Curriculum issues: after group discussion of a theme related to exploration of governmental systems, ecosystems, and the media, complete two-hour written description of the instructional development of a theme drawing on one of the above subjects.</td>
<td>3. Analysis of student writing: analyze set of student papers and discuss analysis with interviewer, making suggestions for improving students’ writing. Videotaped.</td>
</tr>
<tr>
<td>4. Content knowledge: three one-hour written subject examinations.</td>
<td>4. Content knowledge: three two-hour essay assessments on composition, literature, and language. Literature and journal articles are used for the essay prompts.</td>
</tr>
</tbody>
</table>

**Assessment Administration**

The first field tests indicated that candidates felt they did not have enough time to prepare their portfolios, but once NBPTS moved to regular operation, the time allowed for portfolio development was lengthened. The 100 hours teachers spent assembling their portfolios was about twice the time NBPTS had anticipated (Bradley, 1994).
Findings from the first set of assessments (Scriven, 1994) point to several problems. Many candidates said that they were influenced to participate by the absence of a fee, so the cost of the actual examination may discourage some potential candidates. (Since the field test, many states and districts have agreed to pay the cost for teachers who choose to apply, so economic factors will not be that much of a deterrent.) The description of the process needs to be clearer so that candidates know what to expect about the amount of time involved and the content of the exercises. About half of the candidates found the instructions for the portfolio exercises unsatisfactory, and most candidates felt the support provided was inadequate. Teachers sought more specific direction for activities, including, for example, the expected length for written assignments. Peer support groups were judged very successful in helping to prepare materials; however, the help of principals was not useful. About half of the participants found preparation workshops and video support useful.

At the testing center, about 75 percent of the candidates bemoaned the lack of computers for the writing tasks. (NBPTS has since chosen Sylvan Technology Centers as the sites for performance activities, so computers are now available.) The amount of writing required was also thought to be excessive. Observers also indicated that testing coordinators needed to be better trained. The original twelve-hour day was problematic; it was subsequently reduced to eight hours. A particularly troubling finding was that about 40 percent of participants felt that seeking certification placed them at some risk in their schools. Michael Scriven, the evaluator of the test administration, noted that this was “consistent with other evidence that teachers tend to identify efforts to excel as egotistical or undemocratic” (Scriven, 1994, p. 9). He warned that if merit pay was tied to certification, it could increase negative reactions, particularly if principals share this attitude.

Based on the experiences of the first two labs, the draft RFP for subsequent ADLs outlined a streamlined process that would save time and money (NBPTS, 1991). It also addressed many of the concerns raised by Scriven.

---

2 Bradley (1994) notes that of three Fairfax County, VA, candidates for the EA/G, one’s teaching and learning commentary was sixty-six pages, another’s was six pages, and the third’s was one page.
Field Test Network

Once the labs developed tasks, they were field tested (during 1993-1995) through a national FTN of more than 100 school districts. The network included 165,000 teachers, 25 percent of whom were members of a minority group. Districts in the network have a two-year contract to perform a variety of tasks, including reviewing standards, developing staff development programs for candidates, and field testing assessment packages (NBPTS, 1992). The sites participate on an as-needed basis as the examinations are developed. For example, twenty-six sites participated in the first field test.

Scoring Assessments

Scorers are recruited regionally and receive training on the goals of NBPTS, the standards for each certificate, and the scoring guidelines for each activity. During the field tests, exercises were scored independently by two teachers with training in the appropriate subject, teaching experience, and a recommendation for teaching quality (criteria similar to those demanded of candidates). NBPTS is gradually increasing the proportion of scorers for each certificate who are NBPTS-certified teachers from that field, and eventually all will be. If there are differences in the scores, the scorers meet and discuss the evidence, then independently rescore the exercise. If differences still exist, a third scorer is brought in. The final score represents agreement between two of the three scores (NBPTS, 1995).

Overall Progress

Considerable progress has been made toward NBPTS’s long-term goals for increasing the professionalization of teaching. NBPTS has undertaken key steps for developing the certification system and completed many of them. For example, draft standards in twenty-one of the thirty-plus areas have been released for public comment. Yet progress in implementing the assessments has been slowed by the need to revise scoring methods to maintain the board’s high standards for quality. Certification was offered in only two areas in
1994–1995, and six certificates, instead of the projected nine, will be available in 1996–1997:

- Early adolescence/generalist
- Early adolescence/English language arts
- Early childhood/generalist
- Middle childhood/generalist
- Adolescence and young adulthood/mathematics
- Early adolescence through young adulthood/art

NBPTS responded to problems and criticisms that arose in early testing. As a result, both the time required for development and the cost of administration have been greater than anticipated. NBPTS has received over $50 million since its inception in October 1987—$37 million from private donors and foundations, and $19.34 million in one-to-one matching funds from the federal government (Bradley, 1995a)—but this has not been enough to maintain the initial development schedule. In 1995, overall cost overruns caused the organization to cancel three of the seven ADL contracts.

Administering the evaluation to the first group of candidates cost more than $4,000 per teacher (not counting development costs). Bradley (1995a) reported that “most of the expenses . . . went to scoring each candidate’s work, an exhaustive process that in some cases took 23 hours.” Costs for administering the second round of EA/G and EA/ELA tests (1995–1996) were reduced to $3,000 per participant. And costs for the field tests of the early childhood/generalist and middle childhood/generalist certificates were projected to be about $2,500 per participant (Bradley, 1995a). The application fee for the first post-field-test year was $975, but was raised to $2,000 for 1996–1997 and 1997–1998.

The desire to produce innovative assessments (using authentic measures wherever possible) while maintaining very high standards for technical quality has contributed to high costs. NBPTS has implemented many suggestions for reducing costs—e.g., limiting assessments at the testing center to one day instead of two. Consideration is being given to a multiphase process that may also reduce scoring costs and could allow for partial credit (banking of
accomplishments). Despite the substantial time burden (reduced from the first field test) on candidates, those who complete the process generally find it to be extremely rewarding.

TECHNICAL QUALITY

Since it began to develop its certification system, NBPTS has been concerned about producing a high-quality and technically defensible process. After the first round of certifications, NBPTS convened a panel of respected educational researchers to review the development of the certification system and determine whether the process was sound from a technical standpoint. The panel concluded that there were “no technical impediments to the Board’s use of its Early Adolescence/Generalist assessment to award National Board Certification to candidates whose performances satisfy the . . . final recommended performance standard” (Bond et al., 1994, p. 33).

However, the expert panel also recommended that further study be directed as follows:

• To increasing the reliability of the assessment (to resolve problems about scores near the passing standard).

• To determining whether having two content standards assessed less frequently than the others is acceptable.

• To exploring strategies to reduce possible adverse impacts, i.e., the likelihood that the percentage of African American candidates who would be certified would be far lower than that of non-Hispanic white candidates.

• To developing additional forms of the assessment center exercises.

Subsequently, the TAG offered guidance to the assessment developers and evaluated the quality of the certification procedures. During its first year, the TAG either conducted or commissioned eight studies of the technical quality of the 1993–1994 EA/G assessment (Hattie, Sackett, and Millman, 1994; Lloyd and Crocker, no date; Scriven, 1994; Felker, 1994; Heider et al., 1994; Traub, 1994a,b; Jaeger, 1994; Bond and Linn, 1994). The topics of these studies were
1. The development process for content standards
2. Content validity
3. Quality of field-test operations
4. Quality of assessors' training
5. Validity of the application of scoring procedures
6. Consistency of certification decisions and reliability of exercises
7. Recommended performance standards
8. Adverse impacts of differing certification rates of diverse groups

**CONSEQUENCES AND USE OF ASSESSMENT RESULTS**

Assessment developers learned several lessons from the scoring exercises and from debriefing participants (Bradley, 1995b). Scoring of the first set of portfolios indicated that teachers were not skilled at reflecting on their own work; they were also more comfortable describing rather than analyzing teaching practices. In response, NBPTS developed better instructions for each exercise and clearer guidelines for teachers in order to emphasize analysis as well as description. Moreover, interviewing teachers at the assessment center about the work in their portfolios was also problematic. Identifying and training skilled interviewers is difficult and costly, and this proved not to be the most cost-effective assessment method. Classroom videotapes and samples of student work proved to be more reliable measures of teacher practices than teachers' own descriptions.

While point-in-time samples of student work provided little information about how student learning progressed, they could provide worthwhile information about the quality of the specific assignments. Developers also improved test validity over time by linking testing activities more closely to the skill they wanted to evaluate, e.g., a videotaped student presentation of a class project does not help assessors evaluate the teacher.

Formal support for teachers who apply and rewards for those who are certified are gradually growing in states and districts. Merit pay ("pay for knowledge"), mentor status, the right to teach in any state,
and waiver of credential renewal requirements are all rewards that school districts or states can consider. About thirty states have implemented some type of recognition, incentive, or support. North Carolina financially supports teachers who are pursuing the certificate and rewards teachers who obtain it. North Carolina provides the assessment fee, several days’ preparation time, and a 4 percent salary increase for teachers who obtain NBPTS certification (Hunt, 1995). North Carolina, Iowa, New Mexico, and Oklahoma waive state licensing requirements for NBPTS-certified teachers who move to the state. Massachusetts and Ohio accept the NBPTS certificate in lieu of their own state recertification (Richardson, 1995).

NBPTS intends for the assessment process itself to give teachers an opportunity to grow professionally by reflecting on their skills and knowledge, and to measure themselves against objective, peer-developed standards (NBPTS, 1995). This focus in the test on content-specific pedagogy forces teachers to carefully examine this central aspect of practice. Test developers also hope that exposure to new techniques and materials will provide teachers with professional growth opportunities, such as might be gained by exploring the relevance of a computer simulation to their instructional strategies (Cape, Dickey, and Anderson, 1995). For example, an EA/G candidate found that “examining his own professional development was a worthwhile activity that helped him clarify who he is as a teacher. Exploring ways to integrate other subjects into his lessons was particularly exciting” (Bradley, 1994).

**APPLICABILITY TO VOCATIONAL EDUCATION**

The NBPTS experience offers three kinds of lessons for vocational educators. First, the NBPTS experience shows how difficult it can be to use alternative assessments when high stakes are tied to the results. Candidates expect that NBPTS certification will be accompanied by professional recognition and even financial rewards. NBPTS intends for the certification process and standards to drive preservice and inservice training and affect state licensing standards. Consequently, it adopted an approach to insure that the profession endorsed the standards and that the certification process met the highest criteria for quality and fairness. To produce the standards, NBPTS included multiple stakeholders in frequent reviews and
instituted systematic quality control procedures. Its commitments to use alternative assessment strategies and to maintain high standards for reliability and validity have made the process complex and necessitated additional professional review and analysis. All this translates into time and expense.

The experience of the ADLs presents a useful picture of the complexity associated with alternative assessments. Following NBPTS’s lead, the ADLs tried to be innovative and rely on authentic assessments to measure teacher competence. They developed interesting, relevant, and meaningful exercises for teachers, but in some cases, the activities did not necessarily reflect the underlying competencies they were designed to measure; in other cases, it was difficult for raters to agree on the quality of a candidate’s performance. Because ADLs were using new assessment methods, they often could not rely on traditional approaches to monitor quality and thus had to develop new approaches. In fact, NBPTS employed a TAG to consult with and review the work of the ADLs because the test developers were breaking new ground. To its credit, NBPTS set high standards for itself, monitored the assessments carefully, and invested the resources and the time necessary to correct problems. This experience provides a sense of the level of complexity that may be encountered if alternative assessments are used in high-stakes contexts.

These issues should not be unfamiliar to vocational educators, particularly those in the health professions. Requirements for professional credibility and legal defensibility have led to very thorough and comprehensive assessment efforts in the health fields. However, in most cases these professional assessments use traditional techniques, such as multiple-choice and short-answer questions. It is the combination of high-stakes certification and alternative assessments that creates challenges for both quality and defensibility. Vocational educators have used alternative assessments at the classroom level for years, because they link classroom and work experience more closely. The concerns raised here should not weigh too heavily on teachers who want to use more authentic assessments as part of their program. However, quality issues demand greater attention if the assessments become part of a certification system that has important rewards for students.
The second lesson that can be drawn from the NBPTS experience relates to evaluating vocational educators. NBPTS intends to offer a single certification in vocational education (early adolescence through young adulthood). However, it encountered some difficulty in developing the standards because of disagreement about whether general standards could be used for all vocational educators or each occupational area had to be treated separately. In the end, the standards development committee agreed on a single standards framework that recognizes eight distinct areas of industry-specific knowledge: agriculture and environmental sciences; arts and communication; business, marketing, and information management; consumer sciences; health services; human services; manufacturing and engineering technology; and technology education. The framework also says that there is a common core of vocational knowledge appropriate to all fields—understanding workplace basic skills, commanding general industrial knowledge, and integrating vocational content with core disciplines—as well as industry-specific knowledge. The framework is now available in draft form for public comment. The discussion of what general standards covering vocational education should look like is likely to increase in importance as education moves toward greater integration of vocational and academic curricula and as emphasis shifts in vocational programs from specific occupational skills to broader aspects of an industry.

The third lesson from the NBPTS experience focuses on the value of using assessments to encourage teachers to reflect on and improve their practices and to collaborate with their peers. NBPTS has initiated a process in which teachers put together portfolios to demonstrate aspects of their work, analyze and think about ways to improve their practices, and discuss and critique the work of peers in support groups designed to elicit all candidates' best efforts. This process shows promise for changing the historic isolation most teachers face in their work and for enriching professional development, with the ultimate goal of improving instruction.