

Lessons from Ten Years of TEAC's Accrediting Activity

Frank B. Murray

*Teacher Education Accreditation Council (TEAC),
Washington, D.C.*

Introduction

Founded in 1997, the Teacher Education Accreditation Council (TEAC) designed a system that balances three sources of evidence in a single accreditation system: (a) that the program's graduates are qualified, competent, and caring beginning teachers; (b) that the program faculty investigates the factors that improve program quality; and (c) that the program has the capacity for continuous program improvement. One aspect of the third factor is whether that program is offered by a regionally accredited institution and whether the institution is committed to the program (Murray, 2005).¹ In other words, TEAC asks that programs make systematic inquiries into what their graduates know and can do and uses the evidence that they obtain to assess and improve the effectiveness of the program.

TEAC's system was recognized by the Council of Higher Education Accreditation in 2001 and subsequently by the U.S. Department of Education in 2003 and 2005. The Council has 200 institutional members and has accredited 100 teacher education programs in over 15 states.

TEAC's assumption is that the evidence that the program actually

Frank B. Murray is president of the Teacher Education Accreditation Council, Washington, D.C., and H. Rodney Sharp Professor of Education and Psychology at the University of Delaware, Newark, Delaware. His e-mail address is fmurray@udel.edu

relies on to support its claims provides a sufficient basis for accreditation and for public assurance of the program's quality, which is required for accreditation. TEAC accreditation signifies that the program has the capacity for quality, the knowledge of how to enhance and monitor quality, and the skill to ensure that the programs' graduates are competent.

The factors in TEAC accreditation, including evidence of graduate competence, faculty inquiry and investigation, and a capacity for ongoing program improvement, are carefully calibrated so that weaknesses in one can be offset by strengths in others. Overall, however, insufficiencies in any one factor can threaten full accreditation by reducing the term of accreditation to two years. Failure in any two of the three factors leads to the denial of accreditation.

The system also recognizes that none of the currently available measures or assessments in higher education that meet any reasonable standard of validity is alone up to the task of assuring the graduates' competence. Thus, TEAC programs are free to use the measures upon which the programs actually rely in making their claims that their graduates are competent. Because no single measure is adequate, however, programs must employ multiple measures. Additionally, because the validity of all measures is suspect, programs also must provide local evidence of the reliability and validity of the measures that they employ.² Within these constraints, program faculty members are free to use whatever measures on which they actually rely to determine program quality, provided that the evidence meets a scholarly standard for evidence. This standard requires that the preponderance of the evidence be consistent with the claims and promises about its program. With regard to the key and unavoidable issue of the magnitude of the evidence, TEAC employs the following heuristic: absent any other standard accepted by the field, 75% of whatever scale is presented is considered sufficient.

TEAC, in fact, asks the program faculty to take a position on 20 categories of evidence available in the field and declare whether they have that evidence and whether they rely on it. Additionally, if they do not have evidence in certain categories, they need to state their reasoning (e.g., they do not value it, it is too costly or time-consuming to procure, it is confidential, it would be misleading, or they will acquire it in the future).

On the whole, the programs that seek TEAC accreditation have solid evidence that the institution is committed to the program. Because the institution that offers the program must be regionally accredited (or its equivalent), the institution's capacity in the traditional *in-put* areas cited in the federal regulations is established, and the only question is whether the program conforms to the institutional norms with regard

to these input areas of capacity. One line of evidence for commitment is that there is parity between the program and the institution overall.

TEAC's Expectations and Experience

TEAC expected initially to find evidence for the widely held belief that teacher education programs are out of parity with their institutions, that they are *cash cows* (i.e., high-volume programs run "on the cheap," whose considerable profits are used to run the more costly programs that the institution actually values). TEAC has found just the opposite so far in its sample of national programs. Specifically, teacher education programs are more costly than the norm, owing to required clinical experiences throughout the programs, the funding of cooperating teachers, special library and media collections of curriculum materials, and instructional technology.

The concerns that have been uncovered more often involve the faculty's ability to articulate and assess the quality of the program's control systems and the nature and analysis of the evidence on which the programs rely to support their otherwise confident claims that their graduates are competent teachers. Of more concern, at least at the outset of TEAC's work, was the lack of confidence that many faculties had in bringing forth the evidence upon which they actually rely and in acknowledging the weaknesses that their quality control system uncovered. However, the concerns are not of the magnitude of those claimed by the recent Ed School Project report (Honawar, 2006), which asserted that the quality of the nation's teacher education programs is so low that teacher education needed to be redesigned in the United States.³

The report, in fact, was based on surveys and interviews of alumni, school principals, and some deans of teacher education schools of only 28 programs and avoided the usual scholarly conventions in its text (such as precise descriptions of methodology, results, and data analysis). Nor are the concerns⁴ that TEAC uncovered completely consistent with those cited by the U.S. Secretary of Education in several recent speeches at the University of Virginia (University of Virginia, 2009) and Teacher's College, Columbia University (Teachers' College, 2009).

The 100 programs that TEAC has audited and accredited are a reasonable cross-section of the nation's programs and, in contrast to Ed School Project and the Secretary's samples, all had convincing evidence from multiple sources for their claims that their graduates were competent teachers. The sources were typically limited to grades given by faculty in the areas of education and arts and sciences; license test results; ratings given by students, alumni, cooperating teachers, employ-

ers, and clinical faculty members; and faculty evaluations of student portfolios and work samples. There were also many novel and tailored lines of evidence that various faculties provided. In all cases, the faculty had investigated the reliability and validity of the measures that they cited and had found that there was a convergence of multiple measures that were consistent with the program's claims that their graduates understood their teaching subjects and pedagogy and could teach in a caring and effective manner.

The Balance Between Student and Faculty Learning in Accreditation

To illustrate the balance between evidence of student achievement and evidence of program improvement, we need to consider one of the early TEAC audits of the evidence that a program submitted for its claims. In this audit, the auditors were attempting to verify the evidence of the program's claims of student achievement, and the audit trail led to some student folders that revealed remarkably low SAT scores. The program had been silent about its students' SAT scores, presumably because they were embarrassed by how low they were. As often happens in accreditation, and despite TEAC's repeated assurances that only unaddressed weaknesses, not weaknesses themselves, were problematic, the program decided to bury its perceived shortcomings and not to speak about them at all. It turned out that, while the program was disguising what it took to be evidence that weakened its claims that its graduates were competent, it really had denied itself the opportunity to provide stronger evidence that it had a robust system of quality control and inquiry.

What the program had done over the years, in frank recognition of its low SAT scores, was to simply accept the fact that its students, typically first-generation rural college students with English as their second language, were low scorers on the SAT. They first designed their own "study skills measure" in an effort to at least accept into the program students who knew how to study. Their homegrown instrument was a failure, and they shortly gave it up and replaced it with an intensive effort to teach their students how to prepare for and pass standardized tests. They aligned their curriculum with the state license test, ran workshops on test-taking skills, paid for practice tests and, in the end, their students had nearly the highest pass rates on the state's license tests.

By hiding their low SAT scores, they had denied themselves the opportunity to present a convincing case for a robust quality control system. They had sought to improve the program by basing program decisions on evidence of student performance. As such, they shaped their program to

respond to evidence of student success and failure, to discard unproductive approaches, to refine student responses, and to gradually improve the program's structure so that it yielded one of the highest pass rates in their state. Thus, the apparent weakness in the evidence for student competence was actually a strength in the evidence that the program could have advanced in regard to its inquiry and improvement.

In another case, the auditors found that the program's quality control efforts and inquiry were not systematic but rather idiosyncratic to each faculty member's preferences and style. This weakness was compensated for by the very high levels of performance of the program's students, owing to the selective nature of the college as well as by the competence of the individual faculty in their subjects and in their personal and tailored, although not coordinated, advising of their students.

A Balance of Status and Value-Added Claims

A faculty's case for evidence of student achievement is not a value-added case. It only requires evidence about the *status* of graduates, not how well they perform in comparison to some other group or in comparison to how much less they knew at some earlier points in the program. The claims associated with this evidence, in other words, need not be claims about the source of the graduates' competence or how much it changed over the course of the program.

Claims about *cause* and *growth* of student achievement, however, are encouraged and expected in connection with the evidence of the faculty's inquiry into the quality of their program. TEAC expects that the program faculty members are curious about the program's effectiveness and its added value and that the faculty members conduct research into the factors associated with the effectiveness of its program.

The public and employers, in comparison, are largely concerned with only the *status* of the program's graduates. They want to know whether the graduates are competent, caring, and qualified more than they want, or need, to know how that competence was acquired or whether the graduates are more competent than some other group or when they began the program. TEAC wants to know the status, as well, but for a different reason; TEAC uses the information as a key ingredient in its judgment of the quality of the program.

The institution and the program faculty, in contrast, may be more interested in knowing which attributes of the program contributed to the graduates' competence. Those students who enrolled in the program, their parents who paid tuition, and benefactors who funded scholarships might also have a keen interest in whether any value was added by the

program and whether the students showed development over the course of the program. Indeed, in communicating with the public, the program faculty and institution undoubtedly make ambitious claims about the effectiveness of the program and the value that is added from the college experience. This information is captured in their evidence of the program's quality management and control. TEAC wants to know this, as well, but again for a different reason: TEAC uses the information as another key factor in its judgment of the quality of the program

Unanticipated Benefits for Programs of TEAC's Audit

One of unanticipated outcomes of the TEAC's audits to date is that the auditors invariably find better evidence for the program's claims than the program had advanced in support of them. For example, in one case, the auditors found that the program had used pupil evaluations of student teachers but had not seen the potential in the information for evidence of their case that their students could teach caringly and effectively. In another case, the auditors discovered that the institution's arts and sciences departments regularly used teacher education master's students as graduate teaching assistants in their courses because the master's students knew their subjects well and received high course evaluations, a fact not known or cited by the education faculty in its case for accreditation.

Sometimes the auditors find evidence that indicates that some claims of the program are unfounded. For example, one program claimed that the state's license examinations were unrelated to its program's goals and accomplishments. The program prided itself on its personal knowledge of each student and its ability to spot and cultivate teaching talent. The auditors assembled a list of student names and probed the program's claims by having the faculty rate the students by name only on the degree to which they satisfied the program's claims and goals. The auditors found that the faculty's ratings were highly related to each other, to the grades that they and others gave, to the cooperating teachers' ratings, and, to the faculty's surprise, to their students' scores on the state's license tests. Moreover, those students who years later earned national board certification were spotted as talented by this faculty when they had been students in the program. Another program asserted that the faculty members were more demanding of their own advisees than they were of other faculty members' advisees, but the auditors in this case found just the opposite in their closer examination of the evidence; other faculty in the program rated the advisors' students lower than the advisor had rated them.

In another program, the auditors uncovered a flaw in the institutional research practices of the institution that extended well beyond the teacher education program to the entire institution. Specifically, the number of evaluations (repeated measures) was conflated with the number of students so that the results of any analysis done by the institution was uninterpretable. Auditors on more than one occasion found that faculty used the mean of the sub-means as the grand mean without awareness of the errors that were introduced into their inquiry by this practice when the subgroups were of unequal sizes. On another occasion, at least one faculty rater used the same written justification for her ratings, regardless of the student's performance, thus invalidating the rating instrument's assumption of independent assessment.

Regrettably, auditors have sometimes uncovered errors in some faculty members' evaluations of subject matter content in lesson plans. In one case, the lesson had as its objective to explain why stars are seen at night and not in the day, and the student teacher erroneously taught: "When we are away from lots of lights, we are able to see more stars because they are much brighter." This is because "the stars are only seen at night because they are able to shine through the material in the night sky. This material is less dense than during the day."

Auditors often probe the faculty's ability to reliably use their own evaluation rubrics by showing them a videotaped lesson and having them score it. Typically, the raters show high levels of agreement (within one unit of each other) and come to the same conclusion with regard to whether the lesson was below, at, above, or well above the program's standard for a student teacher. The results, however, are more instructive and valuable for the faculty when the auditors find unacceptably high levels of inconsistency between the raters' evaluations or even of the degree to which the raters can agree that evidence was even present for various items on their evaluation instrument.

Occasionally, auditors present faculty and students with classic teaching dilemmas, along the lines of the prompts used by National Board for Professional Teaching Practices, to elicit the sophistication of a candidate's teaching practices and their alignment with the program's goals. Thus, an auditor might ask how the program has guided its students with regard to whether they should seek the IQ scores of their pupils before the school year begins and, regardless of whether they decide to acquire them or ignore them, how the program advised them to think about the issue. The point of these examples is that the auditors' probes reveal information, otherwise unknown to the faculty members, which is valuable to the program faculty as they consider ways to improve the quality of their programs.

A More Optimistic Picture of US Teacher Education

In closing, it is fair to ask whether TEAC's approach to accreditation has shed light on the quality of teacher education and whether the collective evidence from its accreditation work is in line with current investigations and pronouncements of the national quality of teacher education. TEAC now has a wealth of data from the 100 programs it has accredited. Some of the findings from this work are newsworthy as well as counterintuitive. For example, the grades that teacher education students earn in courses in the disciplines of the arts and sciences are invariably equal to or better than the grades that the arts and sciences majors earn in the same courses. This finding holds for all kinds of institutions—flagship research universities and small liberal arts colleges.

As a more troubling example, as seen in Tables 1 and 2, the performance of teacher education students in the clinical portions (capstone) of the program, for two representative programs, indicates that such performance is strikingly unrelated to performance in every other part of the program (including the license test scores). The components of clinical performance (ratings by clinical faculty, cooperating teachers, and student teachers) are, fortunately, highly correlated with each other, but they are not related to license test results or to grades in the teaching subject and in pedagogy (which are themselves also highly related

Table 1
*Correlations among Six Academic Sources of Evidence
and Two Clinical Sources in a Colorado Teacher Education Program*

Variable	Pedag. License	A&S License	SAT-V	SAT-M	ACT	Cooper. Teacher	Faculty Superv.
GPA	.16	.47*	.27	.49*	.51*	-.08	-.17
Pedagogy License Test		.61*	.68*	.24	.30	.24	.20
Arts&Science License Test			.71*	.43*	.49	.00	-.09
SAT Verbal				.39*	.73*	-.03	-.04
SAT Math					.63*	-.51*	-.42*
ACT						-.22	-.46
Cooperating Teacher							.81*

Note. The correlations in **bold** are between clinical and non-clinical items; the correlation between the two clinical measures is .81; and the correlations among the academic measures are invariably positive and statistically significant. $p < .05$.

Table 2
*Correlations among Five Academic Sources of Evidence
 and Four Clinical Sources in a New York Teacher Education Program*

Variable	Method GPA	A&S GPA	Educ. GPA	LAST Test	ATS Test	Clin'1 1	Clin'1 2	Clin'1 3	Clin'1 4
Major GPA	.59*	.63*	.68*	.58*	.45*	.05	.07	.05	.28*
Method GPA		.58*	.91*	.56*	.49*	.02	.05	-.02	.19
A&S GPA			.69*	.37*	.21	.18	.20	.30*	.37*
Education GPA				.55*	.45*	.08	.13	.08	.30*
LAST License					.68*	.07	.00	-.02	.17
ATS License						-.18	-.16	-.19	-.01
Clinical 1							.66*	.70*	.50*
Clinical 2								.80*	.58*
Clinical 3									.68*

Note. The correlations in **bold** are between clinical and non-clinical items; and the correlations among the clinical and non-clinical measures are invariably positive and statistically significant. $p < .05$.

to each other). These findings also hold throughout the country, in large and small programs that TEAC has accredited.

These results present quite a different picture of the health of teacher education than do the typical reports of low-quality teacher education that are found in the recent Ed School Project reports and other alarmist reports on teacher education. The TEAC data show that the nation's prospective teachers are quite able in their teaching fields, or as able as majors in those same fields, and they show that there is another dimension to their competence, one seemingly independent of that captured by the typical academic assessments. This dimension is internally consistent, and consistent with the program's claims that its graduates can teach. This other dimension also indicates, at least preliminarily, that schemes for recruiting new teachers that rely solely on subject matter knowledge expertise are likely to be insufficient. They also show that the typical teacher education program has a striking disconnection between the academic and clinical components of the program. Apparently, significant amounts of what teacher education students are required to study has little influence on their teaching.

These findings also indicate that schemes to recruit teachers solely on the basis of their subject matter expertise, a practice often advocated in times of teaching shortages, are also likely to be inadequate. At the same time, these findings are not supportive of traditional teacher edu-

cation programs, whose many requirements cannot be shown, at least in these TEAC accreditation samples, to be related to the candidate's teaching competence.

Of course, these findings require further inquiry, as it may turn out that the lack of a correlation between the clinical and other program components is more parsimoniously attributed to restricted variance, limitations in the coverage and overlap in clinical and other assessments, or that the lack of a significant linear correlation may be due to a threshold effect in which only a certain modest level of academic accomplishment is required for teaching competence, and accomplishment beyond that threshold value has diminishing influence (i.e., the relationship may be curvilinear).

TEAC, as part of its academic audit, has recently begun to ask that students, faculty, and cooperating teachers respond to a series of survey questions about the adequacy of the program (whether aspects of the program were *inadequate*, *barely adequate*, *adequate*, *more than adequate*, or *excellent*). To date, as Table 3 indicates, these survey results demonstrate that students, faculty, and cooperating teachers, in contrast to the Ed School Project findings, rate nearly all aspects of the programs in the *more than adequate* range.⁵ All but two of the differences in means (Subject Matter and Pedagogy Courses and Subject Matter and Pedagogical Faculty) are statistically significant ($p < .001$). Thus, students see their own teaching skill as superior to their knowledge of their subject matters and pedagogy, but the source of this superiority does not seem to be wholly in their clinical courses or from the clinical faculty.

As Table 4 shows, the students see their own understanding of their teaching subjects, their understanding of pedagogy, and their ability to teach in a caring and effective manner as somewhat independent of their overall grades in the program (3.7/4.0, $SD = .31$) and their ratings of the adequacy of the program faculty and courses. They see the adequacy of the faculty and the adequacy of the courses, by contrast, as highly

Table 3
Means and Standard Deviations of Program Students (N = 568)
from Eight Accredited Programs

Topic	Adequacy of One's Own Ability	Adequacy of Course	Adequacy of Faculty
Subject Matter	4.46 (.71)	4.27 (.85)	4.37 (.82)
Pedagogy	4.38 (.74)	4.32 (.80)	4.34 (.81)
Teaching Skill	4.71 (.55)	4.12 (.93)	4.18 (.91)

Note. 1=Inadequate, 2=Barely adequate, 3=Adequate, 4=More than adequate, 5=Excellent.

Table 4
*Correlations of Student Ratings of their own Knowledge
 and Teaching Skills with their Ratings of their Courses, Faculty, and GPA*

Topic	Own with Course	Own with Faculty	Own with GPA	Courses with Faculty ^a
Subject Matter	.44**	.40**	.15**	.71**
Pedagogy	.45**	.43**	.16**	.71**
Teaching	.34**	.27**	.12**	.75**

Note. a. Correlations are between student ratings of the courses and their ratings of the faculty in each area; $p < .001$.

related to each other. Thus, it is not that there are not highly correlated dimensions in the survey results but rather that the students believe that their own expertise has its sources elsewhere.

These results contrast with those of the Secretary of Education's citation of the 2006 report by Arthur Levine, former president of Columbia's Teachers College, in which 61% of educators surveyed stated that their colleges did not offer sufficient instruction to prepare them for the classroom. While the students see that their courses and faculty are highly similar in adequacy, the adequacy of their own knowledge and skill is relatively less related to the grades that they have earned or to their courses or faculty, particularly with regard to the clinical courses and faculty.

Not all the news from TEAC's accreditation work is encouraging, however. TEAC bases its accreditation decision on what it calls an *Inquiry Brief*, a research monograph in which the program presents the evidence that it has in support of the hypothesis that its graduates are competent (i.e., know their subject matters, understand the pedagogical literature, and can teach in a caring and effective manner). To date, all programs have had some difficulty in writing these monographs, which is somewhat surprising because the faculty otherwise regularly publish their own research findings in the scholarly literature. However, their efforts to turn the tools of their scholarship on their own programs often fall below the standard of acceptable scholarship and reveal serious weaknesses in their grasp of research methodology. This finding may be attributable to the segregation of roles in schools and departments of education between those who take responsibility for the education of the next generation of teachers and those who take responsibility for the education of the next generation of faculty members in education.

Implications

These findings from the audit of accreditation self-studies (TEAC *Inquiry Briefs*) have potentially important implications for the design and rationale of teacher education programs, but require confirmation and deeper analysis. More to the point here is that they also support the view that evidence derived from the coordination of program accountability and program improvement can provide an adequate basis for program accreditation.

Notes

¹ The evidence for institutional commitment is taken by TEAC to be parity between the program and the institution with regard to the capacity dimensions identified by the US Department of Education (e.g., faculty, facilities, resources, student support services). The TEAC argument is that, because regional accreditation demonstrates that the institution satisfies the federal capacity standards for quality, parity demonstrates that the program has satisfied the same standards.

² Standardized test-makers establish the reliability and the validity of their tests for a standardized representative sample. The local program, particularly one that claims that it is distinctive, unique, or of higher quality than most other programs, cannot simply assume that the test-makers' findings about the sample hold for the students in the program. They must make their own determination by investigating the reliability and validity of these standardized tests.

³ The report did not discuss the newer TEAC, claiming that it was too new and too small to evaluate its impact on the field. TEAC was less than a decade old and had accredited fewer than 50 programs at the time of the report, in contrast to NCATE's 75-year history and its accreditation of approximately half of the nation's 1,300 teacher education programs at the time.

⁴ In a speech at Teachers College in New York, Duncan stated, "By almost any standard, many, if not most, of the nation's 1,450 schools, colleges, and departments of education are doing a mediocre job of preparing teachers for the realities of the 21st century classroom."

⁵ While the ratings are generally high (4.0+/5.00), they are not undifferentiated ceiling effects, as there are significant differences among some components in some programs (e.g., technological adequacy, multicultural understanding).

References

- Honawar, V. (2006). Prominent teacher-educator assails field, suggests new accrediting body in report. *Education Week*, 26(4), 1-4.
- Murray, F. (2005). On building a unified system of accreditation in teacher education. *Journal of Teacher Education*, 56(4), 307-317.
- Teachers' College. (2009). *Teacher preparation: Reforming the uncertain profession*. Retrieved November 18, 2009, from <http://www.edgovblogs.org/dun>

can/2009/10/a-call-to-teach
University of Virginia. (2009). *A call to teach*. Retrieved November 18, 2009,
from <http://www.tc.columbia.edu/news/article.htm?is=7179>