ASSESSING EDUCATION CANDIDATE PERFORMANCE

A Look at Changing Practices

Emerson J. Elliott, Project Director

NCATE
National Council for Accreditation of Teacher Education

May 13, 2003
Contents

Foreword iv

Preface vi

Assessment Examples Project Committee Members viii

Introduction to the Assessment Examples Project
  Background on Assessments Submitted for Committee Review
  Committee Perspectives on the Current State-of-Practice in Assessment
  Organization of Assessing Education Candidate Performance

Committee Review Part A: Criteria Applied on Qualities and Uses of Assessments
  Criterion 1—Assessments are appropriate for the standards they are meant to address; the instruments assess meaningful elements of content, cognitive demands, and skill requirements at challenging levels for candidates.
  Criterion 2—Assessments are accompanied by explicit statements of proficiencies candidates are expected to demonstrate in their responses.
  Criterion 3—Assessments are constructed so that different levels of candidate proficiency are clearly distinguished; scoring processes are consistent and reviewers are trained so that judgments can be discriminating.
  Criterion 4—Assessments are used to reach meaningful decisions, including ones that identify candidate needs; judge candidate progression; and evaluate courses, programs, or units. Assessments that are combined to reach a decision are accompanied by explicit and public statements describing the manner of weighting and summarizing results.
  Criterion 5—Assessment systems include some “authentic” forms of assessment in which candidates are asked to perform tasks similar to those they will face in their initial employment as education professionals. These contain assessments of learning by the candidate’s P-12 students.
  Criterion 6—Assessments are systematically evaluated to ensure fairness, accuracy, consistency, and avoidance of bias.

Committee Review Part B: Committee Consideration of Cost, Efficiency, and Innovative Elements
  Consideration 1—Developmental costs and operating costs
  Consideration 2—Effectiveness and efficiency of assessments in gathering important and significant information
  Consideration 3—Innovative features illustrated by assessments, including use
of information and computer technology

Example Locator by Type and Use of Assessment

Excerpts from Examples
Example A—A Performance Assessment of Teaching Effectiveness and Student Learning
Example B—Unit Assessment System
Example C—A Course Assessment for Foundations of Teaching and Learning
Example D—Performance Based Assessment in Teacher Preparation Using a Web-Based System
Example E—Professional Education Program Admission I Requirements
Example F—Impacting Student Learning
Example G—Teacher Work Sample: The Renaissance Partnership for Improving Teacher Quality
Example H—Model of Appropriate Practice (MAP) for Teacher Candidates
Example I—Action Research Project Using NCTE Standards
Example J—Learning Log on the Uses of Assessment in Instruction
Example K—Math Prompt and Scoring Guide
Example L—Admissions Portfolio requiring Common Essays
Example M—On-Line Performance Requirements for Technology
Example N—Assessing the Teaching Characteristics of Preservice Teachers in an Internship Program
Example O—INTASC Principles and the BSU Student Teaching Rubric
Example P—Performance Assessment with Teacher Standards (PATS)
Example Q—Information and Assessment Systems
Example R—English Proficiency Exam

Performance Assessment in Higher Education: Selected Resources
NCATE formally launched its new performance-based accreditation system in September 2001. That system shifted the emphasis in accreditation from what the institution offers its candidates to what the candidates receive. Institutions must now provide evidence that candidates have the knowledge and skills necessary to teach, the dispositions to teach, and the ability to put these attributes into action so that all students will learn.

The assessment of candidates by institutions is hardly a new phenomenon. Traditionally, every educational institution considers evidence as it decides which candidates to admit, monitors their progress through courses and experiences, and decides when to graduate them. Most institutional assessment practices have been designed to make decisions about students in that institution. Faculty grade students in their individual courses to determine whether, in their judgment, candidates have achieved course objectives. When candidates have accumulated credit through a sufficient number and array of courses, they are judged ready to graduate. This system works well within the context of a particular institution where the purpose of assessment is to make institution-specific decisions about candidates. Graduation from an institution conveys the information that candidates have mastered the goals set for them by the institution. This information suffices for some purposes but not others. Thus candidates may be eligible for jobs that require "college graduation." However, "college graduation" by itself does not typically provide sufficient information for graduate school admissions offices that require GRE or ACT scores to complement the idiosyncratic grading practices of individual institutions.

In the case of teacher preparation, faculty grade candidates not only on knowledge acquisition but also on the acquisition of skills, dispositions, and teaching ability. The assessment of teacher candidates must convey information beyond the information needs of the institution itself. Thus the grades of teachers in content and on other attributes have had to be communicable to state credentialling authorities and school districts. One of the roles of program approval and accreditation is to make information about individual candidates meaningful outside the context of a particular university. And, of course, the accreditation process itself now requires that such information be universally comprehensible.

NCATE's expectations for performance assessment (and NCATE defines performance to include all the referenced attributes) are motivated, in part, by this need and, in part, by the need for the teaching profession to have multiple measures about the variety of attributes associated with teaching. These expectations place new demands on the institutional assessment process and broaden the purpose for which assessment is to be conducted. NCATE expects faculty to collaborate in the design and implementation of the conceptual framework. This framework then guides curriculum, instruction, and assessment, all of which must now be comprehensible within the institution and communicable outside the institution.

NCATE understands that a cultural shift in faculty grading practices covering all the desired outcomes of the preparation process is now underway. Together, faculty members must become explicit about the goals of instruction and the means by which they will be measured.
Traditional grading, benchmarked to explicit goals, may be effective in communicating some learning goals publicly. The assessment of teaching performance will almost certainly also require development of new approaches to assessing the acquisition of skills, dispositions, and teaching ability.

Change is underway; the purpose of this project is to capture and describe some state-of-the-art practices that may be useful examples for institutions starting down the path to performance-based accreditation.

Emerson Elliott, as director of this project, led The Committee to produce outstanding work. Our deep appreciation goes to The Ford Foundation, The Atlantic Philanthropies, The William and Flora Hewlett Foundation, and The Carnegie Corporation of New York, for supporting this effort and making this work possible.

Arthur E. Wise, President
Preface

The predominant theme in education policy discussions through the late 1990s and into the 21st century is “standards-based reform.” The idea is simple enough. It is built on long experience of state and national policymakers in fashioning laws, regulations, or processes to “improve” education. From this experience, they identified conditions they sought to change and directed public funds from them. Years later, policymakers decided that the success of policies in education should be judged by whether or not student learning is enhanced. From that shift in perspective came the idea of writing explicit standards stating what students should know and be able to do, then judging “outcomes” through assessments or other evidence of learning instead of judging “inputs” as measures of effectiveness.

There is a parallel experience for accreditation as well. Traditionally relying on indirect indicators of quality such as student faculty ratios, library collections, or the condition of facilities, accreditation evolved over the years to examination of courses and experiences that represented “inputs” to learning. But accreditors, similar to public policymakers, found this too indirect as evidence that candidates are prepared to enter their profession. So, policymakers in accreditation have also moved to emphasize standards that describe intended learning outcomes, and assessments that demonstrate achievement of those outcomes.

The project reported in the pages of Assessing Education Candidate Performance: A Look at Changing Practices is an answer to many requests for “examples” that we have received since performance-based accreditation was set as the NCATE policy in 2000. Faculty want to be responsive to NCATE’s insistence on assessment results as evidence that candidates have mastered unit and program standards prepared by NCATE and its affiliated program specialty associations. Moreover, they are trying to build on the best that we know about use of assessments for learning—such as “alignment” of assessments with standards, use of multiple measures, and evaluating a wide range of knowledge, skills, and dispositions. Faculty who are searching for examples expect that someone, somewhere, in professional educator preparation programs has developed assessments that can serve to demonstrate what is possible. What is needed is not an elegant new design from external testing experts, or the latest research efforts in test development. Instead, what are needed are assessments created by colleagues in other institutions, and actually used in courses and experiences of professional educator preparation programs. That way we can know that the examples are not out of reach, but, are practical for adaptation and as sources for new ideas.

As with any effort of this kind, the project could only be undertaken with the support and contributions of many individuals and organizations. First, we would like to thank the foundations whose support has made this work—an integral part of NCATE performance-based accreditation—possible. Our funders include the Carnegie Corporation of New York, The Ford Foundation, The Atlantic Philanthropies and The William and Flora Hewlett Foundation. We are greatly indebted to our Assessment Examples Project Committee, a group of fifteen members who created review criteria, prepared a letter of invitation, reviewed submissions, added to the composite pool of examples from their own experience, identified examples to illustrate criteria
set for the project, wrote critical sections of the paper, and prepared its bibliography. Special thanks are due to the faculty who responded to the Committee invitation by creating the packages of examples and explanatory material that have become the principal focus of these pages. The Committee and I were especially pleased to have, as our consultant, Jeff Kenney, President of Professional Development Partners, Inc, who brought to us his expertise with assessments in preparation of professionals for fields other than education.

Emerson Elliott, Project Director
Assessment Examples Project Committee Members

Margie Crutchfield
Director for Program Accreditation
Council for Exceptional Children
Arlington, Virginia

Mary Diez
Graduate Dean, Alverno College
Milwaukee, Wisconsin

Charles Duke
Dean, Reich College of Education
Appalachian State University
Boone, North Carolina

Eugenio J. Gonzalez
International Study Center, Manresa House
Lynch School of Education
Boston College
Chesnut Hill, Massachusetts

Penelope Haile
Associate Dean, School of Education and
Allied Human Services
Hofstra University
Hempstead, New York

Marilyn Hala
Professional Programs Manager
National Council of Teachers of Mathematics
Reston, Virginia

Gene Hall
Dean, College of Education
University of Nevada, Las Vegas
Las Vegas, Nevada

Marilou Hyson
Associate Executive Director for Professional
Development
National Association for the Education of
Young Children
Washington, DC

Mieko Kamii
Associate Professor
Wheelock College
Boston, Massachusetts

Richard Mainzer
Assistant Executive Director for Professional
Standards & Practice
Council for Exceptional Children
Arlington, Virginia

Marion Gillis-Olion
Associate Vice Chancellor for Academic
Affairs
Fayetteville State University
Fayetteville, North Carolina

Barbara Plake
Director, Buros Institute of Mental
Measurements
University of Nebraska-Lincoln
Lincoln, Nebraska

Joe Prus
Professor of Psychology and Director, Office
of Assessment
Department of Psychology
Winthrop University
Rock Hill, South Carolina

Cathy Roller
Director of Research and Policy
International Reading Association
Newark, Delaware

Kate Steffens
Professor, College of Education
Saint Cloud State University
Saint Cloud, Minnesota

Project Consultant:
Jeff Kenney
President, Professional Development Partners, Inc.
Washington, DC

NCATE Project Director:
Emerson J. Elliott
Introduction to the Assessment Examples Project

The NCATE Assessment Examples Project was created to identify examples of assessments used in preparation of teachers and other education professionals. These examples are to provide ideas for faculty and education unit/program administrators as they develop assessments for standards-based preparation of teachers and other educators, and also to serve as common reference points for NCATE Board of Examiner members and specialty organization program reviewers.

NCATE standards strongly emphasize what candidates should know and be able to do to be effective teachers or other education professionals. Standards for candidates are conveyed through NCATE unit standard 1 (on candidate knowledge, skills, and dispositions) and through 19 sets of specialty area standards that guide program reviews. Under NCATE’s “performance-based accreditation,” the evidence that standards are met takes the form of aggregations of assessment results demonstrating that candidates know their subject and can teach effectively so that all students learn.

The assessment examples project and this paper are in response to frequent questions NCATE receives about performance-based accreditation—questions such as: What are appropriate assessments to use in preparation of educators? How might an assessment of subject content knowledge differ from assessment of classroom teaching skills? How can learning among a candidate’s P-12 students be responsibly demonstrated? What information can be gathered during an admissions process to help identify candidates with potential to become effective teachers? And, especially, what examples do you have of assessments for any of these questions?

The project was conducted through a committee created by NCATE comprised of representatives from specialty organizations, higher education faculty and administrators, and testing experts. At its initial meeting in the winter of 2002, the committee devised criteria to define sound features of assessments used in preparation of professional educators. On behalf of the committee, NCATE solicited standards-based assessments from faculty in the spring, explicitly stating the committee criteria so that respondents would know how the committee would be reviewing their submissions. In the summer and early fall the committee met twice more to review each example, apply the criteria they had already established, and select examples to fit. The process led the committee to add observations about other attributes of assessments—the

---

1 See membership list on page viii.
cost of developing and maintaining assessments, the efficiency of assessments, and innovative features found in assessments.

The Committee did not attempt to rank assessments, as if in a horse race, proclaiming one or a few as the “best” examples for all purposes in preparation of professional educators. Instead, the Committee selected examples to illustrate values in assessment expressed through each criterion individually. Some examples were selected as illustrations for more than one criterion. The collection of examples illustrates different forms or types of assessments used to evaluate elements of unit or program standards describing candidate knowledge, skills, dispositions, or effects on student learning. The collection also illustrates levels of knowledge or performances that faculty expect from teacher candidates and other education professionals completing their preparation.

The results of the Assessment Examples Project are exhibited in this paper—with its detailing of the criteria and excerpts from examples—and in the full examples that have been placed on NCATE’s website, http://www.ncate.org/resources/m_resources.htm. The selected examples are highlighted in NCATE training and technical assistance activities. Readers are also encouraged to see further supporting material at the same website, titled “Perceptions of Performance-Based Program Review and Assessment Evidence,” excerpted from Program Standards for Elementary Teacher Preparation.

**Background on Assessments Submitted for Committee Review**

The committee received 29 examples from 16 different institutions, and committee members brought seven more examples (for a total of 36) from six institutions (for a total of 22) to complement this pool. These additional examples were ones committee members had encountered in their own experience (not necessarily in their own institutions) with an emphasis on measuring candidate subject content knowledge—an aspect of assessment that was not well represented in the original set of submissions. The submitting institutions included ones with small numbers and others with large numbers of graduates; both public and private institutions were represented, as were all principal regions of the contiguous states.

The largest proportion of assessments submitted for the project addressed aspects of candidate clinical preparation, but there were assessments used in admissions, assessments modeled on the INTASC principles, assessments of candidate subject content knowledge, as well as assessments to examine technological abilities, work characteristics and on-the-job experience. Among the examples were several assessment “systems” that combined assessments for educator preparation with computer or web-based data management capability. As a whole, the collection represents quite diverse forms of assessment, applied at different stages of a candidate’s preparation program.

The committee recognizes that the assessment and assessment system examples it received do not constitute a representational sample of all the assessments currently used in educator preparation. The number of submissions was small. The committee could not say that the array of institutions from which assessments came is statistically representative of all colleges and universities that prepare teachers and other education professionals. Most important for faculty
in NCATE institutions, however, observers who have closely followed the NCATE Board of Examiner visits see the examples as highly consistent with what they are finding in NCATE accreditation reviews. Committee members viewed the examples as reflective—even if not representative in the statistical sense—of the current state-of-practice in assessment for preparation of education professionals.

Committee Perspectives on the Current State-of-Practice in Assessment

Both strengths and limitations appeared as the committee examined the assessment examples. On the one hand, the committee found a widespread and quickly developing use of performance-based assessment, and a variety of assessment forms, including a number of promising and innovative examples. However, many examples provided only limited guidance for judging candidate performance; assessments of candidate subject content knowledge were infrequent; few assessments exploited the capabilities made possible through technology; and few assessments addressed the effects of candidates on P-12 student learning. Here are some details:

- **Rapid evolution**—The committee observed that the assessments they received were time-sensitive. Assessments used for preparation of educators are rapidly evolving. Assessments now in pilot or initial administration stages represent an early phase of development. While faculty in some institutions may be farther along than others, no faculty believe they have reached their goals. Most of the assessments submitted for Committee review were described as in draft, ready for piloting, administered once but with results to be evaluated, or ready for further refinement in academic year 2002-2003. So assessments reviewed by the Committee might be characterized as developing, more successful than a year or two ago, perhaps most successful “to date,” a “snapshot” of assessment in 2002—but certainly not complete or perfect or “for all time.”

- **Use of innovation and variety in approaches to candidate assessment**—Varied, innovative, and authentic forms of assessment were abundantly represented in the examples submitted. These examples included portfolios (with arrays of different sorts of material such as lesson plans, videotapes, assessment of P-12 student learning), reflections, surveys, vignettes, and many others. Several required candidates to write pieces that demonstrate their proficiencies in reflection, or their understanding of content, or their ability to apply what they have learned about instruction. Others created situations for candidates to demonstrate their ability to make oral presentations, to conduct “action research” in their teaching, or to exhibit proficiencies with use of the computer and the internet.

- **Need for clear expectations**—Even innovative forms of assessment often paid little attention to the need for clearly defined expectations (scoring guides or rubrics) for proficiencies that candidates should demonstrate on an assessment. The value of these innovative assessments in providing candidates with constructive feedback and evaluating candidates’ progress was, therefore, sometimes limited. In addition, the committee only occasionally found evidence that the faculty who judged candidate responses had been systematically trained to assure consistency in their evaluations
and reliability in their judgments over time. Yet, training the scorers is an essential element of good assessment systems (see criterion 3, below). Moreover, it should be common knowledge among faculty and candidates not only what the expectations are for any particular assessment, but how summative judgments are reached. From submissions the committee reviewed, it was often not clear how discrete assessment results from critical decision points—admissions, end-of-subject course work, admission to field experiences, conclusion of field experiences, conclusion of program, and post program—were weighted and accumulated for decisions about candidate progress, completion, or early success in education employment.

• **Assessing candidate content knowledge**—Only a few of the submitted examples offered measures of candidate content knowledge in relation to standards, such as subject matter tests, essays, projects, or other material that would demonstrate depth of understanding in content areas. Some examples used completion of the subject major, grade point average in the subject major courses, or a state licensure exam as evidence of subject matter proficiency without indicating how those measures reflected the knowledge and skills expressed in professional standards. It is often observed that educator preparation is an institution-wide responsibility and strong ties must be developed across the boundaries of schools of education and of the arts and sciences faculty. This is true in assessment just as it is in building subject matter competence and strong liberal arts background for candidates. No external licensure exam and no system of assessments can replace the synergy that comes from such collaboration.

• **Technology and candidate assessment**—Despite the pervasive presence of computers and the internet in professional education, the examples contained few illustrations of technology applications to assessment. A few examples did show what units might consider:
  a. One application of technology is to **build candidate portfolios in electronic form**. Using technology in this way, portfolios can show progress and proficiencies of candidates throughout their coursework and clinical experience, and perhaps into induction (e.g., archiving information on candidate experiences, creating artifacts from candidate written and oral work, producing videos of candidate teaching, and developing evidence of professional interactions).
  b. Another technology application is to **manage assessment data** by, for example, creating systems for storage, retrieval, and analysis, and by combining assessment information with other data from candidates, faculty, or administrative officers, and from courses as well as from unit or institutional operations and activities.
  c. A third technology application is to **a resource file for information on assessments** with different properties suitable for courses or for evaluating candidate progress at various points in a program, or during induction into the profession.
  d. Technology can be used as the **mode of delivery for assessments**. It offers especially attractive capability to administer assessments, to create
simulations for use in assessments, or to deliver “adaptive testing” in which candidates are presented with items of lesser or greater difficulty depending on the pattern of responses they have provided. The committee did find submitted examples of each of these uses, but while the portfolio and data management applications appeared in several examples, the resource file and direct use in creating and administering assessments were rare or incipient. On the whole, the committee concluded that technology applications to assessment were not nearly so frequent or robust as might be anticipated in an era that demands high technology literacy (a need that led to embedding technology in NCATE accreditation standards).

- **Assessment of candidates’ effects on P-12 student learning**—A few assessments incorporated candidates’ effects on P-12 student learning. Members of the committee viewed assessment in educator preparation as a powerful means to facilitate learning and as an opportunity to model for candidates how they should assess their own pupils in P-12 settings. Clinical experiences offer occasions for candidates to apply what they have learned about appropriate ways to assess learning by P-12 students. Since the critical goal of educator preparation is improved learning outcomes for students, committee members expect that this dimension of an assessment system should become routine as the transition to NCATE’s performance-based accreditation is completed.

Committee discussions also delved into how assessments are used—not only for monitoring candidate progress but for evaluation of unit and program effectiveness as well. Additional context for this discussion is found in the supporting explanation for NCATE unit standard 2, *Assessment System and Unit Evaluation*:

The unit and its programs are informed by an assessment system that examines the (1) alignment of instruction and curriculum with professional, state, and institutional standards; (2) efficacy of courses, field experiences, and programs, and (3) candidates attainment of content knowledge and demonstration of teaching that leads to student learning.

. . . Evaluations must be purposeful, evolving from the unit’s conceptual framework(s) and program goals. They must be comprehensive, including measures related to faculty, the curriculum, and instruction, as well as assessments of what candidates know and can do. The measures themselves must be sufficient and appropriate to inform the important aspects of faculty, curriculum, instructional, and candidate performance.

Fairness, consistency, accuracy, and avoidance of bias in the assessment system must be considered. In particular, attention must be paid to the potential impact of the assessments on a diverse pool of teacher candidates. In addition, the unit assessments and evaluations must consider how to provide and use information constructively from various sources—the unit, field experiences and clinical sites, general education and subject content preparation courses, faculty, candidates, graduates, and employers.

As committee members reviewed the various assessments submitted for the project, they were unable to answer such questions as:

- What is the role of an assessment in a content course and how does it align with other assessments in professional courses or the unit in general?
- How are assessment results aggregated to monitor candidate progress or make decisions about candidates?
• How are results combined with other data to evaluate program and unit effectiveness?

They did not find, and indeed had not requested in the invitation for assessment examples, information that could answer such questions. Of course, assessments in any program or unit must be aligned with standards and have appropriate scoring guides and decision making processes—as the pages of this paper emphasize. Yet the assessment designs and results must be part of an integrated system for evaluation of units and programs. Otherwise assessment information will remain fragmented and of limited use to the unit or program. The committee concluded that faculty charged with developing and maintaining assessments in any unit or program must attend to questions of both the “whole system,” and the individual “parts,” as detailed in unit standard 2.

And, finally, a committee reflection on the future. Members of the committee found that “the field” of assessment in educator preparation programs “is moving in the right direction.” The assessment examples suggest that programs and units are taking steps that will use assessments to monitor progress, identify needed improvements (for both candidates and programs), and demonstrate that candidates know and can do what is described in professional standards. The committee commends campus administrators and faculty for their efforts to date. Committee members believe, moreover, that the examples selected to highlight in this Project can assist faculty in finding new ideas, and then adapting what they find to their own needs.

The committee strongly advises against simply copying assessments found in these pages. Context matters and what is found here must be adapted to the setting and conceptual framework of each institution or program. However, the committee encourages faculty to be in touch with individuals listed as contact points for the examples. The assessment authors and developers who submitted these examples are, in most cases, continuing their work on piloting, examining, and refining the assessments. They have a wealth of instructive experience to share about the design and functioning of the assessments and relationships of the assessments with professional standards; they have critical information, as well, about the surrounding context and about uses made of the assessments by programs and units.

Organization of Assessing Education Candidate Performance

The Committee did not attempt to certify any assessment as representing a “gold standard.” Rather, members applied the criteria one by one to each assessment they reviewed, and through their discussions and consensus building, selected assessments that clearly illustrate each criterion. Some examples were selected two or more times because of their value in illustrating more than one criterion. The bulk of this resulting report is comprised of four sections:

• *Committee Review Part A*, contains descriptions of six criteria addressing qualities and uses of assessments; it lists specific assessment examples illustrating the criteria.
• *Committee Review Part B*, reports committee considerations on assessment costs, efficiency, and “innovative” features in assessments; it also lists examples.
• *Excerpts from Examples*, is comprised of brief excerpts from assessment examples illustrating the criteria and cost or innovative features referenced in Parts A and B.
The last part is a bibliography of selected resources on “Performance Assessment in Higher Education” that can serve as a general reference source; it also demonstrates that the Committee’s criteria in Part A are derived from a large body of research on, and applications of, assessment for use in instruction.

Committee Review Part A: Criteria Applied on Qualities and Uses of Assessments

The Assessment Examples Project Committee reviewed assessment examples against criteria representing generally accepted attributes of high quality assessments. For that reason the criteria are prominent in this report and are perceived by the committee as the most important part of the report. The criteria are a collective answer the question, what are good assessments for preparation of professional educators? The criteria, together with the selected examples, can serve as guides and sources of ideas for professional education units.

1. Assessments are appropriate for the standards they are meant to address; the instruments assess meaningful elements of content, cognitive demands, and skill requirements at challenging levels for candidates.

   • An assessment should be “appropriate” for, or aligned with, standards, and not be treated as a substitute for standards. Appropriateness implies that:
     a. the same or consistent categories of content appear in the assessment that are in the standards;
     b. the assessments are congruent with the complexity, cognitive demands, and skill requirements described in the standards; and that
     c. the level of effort required, or the difficulty or degree of challenge of the assessments, is consistent with standards and reasonable for candidates who are ready to teach or to take on other professional educator responsibilities.

   • The committee encourages assessments constructed so as to provide information about several elements of a standard, or portions of several standards, that are “meaningful”—that cover major or significant amounts of learning. The committee discourages assessments that attempt to measure atomized bits of learning—those forms of assessment limit the utility of results and make comprehensive understanding of candidate knowledge, skills, and dispositions unmanageable. In addition, sound practice emphasizes that program or unit assessments employ assessments in multiple forms. This is necessary to examine candidate proficiencies in relation to different types of standards (e.g., those addressing knowledge, teaching

---

2 Joe Prus, with the assistance of Mary Diez, has prepared this listing of selected resources on “Performance Assessment in Higher Education;” it begins on page 72 of this paper.
3 See, also, the NCATE paper, “Aligning Assessments with Standards,” under “Performance Assessment Papers” at the NCATE web site: http://www.ncate.org/resources/m_resources.htm
skill, dispositions, or effects on P-12 student learning) and to measure candidate progress at different points in time.

- Examples of assessments well aligned with standards include:
  a. Example A, “A Performance Assessment of Teaching Effectiveness and Student Learning,” Alverno College
  b. Example B, “University Assessment System Design,” University of Mississippi School of Education
  c. Example C, “A Course Assessment for Foundations of Teaching and Learning,” Alverno College
  e. Example R, “English Proficiency Exam,” Hofstra University

2. Assessments are accompanied by explicit statements of proficiencies candidates are expected to demonstrate in their responses.

- Unit and program leaders should be clear and explicit about their expectations for candidate proficiencies in relation to standards, and candidates should know and understand what those expectations are so they can effectively strive to achieve them. Faculty expectations may be conveyed in narrative descriptive material, perhaps including examples, in advance of an assessment. When assessments are administered, clear and explicit expectations refer both to instructions and scoring guides, discussed in the next two points.

- Faculty have a responsibility to provide clear assessment directions covering what candidates are supposed to do, how their responses are to be prepared, and any conditions or constraints such as access to resources or page restrictions.

- Scoring guides should address relevant and meaningful attributes of candidate performance on an assessment task and may occur in several forms, such as:
  a. simple number scales (not explained or anchored),
  b. Likert scales (a form in which the respondent indicates levels of agreement from strongly agree to strongly disagree),
  c. narrated holistic scales (descriptions that characterize what knowledge and skills a candidate would exhibit at a particular point on the scale),
  d. analytic scales (with descriptions of knowledge or skills on specific traits), and
  e. composite or number scores (representing proficiency on knowledge tests).

The c and d forms of scoring guides are sometimes called “rubrics,” a term that NCATE defines as:

written and shared criteria for judging performance that indicate the qualities by which levels of performance can be differentiated, and that anchor judgments about the degree of success on a candidate assessment.

- The committee evaluated scoring guides with reference to their appropriateness for the assessment form to which they applied. Ones the Committee found most
successful were examples that based judgments on explicit statements of the knowledge, skills, and dispositions that could be matched with candidate proficiencies. Similarly, the most successful examples had more than one level of proficiency, rather than a simple criterion, thus facilitating a reviewer’s ability to evaluate the extent of candidate proficiencies more accurately.

- Examples of explicit statements of proficiencies and scoring guides appropriate for their assessments include:
  a. Example D, “Performance Based Assessment in Teacher Preparation Using a Web-based System,” University of Central Arkansas
  b. Example E, “Professional Education Program Admission I Requirements,” University of Wisconsin Oshkosh
  c. Example I: “Action Research Project using NCTE Standards,” University of Charleston, WV
  d. Example J: “Learning Log on the Uses of Assessment in Instruction,” University of Nevada, Las Vegas
  e. Example K: “Math Prompt and Scoring Guide,” Wheelock College, Massachusetts

3. **Assessments are constructed so that different levels of candidate proficiency are clearly distinguished; scoring processes are consistent and reviewers are trained so that judgments can be discriminating.**

- When decisions are to be made about programs, or about candidate needs, progress, or completion, the decisions must be based on strong evidence as to the level of performances achieved. This requires:
  a. Assessments constructed in a way that provides candidates an opportunity to show the proficiencies they have attained across a range of levels, from beginning to advanced;
  b. Scoring guides that distinguish one level of performance from another and development of common procedures to judge responses; and
  c. Training for the raters—including examples of performance at different levels—so that candidate responses are interpreted consistently.

- The *Standards for Educational and Psychological Testing*\(^4\) prepared jointly by AERA, APA, and NCME, provide supporting guidance.
  a. The *Standards* emphasize standardized scoring processes. Standard 3.22 urges that scoring guides be “in sufficient detail and clarity to maximize the accuracy of scoring.” In addition, the standard states that “instructions for using rating scales . . . should be clear.” (p. 47)
  b. The *Standards* also address consistency in scoring. Standard 3.24 calls for “adequate training and instruction” of scorers, and additional steps “for examining scorer agreement and accuracy.” The comment for this standard

goes on to describe use of examples of test takers’ responses at each score level as a way to ensure consistent interpretations of responses. (p. 48)

- Examples of assessments that can make effective discriminating judgments that were selected by the Committee include:
  a. Example F: “Impacting Student Learning,” Augusta State University
  b. Example G: “Teacher Work Sample,” The Renaissance Partnership for Improving Teacher Quality

4. **Assessments are used to reach meaningful decisions, including ones that identify candidate needs; judge candidate progression; and evaluate courses, programs, or units. Assessments that are combined to reach a decision are accompanied by explicit and public statements describing the manner of weighting and summarizing results.**

- Information that comes from assessments of candidate proficiencies is the primary basis for making decisions in preparation of professional educators. The NCATE unit standards emphasize this in standard 2, *Assessment System and Unit Evaluation*, in which the supporting explanation begins:
  The unit has a professional responsibility to ensure that its programs and graduates are of the highest quality. Meeting this responsibility requires using information technologies in the systematic gathering and evaluation of information and making use of that information to strengthen the unit and its programs. The unit and its programs are informed by an assessment system that examines the (1) alignment of instruction and curriculum with professional, state, and institutional standards; (2) efficacy of courses, field experiences, and programs, and (3) candidates’ attainment of content knowledge and demonstration of teaching that leads to student learning.

- Leaders in units and programs need to determine how various candidate assessments—perhaps supplemented with other information such as grade point averages—are assembled into cumulative decisions about candidates’ progress. This would be done by determining what assessment and other information to include, and how much each piece of information is weighted in making decisions. Faculty who are evaluating candidate progress need to be trained so they will reach consistent judgments about factors that discriminate among candidates.

- An example of an assessment well-linked with meaningful decisions is:
  a. Example L: “Admissions Portfolio Requiring Common Essays,” University of Nebraska, Lincoln

5. **Assessment systems include some “authentic” forms of assessment in which candidates are asked to perform tasks similar to those they will face in their initial employment as education professionals. These contain assessments of learning by the candidate’s P-12 students.**

- The components of a complete assessment system should include “authentic” forms of assessment—ones in which the tasks are work in a classroom, direct measures of classroom performance, or similar to daily activities a candidate would face in initial
employment following successful completion of a program. This form of assessment includes measures of the performance of a candidate’s P-12 students. Committee members note that there should be both an authentic situation (for example, a classroom with P-12 students) and a task that calls for an authentic response (for example, teaching a lesson that is video recorded and then reflecting on particular aspects of the teaching performance as recorded).

- Alverno College describes a “principle of performance” that captures this idea of authenticity:

  The principle of performance requires that we assess abilities in action, in the kind of integrated situation in which students will use them in their life beyond campus. This principle insists that if we are to assess our students’ thinking, we must find ways to make their thinking observable. If we are to assess their problem solving ability, we must observe them solving problems. If we are to assess their interactive ability, we must provide a situation and watch them interact.

---

5 This is a frequent topic in assessment literature. One example is an excerpt on “Authenticity,” from a book by assessment expert Grant Wiggins, *Assessing Student Performance*, published by Josey Bass in 1993. It comes from pages 228-230:

What we require, therefore, are more general design criteria that can be useful for framing challenges that are psychometrically useful but also more “authentic”—that is, that require performance faithful to criterion situations. But first we need to be clear about what we mean by *authenticity*. Here is my latest version of a much-revised set of criteria for judging the authenticity of a test. Authentic tests of intellectual performance involve the following:

1. Engaging and worthy problems or questions of importance, in which students must use knowledge to fashion performances effectively and creatively. The tasks are either replicas of or analogous to the kinds of problems faced by adult citizens and consumers or professionals in the field.
2. Faithful representation of the contexts facing workers in a field of study or in the real-life “tests” of adult life. The formal *options, constraints, and access to resources* are apt as opposed to arbitrary. In particular, the use of excessive secrecy, limits on methods, the imposition of arbitrary deadlines or restraints on the use of resources to rethink, consult, revise, and so on—all with the aim of making testing more efficient—should be minimized and evaluated.
3. Nonroutine and multistage tasks—in other words, *real* problems. Recall or “plugging in” is insufficient or irrelevant. Problems require a repertoire of knowledge, good judgment in determining which knowledge is apt when and where, and skill in prioritizing and organizing the phases of problem clarification and solution.
4. Tasks that require the student to produce a *quality* produce and/or performance.
5. Transparent or demystified criteria and standards. The text allows for thorough preparation as well as accurate self-assessment and self-adjustment by the student; questions and tasks may be discussed, clarified, and even appropriately modified, through discussion with the assessor and/or one’s colleagues.
6. Interactions between assessor and assessee. Tests ask the student to justify answers or choices and often to respond to follow-up or probing questions.
7. Involve response-contingent challenges where the effect of both process and product/performance (sensitivity to audience, situation, and context) determines the quality of the result. Thus there is concurrent feedback and the possibility of self-adjustment during the test.
8. *Trained* assessor judgment, in reference to clear and appropriate criteria. An oversight or audit function exists: there is always the possibility of questioning and perhaps altering a result, given the open and fallible nature of the formal judgment.

---

6 *Student Assessment-as-Learning at Alverno College*, Alverno College Faculty, Alverno College Institute, 1994, Jo Ann Schmitz, Editor, p. 19
A more succinct description of “authentic” assessment from Alverno, comes from “Criteria for review of a newly designed performance assessment,” used across faculty at the College:

Does the assessment prompt bring students as close as possible to a personal and/or professional situation within which they will be using the outcomes outside the classroom?

- If an observer is visiting a classroom (where the candidate is performing an “authentic” task), the value of the observation is enhanced when the scoring guide is tailor made to the observation, and explicit so that observers and candidates have the same understanding of the basis for judging performance. Sometimes this can be done with descriptive words, or from a selection of prepared phrases that characterize a candidate’s actions. In its instructions for a teacher portfolio assessment, INTASC (the Interstate Testing and Assessment and Support Consortium of the Council of Chief State School Officers) says: “. . . it is not possible to assess teaching without viewing instruction.” In INTASC portfolio assessments (created for first year teachers in mentoring programs), videotapes are an essential component, along with diverse examples of candidate work; judgments are reached through trained reviewers, and examples of the new teacher’s work at different levels are provided.

- Examples of well developed “authentic” forms of assessment include:
  a. Example F: “Impacting Student Learning,” Augusta State University
  b. Example G, “Teacher Work Sample,” The Renaissance Partnership for Improving Teacher Quality
  c. Example H: “Model of Appropriate Practice for Teacher Candidates,” Columbus State University
  d. Example I: “Action Research Project using NCTE Standards,” University of Charleston, WV
  e. Example M: “On-line Performance Requirements for Technology,” Boston College

6. Assessments are systematically evaluated to ensure fairness, accuracy, consistency, and avoidance of bias.

- The “credibility” of assessments of candidate proficiencies should be systematically evaluated by faculty along several dimensions. First, assessments should measure

---

8 INTASC, op cit., p. 57. INTASC provides the following “evaluation scale” descriptions for level 3 and level 1 performance (on a 4 point scale) for “analysis of learning” as exhibited in the videotape plus written material in the portfolio:

Level 3: The teacher gathers sufficient information to be able to determine what the students are learning. The formal assessments maintain an emphasis on important mathematics, extending beyond procedure, and are aligned in content and process with the instruction. The evaluation of student work is accurate and includes some attention to a diagnosis of the student error. The teacher establishes and communicates clear scoring criteria and/or rubrics for formal assessments. Feedback to students clearly communicates what the student did well and where he or she needs work.

Level 1: The teacher gathers a minimal amount of information to determine what the students are learning. The formal assessments may not emphasize mathematics. The evaluation of student work is inaccurate and/or inconsistent. Feedback to students is inaccurate or absent.
meaningful knowledge and skills that are taught. Assessments are “fair” only if the program and unit provide opportunities for candidates to learn and practice the knowledge and skills examined in the assessment and only when candidates know what will be expected of them. Second, assessments should measure what they purport to measure—that is, they should be accurate. Third, the measures should be robust so that results are dependable and, therefore, consistent over time. And fourth, assessments should be constructed in ways that avoid biases in testing situations or in language.

- The extent of “credibility” evaluation in candidate assessments may vary to some degree from one institution to another. Some institutions make official recommendations for licensure to a state licensing agency, so the decisions are high stakes for the candidates. Other institutions may use the information from candidate assessments more for grading performance in individual courses and overall evaluation of programs, rather than for summative evaluations of candidates. The credibility studies should be focused on the institution’s intended purposes.

- An assessment system will be more credible when it contains some external corroboration of candidate performance. Information from state licensure is an external measure. Another is evaluations conducted by employers. Principals and mentors of new teachers, especially, and school personnel officials, can be sources of descriptive information about early experiences of new teachers. Such information can be used by units to evaluate and revise preparation programs. Some states are trying out “guarantees” so that teacher graduates can return for additional coaching if their employers find shortcomings. Employers can also serve as advisors to education units about what knowledge, skills, and dispositions—important on the job—should be included in assessments.

- Even though faculty may think that establishing credibility of assessments is a time consuming, highly technical, and perhaps even mysterious process, in fact virtually all institutions are already gathering information that can help to ensure credibility of assessments. In the list that follows, there are eleven strategies that institutions can undertake to document the credibility of their assessments. The first seven, a-g, would be addressed in planning of an assessment system, in applying criteria on alignment or scoring guides, or in taking steps to differentiate different performance levels described in criteria 1, 2, 3, and 4 of this paper. Only strategies h-k require additional analyses, evidence, or documentation. Strategies to document credibility include:
  a. providing evidence for linkages of the assessment with the standards it is designed to inform (Criterion 1);
  b. clearly articulating the expectations for student performance (Criteria 2 and 4);
  c. documenting the scoring guides so that raters are clear on the relevant features of the candidate performance that are being evaluated (Criteria 2 and 4);
  d. providing training for raters/scorers so that they have a clear understanding of the meaning of the various score points (Criterion 3);
e. providing advance information to candidates so that they understand what they are expected to do, what they will be evaluated on, what level of performance is expected for scoring decisions, and any context features (availability of support, resources, page limitations, structure of response that may influence a candidate’s performance on the task) (Criteria 2 and 4);

f. gathering evidence of consistency of scoring of candidate performance by different raters and/or on different occasions, or by parallel assessments (see reviewer training in Criterion 3);

g. documenting where in the curriculum candidates have the opportunity to learn the test content (a necessary and usual part of an assessment planning process would be to “map” the standards and assessments onto the courses and experiences in a unit or program9);

h. conducting analyses that address the underlying structure of the assessment, such as internal consistency and match to the intended underlying content or cognitive and skill components10;

i. gathering evidence that the assessment is constructed in such a way as to avoid biases in language or testing situations;

j. gathering evidence of the relationship between performance on the assessment tasks and important program outcomes, such as performance on subsequent coursework, progress of the candidate through the program, and successful completion of the program;

k. documentation of the expected relationships between performance on the assessment task and performance on other related (convergent) or unrelated (divergent) tasks.

• Examples of unit and program efforts to evaluate the credibility of their assessments are:
  a. Example G, “Teacher Work Sample,” The Renaissance Partnership for Improving Teacher Quality

9 For example, see the NCATE website, http://www.ncate.org/resources paper, “Assessment Systems: An Explanation of the NCATE Transition Plan.” Also, at the same address, choose “Commissioned Papers on Development of Assessment Systems and Aggregating Data,” then select E. Lynne Weisenbach, “Making the Case: Marshaling the Evidence About Candidate Proficiency,” (see page 5), and Beth Stroble, “Unit Assessment Systems,” (see page 8).

10 To emphasize the point, the issue in strategy h is whether an assessment is measuring the concept or characteristic that is intended to be measured, or “construct validity.” For example, if the intent is to measure three components of performance, and provide subscores that represent candidate proficiencies on these components, then it is important that the analyses support conclusions that these underlying components are present and provide interpretable scores. Most often this is accomplished through some statistical analyses documenting that the subscales are reasonably homogeneous so that a single score can reliably represent candidate performance on the subscales. Factor analyses are also used to identify (explore) or verify (confirm) that the components are reasonably represented in the test.
Committee Review Part B: Consideration of Cost, Efficiency, and Innovative Features

As was noted earlier, this Part B of the paper addresses Committee consideration of assessment cost, efficiency, and innovative features. It supplements the criteria described in Part A that served as the basis for reviewing inherent qualities and uses of assessment instruments and systems that were submitted to the Committee.

In their submissions, sponsors were not asked for information about the cost of preparing and administering assessments. That would have been a difficult task. In most cases, costs take the form of faculty and administrator time. Even in those cases where development of the assessment is a special project carried out with designated funds—as may have been the case with the assessment system examples—contributed faculty time would be a significant additional component. Nevertheless, Committee members, drawing on their own experience, were struck with the obvious range in costs associated with different assessment examples. Their observations are described in the first “consideration,” below.

In the letter of invitation for submissions of assessments, the Committee had established a review factor for “effectiveness and efficiency of the assessment in gathering important and significant information.” This introduces the notion of the amount of candidate information in relation to the investment. Again, without detailed cost information what can be observed is limited, but Committee members believe that effectiveness and efficiency of the assessment investment is something that faculty should consider so this is the second “consideration” below.

Finally, in the invitation for submission of examples, the Committee had established a review factor for “innovative features illustrated by the assessment, including use of information and computer technology.” As they conducted their review, the Committee identified many innovative elements of assessments to highlight, and those are described in the third “consideration” that follows.

1. Developmental costs and operating costs

- Costs associated with assessments can be addressed from the perspective of when they occur and how they are incurred. The “when they occur” dimension refers to developmental costs on one hand and recurring operational costs on the other. The “how they are incurred” dimension includes large scale campus, multi-campus, or state system arrangements at one end of the cost range, and individual faculty assessment development (as part of course preparation) at the other.

- Assessment systems examples have high initial developmental costs. Once developed, it appears that annual operating costs even for these costly assessments would be more modest, yet there are recurring expenses associated with the computer input, storage, and data analysis at least, probably a continuing need for faculty and staff training, and on-going management expenses for the system as well.
a. An example is the Renaissance project, supported by Federal grants with the goal of building a model that can be adapted and used without charge. Even if the “model” is freely available to others in the future, institutions adopting the Renaissance approach would need a period for developmental training of faculty and staff such as the Renaissance institutions experienced through their participation in the project. Once it is in operation, the Renaissance approach to Teacher Work Sampling may not have costs very different from other forms of clinical experience portfolio assessments.

b. Both the developmental costs and the annual operating costs can be made less burdensome, however, when they can be spread over several institutions and large numbers of candidates—that is, the unit-cost of assessment information for individual candidates can be reduced. State systems might benefit from this lower unit-cost effect. The counter-argument to reducing unit-costs through multi-unit assessments, however, could be loss of “fit”—that is, a group of several institutions may not share goals, standards, or conceptual frameworks for their educator preparation units, so commonly used assessments may not be appropriate.

• At the other end of the cost spectrum are assessments created by individual faculty members as a part of a course. This may involve no add-on costs whatever since a faculty member ordinarily devises ways to evaluate student learning acquired in a course. There are two cautions with this kind of assessment development and operation, however. One is that some faculty may have little training or expertise in assessment development and use. The assessment criteria described in Part A of this paper would be useful guides for these faculty, but further investment in development of faculty expertise would be essential if the resulting assessments are to be of high quality. Another caution is specific to NCATE. The NCATE emphasis in assessment for accreditation is on assessments that are taken commonly by all candidates in a unit, or at least in a program. Assessments taken within courses that are not common across all candidates cannot be readily summarized to represent the performance of a cohort of candidates. But where a course is required of all candidates, and where faculty do have training and experience in creating assessments appropriately linked with standards, then faculty developed assessments can be an excellent alternative at relatively low cost.

2. Effectiveness and efficiency of assessments in gathering important and significant information

• Faculty and administrators should routinely consider the amount and usefulness of assessment information from candidates in relation to costs. Assessments that provide information on multiple standards and serve multiple purposes can be efficient. However, a large investment in a web-based data management system that includes only spotty assessments, or ones failing to represent the scope of standards, may not be a cost-effective source of information. State licensure exams that fail to measure candidate performance in clinical experiences, and assessments in courses that are not commonly elected by all candidates in a unit or program, are also not
cost-effective sources of information by themselves. Both examples selected by the Committee integrate candidate assessment information with instruction (the “action research project” in West Virginia and “learning log on the uses of assessment in instruction” from Nevada).

- Examples of efficient and effective assessments that were selected by the Committee include:
  
  1. Example I; “Action Research Project Using NCTE Standards,” University of Charleston, WV
  2. Example J; “Learning Log on the Uses of Assessment in Instruction,” University of Nevada, Las Vegas

3. **Innovative features illustrated by assessments, including use of information and computer technology**

   - Units and programs represent different stages of developing performance assessments and assessment systems. What may be an innovative assessment for one institution may be common practice in another, while a third institution may have progressed several generations beyond a particular assessment. Committee members selected particular elements of assessment submissions that they thought represented unusual practice, at least, certainly not what they perceived as common or routine. For each example of an innovative assessment, below, the Committee has identified the features or elements that led to the selection.

<table>
<thead>
<tr>
<th>Example</th>
<th>Innovative element</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example B, “A Performance Assessment of Teaching Effectiveness and Student Learning,” University of Mississippi School of Education</td>
<td>Integration of overall assessment system</td>
</tr>
<tr>
<td>Example E, “Professional Education Program Admission I Requirements,” University of Wisconsin Oshkosh</td>
<td>Assessment includes dispositions</td>
</tr>
<tr>
<td>Example F, “Impacting Student Learning,” Augusta State University</td>
<td>Assessment includes learning by candidate’s P-12 students</td>
</tr>
<tr>
<td>Example G: “Teacher Work Sample,” The Renaissance Partnership for Improving Teacher Quality</td>
<td>Illustration of work sample methodology</td>
</tr>
<tr>
<td>Example H: “Model of Appropriate Practice (MAP) for Teacher Candidates,” Columbus State University, Georgia</td>
<td>Conceptual framework modeled on Pathwise; used by instructor for feedback to candidates</td>
</tr>
<tr>
<td>Example I: “Action Research Project using NCTE Standards,” University of Charleston, WV</td>
<td>Content of the assessment</td>
</tr>
<tr>
<td>Example L: “Admissions Portfolio Requiring Common Essays,” University of Nebraska, Lincoln</td>
<td>Use of weights in applying rubrics and inclusion of 50% of score based on dispositions (i.e.,</td>
</tr>
<tr>
<td>Example M: “On-line Performance Requirements for Technology,” Boston College</td>
<td>Direct application of web-based assessment to evaluate candidate proficiencies in technology</td>
</tr>
<tr>
<td>Example N: “Assessing the Teaching Characteristics of Preservice Teachers in an Internship Program,” California State University, Fresno</td>
<td>Assessment center used for part of preservice assessment</td>
</tr>
<tr>
<td>Example O: “INTASC Principles and the BSU Student Teaching Rubric,” Ball State University</td>
<td>Assessment linked with INTASC standards used by the state; also has consistent rubrics with guidelines for reviewers</td>
</tr>
<tr>
<td>Example P: “Performance Assessment with Teacher Standards (PATS),” University of South Florida, St. Petersburg College of Education</td>
<td>Illustrates a computer and web-based way to manage data</td>
</tr>
<tr>
<td>Example Q: “Information and Assessment Systems,” University of Maryland Baltimore and the University of North Carolina at Charlotte</td>
<td>Overall architecture of the assessment system</td>
</tr>
</tbody>
</table>
Example locator by type and use of assessment

In the chart below, the word or phrase following each listed example is intended to describe the principal feature of the assessment example discussed in meetings of the Committee. There is no intent to indicate how the sponsor would describe the example or to represent all the features of an example. State names in this listing indicate that standards are based on state requirements, such as “INTASC” in Indiana and “Accomplished practices” in Florida.

1. By Decision Point in a Preparation Program

<table>
<thead>
<tr>
<th>Admissions</th>
<th>B, multiple assessments and assessment at decision points; E, multiple indicators in an admissions portfolio; L, admissions portfolio with dispositions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical experience</td>
<td>A, clinical completion; F, faculty assessments with common scoring guide, Georgia; G, teacher work sample; H, ETS Pathwise and feedback to candidate, Georgia; O, INTASC-based practicum courses with tasks and scoring guides</td>
</tr>
<tr>
<td>End of program</td>
<td>B, multiple assessments and assessment at decision points; R, English proficiency graduation requirement</td>
</tr>
</tbody>
</table>

2. By Element of Program or Unit Standards

<table>
<thead>
<tr>
<th>Subject content knowledge</th>
<th>E, multiple assessments in an admissions portfolio including content knowledge and dispositions; I, NCTE standards and action research; K, remedial math; R, English proficiency graduation requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedagogical knowledge</td>
<td>C, “foundations” course assessment; F, faculty assessments with common scoring guide, Georgia; J, “log” on assessment knowledge</td>
</tr>
<tr>
<td>Professional knowledge</td>
<td>I, NCTE teacher preparation standards and action research; N, vignettes administered in an assessment center</td>
</tr>
<tr>
<td>Classroom skill</td>
<td>A, clinical completion; F, faculty assessments with common scoring guide, Georgia; G, teacher work sample; H, ETS Pathwise and feedback to candidate, Georgia; O, INTASC-based practicum courses with tasks and scoring guides</td>
</tr>
<tr>
<td>Dispositions</td>
<td>A, clinical completion; E, multiple assessments in an admissions portfolio including content knowledge and dispositions; J, “log” on assessment knowledge; L, admissions portfolio with dispositions</td>
</tr>
<tr>
<td>Candidate influence on P-12 student learning</td>
<td>A, clinical completion; F, faculty assessments with common scoring guide, Georgia; G, teacher work sample; P, standards based, with validation studies, Florida</td>
</tr>
<tr>
<td>Candidate technology</td>
<td>M, candidate computer and web capabilities</td>
</tr>
</tbody>
</table>
### Competence

<table>
<thead>
<tr>
<th>Competence</th>
<th>A, clinical completion; C, “foundations” course assessment; F, faculty assessments with common scoring guide, Georgia; G, teacher work sample; N, vignettes administered in an assessment center; Q, system architecture</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCATE standard 1</td>
<td>A, clinical completion; F, faculty assessments with common scoring guide; H, ETS Pathwise and candidate feedback, Georgia; O, INTASC-based practicum courses with tasks and scoring guides, Indiana</td>
</tr>
<tr>
<td>INTASC principles</td>
<td>A, clinical completion; F, faculty assessments with common scoring guide; H, ETS Pathwise and candidate feedback, Georgia; O, INTASC-based practicum courses with tasks and scoring guides, Indiana</td>
</tr>
</tbody>
</table>

### 3. By Form of Assessment

<table>
<thead>
<tr>
<th>Form of Assessment</th>
<th>C, “foundations” course assessment; L, admissions portfolio with dispositions; R, English proficiency graduation requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Essay or writing</td>
<td>C, “foundations” course assessment; E, multiple assessments in an admissions portfolio including content knowledge and dispositions; J, “log” on assessment knowledge; L, admissions portfolio with dispositions</td>
</tr>
<tr>
<td>Reflections</td>
<td>C, “foundations” course assessment; E, multiple assessments in an admissions portfolio including content knowledge and dispositions; J, “log” on assessment knowledge; L, admissions portfolio with dispositions</td>
</tr>
<tr>
<td>Performance task</td>
<td>A, clinical completion; B, multiple assessments and assessment at decision points; E, multiple measures in an admissions portfolio including content knowledge and dispositions; G, teacher work sample; H, ETS Pathwise and feedback to candidate, Georgia; K, remedial math; O, INTASC-based practicum courses with tasks and scoring guides, Indiana</td>
</tr>
<tr>
<td>Computer administered</td>
<td>M, candidate computer and web capabilities</td>
</tr>
<tr>
<td>Action research</td>
<td>I, NCTE teacher preparation standards</td>
</tr>
<tr>
<td>Multiple assessments</td>
<td>B, multiple assessments and assessment at decision points; E, multiple measures in an admissions portfolio including content knowledge and dispositions; N, vignettes administered in an assessment center</td>
</tr>
<tr>
<td>Vignettes</td>
<td>N, vignettes administered in an assessment center</td>
</tr>
</tbody>
</table>

### 4. By Size of Institution

<table>
<thead>
<tr>
<th>Size of Institution</th>
<th>A, clinical completion; C, “foundations” course assessment; I, NCTE teacher preparation standards and action research</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Small” institutions</td>
<td>E, multiple assessments in an admissions portfolio including content knowledge and dispositions; F, faculty assessments with common scoring guide, Georgia; G, teacher work sample; H, ETS Pathwise and candidate feedback, Georgia; K, remedial math; L, admissions portfolio with dispositions; P, standards based, with validation studies, Florida; Q, INTASC-based practicum courses with tasks and scoring guides, Indiana</td>
</tr>
<tr>
<td>“Medium” institutions</td>
<td>B, multiple assessment and assessments at decision points; D, Danielson framework and ETS Pathwise; G, teacher work sample; J, “log” on assessment knowledge; M,</td>
</tr>
<tr>
<td>“Large” institutions</td>
<td></td>
</tr>
</tbody>
</table>
candidate computer and web capabilities; N, vignettes administered in an assessment center; O, INTASC-based practicum courses with tasks and scoring guides, Indiana; R, English proficiency graduation requirement

### 5. By Location of Assessment Administration

| Administered in a course | B, multiple assessment and assessments at decision points; C, “foundations” course assessment; F, faculty assessments with common scoring guide, Georgia; I, NCTE teacher preparation standards; J, “Log” on assessment learning; K, remedial math |
| Administered in a clinical setting, or course/ clinical | A, clinical completion; F, faculty assessments with common scoring guide, Georgia; G, teacher work sample; H, ETS Pathwise and candidate feedback, Georgia; O, INTASC-based practicum course with tasks and scoring guide |
| Multiple | B, multiple assessment and assessments at decision points; D, Danielson framework and ETS Pathwise; P, standards based, with validation studies, Florida; Q, INTASC-based practicum courses with tasks and scoring guides, Indiana |
| Prior to admissions | E, Multiple assessments in an admissions portfolio including content knowledge and dispositions; L, admissions portfolio with dispositions; |
| Graduation requirement | M, candidate computer and web capabilities; N, vignettes administered in an assessment center; R, English proficiency graduation requirement |

### 6. Assessment Systems

| Assessment systems | B, multiple assessments and assessment at decision points; D, Danielson framework and ETS Pathwise, Arkansas; P, standards based, with validation studies, Florida; Q, system architecture |

NOTE: Readers are encouraged to be in touch with individuals listed as the contact point in the following examples. The authors and developers who submitted these examples are, in most cases, continuing their work on piloting, examining, and refining the assessments. They will have experiences to share about the design and functioning of the assessments and the relationship with professional standards; they will have critical information, as well, about the surrounding context and uses made of the assessment by the program and unit.
A Performance Assessment of Teaching Effectiveness and Student Learning

This example is from Alverno College, Milwaukee, Wisconsin. Alverno falls into the “small” category for the Assessment Examples Project, with “10 to 50” annual graduates. Candidates are admitted at the beginning of the sophomore year.

The assessment is administered at the point of exit from clinical practice for candidates in initial teacher preparation programs. It is intended to evaluate candidate proficiencies under NCATE unit standard 1 on candidate knowledge, skills, and dispositions. The assessment requires candidates to complete components of a “packet” as evidence of teaching effectiveness and the excerpt copied on the next page details tasks to be performed by the candidate. The full example also includes criteria for evaluation of evidence. Those for the Alverno goal, “Teachers know the subject they are teaching” appear in the chart below:

<table>
<thead>
<tr>
<th>EXCEEDS STANDARD</th>
<th>MEETS STANDARD</th>
<th>DOES NOT MEET STANDARD</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-1. Concepts, content, and skills to be taught are clearly stated on lesson plan(s). Interdisciplinary connections are explicitly made.</td>
<td>A-2. Concepts, content, and skills to be taught are clearly stated on lesson plan(s)</td>
<td>A-3. Concepts, content, and skills are not clearly stated on lesson plan.</td>
</tr>
<tr>
<td>B-1. Explanations of content are clear, detailed, thorough, and age appropriate. Examples make content meaningful by relating it to students’ lives or understandings.</td>
<td>B-2. Explanations of content are incomplete for this age learner. Examples did not relate subject matter to students’ lives or understandings.</td>
<td>B-3. Explanations of content are unclear for this age learner. Examples are not relevant for this content or age learner.</td>
</tr>
<tr>
<td>C-1. Feedback on student work extends student understanding of concepts, content, and skills by asking meaningful questions and/or making comments that help students move forward.</td>
<td>C-2. Feedback in the form of questions and/or comments on student work begins to extend student understanding of concepts, content, and skills.</td>
<td>C-3. Feedback is limited and does not move student understanding forward.</td>
</tr>
</tbody>
</table>

The full example has similar decision criteria for five other Alverno goals:
- Teachers use their knowledge of child and adolescent development and of the diverse needs of learners when planning instruction
- Teachers understand and use a variety of instructional strategies
- Teachers know how to assess student learning
- Teachers are able to evaluate themselves
- Teachers know how to create and maintain an inviting learning environment

The assessment is used both for analysis of individual candidate performance during clinical experience and/or internship, and also for evaluation of Alverno’s professional education unit.

The excerpt for this example was selected to illustrate alignment of the performance assessment with the standard (in this case, NCATE unit standard 1). Note that the example includes “assessment of student learning” explicitly as one evidence of candidate effectiveness.

CONTACT: Kathy Lake, Chair, Education Division, Kathy.lake@alverno.edu
Last revised April 3, 2002
Example A—ALVERNO COLLEGE
A Performance Assessment of Teaching Effectiveness and Student Learning

This semester you will have multiple opportunities to demonstrate your effectiveness as a teacher and to document that effectiveness through a process that is integrated with your student teaching experience. Through this process, you will demonstrate how you plan, teach, assess, give feedback, and evaluate your teaching and its effect on student learning. Along with the materials you design and implement, student work will be an important component of this process.

Demonstration of overall teaching effectiveness is the major purpose of this assessment. Emphasized within this purpose is the process of creating learning that considers the range of learning abilities within a classroom group and connects to the content standards used in Wisconsin.

**DIRECTIONS**

Select a unit or related group of lessons that you will be teaching this quarter. Using the descriptions below, collect evidence of your effectiveness as a teacher in an accordion folder provided to you when you purchase your student teaching syllabus. At the end of each 9-week period, you will submit the completed packet in the accordion folder to the Education Department. It will be retained there as evidence of your teaching effectiveness.

You will compile one sample from both the first and second 9 weeks of your student teaching.

The packet should include:

**General Information**

1. **Context Setting**-Describe school, grade level, class size, description of students and school culture. Explain how the lessons/unit you have chosen comprises or fits into the curriculum of this grade/subject.

2. **Students’ current level of understanding**-Describe the current levels of student understanding of the concept, skill, or content to be taught. Be sure to include the range of student ability levels within your classroom in your description. Include a summary of how you determined this information and what you found. Possible resources for this information might be: discussion with your cooperating teacher, assessment information from previous units, an initial activity that is used to determine understanding, or any other creative methods you may choose to employ.

**Instructional Strategies**

3. **Lesson/Unit Plans**-Include the lesson plans, self assessments, and teacher/supervisor feedback that represent the scope of the unit, with a minimum of two plans. Explain how the plans explicitly connect the lesson goals/objectives to district or state P-12 standards.

4. **Videotape**-Collect videotape that represents at least one lesson from the unit.

5. **Assessment of Student Learning**-Describe the strategies used to assess student learning over the course of the lessons/unit. These strategies should assess students’ understanding of skills, concepts, or content that are goals of the unit or lessons. Include an explanation of why you selected or designed this type of assessment. Examples of assessments include focused observations, written products, speeches, panels, projects, and simulations. In addition to this explanation, each assessment should include:
   a. Criteria for student performance
   b. Student self assessment opportunity
   c. Feedback, based on criteria, that you have provided to students

**Analysis of Three Students**

6. **Samples of P-12 Student Work**-Select at least three (3) students who are of differing levels of ability. Include work done by these students throughout the unit or series of lessons, including that which represents initial instruction.

7. **Analysis of Student Learning**-Analyze student gains using work samples to provide evidence of student learning relative to unit/lesson goals and objectives.

**Synthesis of Your Growth**

8. **Self Assessment**-Complete a self-assessment of this process using the attached rubric. Be sure to indicate where to locate the evidence of each standard on your self assessment.

9. **Reflection on Teaching Growth**-Using your self assessment rubric (#8) as a guide, describe your strengths and areas for future development as you reflect on this assessment of teaching effectiveness experience. April 5, 2002
Unit Assessment System

This example is from the School of Education at the University of Mississippi, University, Mississippi. Mississippi falls into the “large” size category for the Assessment Examples Project, with more than 200 annual graduates.

The example was submitted as a comprehensive assessment system that is being developed to evaluate candidate, program and unit performance continuously. The system is created around five specific “outcomes” of the conceptual framework:

- Life-long learners who take responsibility for their own learning and continuously foster their professional renewal.
- Problem solvers who develop solutions to improve the educational environment for all students.
- Communicators who effectively use verbal, non-verbal, electronic, and print modes of communication to establish a positive school environment and promote student thinking and learning.
- Users of technology who integrate multimedia in learning environments as instructional and management tools to enhance student learning.
- Professionals committed to diversity who appreciate diverse contexts and promote and model the values of diversity.

The first part of the excerpt on the following two pages is taken from the University’s description of its education unit assessment system, and, more specifically, the components for “candidate assessment,” “program evaluation,” and “unit evaluation.” That is followed, at the bottom of the page, with examples of course assessments used in each program (only the elementary bachelors program is shown in the excerpt) to measure candidates’ performance in relation to each of the outcomes quoted above.

On the second excerpt page for this example, below, assessments that are used at decision points are listed. The excerpt shows all of the “Summary of criteria for program entry” and about one-half of the “Summary of criteria for program completion review” for the elementary education, bachelor level. The full submission also has criteria for an “intermediate” review of candidate progress. These specify the criteria, assessment tools, and the University “standard” (the required cut score, GPA, or course grade, etc.) that determine success on each assessment. For the “program entry” excerpt, note the topics covered by the items listed in the first column (similar to NCATE unit standard 1 on candidate knowledge, skills, and dispositions), and the specific performance expectations (each suited to the assessment tool to which it applies) stated in the third column.

This example was selected by the Committee to illustrate alignment of assessments with standards, criterion 1, and also use of multiple assessments in a comprehensive system.

CONTACT: Kathleen Sullivan, Chair, Unit Assessment Committee, ksull@olemss.edu

Last revised April 3, 2003
Example B—UNIVERSITY OF MISSISSIPPI

Here are elements of the University’s description of its system:

I. Candidate Assessment
Program faculty assess candidate content and professional and pedagogical knowledge, skills, and dispositions through course-based assessments, as well as through decision-point assessments. Data from these assessments are used to make decisions about candidate performance at the pre-admission, intermediate, and program completion stages. As candidates progress through the educator preparation programs, they are expected to demonstrate increasingly higher levels of knowledge, skills, and dispositions as identified in the unit’s conceptual framework and program knowledge bases. A wide variety of assessment types are used within courses to evaluate candidate knowledge, skills, and dispositions. Examples of these assessments are group presentations, reflective essays, examinations, lesson planning activities, practicum observations, case studies, and videotape-based skill evaluations. Rubrics, checklists, and other scoring tools are used to assess candidate performance. Decision point assessments are used to determine whether the candidate meets the standards required to enter the program, continue in the program, and complete the program. Intermediate and program completion assessment criteria are aligned with SPA standards and with the unit’s conceptual framework to ensure that these assessments consistently measure candidates’ performance in relation to professional and unit standards. Candidates’ demonstration of the knowledge, skills, and dispositions needed to impact K-12 student learning is an important component of each program’s assessment procedure. The candidate assessment database will support the systematic collection, reporting, and analysis of candidate assessment data. The design of the database also supports the creation of standard and custom reports for use in evaluation at the program and unit levels.

II. Program Evaluation
To thoroughly review each program on an annual basis, program faculty and department chairs will examine findings developed through curriculum alignment audits, as well as aggregated internal data on candidate competencies and information from external sources, such as follow-up studies, candidate performance on licensure examinations, employer reports, and state program reviews. Faculty and department chairs also will use results of periodic research studies conducted by the unit using a value-added approach to compare K-12 student achievement gains of classes taught by (and schools administered by) unit graduates with statewide average gains. This program evaluation process results in development of an action agenda for revising the program curriculum for improving instruction, for revising field experiences, and for redesigning other components of the program.

III. Unit Evaluation
Just as candidate data are aggregated for use in evaluating programs, comprehensive analyses of program strengths and weaknesses are aggregated for use in evaluating the unit’s effectiveness. In the unit evaluation process, a unit evaluation committee uses aggregated analyses generated through analysis of evaluation data from all programs. The unit evaluation process also includes analysis of program recruitment and retention data, as well as unit-wide data in the form of faculty evaluations; information on student, staff, and faculty diversity; and unit leadership assessments to identify changes needed to improve unit performance. A unit-level action plan is recommended to the Dean of the School of Education by the unit evaluation committee, which consists of department chairs, designated faculty, and members of the professional community.

Examples of Assessment Types by Program and Outcome Area

<table>
<thead>
<tr>
<th>Program</th>
<th>Level</th>
<th>Lifelong Learners</th>
<th>Communicators</th>
<th>Problem Solvers</th>
<th>Users of Technology</th>
<th>Committed to Diversity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>EDCI 352 – Reflective journal of classroom observation</td>
<td>(self critique)</td>
<td>EDEL 414 – Five-day lesson plan</td>
<td>EDEL 416 – A technology</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>EDELE 417 – Dispositions Checklist</td>
<td>EDLE 417 – Observation by supervisor</td>
<td>opinion paper</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Checklist</td>
</tr>
</tbody>
</table>
### Elementary Education Bachelor Level

#### Summary of Criteria for Program Entry

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Assessment Tool</th>
<th>Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Achievement</td>
<td>Praxis I, SAT, or ACT scores</td>
<td>Praxis I Reading – 170; Writing – 172; Math – 169; SAT – 860; ACT – 21 w/no subtest below 18</td>
</tr>
<tr>
<td>Content Knowledge</td>
<td>Completion of 45-hour Core Curriculum</td>
<td>Successful completion of Core Curriculum (Applicant passed each core course)</td>
</tr>
<tr>
<td>Content Knowledge</td>
<td>GPA for Core Curriculum</td>
<td>2.50 GPA</td>
</tr>
<tr>
<td>Pedagogical Content Knowledge</td>
<td>Completion of EDCI 351 &amp; EDCI 352 w/required grade</td>
<td>Completion of EDCI 351 and EDCI 352 with grade of C or better</td>
</tr>
<tr>
<td>Appropriate dispositions</td>
<td>Dispositions checklist</td>
<td>Satisfactory ratings on dispositions checklist (progress in correcting any serious problems)</td>
</tr>
</tbody>
</table>

---

#### Elementary Education Bachelor Level

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Description</th>
<th>Rating Tool</th>
<th>Outcome Area(s) Addressed</th>
<th>Location of Assessment within Program</th>
<th>Method Used in Providing Performance Feedback</th>
<th>Method Used in Providing Developmental Opportunities*</th>
<th>Timeframe for Readministering Assessment*</th>
<th>Maximum Readministrations Permitted*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student Teaching Assessment Instrument (STAI) Observation #2</td>
<td>Candidates plan and deliver a 10-day lesson plan during the student teaching semester. Two STAI evaluations are completed by University supervisors and clinical instructors to assess candidates’ performance. The STAI is based on standards developed by INTASC. For each of the 38 indicators, candidate must achieve the passing scores (3 or 4, depending on indicator) required for graduation.</td>
<td>The STAI is based on standards developed by INTASC. For each of the 38 indicators, candidate must achieve the passing scores (3 or 4, depending on indicator) required for graduation.</td>
<td>Lifelong Learners; Problem Solvers; Communicators; Users of Technology, Commitment to Diversity</td>
<td>Completion of student teaching.</td>
<td>University supervisors conduct follow-up exit conferences at the end of the program to discuss progress made throughout the professional experience.</td>
<td>A third observation is scheduled.</td>
<td>During the student teaching semester.</td>
<td>One</td>
</tr>
<tr>
<td>Program portfolio evaluation . . .</td>
<td>Using a rubric, the university supervisor . . . The rubric includes each of the ACEI standards and . . .</td>
<td>The rubric includes each of the ACEI standards and . . .</td>
<td>Lifelong learners; problem solvers; communicators; users . . .</td>
<td>Completion of student teaching . . .</td>
<td>University supervisors conduct . . .</td>
<td>Candidates meet with their . . .</td>
<td>During student teaching. . .</td>
<td>One</td>
</tr>
</tbody>
</table>

* Applies to candidates whose performance is judged “Not Satisfactory”
A Course Assessment for Foundations of Teaching and Learning

This example is from Alverno College, Milwaukee, Wisconsin. Alverno falls into the “small” category for the Assessment Examples Project, with “10 to 50” annual graduates. Candidates are admitted at the beginning of the sophomore year. Alverno faculty have attempted to evaluate the fairness of the assessment.

The assessment is administered as part of the requirements for a single course (Foundations of Teaching and Learning) in the pedagogical sequence for initial teacher preparation. It requires teacher candidates to complete a “lesson plan project” in which they write to describe, reflect on, and apply things they have learned through the course to development of daily lesson plans for a unit, then conclude the project with a presentation to colleagues in the course. The excerpt on the next page lists the tasks. The submission describes information about candidate proficiencies that the assessment is meant to provide:

- Pedagogical Knowledge is assessed throughout the different parts of this plan, but in particular, it is covered in Part II. Students discuss motivation techniques, developmental activities, and learning theory. Then they are asked to relate the theories, techniques, and activities to the appropriate domains (cognitive, affective, or psychomotor). They must also connect what they plan in their unit with the five Alverno Education Abilities (Communication, Coordination, Diagnosis, Integrative Interaction, and Conceptualization).

- Professional Knowledge is assessed through the written lesson plans and the mandatory presentation. Each student had to organize her unit by using the components associated with good unit planning. These professional components include a rationale that was focused and goals that rise out of the rationale. Not only are goals and objectives connected to the rationale, the assessments relate directly back to the learning objectives. Professional Knowledge is also an important part of the standards that were used to build each lesson. A basic knowledge of appropriate standards and understanding of their application is a base for each lesson plan.

- Skills, dispositions, and a positive environment for P-12 student learning are implicit within each lesson plan. It is also clearly addressed in the lesson plan format. However, because these students are in their first education course, we felt that these areas were more emerging areas of knowledge.

The excerpt on the next page lists the tasks to be completed by the candidate. These are judged by faculty at “above expectations,” “meets expectations, and “doesn’t meet criteria” levels in the following areas:

- Rationale explains why this unit plan is important in the student’s learning.
- Title, target audience, and time allocations are part of the unit plan.
- Goal/s are general and rise out of the rationale.
- Learning objectives are clear and measurable.
- Key developmental activities are directly related to the goals and objectives.
- The assessment plans for the unit are directly related to the goals and objectives.
- Motivational activities are presented in a compelling manner.
- There is a clear explanation of how your unit plan will meet the needs of diverse students.
- Lesson plans include references to standards.
- Lesson plans also include references to the Alverno Education Abilities.
- You have added your self-assessment.

The full submission also has examples of candidate work and faculty feedback.

The example was selected to illustrate alignment of an assessment with standards—in this case, the NCATE unit standard on candidate knowledge, skills, and dispositions.

CONTACT: Lee Breese and Rita Nawrocki-Chabin, Lee.Breese@Alverno.edu  Last revised April 8, 2003
Example C—ALVERNO COLLEGE
ED 201
Elementary Unit Plan Project

Part I due April 8, 2002
• Title of unit
• Grade level
• Subject areas integrated into the unit
• Estimate of total time required (aim for a block of one hour for 2 weeks)
• Rationale/significance for unit
• Goal/learning objectives
  Goals will cover the broad concepts students will have at the conclusion of the unit
  Objectives will include how students will demonstrate/develop specific skills during the unit
  Attitudes/appreciations students will develop by the conclusion of the unit

Part II due April 10, 2002
• Description of motivation techniques at the beginning and where appropriate in the lesson
• Description of developmental activities from the start to the end of the unit
  Be sure to explain how these activities create a natural building of the objectives of the lesson on the part of
  the students
• Description of the assessment procedures throughout the unit.
• Integration of (explicit description of) two theories (invitational learning, multiple intelligences) and any three
  methods (group instruction, questioning, technology, discovery learning, learning center) into the above
  motivations, activities and assessments
• Discussion of how your unit plan draws on all five Alverno abilities
• Discussion of how your unit plan meets Wisconsin teacher/student standards
• Discussion of how your plan addresses elements in the cognitive, affective, and psychomotor domains
• List of all materials and resources in the unit including but not limited to: printed materials, audio-visuals,
  field trips, community resources, learning centers
• OPTIONAL-Description of activities at the end of the unit that provide opportunity for students to integrate
  the learning they have been doing throughout the unit

Part III due April 24, 2002
• Daily lesson plan for each of the days in the unit using the format of your choice. Address only the
  appropriate goal, objectives, procedures, assessment, and strategies for students who need additional help

Oral Presentation due May 1/May 6
• Present your unit plan focusing on the following components of your unit:
  1. Introductory motivating techniques that invite your students to take ownership for their learning
  2. Visual representation of the developmental aspects of your unit. You will speak from this map
     (inspiration) or powerpoint slides to explain to your audience (your colleagues) how student
     understandings, skills, and attitudes, will develop from the beginning to the conclusion of the unit. Be
     sure to explicitly address the two theories and three methods you chose to include in your unit
  3. How your unit meets established Wisconsin teacher/student standards. Although your unit will probably
     address many of these standards, select any two teacher and two student standards for the oral
     presentation
  4. Plan your presentation to fit a 15 minute time frame
  5. You will be videotaped and you will self-assess your videotaped presentation
  6. One example of an assessment instrument that you will use in the unit. Bring 10 copies of the instrument,
     one for the instructor and one for each member of your feedback group.
Performance Based Assessment in Teacher Preparation Using a Web-Based System

This example is from the University of Central Arkansas. The education unit at Central falls into the “large” size category for the project, graduating more than 200 candidates each year. Candidates are admitted at the beginning of the junior year.

The assessment is part of a comprehensive system that Central Arkansas faculty have developed as a response to all six NCATE unit standards. The intent, as stated in the supporting explanation for NCATE unit standard 2, is to examine

the (1) alignment of instruction and curriculum with professional, state, and institutional standards; (2) efficacy of courses, field experiences, and programs, and (3) candidates’ attainment of content knowledge and demonstration of teaching that leads to student learning.

The system also includes web-based tracking of the information. Candidates know the criteria by which their competence will be evaluated and the University seeks to attain direct evidence of proficiency through examinations, on-demand tasks, and longitudinal assessment of performance. Here is an excerpt from the candidate assessment component of the University’s description for its “Unit Assessment System for Standard 2:”

Level I: Admission to Teacher Education Program
GPA (2.5); writing sample; application; 45 semester hours; specific coursework (C or better); interview; Praxis I; 2 recommendations; committee recommendation

Level 2: Admission to Internship II
GPA (2.5 cumulative, 2.5 major); advisor recommendation; summative evaluation; Praxis II-Content test; Committee recommendation

Exit Requirements
GPA (2.5 cumulative, 2.5 major); summative evaluation; Praxis II—PLT; Praxis II-Content pedagogy; Grade of C or better in internship II; Baccalaureate degree requirements

Other Evidences of Effectiveness (Candidates & Programmatic)
Specialized program reviews; candidate survey (programmatic); candidate survey (dispositions); 1st-year teacher survey; mentor teacher survey; administrator survey; Praxis III

The Committee selected this example as an illustration of criterion 2; the scoring guides provide explicit statements of expectations for candidate proficiencies to be demonstrated. These expectations are adapted from Charlotte Danielson’s book, Enhancing Professional Practice: A Framework for Teaching (ASCD, 1996) and the ETS PATHWISE: Classroom Observation System (1995). The “components of professional practice” are displayed in four “domains” (planning and preparation, classroom environment, instruction, and professional responsibilities) that serve as the standards being assessed. Four levels of proficiency are described for each of the 25 criteria within these domains [unsatisfactory, basic (needs improvement), proficient (competent) and distinguished (accomplished)]. The excerpt on the next page displays two of these criterion and level descriptions: A3, Demonstrates knowledge of content by designing instruction that connects past, present, and future content; and C5, Uses instructional time effectively.

CONTACT: Jane H. McHaney, Dean, College of Education, jmchaney@mail.uca.edu
Last revised February 21, 2003
Example D—UNIVERSITY OF CENTRAL ARKANSAS

Performance Based Assessment in Teacher Preparation, University of Central Arkansas

COMPONENTS OF PROFESSIONAL PRACTICE: Domain A/Goal 1; Planning and Preparation

A3. Demonstrates knowledge of content by designing instruction that connects past, present, and future content

Knowledge of content
Relates current learning to past and future learning
Knowledge of prerequisite relationships

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Unsatisfactory – 1</th>
<th>Basic (Needs Improvement) - 2</th>
<th>Proficient (Competent) - 3</th>
<th>Distinguished (Accomplished) - 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>A3. Demonstrates knowledge of content by designing instruction that connects past, present, and future content</td>
<td>Candidate is unable to design lessons that connect current content to past and future learning. Candidate displays little understanding of the subject.</td>
<td>Candidate displays a simple understanding of content knowledge. Knowledge does not extend to its concepts, prerequisite relationships, connections with other disciplines, or to possible student misconceptions.</td>
<td>Candidate demonstrates solid understanding of the content concepts, prerequisite relationships and connection to other disciplines, but does not extend to possible student misconceptions. Instructional practices reflect current content related knowledge.</td>
<td>Candidate demonstrates an extensive knowledge of content and content-related practices, showing evidence of a continued search for improved practice. Candidate sees connections between prior and future learning and actively builds on concepts, knowledge of prerequisites, and misconceptions when planning instruction or seeking causes for student misunderstanding.</td>
</tr>
</tbody>
</table>

Goal 3/Domain C: Instruction

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Unsatisfactory – 1</th>
<th>Basic (Needs Improvement) - 2</th>
<th>Proficient (Competent) - 3</th>
<th>Distinguished (Accomplished) - 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>C5. Uses instructional time effectively</td>
<td>Structure and pacing of lesson are lacking. Established routines and classroom procedures are missing. Time is wasted with meaningless activities. Groups and transitions are not managed appropriately.</td>
<td>Structure and pacing of lesson are uneven. Routines and procedures are partially followed. Instructional time partially provides for meaningful activities. Transitions and groups are managed with some time wasted.</td>
<td>Structure and pacing of less are suitable. Routines and procedures are followed. Instructional time provides for meaningful activities. Transitions and groups are managed appropriately.</td>
<td>Structure and pacing of the lesson keep students on task. Routines and procedures are followed by students so that no time is wasted. Instructional time is used to provide maximum benefits. Transitions and groups are effectively managed to make maximum use of instructional time.</td>
</tr>
</tbody>
</table>

The Teacher Performance Outcomes Assessment Instrument Rubric serves as both a formative and a summative assessment instrument. It should be used to assess candidate progress in the classroom throughout the candidate’s program. The instrument should be used to assist candidates with growth and development throughout their teacher preparation program; it does not correlate with the A, B, C, D, F grading system. Candidates are expected to demonstrate competence at the Basic level during the early stages of their program. As candidates have more extensive experiences in classrooms, they are expected to move from Basic to Proficient. When candidates complete their final internship, they should demonstrate their teaching effectiveness at the Proficient level. As the candidates begin their teaching careers, they will set goals to achieve the Distinguished level.

NOTE: This assessment system is adapted from Danielson, C. (ASCD, 1996) Enhancing Professional Practice: A Framework for Teaching, and from PATHWISE: Classroom Observation System developed by ETS (1995).
Professional Education Program Admission I Requirements

This example is from the University of Wisconsin Oshkosh. The College of Education and Human Services at Oshkosh accepts candidates for admission to teacher preparation after 40 completed credits, usually in the third or fourth semester in college. The education unit falls into the “medium” size category for the project, with about 150 graduates annually.

The assessment is a portfolio of materials prepared by applicants for admission. It addresses, therefore, a part of the NCATE unit 2 standard, Assessment System and Unit Evaluation, which is described in the supporting explanation as follows:

*The unit uses multiple indicators (e.g., 3.0 GPA, demonstrated mastery of basic skills, general education knowledge, content mastery, and life and work experiences) to identify candidates with potential to become successful teachers or assume other school personnel roles at the point of candidate entry (as a freshman, junior, or post-baccalaureate student).*

The full portfolio of “Admission I Requirements,” not all included in the brief excerpts here, calls for measurement of multiple attributes that are assessed in diverse ways. The results are used collectively to identify individuals who could become talented teachers. Each candidate is asked to prepare a portfolio with prescribed materials in the following areas: eligibility, recommendations, grade point computation, evidence of understanding in the candidate’s subject major courses (see first following excerpt), a personal statement on commitment to becoming “an educator as caring intellectual,” an oral presentation of the portfolio (see second following excerpt), a reflective statement demonstrating knowledge of content related to the teaching profession, a self-evaluation, evidence of understanding issues related to culture and diversity, and a selection of “other” evidence identified by the candidate.

The Committee selected this example to highlight the scoring guides or rubrics descriptions (criterion 2) at three levels indicating what candidates are expected to demonstrate at each level. Note that there is a variety of forms of evidence that Oshkosh uses to inform “potential to become successful teachers,” two of them illustrated in the following pages, and each requirement or task in the portfolio is accompanied with appropriate scoring guides for that task. The Committee also identified the Oshkosh admissions portfolio as an illustration of “innovation,” due to its inclusion of disposition elements such as professional commitment and understanding of culture and diversity.

CONTACT: Michael P. Ford, Associate Dean, College of Education and Human Services, ford@uwosh.edu
Last revised February 21, 2003
In considering your admission in the Professional Education Program dedicated to developing Educators as Caring Intellectuals, the faculty and staff of the College of Education and Human Services will review evidence you submit within your portfolio to demonstrate knowledge of content in your field of study.

One piece of documentation which needs to be included in your portfolio is a “best sample” example of your work from a general education course specifically related to your future area of interest in teaching. Your general education course work sample and associated reflective statement will be evaluated on a three-point scale. This item will comprise a portion of your portfolio which will be evaluated to determine Admission I status. The following directions are to assist you in preparing that evidence for your portfolio:

1) Clearly identify your current cumulative grade point average in general education courses (exclude all PEP required courses).

2) Select at least ONE example of your work (paper, project, examination, presentation, etc) in a general education course.

Remember: The assignment you select should demonstrate your subject matter knowledge in a content discipline (e.g. an assignment that reflects your understanding of science, social studies, math, literature.) Avoid selecting an assignment related to the field of education (e.g. a research paper on school choice, a unit on teaching geometry, a review of “Among School Children”)

3) Prepare that example for your portfolio making it “viewer friendly” for an outside reviewer. This means that your piece of evidence must be presented in a clear manner that will be easily understood by a reviewer examining the evidence on his or her own without initial verbal explanations from you. Also remember this reviewer may not share your background in the area from which you selected your evidence.

4) Use the following guidelines to prepare a reflective statement to discuss that evidence. Keep in mind that your statement will provide evidence of your potential to be a reflective professional.

   a. Describe the evidence you have selected to include in your portfolio: Explain what the evidence is and why you chose this work sample; identify from which course it comes and the guidelines/expectations for the assignment in that course.
   
   b. Discuss how the evidence addresses your ability to become an educator who is a caring intellectual. What does this evidence say about your knowledge of the content area in which you would like to teach? What does this evidence reveal about your own learning techniques, strategies and styles?
   
   c. Discuss any other ways this evidence demonstrates your potential to be an educator who is a caring intellectual.

Rubric for Evidence of General Education experience

Evidence rated as THREE should encompass:

1. Candidates revealed significant reflective thoughts to illustrate an understanding of content of his or her field; learning techniques, strategies and styles; and an ability to be a lifelong learner.

2. Candidates completely presented evidence using an appropriate work sample and clearly responded to all directions.

Evidence rated as a TWO should encompass:

1. Candidates revealed some thinking to illustrate an understanding of content of his or her field; learning techniques, strategies and styles; and an ability to be a lifelong learner.

2. Candidates presented evidence in a somewhat complete manner using an appropriate work sample and responded to all directions.

Evidence as a ONE should encompass:

1. Candidates revealed limited thinking to illustrate an understanding of content of his or her field; learning techniques, strategies and styles; and an ability to be a lifelong learner.

2. Candidates presented evidence in an unclear manner using an inappropriate work sample and/or not responding to all directions.
Directions to Students for Preparing for Admissions I Portfolio Presentation

The purpose of the portfolio presentation is to explain clearly how you are becoming a caring, intellectual educator, which is the kind of teacher the faculty and staff in the Professional Education Program are committed to developing. Please dress appropriate for a professional interview. You must be able not only to explain your ideas, but also point to evidence from your portfolio to illustrate your ideas.

You will have a brief meeting (15 minutes) with a faculty member in which you will explain:

1. What do you think it means to be a caring intellectual? Why do you think this is the kind of teacher needed for schools today?
2. How do your PPST scores, GPA from general education courses, Fundamentals of Speech grade, personal statement, letters of recommendation, Human Relations Plan and Log, and the evidence and reflection from a general education course reflect the progress you are making in becoming a caring intellectual?
3. What examples from your portfolio illustrate your knowledge of content, commitment to lifelong learning, understanding of diversity, and ability to be a reflective professional?

During your presentation you will be evaluated on your ability to:

1. Demonstrate you have thought carefully about the meaning of a caring intellectual and can cite evidence from your portfolio to support your statements.
2. Demonstrate you have thought carefully about and can cite evidence from your portfolio to illustrate how your knowledge of content, a commitment to lifelong learning, an understanding of diversity, and an ability to be a reflective professional contribute to your development as a caring intellectual.
3. Explain your ideas clearly and completely.
4. Show appropriate poise and demeanor during the presentation.

RUBRIC FOR Oral presentation of Portfolio (Faculty Team Interview)

<table>
<thead>
<tr>
<th>Question</th>
<th>1 point rating</th>
<th>2 point rating</th>
<th>3 point rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Understanding of what it means to be a caring intellectual</td>
<td>Revealed little thinking, unable to explain what caring means or intellectual means or why either might be critical in today’s schools</td>
<td>Revealed some thinking either about what caring or intellectual means and why either might be critical in today’s schools</td>
<td>Revealed significant reflective thoughts about what caring means, intellectual means and why both are critical in today’s schools</td>
</tr>
<tr>
<td>Understanding of self as learner, content knowledge and lifelong learner</td>
<td>Revealed little thinking about self as learner, subject matter knowledge and lifelong learning</td>
<td>Revealed some thinking about self as a learner, subject matter knowledge or lifelong learning</td>
<td>Revealed significant reflective thought about self as learner, subject matter knowledge base and a passion for learning</td>
</tr>
<tr>
<td>Understanding of issues related to culture and diversity</td>
<td>Revealed few insightful thoughts or inappropriate thoughts; unable to discuss plans and efforts in addressing human relations code requirements</td>
<td>Revealed some insightful thoughts but was less able to discuss plans and efforts in addressing the human relations code requirements</td>
<td>Revealed significant insightful thoughts in discussing plans and efforts in addressing the human relations code requirements</td>
</tr>
<tr>
<td>Ability to give examples from the portfolio to illustrate understanding of content, lifelong learning, diversity and culture</td>
<td>Unclear in explaining ideas and using examples from portfolio</td>
<td>Somewhat clear in explaining all ideas using examples from the portfolio or very clearly explained only some of the ideas with examples</td>
<td>Clearly and completely explained all ideas using examples from the portfolio</td>
</tr>
<tr>
<td>Ability to explain ideas clearly and completely</td>
<td>Unclear in responding to questions regarding portfolio and conceptual model</td>
<td>Was able to respond clearly to some questions regarding the portfolio and conceptual model</td>
<td>Clearly and completely responded to questions regarding the portfolio and conceptual framework</td>
</tr>
<tr>
<td>Ability to demonstrate appropriate poise and demeanor during the interview</td>
<td>Demonstrated little or no poise and inappropriate demeanor during the presentation</td>
<td>Demonstrated adequate poise and demeanor during presentation</td>
<td>Demonstrated appropriate poise and demeanor during presentation</td>
</tr>
<tr>
<td>Overall rating of portfolio on a six point scale</td>
<td>1 mixture of 1 and 2’s majority 1’s</td>
<td>3 all 2 ratings</td>
<td>5 mixture of 2 and 3’s majority 3’s</td>
</tr>
<tr>
<td></td>
<td>2 mixture of 1 and 2’s majority 1’s</td>
<td>4 mixture of 2 and 3’s majority 2’s</td>
<td>6 all 3 ratings</td>
</tr>
</tbody>
</table>
Impacting Student Learning (ISL)

This example is from Augusta State University, Georgia. The Augusta State education unit falls into the “medium” range of 50 to 200 annual graduates. Candidates are typically admitted at the beginning of the junior year or as graduate students.

The principal feature of this assessment is a common scoring guide. The guide is used by individual faculty members who determine the appropriate outcome targets, graded tasks and performances (and weighted values) for each content pedagogy course or apprenticeship. The Augusta State standards, and related faculty assessments, are constructed around the Georgia Quality Core Curriculum. This curriculum is intended to develop candidate abilities “to organize, implement and assess instruction and related P-12 student achievement.” The faculty assessments are “employed across all programs to monitor candidate proficiencies and to inform program effectiveness” for NCATE standard 1 on candidate knowledge, skills, and dispositions, and standard 2 on assessment systems and evaluation. All candidates in initial teacher preparation complete three or four work samples over the course of that program. The results are used by Augusta State to analyze individual candidate performance, assess candidate readiness at transition points, and evaluate candidates’ accomplishments during clinical experiences. They are also used for evaluation of programs in the unit. Faculty have attempted to evaluate the accuracy and fairness of these assessments.

Here are some excerpts from the submission:

While individual faculty members maintain the prerogative of determining appropriate outcome targets, graded tasks and performances (and their weighted values) for each course, a common scoring scale has been adopted. The INTASC standards are employed across programs to identify primary goals within each course and for lab assessments. . . . An electronic data-base for recording candidate course, ISL, and lab performance scores has been created.

Within each program candidates complete a minimum of three Impacting Student Learning (ISL) components. . . . University and Professional Development School (PDS) faculty collaboratively determined where in each program the ISLs would be embedded. The ISL components are planned within courses and implemented in PDS-based labs, which take place during a designated five-week period of the semester. . . . During lab weeks class meetings on campus are suspended; faculty and candidates are in the PDSs during this time.

The Committee selected ISL as an example of criterion 2 on explicit scoring expectations, and in this case the key feature is that the scoring scale is common even though individual faculty members create their own outcomes, tasks and performances. The following excerpt from Augusta includes these “scoring rubric criteria.” The assessment tasks that candidates are asked to perform also illustrate criterion 5 on “authenticity.” In addition, the Committee selected the example as an illustration of innovation, for its inclusion of P-12 student learning focus in the scoring guides.

CONTACT: Mary Gendernalik Cooper, Professor and Chair, coopermg@aol.com
Last revised February 21, 2003
Example F—AUGUSTA STATE UNIVERSITY
DEPARTMENT OF TEACHER DEVELOPMENT
IMPACTING STUDENT LEARNING (ISL) COMPONENTS BY PROGRAM BY LEVEL OF FOCUS
All ISLs are embedded in Content Pedagogy Courses or culminating Apprenticeship. Labs comprise five dedicated weeks of each fifteen-week semester

<table>
<thead>
<tr>
<th>Level of Focus → Certification Program ↓</th>
<th>Individual/Tutoring</th>
<th>Whole Class Instruction, 4-7 Student Focus</th>
<th>Whole Class Instruction, Whole Class Focus</th>
<th>Varied Levels of Focus, Multiple Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early childhood Program (PK-5)</td>
<td>Semester I: Math content focus, 1-3 students, 3 weeks.</td>
<td>Semester II: Science content focus, 4-7 students’ performance assessed, 2-4 week unit</td>
<td>Semester III: Social Studies, whole class, 3-4 week unit; &amp; Reading diagnosis/remediation, 2-3 students, 4 weeks</td>
<td>Semester IV: Candidate completes multiple units across varied subject areas, fully integrated into classroom curriculum, jointly determined with master teacher</td>
</tr>
</tbody>
</table>

(Nota: the full example has similar descriptions for the middle grades program and for secondary subjects; it also includes an itemized listing of all elements of the ISL components.)

Impacting Student Learning--Scoring Rubric Criteria

4 = exemplary performance on all components of ISL task – goals are well-stated and appropriate for grade level/subject and students; learning profiles are well-developed, complete, and accurate; diagnosis of student needs is based on data; instructional practices and strategies are active, student-centered, and appropriate to diagnosed needs and the task; assessments are appropriate for the type of learning being assessed; reflections reveal reasoned and supported analysis of lessons, student achievement, and future needs; there is explicit evidence of understanding of students as learners; writing and/or speaking are clear, strong, and thoughtful; references and examples are integrated into the analysis; consistently correct usage of standard English, consistent and accurate use of APA style.

3 = proficient performance on all components of ISL task – goals are appropriate for grade-level; learning profiles give adequate but not complete insights; diagnosis of student needs is based on accurate but insufficient data; instructional practices and strategies are subject-centered and appropriate to the task; assessments address the learning task(s); reflections reveal well-supported analysis of lessons and student achievement; there is some evidence of understanding of students as learners; writing and/or speaking are clear and coherent; some evidence of integration of references and examples; consistently correct usage of standard English, consistent and accurate use of APA style.

2 = efforts to address all components of ISL task are in progress; performance on all dimensions of task is not completely accurate or coherent – goals are related to content; learning profiles are insufficient for accurate planning; planning is superficial and lacks thoughtfulness; instructional practices and strategies are subject- or teacher-centered and don’t fit well with objectives; reflections and analysis are descriptive and superficial rather than reasoned and supported; references are dropped-in rather than integrated; inconsistencies in correct usage of standard English, inconsistent and inaccurate use of APA style.

1 = performance is unsatisfactory; not all areas are addressed; performance is inaccurate, incoherent, lacks clarity – goals are incomplete or poorly stated or inappropriate; profiles are incomplete and/or inaccurate; planning is minimal; instructional practices and strategies are not thoughtful or well-planned; reflections and analysis are not at all evident; faulty reasoning; consistently incorrect usage of standard English, consistently inaccurate use of APA style.

Department of Teacher Development, Impacting Student Learning
Elements of the ISL Component

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Developmental characteristics; Background and experience; School, classroom, community contexts; Learning styles, abilities, needs; Interviews of students, teachers, others</td>
<td>Content understanding. Type of learning task: knowledge, concept, skill, application, theory, disposition; Thinking processes; Outcomes, goals and objectives</td>
<td>Developmentally appropriate; Multiple paths to learning/diversity; Cognitively/actively engaging; Performance-based assessments</td>
<td></td>
</tr>
<tr>
<td>Assessment of Student Learning: Tools; Processes; Pre &amp; post; Formative &amp; summative; Authentic; Connections to standards</td>
<td>Evaluation of Student Learning: Artifacts; Analysis; Explanation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation of Student Learning: Artifacts; Analysis; Explanation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Both include: Implications for further learning; refinements/revisions needed; action plan/next steps.
Teacher Work Sample: The Renaissance Partnership for Improving Teacher Quality

This example is from the eleven Renaissance Partnership Project sites: California State University at Fresno; Eastern Michigan University; Emporia State University, Kansas; Idaho State University; Kentucky State University; Longwood University, Virginia; Middle Tennessee State University; Millersville University, Pennsylvania; Southeast Missouri State University; University of Northern Iowa; and Western Kentucky University. Individually, the annual graduates from educator preparation units at these institutions range from 50 per year to 1500 or more per year. Currently at least 1000 candidates complete the teacher work samples each semester.

The idea of “teacher work samples” has been in development and use for many years at Western Oregon University. It follows a concept of performance assessment designed to “provide credible evidence” of teacher candidate “ability to facilitate learning” with P-12 students. The teaching skill areas (or standards) for documentation are similar to those used in NBPTS assessments and in INTASC portfolios, but adapted for use in pre-service teacher education. They include, for example: using information about learning and teaching to plan instruction; setting challenging goals; using multiple assessment modes to assess student learning; meet specific student needs; reflecting on instruction and how to improve teaching. In the Renaissance materials, each teaching process is elaborated with a full statement of a standard, in addition to “indicators,” tasks, prompts, and rubrics that define various levels of performance for each standard. The eleven institution project, funded by Federal Title II grants, has developed and evaluated the assessments in terms of accuracy, consistency of results, fairness (and opportunity to learn), avoidance of bias, generalizability, and scorer reliability.

In the excerpts that follow, the “TWS standards” are copied first, followed by excerpts from the assessment “prompts” and “rubrics” for two of the standards: contextual factors and analysis of student learning. The contextual factors excerpt also illustrates “scoring guides” provided for reviewers. The Committee selected the Teacher Work Sample to illustrate criterion 3 (on making clear distinctions among levels of proficiency and establishing consistent scoring processes with trained reviewers); criterion 5 (on including some “authentic” forms of assessment, together with assessment of the teacher candidate’s P-12 students); and “innovation,” (as an illustration of application of work sample methodology to pre-service preparation assessment).

CONTACT: Roger Pankratz, University of Western Kentucky, roger.pankratz@wku.edu; Project materials are available at the Renaissance Partnership for Improving Teacher Quality website, http://fp.uni.edu/itq. Click on Teacher Work Samples or Materials for prompts, rubrics and exemplars. There are more than 20 teacher work sample exemplars available at the web site.

Last revised March 7, 2003
Example G—RENAISSANCE TEACHER WORK SAMPLE

First example

Assessment for the Contextual Factors example

TWS Standard
The teacher uses information about the learning-teaching context and student individual differences to set learning goals, plan instruction and assessment.

Task
Discuss relevant factors and how they may affect the teaching-learning process. Include any supports and challenges that affect instruction and student learning.

Prompt
In your discussion, include:

- **Community, district, and school factors.** Address geographic location, community and school population, socio-economic profile and race/ethnicity. You might also address such things as stability of community, political climate, community support for education and other environmental factors.
- **Classroom factors.** Address physical features, availability of technology equipment and resources and the extent of parental involvement. You might also discuss other relevant factors such as classroom rules and routines, grouping patterns, scheduling and classroom arrangement.
- **Student characteristics.** Address student characteristics you must consider as you design instruction and assess learning. Include factors such as age, gender, race/ethnicity, special needs, achievement/developmental levels, culture, language, interests, learning styles/modalities or students’ skill levels. In your narrative, make sure you address students’ skills and prior learning that may influence the development of your learning goals, instruction and assessment.
- **Instructional implications.** Address how contextual characteristics of the community, classroom and students have implications for instructional planning and assessment. Include specific instructional implications for at least two characteristics and any other factors that will influence how you plan and implement your unit.

Suggested Page Length: 1-2

Rubric for the Contextual Factors example

<table>
<thead>
<tr>
<th>Rating</th>
<th>Indicator ↓</th>
<th>1 Indicator Not met</th>
<th>2 Indicator Partially Met</th>
<th>3 Indicator Met</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge of Community, School and Classroom Factors</td>
<td>Teacher displays minimal, irrelevant or biased knowledge of the characteristics of the community, school and classroom.</td>
<td>Teacher displays some knowledge of the characteristics of the community, school and classroom that may affect learning.</td>
<td>Teacher displays a comprehensive understanding of the characteristics of the community, school and classroom that may affect learning.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Scoring Guide for the Contextual Factors example

Sources of Evidence

- Contextual Factors; Assessment Plan; Design for Instruction; Instructional Decision Making

Assessment Dimensions and Questions

1. **Knowledge of Community, School, and Classroom Factors**—Does the teacher display comprehensive understanding of the characteristics of the community, school, and classroom that are relevant to the learning goals?
2. **Knowledge of Characteristics of Students**—Does the teacher display comprehensive knowledge of student differences that are relevant to the learning goals?
3. **Knowledge of Students’ Varied Approaches to Learning**—Does the teacher display comprehensive knowledge of the different ways students learn that are relevant to the learning goals?
4. **Knowledge of Students’ Skills and Prior Learning**—Does the teacher display comprehensive knowledge of students’ skills and prior knowledge relevant to the learning goals?
5. Implications for Instructional Planning and Assessment—Does the teacher provide appropriate implications for instruction and assessment based on student individual differences and community, school, and classroom characteristics?

General Considerations

- Must the teacher provide an implication for every contextual factor described? No.
- What are appropriate implications? Those things that will affect planning for instructional assessment.

Second example

Assessment for the Analysis of Student Learning example

Analysis of Student Learning

TWS Standard
The teacher uses assessment data to profile student learning and communicate information about student progress and achievement.

Task
Analyze your assessment data, including pre/post assessments and formative assessments to determine students’ progress related to the unit learning goals. Use visual representations and narrative to communicate the performance of the whole class, subgroups and two individual students. Conclusions drawn from this analysis should be provided in the “Reflection and Self-Evaluation” section.

Prompt
In this section, you will analyze data to explain progress and achievement toward learning goals demonstrated by your whole class, subgroups of students and individual students.

- Whole class. To analyze the progress of your whole class, create a table that shows pre- and post-assessment data on every student on every learning goal. Then, create a graphic summary that shows the extent to which your students made progress (from pre- to post-) toward the learning criterion that you identified for each learning goal (identified in your Assessment Plan section). Summarize what the graph tells you about your students’ learning in this unit (i.e., the number of students who met the criterion).
- Subgroups. Select a group characteristic (e.g., gender, performance level, socio-economic status, language proficiency) to analyze in terms of one learning goal. Provide a rationale for your selection of this characteristic to form subgroups (e.g., girls vs. boys; high- vs. middle- vs. low-performers). Create a graphic representation that compares pre- and post-assessment results for the subgroups on this learning goal. Summarize what these data show about student learning.
- Individuals. Select two students who demonstrated different levels of performance. Explain why it is important to understand the learning of these particular students. Use pre-formative, and post-assessment data with examples of the students’ work to draw conclusions about the extent to which these students attained the two learning goals. Graphic representations are not necessary for this subsection.

Note: You will provide possible reasons for why your students learned (or did not learn) in the next section, “Reflection and Self-Evaluation.”

Suggested Page Length: 4 + charts and student work examples

Rubric for the Analysis of Student Learning example

TWS Standard: The teacher uses assessment data to profile student learning and communicate information about student progress and achievement.

<table>
<thead>
<tr>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indicator</td>
</tr>
<tr>
<td>Clarity and Accuracy of Presentation</td>
</tr>
<tr>
<td>Evidence of Impact on Student Learning</td>
</tr>
</tbody>
</table>

[Note: criteria for “Alignment with Learning Goals” and “Interpretation of Data” are omitted from this excerpt]
Model of Appropriate Practice (MAP) for Teacher Candidates

This example is from Columbus State University in Georgia. Columbus State teacher education graduates fall into the “medium” size category of 50 to 200, and 72 of these graduates completed programs in which the Model of Appropriate Practice (MAP) was used last year. Candidates are typically admitted to the education unit at the beginning of the junior year.

The Model of Appropriate Practice is constructed around the Georgia Quality Core Curriculum—which, in turn, draws on INTASC principles as primary goals across courses and lab experiences. The structure and the scoring guides are similar to those in the ETS Pathwise program, as illustrated in this excerpt:

<table>
<thead>
<tr>
<th>DOMAINS IN THE COLUMBUS MODEL OF APPROPRIATE PRACTICE (MAP)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DOMAIN 1: PLANNING AND PREPARATION</strong></td>
</tr>
<tr>
<td>1a: Demonstrating Knowledge of Content and Pedagogy</td>
</tr>
<tr>
<td>Knowledge of content</td>
</tr>
<tr>
<td>Knowledge of prerequisite relationships</td>
</tr>
<tr>
<td>Knowledge of content-related pedagogy</td>
</tr>
<tr>
<td>1b: Demonstrating Knowledge of Students</td>
</tr>
<tr>
<td>Knowledge of characteristics of age group</td>
</tr>
<tr>
<td>Knowledge of students’ varied approaches to learning</td>
</tr>
<tr>
<td>Knowledge of students’ skills and knowledge</td>
</tr>
<tr>
<td>Knowledge of students’ interests and cultural heritage</td>
</tr>
<tr>
<td>1c: Selecting Instructional Goals</td>
</tr>
<tr>
<td>Value; Clarity; Suitability for diverse students; Balance</td>
</tr>
<tr>
<td>1d: Demonstrating Knowledge of Resources</td>
</tr>
<tr>
<td>Resources for teachers</td>
</tr>
<tr>
<td>Resources for students</td>
</tr>
<tr>
<td>1e: Designing Coherent Instruction</td>
</tr>
<tr>
<td>Learning activities; Instructional materials and resources;</td>
</tr>
<tr>
<td>Instructional groups; Lesson and unit structure</td>
</tr>
<tr>
<td>1f: Assessing Student Learning</td>
</tr>
<tr>
<td>Congruence with instructional goals</td>
</tr>
<tr>
<td>Criteria and standards</td>
</tr>
<tr>
<td>Use for Planning</td>
</tr>
<tr>
<td><strong>DOMAIN 2: THE CLASSROOM ENVIRONMENT</strong></td>
</tr>
<tr>
<td>2a: Creating an Environment of Respect and Rapport</td>
</tr>
<tr>
<td>Teacher interaction with students</td>
</tr>
<tr>
<td>Student interaction</td>
</tr>
<tr>
<td>2b: Establishing a Culture for Learning</td>
</tr>
<tr>
<td>Importance of the content</td>
</tr>
<tr>
<td>Student pride in work</td>
</tr>
<tr>
<td><strong>DOMAIN 3: INSTRUCTION</strong></td>
</tr>
<tr>
<td>3a: Communicating Clearly and Accurately</td>
</tr>
<tr>
<td>Communicating learning objectives</td>
</tr>
<tr>
<td>Directions and procedures</td>
</tr>
<tr>
<td>Oral and written language</td>
</tr>
<tr>
<td>3b: Using Questioning and Discussion Techniques</td>
</tr>
<tr>
<td>Quality of questions</td>
</tr>
<tr>
<td><strong>DOMAIN 4: PROFESSIONAL RESPONSIBILITIES</strong></td>
</tr>
<tr>
<td>4a: Reflecting on Teaching</td>
</tr>
<tr>
<td>Accuracy</td>
</tr>
<tr>
<td>Use in future teaching</td>
</tr>
<tr>
<td>4b: Maintaining Accurate Records</td>
</tr>
<tr>
<td>Student completion of assignments</td>
</tr>
</tbody>
</table>

The system was first used in academic year 2001-2002 and the assessments and rubrics are currently being evaluated. Drawing on course work, practica, and student teaching, candidates construct a web based folio of their products to show competence in the four domains of the MAP. The unique part of this example, on the second following page, is a “feedback” critique from the faculty to the candidate. The Committee viewed this example as illustrative of an “authentic” form of assessment, criterion 5, in which candidates were assigned tasks that might be similar to ones they would encounter in real teaching situations. The example was also perceived as innovative in its linking of INTASC standards with frameworks from the ETS Pathwise materials, and in its built-in remediation feature.

CONTACTS: Deborah Gober, gober_deborah@colstate.edu and Causey_Virginia@colstate.edu
Last revised April 8, 2003
Example H—COLUMBUS STATE UNIVERSITY

TASKS THE CANDIDATE IS ASKED TO COMPLETE

EDSE 4245 Teaching Social Studies in Grades 7-12
UNIT PLAN—Due Date: December 2
Each teacher candidate will develop a unit of six to eight consecutive lessons on a single topic or theme in the chosen social studies subject. A “lesson” is one class period. Each lesson must include the instructional support materials and student assessment instruments used to teach the lesson and evaluate student performance of objectives. The Unit Plan should be word-processed using software that produces an appropriate product. Correct English grammar and usage are required. Place the unit in a folder or three-ring binder. Please do not place the pages in plastic sleeves.

Requirements:
- The unit must have been taught during the practicum.
- At least three different lessons must illustrate three of the following: …
  - decision-making/problem-solving;
  - response groups/inquiry;
  - interdisciplinary approaches;
  - skills builder;
  - simulation/experiential exercises.
- At least one lesson must infuse instructional technology.
- At least one lesson must have diversity as its focus. (See attached definitions.)
- The last lesson in the unit must be a summative assessment.

[NOTE: The full description of requirements provides further elaboration of the summary excerpted above together with definitions of terms.]

SCORING GUIDELINES FOR THE COLUMBUS “MAP” DOMAINS

DOMAIN 1: PLANNING AND PREPARATION
Component 1A: Demonstrating Knowledge of Content and Pedagogy

<table>
<thead>
<tr>
<th>Component</th>
<th>Below Expectations</th>
<th>Meets Expectations</th>
<th>Exceeds Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Knowledge of Content</strong></td>
<td>Teacher Candidate makes significant content errors or does not correct content errors students make.</td>
<td>Teacher Candidate displays basic content knowledge in area of expertise, but cannot articulate connections with other parts of the discipline or with other disciplines.</td>
<td>Teacher Candidate displays solid content knowledge and makes connections between the content and other parts of the discipline and other disciplines.</td>
</tr>
<tr>
<td><strong>Knowledge of Prerequisite Relationships</strong></td>
<td>Teacher Candidate displays little understanding of prerequisite knowledge important for student learning of the content.</td>
<td>Teacher Candidate indicates some awareness of prerequisite learning, although such knowledge may be limited.</td>
<td>Teacher Candidate’s plans and practices reflect understanding of prerequisite relationships among topics and concepts.</td>
</tr>
<tr>
<td><strong>Knowledge of Content-Related Pedagogy</strong></td>
<td>Teacher Candidate displays little understanding of pedagogical issues involved in student learning of the content.</td>
<td>Teacher Candidate displays basic pedagogical knowledge but does not anticipate possible student misconceptions.</td>
<td>Teacher Candidate’s instructional practices reflect current research on best pedagogical practice within the discipline.</td>
</tr>
</tbody>
</table>

[NOTE: The full submission contains descriptions of three levels of performance on 89 “components” or “elements” of the four domains]

SPONSOR'S DESCRIPTION OF THE COLUMBUS CYCLE OF CANDIDATE OBSERVATIONS AND FEEDBACK

- The candidate planned and taught a unit in a middle grades social studies classroom. I observed one lesson in person and one on videotape. The cooperating teacher observed all lessons.
- Based on my observations and the cooperating teacher's evaluation of the candidate, I concluded the candidate had not met expectations for all MAP components, a College of Education requirement for entering student teaching. I generated the feedback . . . and the professional development plan . . . [NOTE: see the excerpt that follows]. The candidate received these simultaneously.
- The candidate created and submitted the new unit. It still did not meet expectations for all components. He then received additional feedback, including a second professional development plan.
- The candidate submitted his second revision. He still did not meet expectations on all components. Because we are working on a mastery basis, the candidate had a third opportunity and at that point he finally met expectations for all components.
SUMMATIVE FEEDBACK REPORT FROM UNIVERSITY SUPERVISOR

This is feedback to clarify the “Final Evaluation” of your teaching in EDCI 4455. This summative feedback is based on the university supervisor’s observation of two lessons (one in person, one videotape), your cooperating teacher’s assessment, your unit plan, and your practicum portfolio. In order to exit EDCI 4455 and enter student teaching, each teacher candidate must demonstrate teaching behaviors that result in being assessed on the “Final Evaluation” at the “Meets Expectations” level on all relevant components. In the Practicum, Component 4c is not relevant. All other components apply to the Practicum. You have not met expectations on all components. Please review comments below, as well as specific feedback in your portfolio and unit plan.

DOMAIN 1 – PLANNING AND PREPARATION - _points/ 15 possible

**Demonstrating Knowledge of Content and Pedagogy – Below Expectations**

You made some content errors or did not correct content errors students made. The content included in lessons was jumbled in chronology and lesson content did not build from one day to the next. For example, you taught a lesson on the Silk Road of 300 B.C. after lessons on modern East Asia. Also, you displayed little understanding of prerequisite knowledge important for student learning of the content. Students need a lot of background on each of the countries of East Asia. Your lesson content was often superficial, such as lesson #5 in which the students basically learn that the Himalayas are the world’s highest mountains and India has two main rivers. Additionally, you displayed little understanding of pedagogical issues involved in student learning of the content. Each lesson followed the same teacher-centered, passive learning pattern for students. You did not implement a single History Alive strategy.

**Demonstrating Knowledge of Students – Meets Expectations**

You displayed knowledge of the interests and cultural heritage of students, and you used this knowledge to enrich your lessons. You displayed generally accurate knowledge of the developmental characteristics of the age group. You demonstrated the value of understanding students’ skills and knowledge; you displayed this knowledge for the class as a whole. However, you did not address different approaches to learning that students exhibit, such as multiple intelligences. For example, throughout your unit, you address only linguistic, visual, and intrapersonal intelligences.

**Selecting Instructional Goals – Below Expectations**

Your objectives were valuable and consistent with the QCCs, but students rarely performed them in the lessons. For example, in lesson #3, your objective is for students to “trace the cultural links between China, Korea, and Japan.” Yet the lesson has students do a round-robin reading from their text. Further, your goals either were not clear or were stated as student activities.

**Demonstrating Knowledge of Resources – Below Expectations**

You did not use resources available through the school, your course packet and text, professional organizations, the . . .

**Designing Coherent Instruction -= Below Expectations**

Your learning activities were not suitable to students or instructional goals. Though goals were often valuable, students did not often actually perform them. Lessons did not follow an organized progression and did not reflect recent . . .

**Assessing Student Learning – Below Expectations**

Your content and methods of assessment lacked congruence with your instructional goals. The proposed approaches contained no clear criteria or standards. In many lessons, there are not assessment strategies for objectives. . .

DOMAIN 2 – THE CLASSROOM ENVIRONMENT _points/ 15 possible etc. etc

To “meet Expectations” for components 1a, 1c, 1d, 1e, 1f, 2b, 3c, 3d, 4a and 4e of the Model of Appropriate Practice, please complete and submit the following:

A. Create a new unit plan on “China” following the guidelines on pp. 13-15 in your course packet. Be sure to:
   > Write objectives for each lesson so they are measurable and have clear cognitive outcomes
   > At least one objective in each lesson must be at the application level or higher on Bloom’s taxonomy.
   > Include at least one QCC content standard and at least one QCC skills standard in each lesson
   > Include at least one affective objective in each lesson.
   > Create instructional strategies where students clearly perform each of your objectives. At least three lessons must apply three of the effective teaching strategies listed in the Unit Plan guidelines.
   > At least one lesson should focus on “diversity.”
   > You should address all intelligence s in the 6-8 lessons.
   > Write the evaluation section of each lesson plan in this way: “Objective 1 will be assessed by ______” and list the specific student performances you will assess. You must collect written data on student performance to provide evidence for each objective. Establish clear criteria for how you will grade student performances.
   > The 6-8 lessons in the unit should build coherently on content and student skills.

I strongly recommend you share a unit outline with me for feedback before creating the lessons.

B. Complete “Planning Questions” for each lesson in your new unit. (See p. 419 in your course packet
About Example I

**Action Research Project Using NCTE Standards**

This example is from the University of Charleston, West Virginia. Charleston is a small private liberal arts college enrolling approximately 800 students, and education is a “program” under the Department of Social Sciences. About 65 candidates are enrolled in the education program and 17 completed during the last year. Candidates are admitted at the end of their sophomore year, or after attaining at least 60 credit hours.

The assessment is in the form of an action research project. All senior teacher candidates enrolled in the student teaching practicum must complete the assessment as part of the course requirement. The assessment addresses the following NCTE program standards:

3.2 The program prepares the candidate in the practices of oral, visual, and written literacy; as a result the candidate will:

- 3.2.3 use the processes of composing to create various forms of oral, visual, and written literacy;
- 3.2.4 use writing, visual images, and speaking for a variety of purposes and audiences

3.6 The program prepares the candidate with knowledge and understanding of the range and influence of print and nonprint media and technology in contemporary culture; as a result, the candidate will:

- 3.6.3 display an understanding of the role of technology in communication

3.7 The program prepares the candidate with knowledge and understanding of research theory and findings in English language arts; as a result, the candidate will:

- 3.7.1 use major sources of research and theory (i.e., books, periodicals, reports, proceedings of professional conferences, videotapes, electronic and non-electronic data bases) to understand the relationship between research and practice
- 3.7.2 use teacher-researcher models of classroom inquiry

4.0 The program enables the candidate to acquire and demonstrate the dispositions and capacities needed to integrate knowledge of English language arts, students, teaching, and practice; as a result, the candidate will:

- 4.12.2 employ a variety of means to interpret and report assessment methods and results to students, administrators, parents, and other audiences.

The full submission from which the excerpt on the next page is taken includes instructions for candidates in the English Language Arts program describing “action research” as “systematic, intentional inquiry by teachers about their own school and classroom work.” The instructions suggest hypothetical questions that teachers might ask; encourage candidates to read literature on the topic; then detail “observation,” “interviews,” “surveys/questionnaires,” and examination of “artifacts/documents/materials” as potential data collection methods to address questions. Additional parts of the instructions address informed consent permissions, and analysis of data.

The excerpt that follows contains the assessment prompt and three of the seven scoring guides. The committee selected this example to illustrate explicit faculty expectations, criterion 2, authenticity of candidate work, criterion 5, effectiveness/efficiency in gathering useful assessment information, and as an innovative way to examine content knowledge.

CONTACT: Cara Turner, cturner@ucwv.edu

Last revised March 12, 2003
Example I—UNIVERSITY OF CHARLESTON, WV

What Should Be Included in Your Research Report?
Your research report will not be a paper but rather a Power Point presentation similar to one you might give at a professional conference. It should include the following information:

1. How you came to this study/topic
2. Purpose of the study or research question(s)
3. Summary of educational literature on your topic (what other research says about your topic)—(minimum of 10 research reports (no more than 5 from the Internet and 5 must be from peer reviewed professional journals).
4. Context of study (type of school, students grade level, a little about the community, etc. keeping in mind that you will need to keep the location and participants anonymous)
5. Participants in study (the number and any other pertinent information)
6. Data collection (the methods you used to collect your data, including some samples)
7. Analysis (how you analyzed data, including some samples)
8. Results (the answers to your question or questions; this can be done in charts or graphs or tables or by using direct quotations from participants, or by using bullets and summaries)
9. Conclusions (what you as the teacher researcher conclude about your study; what the implications are for teachers and teaching; how this study relates to the professional literature)
10. What you would do differently

How Will You Be Assessed?
You will receive a grade for this project based on the following:
Research skills and data collection; Ability to synthesize data and results; Ability to relate results to professional literature; Ability to draw conclusions and make implications for practice; Presentation skills

You will be graded holistically using the following rubric:

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Target=A</th>
<th>Acceptable=B/C</th>
<th>Unacceptable=Below C</th>
</tr>
</thead>
<tbody>
<tr>
<td>To apply research skills</td>
<td>Data collection yields thorough information that fully answers research question(s); data gathering methods (observation, surveys, interviewing, etc.) are a natural match for answering research question(s); data is informative and insightful</td>
<td>Data collection yields detailed, accurate information that effectively answers research question; data gathering methods are appropriate for answering research question(s); information is interesting</td>
<td>Data collection yields superficial, mechanistic, and/or fragmentary information that presents generalizations or an incomplete answer to the research question; data gathering methods are inappropriate or ill-matched for answering research question(s)</td>
</tr>
<tr>
<td>To synthesize and analyze data</td>
<td>Detailed analysis of the meaning and significance of the interview data; synthesis is concise, yet provides rich context</td>
<td>Helpful or plausible analysis of meaning and significance; synthesis makes sense and provides context</td>
<td>Analysis is descriptive, not interpretive; synthesis is sketchy and provides no content</td>
</tr>
<tr>
<td>To present information in a logical and professional manner</td>
<td>Presentation lasts between 15 and 25 minutes; incorporates technology, specifically a PowerPoint slideshow presentation that significantly goes beyond the minimum of 6 slides; Class members are actively engaged; Includes at least one handout for the class with information (e.g., notes from slideshow), and the handout must include a complete bibliography; the handout is effective, neat, well-written and educational sound; [NOTE: remainder omitted]</td>
<td>Presentation lasts a minimum of 15 minutes; incorporates technology, specifically a PowerPoint slideshow presentation (a minimum of 6 slides); The presentation is engaged most of the time; Includes at least one handout for the class with information (e.g., notes from slideshow), and the handout must include a complete bibliography; Presents him/herself as a professional researcher presenting at a conference (stand up straight, no whispering/talking when not presenting [NOTE: remainder omitted.</td>
<td>Presentation fails to meet the 15 minute minimum; fails to incorporate technology; Class members clearly bored and inattentive; Does not include at least one handout for the class with information (e.g., notes from slideshow), and/or the handout does not include a complete bibliography, and/or the handout is poorly written or messy; Fails to present him/herself as a professional researcher present at a conference</td>
</tr>
</tbody>
</table>

NOTE: This excerpt omits “outcomes” on establishing research purposes, relating results to professional literature, summarizing data and results, and stating conclusions and implications for teachers and teaching.
Learning Log on the Uses of Assessment in Instruction

This example comes from the University of Nevada, Las Vegas. UNLV has a large professional education preparation unit with some 600 teacher program completers annually. Candidates are typically admitted at the beginning of their junior year.

The assessment in this example is used in an introductory course on measurement and assessment related to instruction that all preservice teachers are required to take. The “learning log” reflection examines both skills and dispositions in response to NCATE unit standard 1, Candidate Knowledge, Skills, and Dispositions. Faculty have evaluated the assessment’s accuracy and fairness (the assessment is integral to the syllabus for the course).

The assessment requires that candidates prepare a weekly “Learning Log” that is designed to encourage their thinking about the course material (e.g., reading, lectures, notes, and discussions), including reflections on how the material relates to them as future educators. It can be used by faculty, as well, serving as a formative assessment for them to see what is or is not being understood. The candidate entries are used as the basis for small and large group discussions.

The Committee has selected this example to illustrate criterion 2, on explicit expectations for performance, and also to illustrate cost effectiveness—obtaining useful information at a reasonable cost. The excerpt on the next page contains some objectives for the “EPY 451” course, the tasks, and the scoring guide. The full example includes several samples of candidate work at different levels of performance.

CONTACT: Lisa D. Bendixen, Department of Educational Psychology, libendixen@ccmail.nevada.edu
Example J—UNIVERSITY OF NEVADA, LAS VAGAS

The Learning Log task covers the entire span of the course’s content, so several learning outcomes have been developed. Each entry (16 for the semester) is to include three things: (1) One or more pertinent questions for discussion, (2) Concepts/issues that stand out to the candidate and/or are not understood, and (3) A brief discussion of how the information covered relates to the candidate’s future career as an educator, with at least one specific classroom example. Rubrics are discussed in class and examples of the three levels of achievement used are given in the beginning of the semester. The following are examples of specific learning outcomes taken from the syllabus and show where they fit in relation to NCATE unit standard 1 on Candidate Knowledge, Skills, and Dispositions.

EPY 451 Syllabus Examples

**Knowledge**
- Recognize the relationship among, and the conditions affecting, objectivity, reliability, and validity
- Recognize guidelines for selecting/stating appropriate instructional objectives
- Identify appropriate uses, strengths, limitations, and rules for the construction of selected-response items and constructed-response items
- Recognize advantages, limitations, and appropriate uses of performance-based assessments
- Identify the various types of grading systems
- Understand the uses of introductory statistics in educational assessment
- Recognize major characteristics of the normal distribution

**Skills**
- Make value judgments about the appropriateness of different types of assessments
- Make value judgments about the importance of assessment validity and reliability

**Dispositions**
- Develop an appreciation of the role of assessment in student learning and achievement
- Develop an awareness of the uses and misuses of assessment
- Understand the consequences of different assessment strategies
- Appreciate the importance of assessment validity and reliability
- Appreciate the relationship between instructional objectives and assessment
- Understand the importance of communicating the results of assessments
- Appreciate the implications for the uses of the normal distribution in educational assessment

EPY 451 Learning Log Rubric
Levels of Achievement

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Exemplary</th>
<th>Acceptable</th>
<th>Needs Work</th>
</tr>
</thead>
<tbody>
<tr>
<td>One pertinent discussion question</td>
<td>Your entry contains a relevant, thought-provoking discussion question</td>
<td>Your entry includes a relevant discussion question</td>
<td>Your entry includes a definition question</td>
</tr>
<tr>
<td>Concepts/issues that stood out to you and/or that were not understood</td>
<td>Your entry includes a thoroughly detailed description of concepts that stood out to you and/or were not understood</td>
<td>Your entry includes a description of concepts that stood out to you and/or were not understood</td>
<td>Your entry includes a vague description of concepts that stood out to you and/or were not understood</td>
</tr>
<tr>
<td>Brief discussion of how material covered in class relates to YOUR future career as an educator</td>
<td>Your entry includes meaningful discussion and a specific classroom example was given</td>
<td>Your entry includes meaningful discussion and a general classroom example was given</td>
<td>Your entry includes vague discussion and no classroom example was given</td>
</tr>
</tbody>
</table>
Math Prompt and Scoring Guide

This example is from Wheelock College, Massachusetts. The education division at Wheelock enrolls about 450 students, with about 275 undergraduates and another 175 at the graduate level.

The assessment is described in the submission as follows:

. . . {It} is an example of an assignment given near the beginning of the first semester of the year-long course Concepts and Processes. The course is designed for students whose placement test scores show that they lack solid conceptual understanding of elementary mathematical concepts such as fractions. It is not specifically a “Math for Teachers” course; however, given the focus on education and human services professions at Wheelock College, the majority of students in the course are first-year students who will later elect to concentrate in early childhood or elementary education.

Billion Dollar Winner 2 is a paper that fits into the first unit, “Introduction to Problem Solving, Reasoning, and Communication.” The majority of problems attempted in this unit lead to integer patterns; in the context of learning about the National Council of Teachers of Mathematics (NCTM) Standards Process Goals (the three aforementioned plus Representation and Mathematical Connections), students also review or learn for the first time many concepts from number theory and algebra.

The rubric that accompanies the assignment is designed to communicate expectations to students and to help them write better papers. . . . Billion Dollar Winner 2 is a restatement of the well-known Locker Problem, with standing and sitting on chairs replacing opening and closing lockers. The assignment follows an in-class activity, Billion Dollar Winner, which is a restatement of the Josephus problem. In Billion Dollar Winner, students stand in a circle. The first student is skipped, the second sits, the next is skipped, and so on, skipping one standing student and having the next student sit, until there is only one student standing, who wins a billion dollars . . . The two problems, Billion Dollar Winner and Billion Dollar Winner 2, have some similarities in formulation but lead to very different mathematical patterns.

Students solve the problem in their study groups, and the vast majority solve it correctly (receiving threes or fours on the Problem Solving rubric). Many have trouble with the Reasoning portion of the rubric. The course instructors have come to use the “convince a skeptic” language to help students understand the expectation around reasoning. Language we have used in the past such as “how do you know you’re right?” oftentimes reinforced poor reasoning. Many communication problems stem from students’ poor writing skills, and they often take a remedial writing class concurrently.

This example stood out for the Committee because the assessment task so obviously is mathematics and responding to it demands mathematical reasoning. It was selected to illustrate a scoring guide that can be used in self assessment and in a group setting for criterion 2, and also for innovation in assessment (providing a concrete demonstration of the job of a teacher in reflection, and in a content course rather than in a professional education course).

CONTACT: Debra Borkovitz, dborkovitz@wheelock.edu
Last revised March 25, 2003
You have been invited to consult on . . . (a) . . . Billion-Dollar Winner Game. Who knew that somewhere, so much money was available for good problem-solvers? This new game . . . often has multiple winners. Here’s how it works: On her first pass, the host makes sure that all players are lined up standing in front of a chair. On the second pass, the host starts at the beginning of the line, and as she walks by, every other player sits. Thus, if we number the chairs 1, 2, 3, 4, etc., players in chairs 2, 4, 6, etc., sit. On the third pass, the host walks by and players in every third position change, that is, if a player is standing, s/he sits, and if a player is sitting, s/he stands. Thus the player in seat 3 who was standing now sits; the player in seat 6 who was sitting now stands, etc. On the fourth pass, every fourth player changes, on the fifth pass every fifth player changes, etc., until the host has made one pass for every player present.

When the host is done, everyone who is still standing shares in the Billion-Dollar Prize. Your task is to figure out which players share the prize. Of course, you don’t know how many players will show up, but be sure you can answer the question for at least 500 players. As with the first “Billion-Dollar Winner” problem, it will be important that you find a good representation for the problem, organize your results, look for patterns, justify your reasoning, and communicate clearly.

**Extensions:** Make up your own “Billion-Dollar Winner” game and analyze it. You can modify either this problem or the earlier one. Pick some kind of simple pattern (e.g., in the original problem, two people are skipped for everyone who sits). Don’t pick something trivial (e.g., players begin by sitting instead of standing in the second game).

**Scoring Guide**

**Problem solving**

- **Level 5** Does everything in Level 4 especially well. Extends the problem with insight.
- **Level 4** Solves the problem correctly for up to 500 players.
  - Finds pattern that can extend solution for any number of players.
  - Solves at least one substantial extension problem for one of the Billion-Dollar Winner problems.
- **Level 3** Solves the problem correctly for up to 500 players.
  - Finds pattern that can extend solution for any number of players.
- **Level 2** Solves problem for up to 30 players.
- **Level 1** Does work that is relevant to the problem, but is not up to Level 2.
- **Level 0** Work is not relevant to the problem

[NOTE: Excerpt omits scoring guide for “Reasoning”]

**Communication**

- **Level 5** Everything as in Level 4, but done especially well. The kind of paper you could put in a portfolio or show on a job interview as an example of your best work.
- **Level 4** The paper stands on its own. Someone who has never seen the problems before, and who is at your level of mathematics proficiency, could follow it. The problems are clearly restated in your own words or are obvious from your text, without needing to be restated.
  - The writing flows. It moves from one idea to another in a logical fashion. It is easy to understand. Ideas are explained fully; you tell the reader what you mean, rather than expecting the reader to guess. It’s obvious that you’ve put a lot of thought into your answer.
  - Your arguments are convincing. If there is a piece missing or a part you don’t understand, you tell the reader, rather than hoping s/he won’t notice.
  - The paper is written using complete sentences and virtually flawless grammar, punctuation, and spelling.
  - Written representations of the problem are appropriate and clear. Pictures and/or tables are used to make the ideas clearer.
  - Tables and graphs are labeled and presented in a clear fashion.

[NOTE: Excerpt omits text for levels 3 through 0]

**Self-Score Sheet (must be included with paper, feel free to use back side)**

With whom did you work on this paper? List names.

What makes this paper your own? In other words, how did you put your own spin on the material when you wrote it up? What did you contribute towards getting a solution?

Based on the scoring guide, what score do you think you deserve for problem solving? Why? Please refer to the scoring guide in your explanation.

Based on the scoring guide, what score do you think you deserve for reasoning? Why? Please refer to the scoring guide in your explanation.

Based on the scoring guide, what score do you think you deserve for communication? Why? Please refer to the scoring guide in your explanation.

What was the most satisfying part of this assignment? What was the most frustrating part? Why?
Admissions Portfolio Requiring Common Essays

This example is from the University of Nebraska, Lincoln. The University falls into the “large” education unit size, with more than 200 annual graduates.

The UNL assessment example is a set of essays comprising a portion of the admissions “portfolio” requirement for the Elementary Teacher Education Program. It is administered for the junior year, so it follows pre-entry exposure to college course work, including several education courses and a practicum. It is accompanied by required passing scores on standardized basic skills tests and practicum evaluations from candidate’s university supervisor and cooperating teacher.

The sponsors provide this background information about their intent and decisions:

The weights were the decision of the Elementary Teacher Education Program participants during 1996. The intent of the three essays was to assess the candidate’s ability to engage in the kind of reflective thinking and writing that is promoted in the program. Assessing this type of thought and writing would comprise the bulk of the weight. ETEP members also wanted to keep the content and reflectivity of the writing separate from the technical competence of the writing, therefore, writing conventions were separated out, and given their own category. . . . Interpreting written comments can involve a high degree of inference on the part of the reader. This is the line of thinking that supports the agreed upon weights of ¾ devoted to essays with the remaining ¼ split between practicum and writing conventions, with practicum weighted slightly higher.

The assessment has been evaluated by faculty with regard to consistency of results and the 1999 version here is revised from an earlier one used in 1997. Technology is not used in the assessment.

The example was selected as an illustration of an assessment to make significant decisions about a candidate, criterion 4. It was also selected as an illustration of innovation because of its use of weights in the rubrics and its inclusion of dispositions —on “diversity” and “commitment”—as important elements to consider during admissions decisions.

CONTACT: Kris Mickelson, elementary program leader, k Mickels@unlserve.unl.edu
Last revised March 12, 2003
A Note on Reflection: In responding to the essay prompts below, we want you to be reflective in your writing. This means that in addition to describing an experience or activity, you go into some depth explaining what made it meaningful to you. Think deeply about how the experiences reflect your learning as a future teacher, your relationships with children, and your growing understanding of the teaching profession.

Accomplishment
In this section, we want to learn how you think about and analyze an experience you consider a personal accomplishment. An experience becomes a personal accomplishment when it helps us to face new situations, challenges, or subsequent experiences. Think about a personal experience that was strong, intense, or sustained over a period of time. We would like you to choose only one so that you can demonstrate the depth of the effect of the experience and why you consider it an accomplishment.

First, briefly describe the experience. Then, discuss why you consider the experience an accomplishment for your. It may help to consider how you became involved in the experience, obstacles you encountered and how you dealt with them, what you learned about yourself as a result, and your thoughts, now, as you reflect back on the experience.

Remember to think deeply about how the experience affects your learning as a future teacher, your relationships with children, and your growing understanding of the teaching profession.

Diversity
Diversity, as defined by the Teacher College Multicultural Enrichment Plan, includes people who differ from the applicant in racial/ethnic origin, gender, age, mental and physical abilities, and religion.

Issues of diversity are increasingly affecting teachers, teaching, and learning in schools. In this section, we want to learn about your views relating to people different from yourself.

Experiences with diversity are typically times when our initial feeling is uncomfortable because of extending oneself outside ones own comfort zone. We want to learn ways in which you may have interacted with, related to, and thought about differences based on a specific experience.

You might consider an experience that brought you face to face with an issue of diversity. The experience you choose should be one that made you keenly aware of yourself as different from others. It might be a relationship with someone distinctly different from you, one that helped you articulate, perhaps for the first time, your stand on a particular issue of diversity. Ongoing or sustained experiences usually have greater effect, however, sometimes a single incident is life changing.

First, choose an experience (or relationship) and briefly describe it so that readers will understand the context of the situation. Then, discuss what you learned from the experience. What preconceptions did you bring into the experience? How were your beliefs or attitudes changed because of the experience?

Think deeply about how the experience will affect your learning as a future teacher, your relationships with children, and your growing understanding of the teaching profession.

Commitment
In this section, you have the opportunity to present yourself as someone who is committed to teaching, learning, and working with children. We believe that teachers need to be lifelong learners who are continually open to new experiences and new ideas. Teachers need to have high expectations of themselves, as well as of their students. In this section, we ask that you describe and reflect upon two experiences you believe demonstrate your commitment to the profession. Consider including some of the specific things you learned about learning (from observing children), the teaching profession, and your relationships with children. Being specific is one way of demonstrating that you are reflective (that the experience had great meaning for you.)

The first is an experience of your choice. Choose an experience that demonstrates that you have the type of commitment discussed above. Briefly describe the experience and then reflect upon it so that we can learn how you think the experience demonstrates your commitment to becoming a teacher.

The second is an analysis of the experiences you had in your practicum settings during past semesters. Reread your journals and study the practicum evaluations attached to your application. Give specific examples that come from your journals and the final evaluations, and reflect upon those so that we learn about your commitment. Explain how your cooperating teachers and supervisors knew that you were committed to children, teaching, and learning. Discuss your strengths and areas in which you want to learn more. Again, think deeply about how the experiences affect your learning as a future teacher, your relationships with children, and your growing understanding of the teaching profession.
<table>
<thead>
<tr>
<th>Wrty Conv 10%</th>
<th>Accomplishments 25%</th>
<th>Diversity 25%</th>
<th>Commitment 25%</th>
<th>Pract Evals 15%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Usage and composing very weak; spelling/typing errors, incorrect grammar throughout; lacks coherence.</td>
<td>Simplistic; lacks thought and reflection; disorganized; does not support positions with details; speaks in generalities; no connection to self as learner or teacher; first draft quality; may have missed the point.</td>
<td>Simplistic; lacks thought &amp; reflection; disorganized; does not support positions with details; extremely narrow view of diversity. Views diversity as “we-they.”</td>
<td>Consistently poor university supervisor evaluations.</td>
</tr>
<tr>
<td></td>
<td>Not acceptable</td>
<td>Not acceptable</td>
<td>Not acceptable</td>
<td>Not acceptable</td>
</tr>
<tr>
<td></td>
<td>.10</td>
<td>.25</td>
<td>.25</td>
<td>.15</td>
</tr>
<tr>
<td>2</td>
<td>Some errors in spelling or typos, grammar, sentence structure, typing, punctuation; (contains run-ons, fragments), usage and composing somewhat weak. Will require extra effort on written assignments.</td>
<td>Describes experience, but provides minimal discussion; superficial explanation; speaks in generalities; some thoughtful reflections, but generally lacks depth; makes loose connection to self as learner and teacher.</td>
<td>Immature and narrow view of diversity due to lack of experience; may express tolerance as a good thing; may view self as a “rescuer;” describes experience but reflections lack depth; points may meander from topic; lacks cohesion and elaboration; fails to compare preconceptions to his/her change in perceptions.</td>
<td>Below average university supervisor evaluations in one or both practicums. May lack self-initiative.</td>
</tr>
<tr>
<td></td>
<td>.20</td>
<td>.50</td>
<td>.50</td>
<td>.30</td>
</tr>
<tr>
<td>3</td>
<td>Very few errors in spelling, grammar, punctuation throughout text. May have paragraph break problems. Usage and composing good, but not excellent. Text flows well and arguments are coherent.</td>
<td>Experience explained well throughout; states a clear position; thoughtful reflections; clear and concise, connected to self as learner and teacher.</td>
<td>Describes and reflects on experience; takes a broader view of diversity; e.g., values differences; explains how the shared experience affected or changed applicant’s perceptions and views; ideas presented in an organized, reflective fashion; includes reference to self as learner and teacher in diverse classrooms.</td>
<td>Good university supervisor evaluations, consistent across practicums. Applicant meets expectations for students at this stage.</td>
</tr>
<tr>
<td></td>
<td>.30</td>
<td>.75</td>
<td>.75</td>
<td>.45</td>
</tr>
<tr>
<td>4</td>
<td>Excellent composition. Virtually no errors in spelling, grammar, typing and punctuation throughout text; usage and composing excellent.</td>
<td>Meaningful experience, thoroughly explained; in-depth reflection of the processes, challenges, and obstacles of achievement; arguments &amp; positions well thought out; strong connections between examples/supporting details and self as learner and teacher.</td>
<td>Outstanding description and in-depth reflections on sustained lived experience; presents very broad view of diversity, e.g., empowering all children; applicant discusses thoroughly how pre-conceptions were changed due to experience; articulates strong connections between examples/supporting details and commitment to teaching all children in diverse classrooms. May take activist orientation.</td>
<td>Outstanding university supervisor evaluations; consistent across practicums. Applicant exceeds expectations for students at this stage.</td>
</tr>
<tr>
<td></td>
<td>.40</td>
<td>1.00</td>
<td>1.00</td>
<td>.60</td>
</tr>
</tbody>
</table>
On-Line Performance Requirements for Technology

This example is from Boston College, a private institution in Chestnut Hill, Massachusetts. There are about 600 undergraduates enrolled in Lynch School of Education teacher certification programs, and another 800 graduate students.

The assessment highlighted in this example is a set of on-line tasks that allow candidates to demonstrate their technology competencies. The assessment is a graduation requirement, and was developed to make certain that Lynch students are technologically competent.

Here is an excerpt from one of the technology assessment assignments:

**Use of Quest Online Library Search**

BC Lynch School of Education students will benefit from learning how to effectively perform online information searches. Through the completion of this technology competency students will demonstrate knowledge of performing and narrowing searches and gain the ability to find specific information when searching an online database. To demonstrate competency in the Quest Online Library Search, students will complete the following task:

Choose a topic of research related to education and perform an online Quest search with the goal of finding 10 empirical articles (articles describing a research study) on your selected topic. The articles must be from professional journals (e.g., Child Development, Developmental Psychology) – NOT books or popular publications (e.g., Newsweek, Psychology Today).

Students will complete a preliminary search for articles on their topic of choice. It will be the students’ responsibility to perform a focused search of keywords on their topic and then eliminate the articles that are not relevant to their topic. Students will print out the pages from the literature search which contain the descriptions of the studies that are related to their topic.

After students have chosen ten empirical articles relevant to their topic they will print out the abstract of each of these ten articles.

Students must turn in a copy of the printed pages from their literature search as well as their printed abstracts from 10 empirical articles to the Office of the Assistant Dean for Students and Outreach in order to complete this technology requirement.

Options for Completion of Quest Online Library Search requirement:

This competency requirement will be completed in PY030 Child Growth and Development class in the form of a class assignment. Additionally, this competency requirement can be fulfilled independently as long as all of the above requirements are met.

Students who require further assistance are asked to contact the Boston College library staff . . .

The excerpt on the next page summarizes the specific competencies to be demonstrated by candidates. The example is selected to illustrate authenticity in assessment, criterion 5, and also innovation.

Contact: Alec Peck, Boston College, Alec.Peck.1@bc.edu or the direct Boston College web link: [http://www.bc.edu/bc_org/avp/soe/themes/default.html](http://www.bc.edu/bc_org/avp/soe/themes/default.html), click on “Technology Competencies”

Last revised April 2, 2003
LYNCH SCHOOL OF EDUCATION, SPRING 2001

Lynch School students in the Class of 2002 and beyond should demonstrate the competencies listed herein prior to graduation, but students are strongly encouraged to submit each component as soon as they are ready. Many of these competencies will enhance the academic experience of Lynch School students throughout their Boston College coursework.

Outlined below are the technology competencies required of all students during each undergraduate year. More detail on each requirement is provided under “Guidelines for Demonstrating Technology Competencies” covered in the next section.

<table>
<thead>
<tr>
<th>Competency</th>
<th>Ways to Complete</th>
<th>Course/Portfolio</th>
<th>Available Support</th>
<th>Date Completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Email</td>
<td>Submit e-mail to: <a href="mailto:lsoetechcomp@bc.edu">lsoetechcomp@bc.edu</a></td>
<td>Portfolio</td>
<td>Peer Mentoring</td>
<td></td>
</tr>
<tr>
<td>Clip Art</td>
<td>Create a word processing document with 10 pieces of clip art. Submit as an email attachment to: <a href="mailto:lsoetechcomp@bc.edu">lsoetechcomp@bc.edu</a></td>
<td>Portfolio</td>
<td>Peer Mentoring</td>
<td></td>
</tr>
<tr>
<td>Word Processing</td>
<td>Create a research paper or newsletter. Submit as an email attachment to: <a href="mailto:lsoetechcomp@bc.edu">lsoetechcomp@bc.edu</a></td>
<td>Portfolio</td>
<td>Peer Mentoring</td>
<td></td>
</tr>
<tr>
<td>On-line literature search</td>
<td>Generate printed results and Abstracts from Quest literature search. Submit to professor of PY030</td>
<td>PY030</td>
<td>O’Neill Library</td>
<td></td>
</tr>
<tr>
<td>Digital Photo</td>
<td>Create a word processing document with a digital photo pasted into it. Submit as an email attachment to: <a href="mailto:lsoetechcomp@bc.edu">lsoetechcomp@bc.edu</a></td>
<td>Portfolio</td>
<td>THEMES Workshops</td>
<td></td>
</tr>
<tr>
<td>Scanner</td>
<td>Create a word processing document with a scanned image pasted into it. Submit as an email attachment to: <a href="mailto:lsoetechcomp@bc.edu">lsoetechcomp@bc.edu</a></td>
<td>Portfolio</td>
<td>THEMES Workshops</td>
<td></td>
</tr>
<tr>
<td>Spreadsheets</td>
<td>Create an electronic grade book. Submit as an email attachment to: <a href="mailto:lsoetechcomp@bc.edu">lsoetechcomp@bc.edu</a></td>
<td>ED060</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypermedia Authoring Software</td>
<td>Create a multimedia presentation or web page. Submit as an email attachment to: <a href="mailto:lsoetechcomp@bc.edu">lsoetechcomp@bc.edu</a></td>
<td>Portfolio</td>
<td>THEMES Workshops</td>
<td></td>
</tr>
</tbody>
</table>

Submit each of these Technology Competencies as an attachment via email to the following address: lsoetechcomp@bc.edu Please include the name of the competency being submitted in the subject header of the email. (For example: “Subject: Email Competency,” “Subject: Clip Art Competency,” etc.). Each student will be responsible for saving copies of each competency that is submitted via email for their records. It is suggested that each student create a Technology Portfolio by saving each of their completed technology competencies onto a floppy disk that can be presented to various schools during teaching interviews.
Assessing the Teaching Characteristics of Preservice Teachers in an Internship Program

This example comes from California State University, Fresno. There are about 100 post baccalaureate candidates in the program using this assessment—an alternative certification teacher internship conducted jointly with local schools—while the entire Fresno education unit falls into the “large” category, with more than 200 graduates annually.

The principal features of the assessment are its form, vignettes, and its administration through a campus “assessment center.” The sponsors view it as addressing NCATE unit standard 1 on candidate knowledge, skills, and dispositions. Fresno faculty have evaluated the consistency of results and undertaken efforts to eliminate bias in the assessment.

Here are some descriptions of the assessment center activities from the sponsors:

Students in the teacher preparation program at California State University, Fresno (CSUF) participate in a unique program designed to help them and their supervisor fine-tune teacher skills. The characteristics of a successful teacher that each student in the program possesses are evaluated during an assessment day on campus. This assessment, administered by the PreTeacher Assessment Center (PTAC), rates the interns on a number of characteristics important to successful teaching. Ratings are shared with the intern and the supervisor after the assessment is completed. The intern then develops an action plan to improve in at least two areas that score relatively low. This plan must be completed before the end of the intern teaching experience.

The assessment consists of two vignettes that require the intern to think through a problem, develop a plan, and then share that plan in writing, and in one case, through an oral presentation. These are based on assessment models originally developed by Robert Millward, Indiana University of Pennsylvania. Both are completed in a single day, usually a Saturday during the beginning of the semester. Eligible students receive one academic credit for participation in the assessment and improvement plan.

The vignette entitled “Cities” requires the intern to plan a unit of instruction around this theme. They may adapt the content of the lesson to suit their own content area. Each student is provided a packet of materials with information about the unit with planning resources including a small book about the development of cities. For example, a history teacher could focus on social concerns, an art teacher on the esthetics of cities, or a mathematician on the geometric and engineering of structures. In addition to the written report, each intern must also make a fifteen minute presentation, that is videotaped, explaining their lesson. [NOTE: the submission continues with a description of the second vignette, “Education Fair.”]

In general, the intern assessment is an effective way to help identify those who need a great deal of help in most areas and the ones with clear differences in a few dimensions. Supervisors can use this information combined with their own professional observations and judgments to make the intern experience a better learning environment. . . . The combination of performance assessment and classroom practice helps both the intern and the supervisor focus on important aspects of the beginning teacher experience.

The California State University, Fresno example was selected by the Committee as an innovation in its use of vignettes, and of an assessment center with trained evaluators.

CONTACT: Bernard W. Arenz, barenz@csufresno.edu

Last revised March 19, 2003
Example N—CALIFORNIA STATE UNIVERSITY, FRESNO

Cities

Presentation of Ideas For Developing A Teaching Unit; Instructions for Participants

You are a teacher in Rosebud School District. Rosebud is a large urban school district encompassing rural, suburban, and inner-city neighborhoods. Students who attend Rosebud school are very diverse—ethnically, culturally, linguistically, socioeconomically. District teachers are challenged to provide curriculum that is both consistent with district goals and meets the students' needs. As part of your curriculum, you have been asked to develop a unit which focuses on Cities. It is up to you, however, to decide how you could connect this broad topic to a particular grade level or subject area.

You will have two hours to develop goals and objectives, innovative teaching strategies, and an assessment plan for a unit on Cities. This exercise (development of a unit overview and presentation of ideas on how best to teach about Cities) does not assume you have any special academic or educational training.

At the end of this two hour period, you will present your ideas to an individual who will role play a college supervisor of student teachers. The “supervisor” will listen to the presentation about your unit ideas and may ask you questions either during or after the presentation.

You will have 15 minutes to present your ideas. Before you begin your presentation, you will introduce yourself and will designate what grade level and content area(s) your unit has been designed for. It might be ideas for a first grade unit, an eighth grade unit, an eleventh grade unit, et cetera. You decide.

All the materials you need for your 15 minute presentation are included in your packet. The packet contains:

- Background information on Rosebud School District;
- A book on Cities;
- Some material downloaded from the Internet;
- Paper for scribbling and notetaking; and
- A unit outline form.

Extra paper and colored markers are available if you wish to create visuals or graphic organizers to accompany your presentation.

You may not discuss your unit plan with your peers. You have a maximum of two hours to complete your unit overview plan and to decide how you will present your ideas to the role player. (If you complete the written portion of this simulation before the two hour period is over, notify the PreTeacher Assessment Center Director.) When you finish, you will be assigned to a role player and will provide him/her with your notes and/or unit outline to review.

The role player will take a few minutes to review your notes prior to your 15 minute presentation. Your notes will be returned to you and you can use them during your presentation.

Present your unit ideas—goals and objectives, innovative teaching strategies, plan for evaluating student learning—to the “university supervisor.”

You must complete your presentation within the fifteen minute time limit. The role player will alert you when two minutes remain. Try to expand on your ideas and use as much of the time as possible. The more you share about your plans, the easier it will be for scorers to evaluate your skills in this area.

At the conclusion of your presentation, you will be asked to complete a self-analysis of your unit ideas and of your presentation to the role player.

Scoring [NOTE: Detailed guidelines for scoring the assessments are provided in the full submission.]

A trained evaluator rates the interns based on all of the work they do on each vignette. This includes notes written on documents, highlighted sections of the packet, prepared charts, graphs and presentation materials, the oral presentation and final written reports. A detailed scoring rubric helps the evaluators focus on the most critical elements of the assessment exercise. The following dimensions are given a rating from one through five, with five being the high score:

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Level of inference</th>
<th>Characteristic</th>
<th>Level of inference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planning and Organization</td>
<td>Medium</td>
<td>Oral Presentation</td>
<td>Low</td>
</tr>
<tr>
<td>Leadership</td>
<td>High</td>
<td>Written Communication</td>
<td>Low</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>Medium</td>
<td>Innovation</td>
<td>High</td>
</tr>
<tr>
<td>Decision Making</td>
<td>Medium</td>
<td>Tolerance for Stress</td>
<td>High</td>
</tr>
<tr>
<td>Oral Communication</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
INTASC Principles and the BSU Student Teaching Rubric

This example is from Ball State University. Ball State has a large education unit, falling into the Project category of “more than 200 graduates” annually. Candidates are admitted to the education program as sophomores.

The assessment described here is a guided faculty observation of candidate clinical experience during courses taken by special education program sophomores and juniors. The distinguishing features of the assessment are its administration as part of two special education practicum courses and its link to INTASC standards that have been adopted by the state of Indiana to guide teacher preparation.

Here are some notes from the sponsor that describe initial development of the assessment:

After the first semester of practicum supervision, I felt strongly that we also needed to transition from the traditional assessment of that experience to a set of rubrics, tied to the INTASC principles and the BSU student teaching rubric, that describe the performances expected at these initial stages of classroom experience. In conversations with several cooperating classroom teachers, it became apparent that in addition to performances tied to the INTASC principles they were concerned very specifically about four other issues: punctuality and organization, ability to maintain confidentiality of information, attitude and initiative, and flexibility. Although one could argue those qualities are incorporated within the principles, the request among the cooperating teachers was clear and a section for “professional habits” also is part of the assessment rubric.

The assessment is founded on the ten INTASC principles, adopted by Indiana for use in teacher preparation, together with the INTASC descriptions of “knowledge,” “dispositions,” and “performances.” The excerpt that follows on the next two pages begins with a description of what candidates are expected to do to demonstrate their accomplishments, then illustrates the “Guide for Evaluation” that is constructed around the “performance” elements from INTASC; the guide describes “emerging,” “developing,” and “proficient” levels of candidate performance. The next part of the excerpt shows a similar scoring guide for the professional behaviors that have been added by Ball State faculty: “punctuality/organization,” “confidentiality,” “attitude/initiative,” and “flexibility.” The second page of the excerpted Ball State assessment materials illustrates the form used for University supervisor observations of candidate classroom skills in the areas of “communication,” “classroom management behaviors,” and “lesson planning and delivery.” Finally, a set of “debriefing” guidelines is provided that focus faculty conversation with candidates and encourage acquisition of reflective practice skills.

The Committee selected these assessment scoring guides as an example of innovation for their development around INTASC and program courses. A feature of note is that use of these guides to monitor a candidate’s developmental progress because the same rubrics are consistently applied over the years of teacher preparation.

CONTACT: Sharon R. Schultz, instructor, teachers college, Ball State University, sschultz@bsu.edu
Last revised March 25, 2003
**Example O—BALL STATE UNIVERSITY**

**Ball State University sponsor description of candidate work that is evaluated:**
In addition to performance in the classroom during the practicum, teacher candidates engage in discussions on campus and prepare a number of artifacts related to the practicum experience.

Sophomores: 1) Collect, report on and discuss in written format implications of demographics of the school in which they are serving; 2) Discuss the teacher's classroom management plan, write about how it is implemented and what he/she will do to support that plan while in the classroom; 3) Participate in interviews with education professionals [administrators, paraeducators, nurse, other teachers, secretaries, counselor, etc.] in the building to gain a picture of roles and responsibilities; 4) Read and respond to current literature [focus on current journals] in the field of special education; 5) Create educational materials [bulletin board, learning game, assistive technology device, etc.] specific to the classroom and students and with the input of the classroom teacher; 6) submit a written reflection on the practicum experience

Juniors: 1) Discuss the teacher's classroom management plan, write about how it is implemented and what he/she will do to support that plan while in the classroom; 2) Videotape self teaching the class and engages in self reflection; a written document is submitted; 3) Choose one student [with input from the classroom teacher], collects work samples and anecdotal records of observations, analyzes data, creates an Individual Education Plan for that student as though it were time for the student's annual IEP

**Guide for Evaluation of SPCEd 266 and 361 Practicum Student Performances**

- Indicates level of expected mastery at the end of the practicum experience for BSU students enrolled in SPCEd 266

◊ Indicates level of expected mastery at the end of the practicum experience for BSU students enrolled in SPCEd 361

<table>
<thead>
<tr>
<th>Demonstration of Professional Competencies</th>
<th>Emerging</th>
<th>Developing</th>
<th>Proficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTASC Principle #1: The teacher understands the central concepts, tools of inquiry, and structures of the discipline(s) he or she teaches and can create learning experiences that make these aspects of subject matter meaningful for students</td>
<td>• Assists students in acquisition of content knowledge while participating in one-to-one or small group instruction</td>
<td>Assists students in acquisition of content knowledge and actively assists in creating connections which are meaningful for students; connects to other content areas, connections to real-life situations</td>
<td>◊ Routinely assists students in making connections between and among content subject matter and real-life situations.</td>
</tr>
<tr>
<td>INTASC Principle #2: The teacher understands how children learn and develop, and can provide learning opportunities that support their intellectual, social and personal development.</td>
<td>• Supports student learning by engaging students in conversations that draw on prior knowledge and background information possessed by the student</td>
<td>Supports student learning along a developmental continuum by providing both student and teacher connections to prior knowledge and experiences, current knowledge and skill acquisition, and future goals</td>
<td>◊ Develops and implements lesson plans based on knowledge of current developmental levels of the students served</td>
</tr>
</tbody>
</table>

[NOTE: The remaining eight INTASC principles are omitted from this excerpt]

<table>
<thead>
<tr>
<th>Demonstration of Professional Habits</th>
<th>Emerging</th>
<th>Developing</th>
<th>Proficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Punctuality/Organization</td>
<td>Arrives on time; is developing organizational habits</td>
<td>• Arrives on time; fulfills schedule requirements; consistently prepared for classroom duties; required paperwork is well-organized</td>
<td>◊ Arrives on time; fulfills schedule requirements; consistently prepared for classroom duties; lesson plans are prepared on time; paperwork is well-organized</td>
</tr>
<tr>
<td>Confidentiality</td>
<td>Is unsure of confidentiality issues</td>
<td>• Maintains confidentiality regarding student information</td>
<td>◊ Maintains confidentiality regarding student information; contributes to creating and maintaining an atmosphere of professionalism in the classroom and with all staff members</td>
</tr>
</tbody>
</table>
| Attitude/Initiative | Approaches the practicum experience with routine expectations and output | Demonstrates commitment and enthusiasm for teaching | ◊ ◊ Demonstrates commitment and enthusiasm for teaching; goes beyond the minimum expectations of the practicum experience  
Expectation level for both SPCEd 266 and 361 |
| Flexibility | Having difficulty respond to the diverse needs and situations which arise in the classroom | • Exhibits flexibility in meeting the needs of students in tutoring and/or small group situations | ◊ Exhibits flexibility in meeting the needs of individual students within the context of the classroom |

**Overall Impression of Potential as a Future Teacher**

| Potential based on observations to date | Should consider another practicum experience at the SPCEd 266 level | • Is prepared to move on to practicum requirements of SPCEd 361 | ◊ Is prepared to move on to requirements of student teaching |
Communication Skills

<table>
<thead>
<tr>
<th>Skill Description</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Mostly</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>Listening skills: attends to all students; listens without interruption or interjection; smoothly transitions when need to move on with lesson or activity</td>
<td></td>
<td></td>
<td>◊</td>
<td></td>
</tr>
<tr>
<td>Oral language in use: grammatically correct structures are exhibited in both presentation and conversation</td>
<td></td>
<td>◊</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Written language in use: prepared handouts and spontaneous writing exhibit correct use of grammar and spelling</td>
<td></td>
<td>◊</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body language: the spoken message and the visual message are congruent and supportive of student learning</td>
<td></td>
<td></td>
<td>◊</td>
<td></td>
</tr>
</tbody>
</table>

[NOTE: This excerpt omits “Classroom Management Behaviors/Skills” and “Lesson Planning and Delivery” which are included in the full example.]

Guidelines for SPCED 351
Practicum Observation and Debrief

I. Preparation
Practicum Student (PS) sets the stage for a productive observation process by employing organizational skills and practicing lesson presentation before class.

A. Schedules in a timely manner
B. Reviews “Observation of Performance” Form
C. Plans best place for observer to be seated
D. Provides copy of lesson plan upon arrival of university observer

University Observer (UO) sets the stage for a productive observation process.

A. Arrives on time
B. Reviews written lesson plan
C. Observes practicum student for 30 minutes

II. Practicum Student initial reflection on lesson
To support reflective practice of the practicum student it is important for him/her to begin the session, rather than the university observer to begin with his/her observations.

A. UO withholds judgment during this activity
B. UO encourages PS to state things/items/events that are positive/meeting standards as reflected in the “Observation of Performance” Form
C. Using active listening, paraphrasing, and direct questioning, UO encourages elaboration and presentation of data to support PS opinions relative to positive statements
D. UO encourages PS to state things/items/events that are in need of more attention to meet standards
E. Using active listening, paraphrasing, and direct questioning, UO encourages elaboration and presentation of data to support PS opinions relative to need statements

III. University Observer responds to initial reflection
To positively support professional growth of the practicum, student focuses on data in an objective manner.

A. Validate accurate points in PS’s initial reflection
B. Present observable data (observation of lesson presentation based on items as listed on “Observation of Performance” Form
C. Check for understanding of offered information—observe body language; ask for and encourage questions
D. Clarify comments as needed

IV. Analysis of lesson
Both work collaboratively throughout analysis and application of new knowledge.

A. Summarize impressions
B. Recall data and supporting information
C. Compare planned results (meeting a standard/movement on a rubric, etc.) with achieved results
D. Analyze, infer, and determine cause and effect relationships
E. Draw conclusions

V. Application of new knowledge
A. PS design new goals for successive lesson planning and presentation
B. PS determines action steps to reach goals
Performance Assessment with Teacher Standards (PATS)

This example of an assessment system is from the University of South Florida St. Petersburg. The system was developed with assistance from the Florida Department of Education and is a web-based relational database used to design standards-based assessments, as well as to collect and aggregate performance data. The PATS system was developed to assist Florida colleges of education and sponsors of alternative programs in meeting their responsibilities under the “Florida Educator Accomplished Practices.” The St. Petersburg unit falls in the “medium” size range of 50 to 200 annual graduates. Here are excerpts from the authors’ descriptions:

Standards Alignment: At the core of the PATS system is the standards alignment feature, which allows faculty to select standards and indicators measurable in specific tasks. Various standards (Florida Educator Accomplished Practices, INTASC Principles, Florida Content Standards, Florida ESOL Standards, standards of national Specialty Professional Associations, and institutional standards) are aligned around a common core of “base standards” (e.g., assessment, communication, knowledge of subject matter).

Content Sampled: A test blueprint or sampling plan was developed prior to the creation of task directions and scoring instruments for validity reasons in both systems. As tasks are entered in PATS, the developer classifies the type of competency being assessed: knowledge, skills, dispositions, and/or impact on K-12 learning. Specific attention was focused in the blueprint on ensuring that all four of these competency types were assessed. In the USFSP program there are five tasks with a strong emphasis on impact on K-12 learning, including work with an individual student who shows improvement, analysis of assessment results for an entire class on a unit, documentation of the results of accommodations, and a K-12 Student Portfolio (in the excerpt on the next two pages). Some of the disposition-related assessments attend to issues of bias and ethics, cultural awareness, and ability to work with diverse populations. Content knowledge is assessed not only through an initial transcript analysis and a passing score on the subject matter component of the Florida Teacher Certification Examination but also through 20 tasks in the system. These tasks include course-based tests, lesson plans assessed for content knowledge, and sets of mini-performance tasks and activities demonstrating content knowledge. The K-12 Student Portfolio is also assessed for candidate’s demonstration of ability to teach the content areas with results achieved through student products. Skill-based decisions permeate the tasks, and there is a mix of product and performance assessments incorporated in both systems.

Administration and Implementation: Assessments within degree programs are administered within required courses and field experiences. . . . Faculty are currently identifying specific points in the program when each candidate’s progress will be formally evaluated – probably each practicum and prior to program completion.

Data Entry and Manipulation: . . . The computer aggregates the data for reporting at the student, program, and unit levels for each set of standards in the system. With the vast amounts of time saved by computerization, faculty can then turn their attention to remediation of student difficulties and needs for program improvement. The system also produces the reports necessary for accreditation.

The Committee selected St. Petersburg as an innovation in assessment systems, and, as with other “systems” examples, much of the important information for readers in incorporated into the descriptions of what the system is and how it functions. In addition to the paragraphs above, on this page, additional information follows at the end of the second following page. This addresses a particularly significant feature of the St. Petersburg example—the efforts of its sponsors to conduct validation and psychometric studies for this high stakes system.

CONTACT: Judy R. Wilkerson, project director, wilkerso@tempest.coedu.usf.edu
Last revised March 25, 2003
Task 04E: Portfolio of K-12 Student Work, EDG 6947: Internship

Standards Infused and Assessed

Base Standards: Critical Thinking and Problem Solving and Knowledge of Subject Matter

Actual Standards/Indicators Assessed:
- Accomplished Practice Indicators: 4.01, 4.08; 8.01, 8.02, 8.03, 8.04
- Florida Elementary Standards: 6, 12, 17
- INTASC Indicators: 4.01, 4.07, 4.08; 1.01, 1.11
- ACEI Indicators: 2.1; 2.2.01, 2.2.11, 2.2.16; 2.3.01, 2.3.03, 2.3.04; 2.4.01; 2.5.03, 2.5.06; 3.3.01, 3.3.03
- USF Critical Skills: 4.02, 4.03; 8.01, 8.03

Directions

Your students’ work is the best evidence of your success as a candidate. If they have learned the content well from you, then there is clear evidence that you not only know your content but can teach it and you can teach them to use knowledge to problem solve and think critically. After all, the proof is in the pudding! For that reason, this task is one of the most important tasks in this assessment system. In this task, you will create a showcase portfolio of student work over the semester. Some students are so good that anyone could teach them. For that reason we are asking you to show us not just the best work from the best students but also some work that caused you to do a little extra with students who were not as easy to teach. You should do the following:

1. Select eight samples of work from students who have exceeded your expectations with regard to acquisition of knowledge in the content area – your best students. These samples should show not only mastery of content but also the ability to think critically/creatively and/or solve problems. Include two each from each of the content areas: language arts, mathematics, social studies, and science.

2. Select two more samples of work from students who did not meet your expectations with regard to acquisition of knowledge in the content area when they first attempted the work. These are the students who were a challenge to you and required some extra effort on your part. Ideally, you will be able to include their first and final attempts on a project so you can demonstrate how much you have helped them improve.

3. Select assignments that were targeted to achievement of the Sunshine State Standards. The work must be from at least six different students.

4. Write a reflection with each work sample that includes the following:
   - Why you selected the piece
   - How it demonstrates the student’s ability to use knowledge and skills in the content area.
   - The Sunshine State Standard is addressed
   - How the work demonstrates the student’s ability to think critically or creatively and to solve problems.
   - How you linked the work to real world experiences?
   - Source materials you used to create the assignment or test, including the Internet, materials located through a search of a database such as ERIC, and/or materials from professional content organizations, e.g., NCTM, NCSS, NCTE, NSF, IRA, or others.
   - For those work samples selected for not meeting expectations – at least initially, describe the actions you took to help the student improve.

Scoring Rubric

Name: _______________ Submission #: _____

(Rating Scale Key: T = target; F = flawed; U = unacceptable)

Decision for A.P. 8 (Subject Matter) on this Task (check one):
- Demonstrated: 0-3 ratings are flawed; none are unacceptable.
- Partially Demonstrated: 4 or more ratings are flawed; none are unacceptable.
- Not Demonstrated: 1 or more ratings are unacceptable.

<table>
<thead>
<tr>
<th>Element #</th>
<th>Criterion for “target” rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>The K-12 students have generated knowledge and/or tested hypotheses according to the methods of inquiry and standards of evidence used in the discipline.</td>
</tr>
<tr>
<td>8</td>
<td>The students have demonstrated acquisition of knowledge and skills in language arts.</td>
</tr>
<tr>
<td>9</td>
<td>The students have demonstrated acquisition of knowledge and skills in math.</td>
</tr>
<tr>
<td>10</td>
<td>The students have demonstrated acquisition of knowledge and skills in science and technology.</td>
</tr>
</tbody>
</table>
System Credibility: Attention to Legal and Psychometric Issues

Both ACP and USFSP systems have been designed with careful regard to the legal and psychometric requirements of programs leading to certification. These tasks constitute a certification “test,” as defined in the APA/AERA/NCME Standards. As a high-stakes test, the psychometric properties need to be examined during both the design and implementation stages.

Initial Validity Study: Authors analyzed issues of construct representativeness and construct relevance in the design of the set of tasks. This was possible through the design (standards selection) and reporting features (coverage reports) of PATS. Also used was a modified Jaeger procedure, consistent with the APA/AERA/NCME Standards, to ensure that the each task is job-related. This was the first step in developing a cut-score that will lead to valid and reliable decision indices for inferences about teacher competence. As each task was developed, two school system practitioners (one former, one current) were asked to judge the critical nature and job-relatedness of each task. This should serve as an initial protection against potential legal challenges under the Civil Rights Act of 1964 (Titles VI and VII) and the Fourteenth Amendment of the U.S Constitution, the two primary sources of legal challenges in certification. While an initial attempt to establish appropriate cut-scores to define minimal competency has been made, this, too will require further exploration, especially with regard to the social consequences of the scores.

In order to maximize the fairness of the system, we have adopted a standardized format for all tasks. This format includes detailed instructions with a corresponding scoring instrument. Criteria in the scoring instrument are based on both the directions and the language of relevant standards, thereby increasing the validity of the tasks as well as the fairness to candidates. All candidates have three opportunities to complete a task, allowing for remediation. The curriculum is currently being reviewed to ensure adequate learning opportunities and formative assessments so that candidates have adequate opportunity to learn. These two efforts are designed to ensure the instructional/curricular validity of the system. Next steps include the development of policies, procedures, and advising materials to ensure that all candidates know their rights and responsibilities. Students and faculty will have access to assessment results in PATS in order to identify and address remediation needs as soon as they surface.

Planned Psychometric Studies: Both systems (ACP and USFSP) are in field test mode this semester with data being gathered for initial item analysis studies. In addition to continuing the validation studies started in the design stage, authors expect to apply some of the latest methods of Item Response Theory (using Andrich’s RUMM 2010 software) to begin the development of logistic rulers for each of the domains being assessed. RUMM 2010 software produces data in both classical test theory and item response theory. Sources of measurement error will be identified. Over time, using these rulers and other assessment tasks developed, USFSP hopes to develop equivalent alternative tasks to meet teacher candidate needs. Additional examination for bias (gender, race, culture) and construct fit will fine-tune both the tasks and the task scoring.
Information and Assessment Systems (IAS)

This example was submitted jointly by the University of Maryland, Baltimore County, and the University of North Carolina at Charlotte. The example is a group of web-based information and assessment systems (IAS). At Baltimore County, teacher candidates are admitted at the beginning of the junior year or as graduate students. The institutions education units fall into the “medium” size category with 50 to 200 annual graduates.

The standards being assessed include NCATE standard 1, candidate knowledge, skills, and dispositions, and standard 2, Assessment systems and unit evaluation. State standards for teachers are also addressed. Technology is an important element of the systems: candidates are required to have technological capability to complete their responses; the assessments are administered or scored by computer; and the unit manages assessment results through data system technology. The sponsors are conducting reviews of the fairness of the assessments and consistency of results. Some 50 candidates have been involved in the pilot phase, and all programs will begin to use the IAS in 2003.

Here are some descriptive notes about the IAS (and more are provided on the next page):

Information and Assessment Systems

The Department of Education at UMBC, with the involvement of its professional community and assistance from the University of North Carolina at Charlotte (UNCC), is piloting the Information and Assessment Systems (IAS) to ensure continual improvement of candidate performance, program quality, and unit operations. The IAS documents both qualitative and quantitative performance data through five integrated production systems: (1) Student Information System, (2) Performance Assessment System, (3) Electronic Portfolio System, (4) Electronic Learning Community, and (5) EDUC Alumni Association.

. . . The Performance Assessment System and the Electronic Portfolio System . . . are used to document and enhance performance at each of the five critical stages across the developmental sequence of the teacher education program: (1) Program Entrance Assessment, (2) Course and Field Experience Assessment, (3) Clinical Practice Assessment, (4) Program Exit Assessment, and (5) Post Graduation Assessment.

With the innovative adaptation of cutting edge technologies and the strong support from the UMBC education community, we plan to fully implement the two systems at both the unit and program levels (Early Childhood, Elementary, ESOL and Secondary Education at the undergraduate and graduate levels) by the 2002-2003 academic year; and expect to implement all five production systems of the IAS by 2004-2005.

The example was selected by the Committee as an innovation illustrating the architecture of a technologically advanced assessment and data management system.

CONTACT: Yi-Ping Huang, Department of Education University of Maryland, Baltimore County: yhuang1@umbc.edu or the website dedicated to this project at http://www.umbc.edu/education/ias

Last revised May 27, 2003
Description of the Information and Assessment Systems (IAS) performance assessment system

- **Candidate assessment** emphasizes measurements of our teacher candidates’ knowledge, skills, and dispositions, and will examine the impact of our teacher candidates’ work on learning of children in P-12 settings. **Program and unit assessment** emphasizes efficacy in aligning instruction and curriculum with professional, state and instructional standards; efficacy of courses, field/clinical experiences and programs; and efficacy of candidates’ content, pedagogical and professional proficiencies that lead to student learning.

- Decisions about candidate performance are based on **multiple sources, multiple assessments with multiple indicators** made at each of the five critical stages across the developmental sequence of the teacher education program. Assessments at the first few stages focus primarily on content, pedagogical and professional knowledge and skills as measured by Praxis test scores, GPA on all coursework, and professional ePortfolio development. A **triangulated candidate evaluation** was designed to assess candidate’s field experience and clinical practice through clinical instructor, university supervisor and candidate’s reflective assessments; and to provide students with continual feedback on all three areas of knowledge, skills, and dispositions. A **triangulated program and unit evaluation** was also designed to assess the efficacy of faculty, curriculum, instruction and candidate performance through the clinical instructor, university supervisor and teacher candidate. Assessments near completion of the program and later stress content, pedagogical and professional skills and dispositions. The different emphases of assessment are consistent and coherent with the Theory into Practice Continuum outlined in the UMBC Conceptual Framework for Teacher Education.

Description of the Electronic Portfolio System

- The **Electronic Portfolio System** is a dynamic and context sensitive system designed to facilitate the documentation of knowledge, skills, dispositions and student learning; and to demonstrate standards-based proficiencies across the five developmental sequences of the teacher education program (see next page). The three benchmarks of Developmental ePortfolio, Showcase ePortfolio and Professional ePortfolio are closely associated with the critical transitions from university student to teacher candidate and teaching professional.

- Developmental ePortfolio serves as a consortium for a teacher candidate to collect, select and reflect on evidence and documentation throughout the teacher preparation program, particularly during field and clinical practice. As a program exit criterion, students are required to prepare a showcase portfolio for both oral and online presentation to the portfolio committee. The Showcase ePortfolio is a “best evidence” collection of documentation and artifacts demonstrating competencies on the standards and on the candidate’s professional growth. Candidates are expected to reflect on each of the artifacts, and validate why and how the artifact demonstrates growth or competence. Professional ePortfolio is useful in documenting and demonstrating quality performance and continual professional growth. The Department of Education provides all UMBC graduates with a wide variety of support services, including hardware, software access, training, publishing (via CD Rom and the Internet) and hosting of their Professional ePortfolio.

- Candidates and graduates are expected to address the following four areas in their ePortfolio:
  1) Personal and Professional Information (Personal information, professional information, field experience and clinical practice, site descriptions)
  2) Philosophy and Dispositions (Philosophy and Dispositions, Professional Development Plans, and Progress Journal)
  3) Standards-Based Achievements (INTASC Proficiency, Maryland Teacher Technology Proficiency, Professional Standard-Based Proficiency, and Capstone Experience—an integrated instructional unit that addresses content, pedagogical, professional and technology proficiencies)
  4) With Honor (Professional achievements, credentials, resume and caliber)
# Performance Assessment System: Measures at Each of the Five Developmental Stages

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Admission Assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pass Praxis I</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td>80%</td>
<td>90%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Minimum Overall GPA of 2.75</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Professional Development Plan</td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Education Foundation Courses</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td>80%</td>
<td>90%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Candidate Interview</td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Online Application</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>2. Course and Field Experience Assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minimum Overall GPA of 2.75</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Continuing Advisement</td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Criminal History Disclosure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Developmental ePortfolio</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>60%</td>
<td>80%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Survey of Field Experience</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>3. Clinical Practice Assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minimum Overall GPA of 3.0</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Continuing Advisement</td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Developmental ePortfolio</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>60%</td>
<td>80%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Surveys of Clinical Practice</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>4. Program Exit Assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minimum Overall GPA of 3.0</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Coursework Evaluation and Licensure Application</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Pass Praxis II</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td>80%</td>
<td>90%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Showcase ePortfolio</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>50%</td>
<td>75%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Maryland Technology Standards</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>60%</td>
<td>80%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Candidate Graduation Surveys</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>5. Post Graduation Assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Survey of First Year Practice—to be completed by UMBC graduates</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>70%</td>
<td>85%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Program and Graduates Survey—to be completed by employers</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>70%</td>
<td>85%</td>
<td>100%</td>
</tr>
</tbody>
</table>
English Proficiency Exam

This example comes from Hofstra University, New York. Hofstra is a large private university that graduates around 600 teacher candidates annually.

The assessment is an English proficiency exam created and administered by the English Department. Success on the assessment is a degree requirement for all undergraduate degrees, and passing it is necessary for graduation. Students are exempted from the requirement only if they have “outstanding ability in English” as demonstrated by high SAT verbal scores or high school grades in English of 90 or above. Students not passing the exam are required to work with individual tutors, retake the exam, and then if they still fail, must enroll in another English course.

Faculty in the Hofstra English Department view this assessment as a “work in progress” and currently plan to make revisions. Nonetheless, Committee members selected the example of criterion 1, an illustration of an assessment aligned with subject content standards. In addition, it serves as an example of criterion 6 on conducting studies to assure assessment credibility. In the criterion 6 listing of eleven “strategies” to document credibility, Hofstra faculty state that they have addressed those described under a, b, c, d, and e, as well as g (see pages 17 and 18).

CONTACT: Penelope Haile, Associate Dean, Hofstra University School of Education,
edapjh@hofstra.edu
Last revised March 18, 2003
SUCCESSFUL COMPLETION OF THE PROFICIENCY EXAM IS A GRADUATION REQUIREMENT.

INSTRUCTIONS:
After summarizing the author’s thesis or main point, identify, analyze and evaluate the argumentative or persuasive strategies which the author uses to support his or her thesis or main point.
Do not merely summarize the work or argue with the writer. You may present your opinion on the issue, but only insofar as it is relevant to your evaluation of the author’s specific argumentative or persuasive techniques. Ultimately, the essay will be graded on how well you present and support your opinion of the author’s argumentative strategies, not the author’s beliefs.
Take enough time to read the passage carefully and sketch an outline of your essay before you begin to write. Your essay will be graded on thoughtfulness of argument, organization, development, and grammar. You have 90 minutes.

Sample PE Essay

Smokers get a Raw Deal, by Stanley S. Scott

Stanley S. Scott (1933-1992) was vice president and director of corporate affairs of Philip Morris Companies Inc. This essay originally appeared in December 29, 1984, in the op-ed page of the New York Times.

The Civil Rights Act, the Voting Rights Act, and a host of antidiscrimination laws notwithstanding, millions of Americans are still forced to sit in the back of planes, trains, and buses. Many more are subject to segregation in public places. Some are even denied housing and employment: victims of an alarming--yet socially acceptable--public hostility. This new form of discrimination is based on smoking behavior. If you happen to enjoy a cigarette, you are the potential target of violent antismokers and overzealous public enforcers determined to force their beliefs on the rest of society.

Ever since people began smoking, smokers and nonsmokers have been able to live with one another using common courtesy and common sense. Not anymore. Today, smokers must put up with virtually unenforceable laws regulating when and where they can smoke--laws intended as much to discourage smoking itself as to protect the rights of nonsmokers. Much worse, supposedly responsible organizations devoted to the “public interest” are encouraging the harassment of those who smoke. This year, for example, the American Cancer Society is promoting programs that encourage people to attack smokers with canisters of gas, to blast them with horns, to squirt them with oversized water guns, and burn them in effigy.

The basic freedoms of more than 50 million American smokers are at risk today. Tomorrow, who knows what personal behavior will become socially unacceptable, subject to restrictive laws and public ridicule? Could travel by private car make the social engineers’ hit list because it is less safe than public transit? Could ice cream, cake, and cookies become socially unacceptable because their consumption causes obesity? What about sky diving, mountain climbing, skiing, and contact sports? How far will we allow this to spread?

The question all Americans must ask themselves is: Can a nation that has struggled so valiantly to eliminate bias based on race, religion, and sex afford to allow a fresh set of categories to encourage new forms of hostility between large groups of citizens? After all, discrimination is discrimination, no matter what it is based on.
# Grading Standards

## English 1

<table>
<thead>
<tr>
<th>QUALITY</th>
<th>Excellent (A)</th>
<th>Good (B)</th>
<th>Satisfactory (C)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Thesis</strong></td>
<td>Expresses a complex, comprehensive understanding of the topic while presenting thought-provoking, sophisticated analysis &amp; argument</td>
<td>Clear and specific; thoughtfully delineates the major elements of the argument</td>
<td>Responsive to assignment; arguable; reader gains a general understanding of the paper’s overall argument by the end of the first paragraph</td>
</tr>
<tr>
<td><strong>Organization</strong></td>
<td>Introduction that indicates the issues to be dealt with and justifies the approach taken. Rhetorically sophisticated arrangement of paragraphs. Conclusion that considers the implications and consequences of the argument.</td>
<td>Introduction that provides relevant background for the thesis. Logical development that necessitates the particular ordering of paragraphs. Smooth transitions that clearly articulate the interrelations between paragraphs.</td>
<td>Clearly identifiable introduction, body, conclusion. Proper paper format.</td>
</tr>
<tr>
<td><strong>Paragraphing</strong></td>
<td>Sophisticated use of a variety of forms or paragraph organization. Concluding sentences that advance the discussion in a way that sets up the smooth transitions to the next paragraph.</td>
<td>Topic and point of body paragraph communicated skillfully, usually at the beginning of each paragraph. Unified paragraphs. Ideas &amp; information flow logically from sentence to sentence.</td>
<td>Body paragraphs based on a particular topic pertinent to the thesis. Topic sentences.</td>
</tr>
<tr>
<td><strong>Development &amp; Support</strong></td>
<td>Sophisticated handling of the complexities of development and support. Fresh detail and illuminating analysis; argument explored and advanced with original insight &amp; intellectual rigor.</td>
<td>Thoughtful use of different argumentative strategies. Good selection of evidence including significant details and examples. Quotation and paraphrase integrated, contextualized, and analyzed. Good deployment of logical reasoning. Incorporation of other points of view.</td>
<td>Treatment of topics directly related to thesis. Some use of example, anecdote, explanation, and/or logical reasoning to substantiate assertions. Proper citation.</td>
</tr>
</tbody>
</table>

**Grading Standards**
The following categories attempt to capture and organize the variety of elements we consider when we assess the student’s attempt to grapple with a particular writing assignment. In general,

- “C” indicates a rudimentary understanding and application of the conventions and principles of analysis and argument;
- “B” indicates a skillful handling of the conventions and principles of analysis and argument;
- “A” indicates a consistently sophisticated, thought-provoking deployment of the conventions and principles of analysis and argument.

Higher grade levels should incorporate, while refining, the skills and competencies of basic levels of achievement. While there should be some flexibility in how these factors are weighed, each should constitute a substantial portion of the student’s grade on a final draft.

**Thesis Statement**

...
Performance Assessment in Higher Education: Selected Resources

The following selective resources, while not an exhaustive compilation of references, include useful information on assessment in higher education. The list was created by Joe Prus, with assistance from Mary Diez. Both were members of the Assessment Examples Project Committee.


The committee has identified three web sites to supplement the selected resources. The first is the American Association for Higher Education, which conducts an annual assessment conference and maintains a continuing assessment forum. Its website is:  
http://www.aahe.org/assessment/joint.htm

In addition, many institutions also have assessment sites with useful links. The following are two examples of such sites, the first at North Carolina University, and the second at the University of Missouri, Rolla:

http://www2.acs.ncsu.edu/UPA/assmt/resource.htm
http://web.umr.edu/~assess/edpedres/websites.htm
Appendix: Perceptions of Performance-Based Program Review and Assessment Evidence

A. THE COMMITTEE’S PERCEPTIONS

This Appendix to the Program Standards for Elementary Teacher Preparation has been prepared as an encouragement for faculty and program heads in elementary teaching preparation programs to develop their own strong performance evidence systems. Such systems provide information that can serve to judge the proficiencies of candidates and improve the institution’s own performance. A sampling and summary of that information can also provide necessary information for NCATE’s performance-based program review.

In these introductory pages of the Appendix, the Drafting Committee makes a statement about the significance of student learning in preparation of elementary teachers, and the challenge of developing and implementing a performance-based program review as part of the NCATE accreditation process. The second part of the Appendix is structured around the qualities of sound assessment systems listed in Part II, and provides additional, and in some cases more technical, detail.

A.1 Student learning and performance-based program review

Student learning is the goal. The process for quality review of teacher preparation programs should focus clearly on preparation of new elementary teacher candidates who help students learn. The previous practice, basing program review decisions on course offerings and experiences of candidates, is remote from elementary student learning, and only indirectly—although instrumentally--related to what teachers of elementary students need to know and be able to do to foster student learning.

The overall approach for performance-based program review adopted by the Committee follows the same strategy that underlies work of the National Board for Professional Teaching Standards in certifying accomplished teachers, efforts of the Interstate New Teacher Assessment and Support Consortium in developing models for teacher licensure, and the current K-12 reforms in most states—that is, set standards and measure performance in relation to those standards. Programs will be judged by their success in producing new elementary teacher candidates who can demonstrate that they know their subject and can teach so that students learn.

One might posit an “ideal” measurement system for a performance-based NCATE program review: find candidate results on commonly used evaluations that provide measures of subject knowledge and teaching accomplishment expected of new elementary teacher candidates. Such a system would undergird consistent decisions about program quality and permit fair comparisons across institutions. It would reflect consensus among professionals about the levels
of proficiencies that should be demonstrated by candidates on these commonly administered and multi-dimensional evaluations. Of course, no such “ideal” has been realized. In a nation of 50 sovereign states and 1300 colleges, departments and schools of education, the level of congruence implied by this description may never be achieved and, perhaps may not even be desirable.

What do we have instead of an “ideal?” Current teacher licensure assessments examine only part of the knowledge and skills that new teachers should acquire. State licensure tests (1) assess basic skills most frequently, (2) often examine content knowledge, and (3) sometimes measure pedagogical knowledge. The mix of components in licensure testing varies from state to state. Moreover, even when common tests are adopted, states set their own “pass” scores. There are some pilot projects that attempt to demonstrate teaching proficiencies of prospective teachers as part of the licensing or induction system, but most states have not even considered such a dimension among their licensing requirements. Only a few states and institutions have attempted to gather evidence that teacher candidates have positive effects on student learning. In summary, current testing for state licensure is a critical element in state accountability practices, one in which candidates must succeed if they are to enter teaching careers. These tests can provide necessary information about candidate knowledge, but they are not sufficient, alone, to inform about proficiencies across all the standards set for candidates—knowledge, teaching, dispositions, and ability to have positive effects on student learning.

The Committee is optimistic that higher quality teacher candidate assessments and at least some greater comparability across these evaluation instruments will be achieved over the next five to ten years. Many states are already grappling with this issue and concluding that performance-based licensure is the way to ensure that new teachers are competent. But a particularly compelling reason for the Committee’s optimism in this regard is found in the large advances in measurement of teaching competencies that have occurred over the past decade. These are partly due to investments in private sector testing but are most visible in the combined developmental efforts of the State of Connecticut, the National Board for Professional Teaching Standards and the Interstate New Teacher Assessment and Support Consortium. Another development over this same period is found in the State of Oregon, which asks that teacher candidates demonstrate “student progress in learning” as one of five standards for an initial teaching license. For many years Western Oregon University has been developing assessment technology to assess student learning during teacher preparation.

How, then, will decisions be made in the NCATE performance-based program review while state licensing requirements continue to evolve? In short, decisions will be made through comparisons: information on candidate proficiencies from an elementary preparation program will be compared with exemplars of such proficiencies identified through professional judgment. As the Drafting Committee sees it, NCATE and the specialty organizations will identify and evaluate numerous examples of assessment exercises, each with explicit descriptions indicating the level of proficiency that is appropriate to expect of candidates completing their program. These examples, which the Committee calls “benchmarks,” will be available on the NCATE and specialty organization web sites, both alerting faculty to the expectations for program review,
and guiding individuals who conduct those reviews. The Committee anticipates that institutions will gather candidate proficiency information from a variety of sources: the teacher preparation courses and field experiences, including samples of student work in classes where candidates teach; and also from external origins such as arts and sciences courses, state licensure exams and employer evaluations. Institutions will establish rubrics or criteria by which to judge candidates, then provide sampled and summarized results for the NCATE program review. Program reviewers will make holistic comparisons between the overall results from the institution and the proficiency “benchmark” levels.

Every part of accreditation makes use of professional judgment, including the current program review of courses and experiences. Web site display of the benchmark examples will permit a forum for debate about professional judgments on what should be expected of new teacher candidates. By organizing this debate in search of a professional consensus, NCATE will be able to determine both what performances are “good enough” in relation to the elementary standards and how professional colleagues will know they are good enough.

A.2 Evolution of performance concepts in NCATE

States are adopting subject content standards stating what students in grades K-12 should know and be able to do. States are also creating examination systems and establishing levels of achievement that are expected for graduation. It is not unreasonable that a similar re-examination might be necessary in teacher preparation programs that have as their goal the provision of competent teachers for every child’s classroom.

Many faculty members in institutions that meet the 1995 NCATE unit accreditation standards will find this Appendix material on assessments and monitoring of candidate performance familiar. NCATE’s standard II.C, on assessing the progress of candidates, includes accompanying indicators that state:

The progress of candidates at different stages of programs is monitored through authentic performance-based assessments using systematic procedures and timelines.

Assessment of a candidate’s progress is based on multiple data sources that include grade point average (GPA), observations, the use of various instructional strategies and technologies, faculty recommendations, demonstrated competence in academic and professional work (e.g. portfolios, performance assessments, and research and concept papers), and recommendations from the appropriate professionals in schools.

NCATE’s 1995 standard II.D on ensuring the competence of candidates extends the performance concept with the following indicators:

The unit establishes and publishes a set of criteria/outcomes for exit from each professional education program.
A candidate’s mastery of a program’s stated exit criteria or outcomes is assessed through the use of multiple sources of data such as a culminating experience, portfolios, interviews, videotaped and observed performance in schools, standardized tests, and course grades.

The Program Standards for Elementary Teacher Preparation are an application and extension of the 1995 NCATE expectations for unit monitoring and assessments. Moreover, they are an integral part of the new NCATE 2000 performance-based standards for teacher education units. Under the performance-based program review, NCATE uses results from elementary teacher candidate assessments as evidence to evaluate an institution’s elementary teacher preparation program. Assessments employed in preparation of elementary teacher candidates should provide opportunities for candidates to demonstrate what they know and can do.

Some faculty members may not be familiar with the 1995 NCATE standards or the coming NCATE 2000 unit standards. They may find that both the elementary standards and the forthcoming NCATE 2000 unit standards will encourage them to consider how well-focused their elementary preparation activities are in a performance-based context. They may even consider the possibility of making significant revisions in their instructional programs, or in the way they make use of assessments and evaluations.

Performance-based program review answers the question: “Is the institution preparing elementary teacher candidates with appropriate knowledge, teaching strategies, and dispositions to teach elementary students so those students learn and achieve standards?” Shifting NCATE’s program review to a performance-base does not lessen the importance of what institutions do to prepare candidates—to the contrary. It is not responsible to hold candidates accountable for mastery of knowledge and skills in the Program Standards that they have had no opportunities to learn and practice. Of course, opportunities to learn must be in place for both candidates and their students. Successful programs offer candidates occasions to examine relationships between the routines and activities they organize for their students and the kinds of intellectual development and learning these routines foster. Members of the Drafting Committee for the Program Standards recognize (1) that teacher candidates are required to make many on-the-spot judgments about what their students are actually learning and about what that implies for the teacher's next decisions, and (2) that student work is often unpredictable and classroom activities can be complicated. Preparing candidates to make on-the-spot judgments, to deal with unpredictable situations, and to teach content not merely for exposure, but for understanding and conceptual meaning—as many K-12 student content standards imply—are enormous challenges to any faculty. These challenges will continue to influence every decision about course structure, course content, field experiences, and advisement provided by the institution.

B. A SOUND SYSTEM FOR PERFORMANCE EVIDENCE

This section of the Appendix is constructed to assist faculty as they contemplate the implications of performance-based program review in their own institution. It begins with a statement of
principles for performance-based assessment systems from the NCATE Specialty Areas Studies Board (SASB). That statement reads as follows:

Principles for Performance-Based Assessment Systems in Professional Education Programs

Assessing what professional educator candidates know and can do is critical to implementing the performance-based standards of the National Council for Accreditation of Teacher Education (NCATE) and its affiliated national professional specialty organizations. Given the complexities of teaching and other educational professions; the range of knowledge, skills, and dispositions to be assessed; the multiple purposes for which assessment results are used; and the stakes associated with the outcomes, assessment in professional education programs and units needs to include multiple measures implemented on a systematic and ongoing basis as part of a comprehensive system. This document outlines principles set forth by the NCATE Specialty Areas Studies Board for performance-based assessment systems at the program level.

Although assessment systems will vary across programs and units, they generally should:

(a) address the knowledge, skills, and dispositions to be acquired by professional educator candidates as set forth in program goals;
(b) be consistent with the standards of relevant national and state accrediting/approval bodies;
(c) have multiple means for measuring candidate performance and impact; and
(d) provide on-going, systematic information useful for decision-making.

It is particularly critical that assessment systems provide credible results that are collected and used in a fair, valid manner consistent with their intended purpose(s).

An appropriate assessment system for a program or unit has the following characteristics:

1. The system is driven by a conceptual framework and program values that espouse assessment as a vehicle for both individual and program self-evaluation and improvement. Assessment is planned and implemented by key stakeholders in a manner consistent with the method of inquiry in the discipline and is considered a means to an end rather than an end in itself.

2. The system includes components that work together in a synergistic manner to address the knowledge, skills, and dispositions of candidates across program goals, objectives and curriculum consistent with the performance-based standards of the respective national professional specialty. Assessment is a goal-oriented process linked to program purposes/goals and national standards.

3. Multiple measures are planned and administered on a systematic, ongoing basis throughout the program beginning with the admissions process. The system includes quantitative and qualitative measures useful for formative and summative assessment. One or more measures designed to yield evidence of positive candidate impact on students are included in the system.

4. The system includes one or more measures that have been created, reviewed, and/or scored by specialty professionals external to the program. Such professionals include those with relevant specialized expertise whose primary responsibility is not to the program/unit, such as field-based master teachers, clinical teachers, intern supervisors, and/or supervisors/employers of program candidates/graduates.

5. The system is clearly delineated. Measures and associated criteria or rubrics (including minimal proficiency levels), as well as policies and practices for obtaining and using results, are described in program documents in a manner that candidates and other stakeholders can understand. Candidates are made aware of program standards and assessment requirements to which they will be held and are provided with models and/or examples of performance and the instruction and support needed to attain such levels.
6. The assessment methods and corresponding criteria included in the system are sufficiently comprehensive and rigorous to make important decisions about the proficiencies of candidates and to safeguard those they may potentially serve. Critical decision-making points are delineated in the system. Decisions that are made reflect the application of relevant criteria and use of results in a manner that discriminates acceptable versus unacceptable performance.

7. The system includes policies and procedures for the gathering, use, storage, and reporting of individual results. Such policies address the rights of individuals (e.g., those afforded candidates by the Family Educational Rights and Privacy Act; confidentiality/anonymity of survey responses). Individual candidate results are reported in a clear manner that acknowledges the source(s) and limitations of the data, individual strengths, and areas of needed or potential improvement.

8. The system includes a structure and procedures for sampling, analyzing, summarizing, and reporting aggregated results. Data are gathered on an ongoing basis and are summarized in a manner that reflects pass rates, the range of performances, and/or the “typical” or "average" performance (e.g., mean, median, or modal performance) as appropriate to the types of measures. Summaries of results are provided to key program stakeholders in a clear manner that acknowledges the source(s) and limitations of the data, data collection and reporting time frame, program strengths, and areas of needed or potential improvement.

9. The program and its assessment system foster the use of results for individual candidate and program improvement. Assessment results are regularly reviewed in relation to program goals and objectives as well as to relevant state and national standards and stimulate changes designed to optimize success.

10. The system has a mechanism and procedures for evaluating and improving itself and its component assessment methods. Evidence of the reliability and validity of the system and its component measures is gathered and used to make decisions about their ongoing use and/or revision. Evidence should address the ability of the system to comprehensively assess performance in a credible manner that is valid, fair, and unbiased.

In the remainder of this Appendix, the Drafting Committee has provided an elaboration of the six qualities of sound evidence listed in the Part II requirements for institutional submission. These are consistent with the Specialty Areas Studies Board principles, but represent those aspects of assessment systems of special concern to the Committee. To restate them, and to note their relationship with the SASB principles, the Committee describes sound evidence that is derived from assessment systems exhibiting several qualitative characteristics. The evidence:

- Results from planned, purposeful, and continuing evaluation of candidate proficiencies, drawing on diverse sources (SASB principle 1);
- Represents the scope of the standards for elementary teacher preparation (SASB principle 2 and the first part of principle 6);
- Measures the different “attributes” of standards in appropriate and multiple ways (SASB principle 3);
- Results from rigorous and systematic efforts by the institution to set performance levels and judge accomplishments of its candidates (SASB principle 5 and the last part of principle 6);
• Provides information that is credible—consistent and accurate for its intended use (SASB principle 10 and also the privacy rights portion of principle 7);

• Makes use of appropriate sampling and summarizing procedures (SASB principle 8).

The Committee has not prepared these notes as a how-to-do-it manual. The intent, rather, is to disclose to readers the understandings of Committee members about what the current assessment state-of-the-art makes it possible to achieve in responsible assessments of elementary teacher candidate proficiencies.

B.1 Results from planned, purposeful, and continuing evaluation of candidate proficiencies, drawing on diverse sources

Sound assessment systems are integrated with learning experiences throughout the teacher candidates' development and are not merely a series of unrelated "off-the-shelf" measures. They are embedded in the elementary preparation program and conducted on a continuing basis. Candidate monitoring is planned in response to faculty decisions about the points in the elementary preparation program best suited to gathering candidate performance information, consistent with the institution's own context and mission. Typically such information is gathered at candidate entry, in coursework, in connection with field experiences associated with teaching methods courses, prior to the start of practice teaching and at completion of the program. The Drafting Committee is primarily concerned that institutions prepare elementary teacher candidates to have a positive and meaningful effect on student learning. It discourages testing or use of performance measures that are administered merely to supply information for NCATE program review.

Institutions will usually begin their assessment planning around activities within the education unit. Examples of types of education unit assessments include end-of-course evaluations; tasks used for instructional purposes such as projects, journals, observations by faculty, comments by cooperating teachers, samples of student work from the candidate's teaching; and other information that would commonly be available for faculty use in determining the adequacy of the candidate's accomplishments in a course. Monitoring information from the elementary teacher preparation program can be complemented by candidate performance data originating from external sources. Examples are candidate performance evaluations during induction years and follow-up studies; performance on state licensure exams that assess candidates' knowledge of their subject content and of pedagogy, especially ones constructed to evaluate classroom teaching and effects on student learning; and academic subject knowledge end-of-course examinations, essays, projects, or other demonstrations of achievement.

Together, all information about candidates' proficiencies, from all sources, can be drawn on by the unit for continuous evaluation of candidate progress and program success. Excerpts,
summaries and samples from this array of information can be provided for use by NCATE in its program quality reviews.

B.2 Represents the scope of the standards for elementary teacher preparation;

In sound assessment systems, candidate performance evidence is congruent with the knowledge and skill standards in Part I, or equivalent ones the program sets for elementary teacher candidates.

Institutions determine the best way to demonstrate that all aspects of the standards are covered, but avoid treating each individual statement in the Part I standards and supporting explanations in an individual, serial, and fractionated way. Instead, faculty think through how all their existing assessment information can be marshaled, and what additional information they should gather, to demonstrate candidate proficiency across the standards.

Too often in citing the results of tests, both educators and the lay public report results—number scores or the proportion of test takers who “passed”--absent any reference to the content and appropriateness of the test instrument itself. It is better professional practice to align tests with instruction and to be explicit about situations where there is lack of fit. Test takers should not be held accountable for performance on assessments for which their instructional experiences have failed to prepare them. The usefulness and value of information derived from tests are the key determinants in decisions to use or exclude them from an institution’s performance measurement system. That usefulness and value depend, in turn, on the relationship of the tests to the instructional goals and anticipated candidate learning.

B.3 Measures the different “attributes” of standards in appropriate and multiple ways

The standards for elementary teacher candidates have different attributes, each of which should be measured in appropriate ways. The attributes are:

- Knowledge that candidates should possess about subject content, pedagogy, child development and learning, motivation, instruction, assessment and the qualities of a professional;
- Ability to apply that knowledge effectively in the classroom and other professional teaching situations, including collaboration with colleagues;
- Dispositions usually associated with candidates who go on to successful teaching careers; and
- Candidates’ ability to have positive effects on student learning.
Sound evidence systems construct assessment exercises that appropriately demonstrate skills described in the standards. However, one conclusion about the current state-of-the-art in teacher assessment is that no single test or measurement of teacher candidates is sufficient by itself to sample that range of skills. Multiple measures provide wide opportunities for candidates to demonstrate their accomplishments in relation to the standards. Through multiple measurements, institutions are able to combine the characteristic of “face validity” found in performance assessments with the strong psychometric properties of more traditional evaluation instruments.

Sometimes the workhorses of testing, multiple choice tests, are most appropriate. They are particularly appropriate when content knowledge and understanding, or pedagogical knowledge, are under investigation. These tests, built on decades of psychometric development, are efficient and highly reliable. Moreover, they have evolved in recent years to include vignettes with follow-up questions, and sometimes written responses of varying length, permitting more sophisticated analyses of knowledge and analytic ability.

Still, multiple-choice tests are not appropriate to measurement of other skills, such as candidate classroom instruction and other teaching abilities. The state of Connecticut, the National Board for Professional Teaching Standards, and the Interstate New Teacher Assessment and Support Consortium have been leaders in development of assessments of teaching. The methodology they have pioneered, frequently described as “teaching portfolios,” includes creation of evidence of positive effects on student learning. In documenting the act of teaching the portfolios contrast with measuring knowledge about teaching. The methodology can be adapted for use in field experiences and clinical teaching of elementary teacher candidates. Some states and a few institutions are making such adaptations. In outline, a teacher candidate portfolio would ask the candidate to do tasks similar to the following:

- Describe the students in the class, along with the school and community, to establish a context for the teaching.

- Identify three or four particular students with different learning needs for follow-up in greater detail, and describe the specific abilities and needs of these students.

- Plan a unit or set of lessons around a specific concept to move the students in the class beyond their current understanding. Describe the expectations for student learning and the subject content that the lessons are to teach. (This reveals the depth of subject content understanding of a particular topic on the part of the candidate.) Summarize the instructional approach and say, specifically, what students will be asked to do.

- Teach the unit and provide videotaped clips of sufficient length to capture different situations (e.g., lecture to the class, class discussion, small group with candidate interactions) including candidate interactions with the identified students.
• Show samples of student work for the class, and for the identified students, and analyze that work.

• Evaluate the student learning for the class with appropriate instruments and analyze the responses for the identified students specifically.

• Reflect on the experience, especially with regard to those aspects of the teaching and assessment that worked as expected, those that did not, and what candidate actions might be taken at various points to improve student learning.

Western Oregon University has developed a methodology for “teacher work sampling” containing several steps similar to the “tasks” in this list. However, Western places more emphasis on identifying learning outcomes to be accomplished in the sample of teacher work, aligning instruction and assessment so that all students will be monitored, and on the relationship of pre and post instructional measures. While the INTASC, Board and Western Oregon methodologies serve as examples of comprehensive assessments of teaching, adaptable for capstone evaluations of teacher candidate proficiencies, institutions have many opportunities throughout the course of the elementary preparation program to gather and synthesize information on candidate performances.

The remainder of section B.3 provides additional comments about each of the four attributes of the standards, excerpts examples of each attribute from the standards and supporting explanations of Part I, and lists illustrative types of assessments that might be used to examine proficiencies for each attribute. Note, however, that the types of assessments are frequently repeated across the illustrations for different attributes. For example, lesson plans, videotapes, and vignettes are included several times. Note, also, that high quality assessments will frequently provide information about candidate knowledge and skills in relation to several standards, as the “portfolio” outlined above suggests. Even less complex assessments, such as vignette exercises in a methods course, can have this cross-cutting quality. For example, a vignette might ask for essay responses to a described situation involving a concept in mathematics, presented in a way that is consistent with the development of children in a third grade classroom in which students represent widely divergent previous knowledge, and where the focus is on appropriateness of both the instruction and the assessment used to evaluate student progress. This multi-dimensional quality of assessments helps to underscore the interconnectedness of teaching—curriculum with development with instruction with assessment.

3.a Teacher candidate knowledge

Effective teaching requires mastery over the subject content that new teacher candidates will impart to their students. “Knowledge” is an area in which measurement tools such as essays, oral examples, multiple-choice tests, and semester projects can be especially useful. It is also an
area where standardized tests are available commercially that may be appropriate in circumstances where the content of the tests is aligned with the elementary program instructional experiences.

Part I includes many standards describing what new elementary teacher candidates should know. Note that the language of the Committee is “know and understand,” words used to indicate that knowledge is essential, but to understand implies an ability to analyze, use, build on, or relate that knowledge to other knowledge. Examples are:

- **From standard 1**—Candidates know (and) understand . . . major concepts, principles, theories, and research related to the development of children and young adolescents.
- **From standard 2b**—Candidates demonstrate a high level of competence in use of the English language arts . . .
- **From standard 2d**—Candidates know (and) understand . . . major concepts, procedures and reasoning processes of mathematics that define number systems and number sense, geometry, measurement, statistics and probability, and algebra . . .
- **From the supporting explanation of standard 3a**—Candidates understand learning theory, subjects taught in elementary schools (described in sections 2a through 2i of the Program Standards), curriculum development, and student development . . .
- **From standard 4**—Candidates know (and) understand . . . formal and informal assessment strategies to plan, evaluate and strengthen instruction that will promote continuous intellectual, social, emotional, and physical development of each elementary student.

Below is an illustrative and diverse list of types of evidence that can be tapped to demonstrate teacher candidate knowledge. Note in all the examples, however, that particular assessment tasks frequently provide information about candidate proficiency for more than one standard.

<table>
<thead>
<tr>
<th>Evidence Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lesson plans</td>
<td>Demonstrate the teacher candidate understands the material being taught</td>
</tr>
<tr>
<td>Multiple choice tests</td>
<td>Probe for information about concept knowledge and applications</td>
</tr>
<tr>
<td>Project reports</td>
<td>Demonstrate an understanding of concepts and problem-solving ability</td>
</tr>
<tr>
<td>Videos or feedback</td>
<td>Help teacher candidates identify naive interpretations and develop more appropriate interpretations</td>
</tr>
<tr>
<td>Vignettes</td>
<td>Exhibiting naive interpretations and lesson plans on which the candidate is requested to provide comments indicating how he/she would proceed in the lesson</td>
</tr>
<tr>
<td>Transcripts</td>
<td>Performance in appropriate course work (e.g., from courses undertaken in subject specialty) indicating level of accomplishment such as exam scores, projects completed, essays prepared</td>
</tr>
<tr>
<td>Written essays</td>
<td>On the content matter demonstrating abilities to develop a topic, write well, clarify questions</td>
</tr>
</tbody>
</table>
Examples of assignments that teacher candidates would prescribe as a consequence of their teaching (and perhaps elementary student work, including teacher feedback)

3.b Teaching performances

Elementary teachers not only know and understand content but are able to relate it to ideas, information, and knowledge previously learned. They know how to teach using a variety of methods, how to adapt their teaching to the subject being taught, and how to engage diverse students in the subject. Candidates who meet these performance competencies will be able to provide evidence of positive effects on student learning.

Here are some examples of teaching standards from Part I:

- **From standard 2b**—Candidates. . . use concepts from reading, language and child development to teach reading, writing, speaking, viewing, listening, and thinking skills. . .
- **From the explanation for standard 2i**—In their instruction, candidates make connections in their instruction across the disciplines and draw on their knowledge of developmental stages to motive students, build understanding, and encourage application of knowledge, skills, and ideas to lives of elementary students across fields of knowledge and in real world situations.
- **From standard 3a**—Candidates plan and implement instruction based on knowledge of students, learning theory, subject matter, curricular goals, and community.
- **From the explanation for standard 3b**—Candidates know how to seek assistance and guidance from specialists and other resources to address elementary students' exceptional learning needs. . . .
- **From standard 5b**—Candidates are aware of and reflect on their practice in light of research on teaching and resources available for professional learning; they continually evaluate the effects of their professional decisions and actions on students, parents, and other professionals in the learning community . . . .

There are many ways to demonstrate that elementary teacher candidates meet the performance aspects of the standards. In addition to the “portfolio” approach described above, here are some illustrative types of teacher performance measures:

**Elementary student impact**
- Evidence of elementary student learning
- Evidence of “surface” and “deep” elementary student learning
- Case studies that candidates are asked to analyze
- Elementary student projects showing evidence of ability to make use of information
- Student essays demonstrating ability to state and elucidate ideas

**Graduates’ success**
| State licensure test pass-rates and results on induction year portfolio assessments |
| Graduate surveys |

Artifacts produced by the teacher candidate
- Lesson plans
- Teaching aids
- Videos
- Feedback on student work
- Vignettes
- Case studies
- Assessment materials

Reflective essays on candidate-prepared lessons
- Written essays
- Reflections on student work emanating from the lessons
- Examples of assessments
- Journal entries
- Interviews

Attestations of teaching accomplishment
- By cooperating teachers
- By students
- By parents
- By principals
- Summative evaluations
- Transcripts of course performance
- Multiple choice tests

Observations
- Videotapes of classroom instruction
- Micro-teaching
- Reactions to videos
- By cooperating teachers
- By teacher preparation faculty
- By other teachers
- By principals
- Videos of talks to various audiences

Professional artifacts
- Parent communications
- Letters to parents
- Surveys of elementary students, parents and colleagues
- Record of professional activities
- Community activities
- Awards, recognitions

### 3.c Dispositions

Dispositions refer to values and commitments and often make the difference between what elementary teacher candidates understand and how they perform in a classroom. Examples of dispositions that are significant in teacher preparation might include:
A belief that all students can learn, a vision of high and challenging standards, a commitment to personal professional development and to a safe and supportive learning environment, an ability to accept responsibility, an understanding of school operation as an integral part of the larger community, an acceptance of families as partners in the education of their children, and a dedication to bringing ethical principles into decision-making processes.

Such values and commitment are important elements in successful teaching and may frequently be of particular relevance at the unit level rather than for individual preparation programs.

One example of dispositions critical to effective teaching was the focus of a paper prepared for the Committee by Mary Kennedy. In stating the importance of “changing one’s conception of teaching,” Dr. Kennedy claims that:

The unusual nature of teacher learning is such that students entering teacher education already “know” a great deal about their chosen field. Moreover, they will use what they already know to interpret any new skills or new theories they acquire during the formal study of teaching. This fact means that the simple acquisition of new skills or theories is not adequate to alter teaching practices. Therefore, the central task of teacher learning must be to change these conceptions.

(Candidates) need to be persuaded that school subjects consist of more than the facts and rules they themselves learned as children. Teachers conceptions of subject matter as fixed, indisputable, and factual, need to be replaced with conceptions that recognize ambiguous concepts and tentativeness, and that acknowledge that even young children are capable of reasoning about and arguing about ideas in each school subject. Unless teachers envision subject matter as conceptual, they cannot teach it conceptually. And . . . teacher educators need to address and alter teacher candidates’ strong desire to control student behavior, for the desire to develop management routines that keep students on task and in line is frequently a stumbling block to implementing conceptual approaches to teaching.

As evidence for her expression of dispositions, Kennedy suggested the following indicators:

- Evidence that teacher candidates adopt conceptual goals for teaching school subjects;
- Evidence that teacher candidates justify their lesson plans and their approaches to teaching according to the concepts they want students to learn, and that these concepts are included in national subject matter standards;
- Evidence that teacher candidates are aware of teaching practices such as reciprocal teaching, cognitive apprenticeships, and the writing process, which are defined according to what is learned rather than only according to how teachers behave; and
Evidence that programs monitor the teaching practices of cooperating teachers.

It is difficult to identify measures for such indicators, and, indeed, it is challenging to establish measures of proficiency for any dispositions. Over time, however, institutions in which dispositions are explicit and important will find ways to demonstrate that candidates have achieved them. The Committee encourages faculty in elementary preparation programs, and the leadership of the unit, both to state values and commitments toward which completing elementary candidates should be disposed, and to search for appropriate ways that candidates’ achievements in these areas can be exhibited.

There are few explicit references to dispositions in the Part I standards as written by the Committee, largely because of the measurement problems noted above. But here are some:

- *From the introduction to Part I*—New candidates for elementary teaching must be committed to elementary students and their learning. They must be prepared to act on a belief that all elementary children have potential for learning rigorous content and achieving high standards.
- *From the supporting explanation of Standard 1*—(Candidates) consider diversity an asset and respond positively to it.
- *From the supporting explanation for standard 2e*—Candidates are able to use knowledge, skills, and dispositions from social studies to organize and provide integrated instruction in grades K-6.
- *From the supporting explanation for standard 2h*—Teacher candidates recognize the critical importance of physically active life styles for all students. Teacher candidates appreciate the intrinsic values and benefits associated with physical activity.
- *From the supporting explanation for standard 5b*—Candidates respect parents’ choices and goals for their children and communicate effectively with parents about curriculum and children’s progress.

3.d Positive Effects on Student Learning

Elementary teacher candidates are expected to gather examples of their elementary students’ work. Those examples can illustrate that candidates’ knowledge and teaching performances result in evidence of positive effects on students’ achievement.

The Committee seeks a focus on student learning and it expects that program review submissions will sample and summarize what faculty monitoring has disclosed about learning among students of candidates in the elementary preparation program. It does not expect “representative samples” of work of a candidate’s students. Nor does it expect Part I standards to be interpreted as efforts to hold elementary teacher candidates responsible for specific student gains on state achievement tests. In their field experiences, teacher candidates often are placed in other teachers’ classes for short periods, are given limited control over the choice of curricula, must
adapt to the teaching style of the cooperating teacher, and rarely have students for sufficient time to see other than very short term effects.

But student learning, as noted elsewhere, is the goal. It is appropriate to devise standards that direct attention to student learning. It is also appropriate to know whether teacher candidates can find ways to understand the level of accomplishment of their students, to use that knowledge as a basis for design of instruction for a particular objective, to identify and apply suitable measures of effects on student learning as a result, then finally, to reflect on the whole sequence and hypothesize how the instruction might have been more effective.

Here are some examples of Part I standards that call for positive effects on student learning:

- **From standard 2a**—Candidates . . . can create meaningful learning experiences that develop students’ competence in subject matter and skills for various developmental levels.
- **From the supporting explanation for standard 2b**—(Candidates) teach students to read competently . . . . Candidates teach students a variety of strategies to monitor their own reading comprehension . . . . They help students think critically about what they read.
- **From standard 2d**—Candidates . . . foster student understanding and use of patterns, quantities, and spatial relationships that can represent phenomena, solve problems and manage data.
- **From the supporting explanation for standard 3d**—(Candidates) create learning communities in which elementary students assume responsibility for themselves and one another, participate in decisionmaking, work collaboratively and independently, and engage in purposeful learning activities.

In the box below some illustrative types of evidence to examine positive effects on student learning are listed:

<table>
<thead>
<tr>
<th>Evidence Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student work indicating material was learned and understood by elementary students</td>
</tr>
<tr>
<td>Student work and reflections on how this work relates to the lesson as taught</td>
</tr>
<tr>
<td>Teacher lesson plan based on understandings of the content domain, and teacher awareness of the common misconceptions that need to be addressed</td>
</tr>
<tr>
<td>Feedback on students’ work that illustrates how the student could improve and advance</td>
</tr>
<tr>
<td>Elementary students’ expressions that the activities are personally relevant, are connected to previous information learned, are interesting and challenging.</td>
</tr>
<tr>
<td>Examples of elementary student work demonstrating skills in integrating learning and generalizing</td>
</tr>
<tr>
<td>Student scores on achievement tasks throughout the time the teacher has been teaching and indications of how they are used to improve and enhance teaching</td>
</tr>
</tbody>
</table>
Evidence about the performance of previous teacher candidates (graduates) after they have taught a few years

B.4 Results from rigorous and systematic efforts by the institution to set performance levels and judge accomplishments of its candidates

Institutions that systematically conduct evaluations of candidate proficiencies also determine performance levels by which candidate success can be judged. They address the question: “How good is good enough?” The terms “rubrics” and “criteria” are frequently used to designate levels of performance. Rubrics and criteria are narrative descriptions of what is valued in a candidate’s response to an assessment—the qualities by which levels or elements of performance can be differentiated—and serve as anchors for judgments about the degree of candidate success. They may be stated in generic terms or may be specific to particular assessment tasks. They may define acceptable levels of performance for the institution and one or more levels below (such as borderline, or unacceptable) and above (such as exemplary), or they may be in the form of criteria defining the institution’s conditions or expectations for success\textsuperscript{v}. The rubrics or criteria are “public,” shared with candidates and across the faculty.

Institutions with sound assessment evidence systems judge individual candidate proficiencies, and also summarize and analyze the proportions of new teacher candidates who reach levels or conditions expressed in the rubrics or criteria. These results are used both for advisement of individual candidates, and also for strengthening of the courses and experiences offered by the institution to prepare elementary teacher candidates. The summary of results from the faculty judgments in applying the rubrics or criteria are used for the NCATE submission.

B.5 Provides information that is credible—accurate, consistent, fair and avoiding bias

The Committee has specified in Part II that the program and unit needs to demonstrate the “credibility” of their assessment information. That term was intended as less technical than the traditional “valid and reliable” (or accurate and consistent) but was expected to encompass those qualities, along with fairness, and avoiding bias.

Accuracy, or validity, is an expectation that the assessment information measures what is important for the decision to be made and represents the performances, competencies, and dispositions that are intended (that is, included in standards for elementary teacher candidates). Institutions with sound assessment systems gather and make use of defensible evidence that their assessment activities validly portray the proficiencies of their elementary teacher candidates. Linn and Miller (1986) made the following important comment about validity:

> It is widely agreed that validity is the most important consideration in the evaluation of the use of a test. Validity is always a matter of degree. It is not a single all-or-none
concept. Rather, the concern is with the degree to which the accumulated evidence supports a particular test use. Many forms of evidence may be relevant in evaluating the validity of a particular test use, and it is not possible to give a simple prescription for the forms or adequacy of the evidence in the abstract. Professional judgment is required to determine the forms of evidence that are most appropriate in a given situation and to judge the adequacy of the support for the intended purpose.

A core concept to the notion of validity is representativeness, that is, the degree to which the assessment task models the construct, and the degree to which it samples the universe of the construct. When choosing assessment procedures for whatever purpose—evaluating candidate achievement, determining program or course effectiveness, preparing for NCATE program approval—it is necessary to address the extent to which the assessments “represent” the performances, competencies, and dispositions, such that the process does not lead to:

- Construct under-representation (the assessment does not capture the important aspects of the construct), or
- Construct irrelevance (the assessment measures something other than what was intended by the construct).

Consistency, or reliability, in institutional assessment systems is an expectation that successive samples of performances from the same candidate are reasonably related. Assessment systems must also be fair; for example, they must be based in opportunities to learn provided by the curriculum and those, in turn, must reflect the standards for teacher candidates. They should avoid bias, providing equitable treatment for all candidates. Making judgments about these matters requires professional expertise and is often determined through peer review, evaluations by external experts, or formal validation studies.

B.6 Makes use of appropriate sampling and summarizing procedures

In preparing the elementary program submission, the institution samples and summarizes information about candidate proficiencies.

Sampling refers both to representing the domain of the standards and representing the full range of the program’s candidates. The candidate sample might be taken from the cohort of teacher candidates completing the program in a specific academic year and previous completers so that information about performance of candidates from their entire preparation experience and into employment can be available for demonstration of candidate proficiency. Of course, anonymity of individual candidates and the students of those candidates must be protected.

Candidate proficiency results can be summarized through averages, spread of scores, and distributions of rubric scores. Summary results are requested because the NCATE interest is in making decisions about program quality, rather than decisions about individual candidates.
These summaries are made meaningful through illustrations such as samples of exam questions, examples of written responses, and analytic materials intended to inform reviewers of the proficiencies that candidates achieve in relation to the standards.

In demonstrating candidate proficiencies as the basis for program review, the guide is: ask enough, but not too much. Two key assumptions underlie the discussion of sampling evidence:

- It is reasonable to assume that all institutions will be able to demonstrate excellence. The key issue is whether this excellence is present among all, or most, of its teacher candidates. Merely asking an institution to provide evidence of excellence is insufficient, as there may be pockets of excellence, as opposed to excellence across all, or a sufficient number of, students. Thus, attention needs to be given not only the nature and level of performances desired, but the satisfactory sampling of this excellence.

- Given that the aim of *Program Standards for Elementary Teacher Preparation* is to raise expectations, then considering only teacher candidates within an institution who are at the “cut-score” point of excellence may lead to overemphasis of the minimally qualified candidate at the expense of considering and promoting the very best.

It is not necessary to sample all aspects of the standards. Sampling some of the standards relative to some of the students in a systematic way is advantageous. The amount and degree of sampling may be related to prior performance by the institution. For example, institutional programs previously designated as “nationally recognized” may require less sampling, whereas those considered “borderline” in previous program reviews may require more sampling.

Whatever sampling methods are used, a series of performance indicators could be constructed to evaluate the quality of the assessments, such as:

- Which content knowledge and skills are intended to be assessed by the performance assessments(s)?

- How adequately do the assessment procedures cover knowledge and performances?

- What evidence is provided to demonstrate that faculty appropriately assess candidate work at levels expressed in the institution’s rubrics for “acceptable” proficiency?

- What is the level/depth of the very best, as well as marginally acceptable, candidate work, with perhaps a sample of other work?

---

State School Officers, INTASC.

2. A publication that both describes and evaluates this topic is *Grading Teachers, Grading Schools: Is Student Achievement a Valid Evaluation Measure?* Edited by Jason Millman, Corwin Press, Inc., 1997. Among the examples of teacher and school evaluation in this book is the Oregon “Teacher Work Sample Methodology” as it has been created and nurtured by Western Oregon University.


4. Two specific approaches for evaluating levels of performance—or judging how good is good enough—came to the attention of the Committee in the course of its work. These are “SOLO” and “scoring rubrics.” These two methods provide some sense of convergence about establishing levels of performance, while each of may be more applicable to particular standards or institutional interests.

a. SOLO taxonomy.

The “Structure of the Observed Learning Outcomes” model, or SOLO, describes four consistent sequences, or levels. These levels are ordered in terms of various characteristics: from the concrete to the abstract; an increasing number of organizing dimensions; increasing consistency; and the use of organizing or relating principles. “SOLO” was developed to assess qualitative outcomes, which may be applied to evaluate learning quality in a wide variety of school and college situations, in most subject areas. The four levels are:

- **Unistructural:** One aspect of a task is picked up or understood serially, and there is no relationship between facts or ideas
- **Multistructural:** Two or more aspects of a task are picked up or understood serially, but are not interrelated
- **Relational:** Several aspects are integrated so that the whole has a coherent structure and meaning
- **Extended abstract:** The coherent whole is generalized to a higher level of abstraction

Here is an example of a performance standard based on the SOLO taxonomy: *It is expected that elementary teacher candidates know the content material to be taught, and can teach that content to elementary students in an integrated and coherent way.*


b. Scoring rubrics

A second approach to setting levels of performance that can define how good is good enough is through scoring rubrics for each assessment in relation to the associated standard(s). The example here is based on work of the National Board for Professional Teaching Standards.

Elementary teacher candidates can be classified on, at least, three levels:
• **Accomplished teacher candidates:** Elementary teacher candidates have accurate and deep understanding of the content relevant to each standard as exemplified in their performances as teachers—for example,

The “accomplished” or level 3 performance provides clear, convincing, and consistent evidence that the elementary teacher candidate has knowledge of the content described in each standard, has the proficiencies to apply that knowledge to teaching situations, has enthusiasm and attitudes appropriate to successful teaching, and can have a positive impact on the learning of all his/her students with respect to the content specified in the standard.

• **Satisfactory teacher candidates:** Elementary teacher candidates demonstrate accurate understanding of the content relevant to each standard as exemplified in their performances as teachers—for example,

The “satisfactory” or level 2 performance provides clear evidence that the elementary teacher candidate has knowledge of the content described in each standard, has the proficiencies to apply that knowledge to teaching situations, has enthusiasm and attitudes appropriate to successful teaching, and can have a positive impact on the learning of all his/her students with respect to the content specified in the standard.

• **Unsatisfactory teacher candidates:** Elementary teacher candidates demonstrate limited and surface understanding of the content relevant to each standard as exemplified in their performances as teachers—for example,

The “unsatisfactory” or level 1 performance provides limited or little or no evidence that the elementary teacher candidate has knowledge of content described in each standard, has the proficiencies to apply that knowledge to teaching situations, has enthusiasm or attitudes appropriate to successful teaching, or can have a positive impact on the learning of all his/her students with respect to the content specified in the standard.

An example of a performance standard based on scoring rubrics is: *It is expected that elementary teacher candidates have sufficient knowledge about the content material to be taught to elementary students that they would be classified at Level 2, Satisfactory teacher candidates.*