



NCHEMS

Piloting a New Approach to Accreditation
in Teacher Education

An Evaluation of the TEAC/FIPSE Project

Peter T. Ewell

December 20, 2001

**National Center for Higher Education Management Systems
P.O. Box 9752 • Boulder, Colorado 80301-9752 • (303) 497-0301**

PILOTING A NEW APPROACH TO ACCREDITATION IN TEACHER EDUCATION

An Evaluation of the TEAC/FIPSE Project

Introduction and Overview

On January 1, 1998 FIPSE funded the Council of Independent Colleges (CIC) to develop and test a new model of specialized accreditation for teacher education programs, based on student outcomes. The model was to be developed and implemented through a newly-created accrediting association—the Teacher Education Accreditation Council (TEAC)—which would eventually seek recognition by the Council on Higher Education Accreditation (CHEA) and the U.S. Department of Education (USDOE). Funding for FIPSE was sought to help underwrite the process of designing a new approach to review, and to support a pilot-test of this approach at six institutions. This process was expected to take three years, but for a variety of reasons, the grant period was extended at no additional cost through December 31, 2001.

The National Center for Higher Education Management Systems (NCHEMS) was engaged to provide an external evaluation of the TEAC pilot, and began tracking its activities formally and collecting data in the summer of 2000. In the course of the evaluation, NCHEMS staff engaged in three distinct kinds of activities. First, staff participated in TEAC meetings to become more familiar with the origins of the review process and the evolving reactions of participants and potential participants. Second, beginning with a site visit to the TEAC office at the University of Delaware, NCHEMS obtained and reviewed in detail all documents generated in the course of the project to that point. These included multiple drafts of training materials for reviewers and for candidate programs, as well as correspondence and documents generated in the course of reviewing the first three pilot institutions to complete the process (the University of Virginia, Fort Lewis College, and Texas Lutheran College). Documents and correspondence related to approximately ten additional institutions that were not as far along in the process as the initial three, selected randomly, were also examined at this time. At a later point in the fall of 2001, revised TEAC documents were reviewed that reflected lessons learned during the pilot. These included materials associated with TEAC workshops on auditor training and preparing the Inquiry Brief, as well as documentation associated with a fourth institution (Western State College) which had made considerable progress by that point. Finally, staff conducted interviews with TEAC staff on site and interviewed key participants at each of the three pilot institutions to obtain their impressions of the process.

Findings of the evaluation are presented in four subsequent sections of this report. The first reviews the major elements of the TEAC review process by way of background. The second compares the design and documentation of the TEAC review process with established “best practices” for quality assurance review both nationally and internationally. The third examines the experiences of the three “first-wave” pilot institutions in some detail to determine both how well the review process actually functioned in the field and the primary lessons learned from this exercise. The fourth and final section addresses the implications of TEAC’s experience for other forms of quality review in higher education including both voluntary accreditation and state-level accountability programs.

Origins of TEAC and the TEAC Review Process

The origins of TEAC lay in growing dissatisfaction on the part of a significant group of colleges and universities with the established process for accrediting undergraduate and graduate teacher-education programs in the U.S.—that administered by the National Council for Accreditation of Teacher Education (NCATE). NCATE’s standards, like those of most established accrediting bodies, rest heavily on criteria such as facilities, faculty credentials, research production, and the existence of independent governance arrangements for Schools of Education. Smaller institutions—and especially the private liberal arts colleges that count many classroom teachers among their graduates—cannot meet such requirements even if they wished to. Many larger (and often highly recognized) Schools of Education could do so, but simply find the NCATE process not worth engaging in. The result is that some 60% of the more than 1100 Schools of Education in the country remain unaccredited. To help address this problem, representatives of unaccredited institutions gathered under the auspices of the Council of Independent Colleges (CIC) approximately six years ago. Their conclusion, after several attempts to work with NCATE, was that an alternative accrediting body was needed.

Creating an accreditation process from the ground up gave the founders of TEAC an unparalleled opportunity to respond to growing demands that accrediting bodies pay greater attention to student learning outcomes in their review processes. Admittedly, accrediting bodies were no strangers to “assessment” by 1995. Indeed, all six regional bodies and many specialized accreditors in fields like engineering and health care had adopted assessment-related standards and review processes by the time TEAC was established. But most such assessment requirements were implemented as “add-ons” to the existing accreditation process rather than being integral to it.

Like the leaders of other accrediting bodies, the founders of TEAC were also intrigued by “academic audit” processes as advocated by a number of higher education observers in the mid-1990s (e.g., Graham, Lyman and Trow 1995; Dill, Massey, Williams and Cook 1996). These approaches emphasize the use of clear review protocols and highly-trained visiting teams to examine institutional quality-assurance practices on site, in preference to comprehensive narrative self-studies prepared by the program or institution and a subsequent site-visit undertaken by largely untrained peer reviewers. But proposed “academic audits” rarely centered on student learning, concentrating instead on internal institution or program-level quality assurance processes like course and curriculum approval and review, or the review of teaching. Recognizing the need to deal with student learning more directly, the architects of TEAC’s approach decided to design a set of audit-like methodologies directed specifically toward validating the actual evidence of student learning advanced by candidate programs.

The TEAC review process thus begins with the preparation of an “Inquiry Brief” by the program’s faculty structured around three “quality principles.” These include a) the establishment of clear goals for student learning (and their conscious application in curriculum design and delivered pedagogy), b) the quality of the evidence that the program advances to support its claims for student learning, and c) the program’s application of this evidence in the form of continuous improvement. In an attempt to harness the language and habits of scholarship, the Inquiry Brief is deliberately modeled after a publishable journal article in educational research. Consequently, its original contents included a “literature review” to

explain the program's rationale and grounding philosophy and intended outcomes, a "methodology section" in which the various methods used to assess student learning are presented and analyzed, a "results" section that presents the evidence, and a "discussion/implications" section in which the program describes the uses that it has made of this evidence. The centerpiece of the presentation, however, is the program's claims for student learning, substantiated by an advanced and analyzed body of research-based evidence. In addition, an appendix to the Inquiry Brief provides results of a mandatory "self-audit" of the program's internal "Quality Control" system undertaken by its own faculty. Ideally, this begins with a set of randomly-chosen students and involves examining actual departmental records to see if established standards and procedures are actually being followed. Examples of the kinds of things that are supposed to be addressed in the self-audit range from the narrowly bureaucratic (for example, whether student academic records are complete and accurate), to the broadly academic (for example, whether course assignments are capable of eliciting intended program outcomes or whether grading standards are appropriate to assure desired levels of attainment).

Following submission of the Inquiry Brief, the program is visited by a team of at least two trained auditors, who, in the future, will be full-time members of TEAC's staff. In an initial meeting, the substance of the visit is defined by clearly delineating the program's claims and the evidence on which these claims are advanced. In preparation for this meeting, auditors annotate the Inquiry Brief to identify potential claims and note whether or not concrete evidence for each of them is provided. The audit team then attempts to validate the evidence submitted through in-depth inspection of the actual methods used to collect and analyze the data. This process can involve directly examining instruments and administration procedures, reviewing raw data and data transcriptions, interviewing participants (including both faculty and students), re-running statistical analyses on site, or examining records of curricular changes or other claims the faculty have made about the use of results. At any point, the team can request additional information—though most of the needed materials have already been assembled based on prior communication with the program's faculty. As they review the evidence, moreover, TEAC auditors use formal evaluation guides and rubrics to rate and comment upon what they find. But the standards on which these evaluations are based are drawn directly from the literature on educational research. As such, they include criteria of validity, reliability and representativeness, as well as the appropriate application of relevant analytical and interpretive techniques. At the conclusion of the on-site review, audit findings are forwarded as a written report (together with the original Inquiry Brief) to an independent panel, which makes a formal recommendation regarding accreditation to the TEAC Board.

Several features of this process contrast markedly with standard accreditation practice. First, the process is focused almost exclusively on student learning and the evidence that it is taking place. Although TEAC does look at each program's resource-based capacity (largely as defined by and at the behest of the federal government), the vast majority of its attention is devoted to this single topic. Second, review roles that might otherwise be in conflict are strictly segregated. A TEAC staff member serves as "coach" to the program and consults on the preparation of its Inquiry Brief. But this individual is a member of neither the audit team nor the accreditation panel. Auditors, in turn, do not make an accreditation recommendation to the panel, except in the area of the evidence of the institution's commitment to the program. They simply report whether or not they found the evidence advanced by the program credible. This segregation of duties is designed to avoid some of the natural confusions inherent in the "consultant-evaluator" role that

is a typical characteristic of peer review. Third, the use of review protocols and audit-trail methods renders the process remarkably free of rhetoric. Indeed, the principal purpose of the initial audit meeting is to extract precise and verifiable claims about learning from the Inquiry Brief's surrounding prose and to allow the program a chance to clarify or to take such claims off the table if they are unable to be substantiated. Finally, the entire process rests on standards of evidence that ought to be widely recognized by faculty in the field of education. In this respect, of course, teacher education has a built-in advantage over other professional fields because practitioners are presumed to be schooled in such standards and to routinely apply them in their own scholarship.

Comparison with Established “Best Practice”

TEAC's creators did not base the design of its review process on “any existing model,” though as noted above their thinking was heavily influenced by the principles of self-regulation that began to emerge in the higher education literature in the mid-1990s. Nevertheless, many aspects of the resulting design strongly resemble leading-edge quality assurance practices for higher education in other nations, as well as quality assurance practices in fields other than education in the U.S. One approach to evaluating the FIPSE pilot, therefore, is to examine how well the process TEAC designed and implemented corresponds with approaches of proven value and effectiveness that are already in place in other settings. The most important of these are a) the “academic audit” approach to quality assurance in higher education developed in the U.K. and now practiced in much of Western Europe and Australasia, and b) the Malcolm Baldrige National Quality Award for industries and organizations in the U.S. administered under the auspices of the U.S. Department of Commerce. Key attributes of these approaches are listed briefly under separate headings below, together with how the TEAC design compares:

- **Focused Scope.** Rather than attempting to examine every facet of an institution or organization, “best practice” review techniques are carefully focused on only a few dimensions of performance. In the case of the academic audit, the primary focus is placed on the institution's internal mechanisms for ensuring the quality of instruction. In the case of the Baldrige, the primary focus is on internal monitoring of quality and how the organization serves its customers. TEAC's review process is focused almost exclusively on student learning and on the quality of the evidence advanced in support of the program's claim that its students are achieving established learning outcomes. Its three “Quality Principles”—Evidence of Student Learning, Valid Assessment of Student Learning and, Institutional Learning—are particularly tightly focused when compared to the standards of other accrediting bodies. At least in theory, this allows the limited attention and energy available for conducting a review to be devoted to looking at a few important things in depth, rather than attempting to “cover” all aspects of program functioning. To even further emphasize this focus mid-way through the grant period, TEAC moved those few “resource and process” topics originally included in its design as a fourth quality principle to an Appendix of the Inquiry Brief.
- **Use of Explicit Protocols of Inquiry.** Standard accreditation practice relies primarily on the instincts and experience of peer reviewers to determine what lines of inquiry to pursue with a given institution and what particular questions to ask on site. In contrast, academic audit methods and the Baldrige follow written and field-tested review protocols. On the one hand,

this promotes equity because all those examined go through the same process. On the other, it ensures that all areas of importance are addressed, and signals the relative priorities that reviewers should attach to particular topic areas. The TEAC process uses several different protocols of this kind. The first is in the form of “clarifications” which are addressed to the institution at the outset of a site visit to establish the precise nature of the claims being made about student learning. “Clarifications” are developed by auditors in advance of the visit, based on an examination of the institution’s Inquiry Brief. They are then recorded in the form of a special worksheet indicating where in the document the need for clarification arose and a specific question or set of questions to be posed to the institution; the worksheet also includes a column to record the institution’s response. A similar use of protocols is in the form of “Precision Probes,” which are used by auditors to examine particular claims in greater depth—and, in particular, whether or not statements made can be justified in the light of available evidence. These are also prepared in advance, specifically for the institution being examined, based on careful prior analysis of the submitted Inquiry Brief.

- **Use of Audit Methods and Audit Trails.** As suggested by the term “academic audit,” reviewers in Europe and Australasia spend most of their time on site visits physically verifying specific forms of evidence. This contrasts strongly with the largely narrative and testimony-based approach typical of U.S. accreditation processes, in which a review team first reads a comprehensive self-study then conducts a wide range of interviews to ask questions about particular topics raised in the self-study. Audit teams in the U.K. and Hong Kong spend their time instead selecting a limited sample of actual instances of the processes they wish to audit. For example, to verify the accuracy of claims made about the completeness of student records, the team might select and examine twenty-five sets of records chosen randomly. Similarly, to verify claims made about the seriousness with which individual academic departments engaged in a review of teaching, the team might select and examine teaching-review reports drawn from five randomly-selected departments, then interview only the faculty associated with those departments. Such review processes have the considerable virtue that they are almost impossible to prepare for except by actually *doing* what is intended: the institution has no advance warning about which particular entities will be sampled for audit. But they do require institutions and organizations to have relatively complete records in hand in order to have “something to audit.” TEAC uses audit methods of this kind in two ways. The most prominent is the process of directly evaluating the veracity of submitted evidence about student learning. Here TEAC auditors examine the instruments and methodologies used to gather assessment evidence, and frequently re-calculate statistics and verify administration procedures through direct inspection of individual student performances. It is important to note here that no other quality assurance agency in higher education in the U.S. or abroad engages in this level of detail when examining assessment-based evidence of student achievement. A second way in which “audit” methodologies arise is in TEAC’s requirement that each program conduct a self-audit of its own quality-assurance processes. This is generally accomplished by selecting a representative sample of student cases and following associated “audit trails” to determine if all appropriate and required quality-process procedures were actually carried out for each of the cases selected. TEAC auditors then “audit the audit” by verifying that the institution chose cases appropriately and carried out its self-audit responsibly and systematically.

- Use of Rubrics to Help Make Judgements.** Rubrics are the counterpart of protocols in the process of drawing summative conclusions about organizational performance. They are extensively used in the Baldrige review process and, indeed, much of a Baldrige reviewer’s time is spent “scoring” the organization under review multiple times (both individually and as a team). Rubrics of this kind contain concrete descriptions of various attributes of an organization (structures, processes, behaviors, results, etc.) arrayed at multiple levels of performance. Reviewers typically consult such rubrics frequently, assign an initial score, discuss the result with fellow reviewers, and revise and update their initial scoring based on the accumulating body of evidence gathered in the course of a review. TEAC’s use of rubrics follows the strict segregation of duties that the process establishes among staff “coaches” who help the institution prepare for review, “auditors” who visit the institution to clarify claims and verify evidence, and “review panels” who examine the evidence provided and render a recommendation regarding accreditation. Because only the latter role involves making a summative judgement about performance, the panel is the only set of actors in the process that requires the use of rubrics and heuristics. The heuristics and rubrics actually employed by TEAC are relatively simple compared to those used by such processes as the Baldrige, consisting of charts that note only the attributes of programs that are “below standard” and “above standard.” But they are applied to eleven distinct aspects of the program (and over sixty distinct sub-attributes within these). They are accompanied by a Table of Decision Outcomes, which is used to summarize rubric results on each dimension in order to guide the Panel to an appropriate accreditation decision.
- Team Training.** “Best practice” approaches including both academic audits and the Baldrige award devote considerable attention to training reviewers. In the former, auditors undergo a minimum of two days of intensive training that includes case studies, simulated interviews and audit tasks, and practice in applying rubrics. In the case of the Baldrige, training is even more extensive and results in official certification as a Baldrige Examiner. This situation contrasts markedly with typical accreditation practice in which primary reliance is placed on the experience and instincts of peer reviewers. Training in both regional and specialized accrediting bodies in the U.S. is typically rudimentary—involving less than a full day spent reviewing the agency’s standards—and it is not required that reviewers complete the training at all. TEAC’s approach to this matter is thorough and consistent with “best practice” standards. Only trained auditors are used to staff audit visits, and all are compensated by TEAC. For the pilot, eight auditors were available, many of whom had helped design the audit process so were highly familiar with it. Auditors are trained in two-day workshops that involve simulated tasks and case studies, together with practice using (and constructing) the required rubrics and rating tools. Particular emphasis in these workshops is placed on actually *practicing* the techniques of audit—for example, how to read an Inquiry Brief to identify embedded claims, how to develop effective precision probes, or how to validate submitted evidence about outcomes by examining source datasets or re-calculating results using the original data. These workshops were fully developed and tested during the grant period. In the case of the TEAC accreditation panel, a training program was also developed and implemented in the form of a two-day workshop containing all of these features, together with practice in applying audit forms and decision rubrics. A particularly important aspect of this training—also present in TEAC’s workshops for institutions directed toward preparing Inquiry Briefs—is its hands-on, practical quality. Participants are engaged

in multiple opportunities to actually use the tools of audit in applied “case-based” situations. Training is thus focused not on learning *about* the review process, which is the primary content of most accreditors’ team-training activities, but on actual practice in *doing the work* of review.

In these five areas, the design of TEAC’s review process strongly resembles “best practice” in quality reviews abroad and in other sectors, and differs markedly from typical U.S. accreditation practice. But TEAC’s review process also differs from these “best practice” models themselves in some important ways. Furthermore, TEAC’s approach has been heavily criticized by a number of observers on the basis of its perceived conceptual shortcomings. Both topics deserve some comment as part of this evaluation.

- **Differences from Established “Best Practice” Models.** Unlike the “academic audit,” TEAC’s review process is focused on the evidence that the program presents to sustain its claims about student learning outcomes. Audit methods are thus employed to examine the program’s claims about student academic achievement *directly*, rather than being applied to “quality processes” that are assumed to be associated with learning such as curriculum review, course approval, and teaching evaluation, as is typical in the academic audit. The Baldrige and the academic audit also have a heavily “managerial” feel to them in both language and in the ways in which auditors actually behave. In contrast, TEAC’s process (at least conceptually) attempts to incorporate the values, language, and practices of educational research—a topic about which, presumably, teacher education faculty are knowledgeable. Both of these features are theoretically admirable, moving TEAC’s approach beyond current “best practice” elsewhere. But both were also heavily tested by the pilot reviews undertaken at the three institutions.
- **External Criticisms of the TEAC Process.** Two major perceived conceptual shortcomings of the TEAC review process have been prominently identified by critics. The first is essentially that the TEAC process ignores the important role that sufficient resources play in delivering a high-quality teacher education program. This position is usually associated with a claim that resource-based standards—as required by NCATE, for example—should be an important part of any accreditation process. Two responses to this criticism can be identified, and both have been used by TEAC. First, TEAC *does* take inputs into account in the form of a “Statement of Commitment,” now expanded to include evidence that the institution has the capacity to offer a quality program, which must be submitted by all institutions as a required addendum to the Inquiry Brief. This addendum requires the program to briefly address a set of relatively traditional “standards” used by the Department of Education on such matters as curriculum, faculty, facilities, fiscal and administrative capacity, and student support services. Given TEAC’s focus on direct evidence of quality, the placement of this material as an addendum is appropriate; but it is not accurate to say that attention to resources is absent or unconsidered. Indeed, all three audit reports prepared as a part of the pilot referenced this material in some way. More importantly, TEAC is correct in contending that the relationship between such resource-based factors and high-quality learning outcomes has never been empirically demonstrated. And TEAC has repeatedly offered (e.g., Frank Murray’s letter to the editor of *Education Week* of July 18, 2002) to adopt any such capacity standard, if its validity can be substantiated by available evidence.

A second claimed conceptual shortcoming is that by basing the accreditation process on the substantiation of claims made by the program under review TEAC “has no standards.” Again, two responses are warranted. First, “goal-based” accreditation is far from uncommon in the U.S. All of the regional accrediting bodies essentially follow this approach and have rarely been accused of “lacking standards” for doing so. Like these bodies, TEAC’s assumption—and it seems a reasonable one—is that a candidate program would be foolish to make claims to “sub-standard” performance simply to ensure that such claims could be easily demonstrated. Peer review processes all over the world bear out this assumption: the feared phenomenon simply does not occur. More importantly, in its demand that candidate programs show evidence that graduating students demonstrate subject matter knowledge (including learning how to learn, multi-cultural perspective, and knowledge of technology), liberal education, pedagogical knowledge, and teaching skill, TEAC is actually quite specific with respect to the “standards” it expects. Experience with the three pilot reviews demonstrated that institutions will exercise considerable rigor to make the case that their graduates exhibit these qualities. At the same time, TEAC has shown considerable forethought and competence in developing consistent rubrics to make judgements about program performance. As noted above, summative rubric-based judgements of this kind are characteristic of current best practice in other fields and abroad. And in such settings, they operate in exactly the same fashion as traditional “standards.”

Lessons of Implementation

Three institutions formally piloted the TEAC review process as part of the FIPSE project in 2000-2001: Fort Lewis College, Texas Lutheran University (TLU) and the Curry School of Education at the University of Virginia (UVA). All three completed Inquiry Briefs in the fall of 2000, were visited by an audit team in the spring of 2001, and received an accreditation decision from TEAC by June of 2001. [Note: Fort Lewis College changed its objective to obtaining “Pre-Accreditation Status” in the course of the review as a result of discussions with TEAC staff on the program’s changing curricular requirements made necessary by the State of Colorado.] Though few in number, these institutions were remarkably diverse, comprising a small private institution, a medium-sized public comprehensive university, and an already highly-regarded research institution. The first two represented members of TEAC’s “natural” constituency, as neither would normally seek NCATE recognition, nor would they likely be successful had they done so. Their experiences provide a reasonable test of TEAC’s approach at the type of institution that CIC intended. The University of Virginia, in contrast, already possessed NCATE accreditation and was thus in a position to compare directly the relative merits of the two processes. Though short of the originally planned six FIPSE-project pilot cases, the experiences of these three institutions thus provided an excellent test of TEAC’s methods in the field.

Experiences with implementation are presented in three separate sub-sections below. A first examines the overall reported benefits of the process as seen by pilot participants. The second looks at the implementation of each major component of the review process in greater detail including preparation of the Inquiry Brief, the Self-Audit of Internal Quality Control Processes conducted by each program, the evidence advanced in support of claims about student academic achievement, and the conduct of the audit itself. The third sub-section draws some conclusions based on this experience in the form of implications and lessons learned.

Reported Benefits of the Process. Though challenging implementation questions arose in all three cases, pilot institutions reported that substantial overall benefits resulted from participating in the TEAC process. Among them were the following:

- **Opportunity to Highlight Distinctiveness.** Rather than emphasizing a uniformly-applied set of resource-based standards that tended to “homogenize” institutions, pilot participants welcomed the opportunity that the TEAC process accorded to them to emphasize their own areas of distinction and value perspectives. In fact the approach emphasizing rigid standards applied in detail that is typical of the traditional approach to program accreditation was seen as having precisely the opposite effect from the one intended. As one pilot representative with experience with NCATE (UVA) put it, “Too much prescription means that we establish malleable internal standards because of the need to be everything for everybody...as a result it leads us directly away from our own identity.” TEAC, in contrast, allowed participants to “tell their own stories” as part of the review.
- **“Value-Added.”** Participants in the TEAC process reported that this approach allowed them to use accreditation to address some of their own questions about themselves. As a result, rather than being seen merely as a necessary burden, accreditation actively “added value” to participating programs by enabling them to identify (and subsequently address) their own areas of strength and weakness. This ability to conduct a useful self-investigation was, in the one case where the comparison with NCATE could be made, commented upon especially favorably. According to the University of Virginia representative, “NCATE was a waste of time related to the outcome/benefit we received.”
- **Impetus Provided for Learning Assessment.** Pilot participants also noted the benefits of centering the TEAC process on student learning and the evidence for it. As the Fort Lewis representative put it [TEAC’s emphasis] “enabled us to get assessment front and center...[it] made faculty aware of the importance of assessment.” Such benefits are not unique to TEAC, of course, as regional accreditors have noted the same phenomenon as they shift the emphasis of their review processes toward assessment. But TEAC’s clear focus on student learning and the evidence for it clearly helped to reinforce the kind of attention to assessment that many Deans or Department Chairs know that they must eventually instill in their faculties.
- **Focus on Own Quality-Control Processes.** Though all three pilot participants had some difficulty executing the required “self-audit” of their own internal quality assurance processes (see below), all felt that this aspect of the TEAC process was especially valuable in uncovering areas for improvement. As the representative from the University of Virginia put it, this aspect of the review was “the most revealing part of the whole process,” and provided the School with real surprises regarding the kinds of records that were and were not being maintained. “Mapping” internal quality control processes in order to carry out the audit also enabled this institution to really understand these processes for the first time, and to see how “one weak link can bring the whole system down.” As a result, random checks of the record-keeping system will continue to occur.
- **Enhancement of Data and Information Capacity.** Participants finally noted that the TEAC review process made them far more conscious of the need for solid and accessible

data and information resources than they had previously been. The need to assemble data to support claims was sometimes challenging. But once accomplished, the resulting gains in information capacity could be used subsequently for many other things. As the representative from Fort Lewis noted, the TEAC review process helped stimulate the creation of the teacher education program's own "internal institutional research process." The stimulus provided by TEAC also forced greater attention to the way data and information resources were structured. For example, participants at the University of Virginia noted that they were moving away from "*ad hoc* reporting" in which required information was generated from scratch to meet each individual need or request, toward a "more general purpose database that can allow us to look at things in combination and respond to requests more flexibly." The TEAC review process also forced UVA to more effectively separate "relevant from irrelevant data," and to focus the faculty's attention on only gathering and archiving data that were "really useful for improvement."

These benefits were noted repeatedly in reports and de-briefing interviews, and constitute a substantial experiential validation of the TEAC process. Despite their considerable (and, in some cases, greater than expected) efforts to implement a new process, pilot participants found their investment in the process to be more than worthwhile.

Implementing Each Component of the Process. Although the overall reactions of pilot participants to the TEAC process were overwhelmingly favorable, implementing it involved many challenges. Foremost among them was the need to master a whole new set of terms and procedures which forced them to "unlearn" many habits acquired in previous accreditation experiences. Faculty at Texas Lutheran, for instance, had just completed an institutional re-affirmation review for the Southern Association of Colleges and Schools (SACS), and found it hard to "get into the language" of TEAC. In addition, both Texas Lutheran and Fort Lewis initially found it difficult to organize for review because the audit process rendered inadequate the usual mode of simply creating separate committees to prepare self-study chapters for each standard. As one participant put it, engaging TEAC meant "adopting a whole new philosophy" and this was not always easy.

Challenges of this kind were manifest in piloting every component of the review process at the three institutions. While in all cases they were overcome, the issues that they raised for TEAC, and the modifications that they suggested, were extremely important to the further development of the process (see Lessons Learned, below).

- **The Inquiry Brief.** Following the strong metaphor of educational research, TEAC designed the Inquiry Brief as a "scholarly exercise" from the outset. In both format and content, it was originally intended to resemble as closely as possible a publishable piece of research in the field of education. The table of contents for the document, for instance, requires text to be prepared around the traditional headings of a scholarly research article including an "introduction" modeled on a literature review, a methods section, a section on results, and a section on discussion and plan section modeled on the "discussion" section of a research monograph. Similarly, the methods used to provide evidence of student academic learning were expected to meet the empirical research standards of the discipline, demanding both a review of the instrumentation used and an analysis of validity and reliability for all methods used. The presumption underlying this approach was that faculty in Schools of Education

would be familiar and comfortable with the research canons of the discipline, and would consequently both welcome this approach over the decidedly “administrative” character of a traditional accreditation self-study and be able to do it.

Contrary to this initial expectation, all three pilot campuses had considerable difficulty with the Inquiry Brief as originally designed. All generated multiple drafts that were submitted and critiqued by TEAC (in one case, four times) before producing an effective presentation. Typical of a first attempt to work through a brand-new exercise, moreover, the resulting documents varied widely in size and scope (from a minimum of 27 single-spaced pages with appendices to a maximum of 78 pages). Despite a common table of contents, moreover, each of the three had a decidedly different “look and feel.” The Fort Lewis narrative, for instance, adhered strictly to the required format and looked the most like a “research brief” with separate sections addressing each hypothesis and reasonably formal language. The University of Virginia narrative, in contrast, looked like a more evidential version of a traditional self-study typified by fluent discourse, embedded claims and evidence, and complex sections of text that treated multiple topics simultaneously.

The prescribed Inquiry Brief format begins with a literature review, which was originally intended to provide faculty with a place to provide a grounded rationale for their particular approach to curriculum and pedagogy. In their initial attempts to produce this section, faculty often had difficulty locating primary research to cite—basing their discussion of program design elements largely on general works in educational philosophy. And even when they did locate “appropriate” material to cite, it proved a challenge to provide a straightforward and understandable description of their program’s goals, structure, and pedagogy within the constraints of the “research article” format. Indeed, the most successful case—the University of Virginia—exhibited the most traditional-looking approach to this task. Institutions also had trouble with a format that seemed to demand that they discuss all methods first, and only then examine results and implications. Because many methods of quite different types and precision were used, a more flexible form of presentation in which the methodological merits of each strand of evidence could be discussed *together* with the results in order to help make sense of them would have been more effective. And indeed, two of the three pilots actually ended up using such a narrative approach. TEAC’s Guidelines for the preparation of the Inquiry Brief do permit the program to use any format that would allow the program to make its case effectively. But the Guidelines display only the research report format and the writing workshops TEAC runs for Inquiry Brief authors emphasize the research report format.

The strong analogy with a publishable piece of empirical research in education also appeared to demand an excessively narrow approach to presenting and verifying claims. In two of the three pilots, for example, a classic “hypothesis-testing” approach was taken in which data were analyzed largely to verify whether or not a particular claimed level of student performance was reached. Such an approach—while appropriate to build new knowledge—did not induce faculty to disaggregate results in order to explore patterns of strong or weak performance, or to get behind straightforward claims of “target achieved” in order to discover areas where improvements might be made. Many of the marginal comments appended to early drafts of Inquiry Briefs by TEAC staff reflect this difficulty: while reinforcing the “empirical research” model implied by hypothesis testing, staff also appeared frustrated by

the fact that this approach did not provide a good way for faculty to demonstrate genuine depth of programmatic understanding.

A final issue that made preparation of the Inquiry Brief difficult for pilot institutions was the matter of validity and reliability. In many cases, the demand that all evidence-gathering methods be validated in accordance with the canons of educational research simply could not be met—either because the methods themselves were “soft” or because the requisite data was proprietary or otherwise unavailable. Lack of adequate validation, for example, was particularly an issue for affective claims—for instance, being a “caring and compassionate teacher.” This condition is, of course, present throughout current assessment practice in higher education: some of the most interesting and revealing methods simply cannot at this point be strictly validated according to research-based standards. As a result, investigators who stick to classic canons are faced with the prospect of throwing out the very things from which they are likely to learn the most. That said, some revealing information *did* emerge from attempts to validate previously unexamined measures. Among the most striking were consistent findings across pilot institutions suggesting that test scores and teacher-performance evaluations were essentially uncorrelated. These findings occasioned much useful discussion, and certainly would not have arisen at all if topics of validity and reliability had not been on the table.

- **Self-Audit of Internal Quality Control Mechanisms.** The self-audit concept is a centerpiece of TEAC’s approach, derived as noted from both the “academic audit” and the case for accreditation’s role in validating internal quality control processes advanced by Graham, Lyman, and Trow (1995). To discharge this task, institutions were given a sample “map” of internal quality assurance processes and were asked to follow a limited number of “audit trails” (selected so as to be representative) in order to verify that the established procedures shown in their map were actually in place. Results of the self-audit were to be presented by the institution as an integral part of the Inquiry Brief.

Many of the same implementation difficulties were experienced by early pilot participants in carrying out this exercise as in preparing the Inquiry Brief itself. First, the self-audit was an unfamiliar exercise and required systematic inspection of records at a level of detail that program faculty had not previously experienced. Second, as in the validity/reliability question, initial attempts to operationalize the audit method appeared a bit methodologically rigid—especially with regard to the methods TEAC insisted upon to draw a representative sample. Despite these relatively minor difficulties, all three campuses conducted a successful self-audit, the results of which were generally quite revealing for improvement. Indeed, for some, the self-audit proved the most valuable component of the review for self-improvement.

But the three institutions varied considerably with respect to both the scope and level of detail involved in their self-audits. For Fort Lewis, the audit ultimately became a pilot study only, sampling the records of twelve students in three courses to determine if syllabi, instructor records, and student grade records were present and in order. For Texas Lutheran, the kinds of materials sampled in a similar design (three students only) was more extensive—adding results of student evaluation of instruction, a faculty profile, teaching evaluations, and actual examples of student work. The University of Virginia self-audit also involved only three randomly-selected student records, but appeared considerably more thorough with

regard to what was examined. This self-audit not only traced each student's entire enrollment history (including the admissions process), but also looked at faculty hiring and promotion records as well as selected examples of student work. The UVA self-audit also uncovered some important differences in the interpretation of school policies, mostly involving standards and the supervision of field placements. None of the three campuses however fully examined the heart of the matter: the degree to which established faculty standards of academic achievement (grades or rubric-based scores) could be independently validated in the light of the student work presented.

A second difficulty encountered by programs as they undertook self-audits was the extent to which faculty in teacher education programs actually "own" the quality control mechanisms that they are auditing. At the University of Virginia, for instance, many of the record-keeping and evaluation functions being audited were university or School-wide so that even if problems were detected, the teacher education faculty could not fix them alone. At a liberal arts institution like Texas Lutheran, moreover, it was even more difficult to establish boundaries between faculty in teacher education and faculty in other departments with respect to "ownership" of different aspects of the program; this issue never did get resolved and, as a result, preparation for and participation in the TEAC review was confined to only the four dedicated teacher education faculty.

All told, however, experience with the pilot revealed the self-audit process to be one of the most successful components of the TEAC design. Institutions executed it responsibly and as rigorously as they were able, it was heavily evidential (though in some cases not very deep), and faculties learned things about their own operations that they would otherwise not have known.

- **Assessment Evidence and Use.** Although it has many other innovative features, the centerpiece of TEAC's approach to accreditation lies in its insistence that each program under review provide solid and direct evidence of student learning consistent with the program's goals. Teacher education programs should be at an advantage in providing such evidence, as most undergo standardized testing for purposes of licensure or certification (with many tested on entry to the program using standardized instruments as well), and all students are placed in field situations in real classrooms where their performance is directly evaluated by program faculty or practicing professionals. This level of "embedded" program assessment is matched by only a few other fields (primarily in the health professions). TEAC therefore had a right to expect candidate programs to produce solid direct evidence of student learning and all three pilot institutions were able to do so.

Not surprisingly, the bulk of the evidence presented consisted of standardized test scores, course grades, and student teaching evaluations. All three programs used these sources prominently to support the majority of their claims. As noted already, a striking finding when these two kinds of evidence were presented together was the essential lack of correlation between them—a fact that should have occasioned (and in some cases did) considerable discussion about what this meant. Other sources used included student portfolios, faculty-made rating rubrics, and alumni surveys. As stressed above, though, it proved difficult to provide publication-standard reliability and validity results for these kinds of measures, so the information drawn from them was probably not emphasized as much as it

could have been. Programs also varied with respect to the amounts and kinds of additional information used. Texas Lutheran took care to ensure that each claim was backed by evidence from at least two different sources, for example. And in addition to “harder” forms of evidence, the University of Virginia’s Inquiry Brief was filled with embedded quotations from currently-enrolled and former students that provided a different look at success. Although no information was provided about how these quotations were selected or how representative they were, the audit later established that their source was comments made during phone solicitations for gifts.

Less clear from the Inquiry Brief was in many cases what the programs had *learned* from what they had found. Again, this may in part have been due to the initial constraints of language—and, in particular, to the whole construction of the Inquiry Brief as a “research study” rather than a highly evidential form of self-analysis. In part, it may have been due to TEAC’s simply not putting as much emphasis on the role of self-analysis and improvement embodied in its Quality Principle III as it placed on the more novel and difficult topic of evidence of student learning when communicating with pilot institutions. But for whatever reason, the three programs differed markedly with respect to presented evidence about institutional learning. Fort Lewis largely let the “evidence speak for itself” without much interpretation provided. Texas Lutheran incorporated a separate sub-section on what the program was doing to address the deficiencies revealed by the evidence under each of its nine claims, and these narratives exhibited a good deal of self-reflection. But the separate discussion of QPIII that it provided was comparatively thin. In contrast to both, the University of Virginia’s discussion of implications and use of the evidence was sophisticated and seamless. On the one hand, strong evidence of institutional learning and programmatic reflection were apparent throughout the UVA narrative, which included many lessons learned and implications for action. On the other, because the UVA narrative was not organized around clearly-identifiable “claims” the way the other two were, it was sometimes difficult to *find* the lessons learned at first. These contrasting styles reveal a clear trade-off inherent in TEAC’s approach: the advantage of a “research/audit” model is that claims are clear-cut and verifiable, while the advantage of a more reflective, open-ended discussion is that action implications and depth of programmatic self-understanding is more apparent. Both are clearly important, but it seemed hard for faculty at the three pilot institutions to embody both at the same time.

In spite of these variations and challenges, the amount and quality of evidence about student learning presented, reviewed, and discussed in the course of all three pilots easily exceeded that present in virtually any other extant accreditation process. Given mainstream accreditation’s ongoing inability to make real progress on assessment, TEAC has without question succeeded in creating a review process that puts—and keeps—evidence of student learning at the center.

- **The Audit Visit.** Many aspects of the TEAC review process are distinctive, but the conduct of the institutional site visit probably differs most fundamentally from current accreditation practice. Trained auditors are contracted directly by TEAC (in contrast to essentially voluntary peer reviewers) and engage in a carefully-scripted set of site-visit activities. This renders the TEAC approach very different from what usually occurs in such encounters. The fact that the TEAC approach to site visits is “claims oriented” also makes the process easy to

structure and manage: most site-based activities can be carefully targeted to address specific claims made by the program.

Experience with the three pilot campuses demonstrated both the feasibility and efficacy of the audit visit as originally designed. Each institution was visited by two TEAC auditors (accompanied in one case by a number of observers) and the schedule proved sufficient in time and structure to fully accomplish the goals of the process. Audit visits begin with a period in which the auditors essentially “play back” to the faculty their own understanding of the program’s “story”—that is, its structure, curriculum, student body, pedagogy, etc. as outlined in the first section of the Inquiry Brief. This session proved important in all three cases in building confidence and in demonstrating that the auditors had carefully read and understood the Inquiry Brief. As one auditor put it, “this session gives them the chance to criticize us.” In at least one case, the initial session also provided the program with an opportunity to bring the auditors up to date on important changes that had occurred in the program’s environment since preparing the Inquiry Brief. A second introductory portion of each visit centers on identifying and clarifying claims through the “precision probe” process. This again worked effectively, as programs had the chance to supply additional information, change the nature of what was being claimed, or (in some cases) to withdraw certain claims or statements entirely. Auditors recorded results in the form of worksheets for later inclusion in the audit report. This highly-structured approach makes the ground rules for the audit extremely clear, in contrast to the typical “let’s go see what we can find out” or “check off the standards” approaches taken by traditional accreditation site visitors.

The bulk of the audit visit in each case was subsequently spent in verifying evidence associated with the program’s claims. Here the auditors’ experiences differed somewhat at the three institutions. In one case, supporting information proved difficult to locate though the auditors did find enough to examine directly; in this case the self-audit of quality assurance processes was also not fully completed, and the audit team had to essentially undertake this task as well. In the words of one auditor this campus “was not really ready for audit.” The incident highlights the dependence of the audit process on careful preparation of supporting evidence by the institution and its physical presence during the site visit. In the other two site visits, such organization was manifest and auditors were able to quickly locate and check what they needed to examine (though in some cases, this evidence had *not* been previously assembled). These examinations were unusually thorough, even to the point of re-checking statistical calculations, and auditors found that the protocols developed by TEAC to check evidence were helpful in disciplining the process. Throughout the audit visit, auditors were careful to maintain their assigned roles, not offering opinions or feedback directly to program faculty and not socializing with them after hours. Indeed, the overnight period built into each site visit provided the auditors with an important opportunity to compare findings and discuss any issues that may have arisen.

The resulting audit reports were, not surprisingly, far more uniform than the three Inquiry Briefs. All were prepared according to a standard format addressing each component of the visit (Summary, Clarifications, Precision Probes, and Audit of Submitted Evidence) and all concluded with a standard finding—essentially that the “data can be trusted.” The only major difference among these reports was that there were a good deal more potential precision probes identified for the University of Virginia than for the other two cases—

possibly a result of the more open writing style that characterized the Inquiry Brief prepared by this institution.

Implications and Lessons Learned. The pilot process without question demonstrated the feasibility of TEAC’s distinctive approach to review. All three institutions were able to engage the process, despite its novelty, and audit visits to all three went quite successfully. Experience with the pilots, however, suggest a number of conclusions that TEAC should bear in mind as it continues to develop its approach to review. And inspection of more recently-prepared audit and Inquiry Brief training materials suggests that many of these lessons have already been used to modify TEAC processes and approaches. The most prominent of these lessons are as follows:

- **The Metaphor of Scholarship.** Evidence from the pilots suggests that TEAC may have been expecting too much in modeling the Inquiry Brief on the standard of educational research. Faculty at two of the three participating institutions (both of which represent TEAC’s primary intended constituency) had a good deal of difficulty with this approach, largely due to sheer unfamiliarity with the “research” paradigm. More importantly, the research-article format for presenting data often appeared to get in the way of reflective commentary on the *meaning* of presented data and what might be done in response. A better guiding metaphor, consistent with TEAC’s intent, might be “reflective scholarship.” Invoking broader notion of scholarship—consistent, for example, with Boyer’s *Scholarship Reconsidered* (1990)—might allow TEAC to continue to signal the essentially academic foundation of the inquiry and to reinforce its serious, evidential, and reflective qualities without confining participants to an overly-restrictive method of presenting data and drawing conclusions. Such an approach would encourage institutions to prepare the first sections of their Inquiry Briefs in a manner broadly grounded in specific traditions of educational philosophy without necessarily requiring them to engage in a formal literature review of relevant empirical studies. (TEAC has already taken this particular step.) It would also allow institutions to present and discuss evidence of student learning from the framework of exploratory data analysis—designed to uncover patterns of strength and weakness—rather than steering them toward a “hypothesis-testing” mode that emphasizes the use of evidence largely to prove that a given standard has been achieved.
- **Models of Inquiry Briefs.** The three documents prepared by pilot programs differed markedly in scope, tone, and format. Partly, of course, such variation was natural as no one had ever written a TEAC Inquiry Brief before. But it was also clear that participating programs had little to go on in deciding what a “good” document was supposed to look like. As experience grows, TEAC will undoubtedly be able to show institutions several different examples of successful submissions together with commentary about what makes them good—much like a writing program is able to show multiple contrasting examples of effective student work. Indeed, some of this modeling is already happening, as evidenced by the most recent round of TEAC Inquiry Brief workshop sessions. As it builds a library of such examples, though, TEAC should be explicit about the particular qualities (not just formats) that ought to characterize such presentations. Based on the pilot experience, the following seem reasonable candidates for such criteria.
 - Clearly-Identified Claims. While institutions should be accorded a good deal of flexibility in how they prepare the Inquiry Brief, they should be strongly encouraged to

make their claims about student learning very clear in the text. While highly readable and engaging, for example, it was sometimes difficult to find such explicit claims in the University of Virginia's submission.

- Complete Documentation of Evidence. Presented evidence should be accompanied by some discussion of *why* a given measure was chosen, as well as a discussion of validity and reliability issues wherever possible. But this might also be accomplished flexibly. Sometimes, for instance, it will be more effective to adopt the approach taken by Texas Lutheran in which each claim is backed by a body of evidence which is discussed *together*, rather than placing all methodological discussion of all measures in a single section. Also, institutions should be encouraged to employ less-than-perfect measures that appear more appropriate to their goals, so long as they a) are demonstrably aware of the caveats involved in doing so, and b) have several strong “core” measures in place in their design to triangulate their conclusions.
- Reflection on the Meaning of Evidence. It should be clear from the narrative that the program's faculty has *thought about* the evidence and its implications for improvement. Again, this can be done appropriately in many ways—including a separate section on improvement—but simply reporting the evidence should not be sufficient.

Many similar criteria can be imagined, but the primary point is that there should *be* some, presented together with appropriate and varied illustrations drawn from actual Inquiry Briefs. But the examples provided should also demonstrate the value of succinctness. One of the main virtues of the TEAC approach is that its focus on evidence should force “self-study” narratives to be claims-oriented and to the point.

- **The Role of QPIII**. The emphasis on student learning and the quality of the evidence for student learning in TEAC's first two “quality principles” is admirable and should remain the centerpiece of its distinctive approach. But more attention perhaps needs to be given to communicating and illustrating the *third* quality principle: Institutional Learning. Clearly this was the strength demonstrated by the University of Virginia. Indeed, the ultimate accreditation panel decisions regarding the three pilot institutions (unanimous approval for UVA, a split decision resulting in provisional accreditation for TLU, and approval of “Pre-Accreditation Status” for Fort Lewis) completely paralleled the degree of “reflectiveness” demonstrated in their final narratives. This suggests that more attention should be given to this topic in the Inquiry Brief workshops and that these qualities be prominently illustrated in the examples provided—without, of course, diminishing the role of evidence of student learning itself. The Evaluation Rubric used to assess Inquiry Briefs (Table 2 on p.21 of the *Prospectus*) might also be revised somewhat to highlight the role of QPIII. The North Central Association's “Levels of Implementation” rubric might be used as a source of ideas for this revision, as it is focused primarily on how evidence is reflected upon and used in the decision process (North Central Association 2001).
- **Segregation of Duties**. One of the more notable aspects of the TEAC review process is the manner in which the role of the auditor is confined to “inspecting the evidence” without making any substantive judgements about the actual quality of the program under review. While admirable in concept, pilot experience suggests that such a strict segregation of duties

between auditors and accreditation panel is difficult to maintain in practice. It also has the unfortunate side effect of preventing those members of the TEAC review team who have the most actual experience with the program under review from sharing any conclusions that they may have come to through direct inspection and interaction (although the lead auditor is present as non-voting member of the accreditation panel). It is important to note in this regard that academic auditors in the U.K., New Zealand, and Hong Kong *do* make substantive judgements about institutions in the course of their reviews; that is, their role is more than simply to “verify the evidence.” And even financial auditors go beyond simply “inspecting the books.” Through the mechanism of the management letter, they are also charged with making substantive suggestions about financial accounting (and frequently other) policies and procedures. While not eliminating the role of the accreditation panel—which constitutes an excellent third-party check on the summative aspect of review that is not present in traditional accreditation decision-making—some thought might be given to allowing TEAC auditors a broader substantive voice in the process. This could be accomplished by asking them either to prepare a private evaluation for the Panel or to provide oral testimony about substantive issues of quality, in addition to their conclusions about the quality of the evidence. Recent revisions in TEAC’s descriptions of the auditor’s role have in fact gone in this direction.

The process of testing TEAC’s processes through pilot implementation was successful in two important ways. First, the practicability of the process was conclusively shown. Three very different institutions were able to complete reviews in a satisfactory manner. While the task of preparing Inquiry Briefs and assembling adequate evidence was not always easy, results suggest that most institutions that might potentially seek TEAC accreditation can complete the review and will learn something from it. Second, many lessons about how to improve the process were apparent in the pilots, and these have already been used by TEAC to improve its workshops on auditor training and the preparation of Inquiry Briefs. Most importantly, TEAC staff have learned to be a bit more flexible in implementing the process while maintaining its central rigor. This will be important as TEAC seeks to recruit a wider range of institutions, and decisively demonstrates its own commitment to becoming a Learning Organization in its own operations.

References

- Dill, D.W., Massy, W.F., Williams, P.R., and Cook, C. M. (1996). Accreditation and Academic Quality Assurance: Can We Get There from Here? *Change Magazine*, 28, 5, September/October, 1996.
- Graham, P.A., Lyman, R.W., and Trow, M. (1995). *Accountability of Colleges and Universities: An Essay*. New York: Columbia University.
- North Central Association (2001). *Levels of Implementation: An Addendum to the Handbook of Accreditation, Second Edition*. Chicago, IL: North Central Association, Commission on Higher Education.
- Boyer, E. L. (1990). *Scholarship Reconsidered: Priorities of the Professoriate*. Princeton, NJ: Carnegie Foundation for the Advancement of Teaching.