

## Section

### 3

# Task Development

*Section Overview:* This section focuses on validity and reliability of the task and rubric provided in section 2. From the beginning of the project in June 2002, the ECSU workgroup was aware that designing a valid and reliable assessment task and rubric would be a difficult task. To ensure validity and reliability, it was imperative that the workgroup consist of experts from every field (Biology, English, History, and Math) that included education faculty, arts and sciences faculty, and 7-12 educators. Furthermore, the ECSU work group included an assessment expert. In addition to the expertise of workgroup members, the development of assessment task and rubric at ECSU received constant support from the Connecticut State Department of Education.

As mentioned in Section 2 of this report, the assessment task, rubric, and response packet were pilot tested with teacher candidates in four subject areas: 2 candidates in Biology, 8 candidates in English, 13 candidates in History/Social Studies, and 7 candidates in Mathematics. The data from each of the 30 secondary teacher candidates who participated in the pilot study were then analyzed and the findings were used to further improve the design of the deliverables. (Note that 8 of these 30 pilot responses are provided in Section 2.) The assessment tools and their rubrics were finalized after everyone in the team had the opportunity to review them and provide feedback. Because the project team was comprised of five education faculty members, four arts and sciences faculty members, and four educators from local schools, widespread collaboration was ensured.

Eastern faculty members, from both the Education Department and the School of Arts and Sciences, who participated in this project, have been involved in preparing various NCATE documents in one form or another. Hence, the team consisted of people who were familiar with NCATE standards as well as discipline-based professional standards (in particular, the standards from NCTM (Math), NSTA (science), NCTE (English), and NCSS (History/Social Studies)). In addition, the four participating 7-12 educators brought to this project their considerable experience teaching in Connecticut middle and high schools. All workgroup members reviewed each deliverable to ensure that the intent of the developed deliverables was consistent with the goals of secondary education, thus enhancing the validity of the deliverables. In addition, all the deliverables developed by the team were further reviewed by Dr. Kathy Lake from Alverno College, a nationally known expert in competency-based assessment. Reliability was enhanced because all three members in each subgroup (Biology, English, History, Math) individually scored candidate responses and discussed the differences among their scores. Actually, the team found that the greater the clarity of the task and rubric the better the validity and reliability. Extensive discussion following the pilot test and subsequent improvement from the pilot task and rubric to the current task and rubric, have vastly improved the validity and reliability of these deliverables.

**Validity:** Validity of an assessment task and rubric should be measured in terms of the following questions:

- Do the assessment task and rubric measure what they intend to measure?
- Do they provide valuable data about candidates' content, pedagogical, and professional knowledge, skills, and dispositions necessary to become effective student teachers in the secondary subject areas of Biology, English, History, and Mathematics?
- Are these deliverables defensible in terms of their usability and value?

In order to increase their validity, the assessment task and rubric went through several cycles of development and revision. This was a tedious process, but the enhanced validity of the deliverables made it worth the effort.

In addition to questions of validity, the team continually asked whether drafts of assessment tasks meshed with NCATE and discipline-based professional standards and whether the task was appropriate for teacher candidates prior to student teaching. The team also studied ECSU's conceptual framework for its teacher preparation program. The ECSU conceptual framework defines six critical strands in which each teacher candidate is required to be proficient: content knowledge, pedagogical knowledge, integration, technology, diversity, and professionalism (See diagram on the next page.). As a result, the proposed assessment task is closely aligned with ECSU's conceptual framework, NCATE standards, and discipline-based professional standards. This connection is expanded upon in Section 4 of this report.

The following are the specific outcomes that the proposed assessment task intends to measure.

### **ECSU Conceptual Framework Competencies Measured Through this Assessment Task**

#### **1: Content Knowledge (CNK)**

- 1.1 Candidates/Graduates demonstrate in-depth understanding of content knowledge including central concepts, principles, skills, tools of inquiry, and structure of the discipline(s) they teach.
- 1.2 Candidates/Graduates are able to formulate clear and meaningful questions about the content to engage students in learning.

#### **2: Pedagogical Knowledge (PDK)**

- 2.1 Candidates/Graduates are able to identify developmentally appropriate learning goals and objectives for students based upon knowledge of subject matter, students, the community and curriculum goals (both state and national) and to plan instructional activities which foster individual and collective inquiry, critical thinking, and problem solving to facilitate learning for all students.
- 2.2 Candidates/Graduates use methods, activities, and grouping arrangements appropriate for lesson goals and objectives.

# ECSU Education Unit Conceptual Framework



- 2.3 Candidates/Graduates conduct learning activities in a logical sequence, which is flexible and developmentally appropriate to the needs interests, ability, and background of students.
- 2.4 Candidates/Graduates use various assessment techniques to evaluate student learning and modify instruction as appropriate to ensure the continuous intellectual, social, moral, and physical development of the learner.
- 2.5 Candidates/Graduates demonstrate enthusiasm, patience, acceptance, and caring about the well-being of students and promote life-long learning, perseverance, self-motivation, and scholarly habits of mind.
- 2.6 Candidates/Graduates appreciate individual variation within each area of development, show respect for the diverse talents of all learners, and help them develop self-confidence and competence.
- 2.7 Candidates/Graduates value the development of students' critical thinking, independent problem solving, collaborative inquiry, and performance capabilities as important tools for success.
- 2.8 Candidates/Graduates are committed to using multiple assessment techniques to identify student strengths and promote student growth rather than to deny students access to learning opportunities.

### **3: Integration of Knowledge (INT)**

- 3.1. Candidates/Graduates promote connections between content knowledge and pedagogical knowledge to help students learn concepts, principles, skills, tools of inquiry, and structure of the discipline(s) they teach.

### **4: Infusion of Educational Technology (TEC)**

- 4.1. Candidates/Graduates use a variety of print, visual materials, manipulatives, media, and electronic resources for exploration and development of concepts, principles, and skills associated with the content they teach.

### **5: Diversity (DIV)**

- 5.1. Candidates/Graduates demonstrate acceptance of, and respect for, individual differences and talents among students, including students from different gender, ethnic, racial, socioeconomic, language, and religious groups and create supportive learning environments for all students to maximize their learning and develop independence, social competence, and positive self-concept.
- 5.2. Candidates/Graduates believe that all students can learn at high levels and persist in helping all students achieve success.

### **6: Professionalism (PRF)**

- 6.1. Candidates/Graduates display professional and ethical behavior in their work with students, colleagues, parents, families, and communities.

**Table 1. Alignment of Unit’s Conceptual Framework With State and Professional Standards**

<b>Eastern’s Conceptual Framework</b>	<b>CCCT</b>	<b>INTASC Principles</b>	<b>NBPTS Propositions</b>	<b>NCATE Standards</b>
<b>Content Knowledge (CNK)</b>	Content 3, 4	Principle 1, 7	Proposition 2	1.1, 1.2, 1.6, 3.3
<b>Pedagogical Knowledge (PDK)</b>	Students 1, 2 Pedagogy 5, 6 Planning 1, 2 Instructing 3, 4, 5, 6 Assessing and Adjusting 7	Principle 1-10	Proposition 1, 2, 3	1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 3.1, 3.2, 3.3
<b>Integration (INT)</b>	Content 3,4 Planning 1, 2	Principle 1, 4, 7,	Proposition 1, 2, 3	1.1, 1.2, 1.6
<b>Technology (TEC)</b>	Instructing 5	Principle 6		1.6, 3.3
<b>Diversity (DIV)</b>	Students 2 Pedagogy 6 Professional and Ethical Practice 1, 2	Principle 3, 5, 7, 8	Proposition 1, 2, 3, 4	1.6, 1.7, 3.3, 4.1, 4.2, 4.3, 4.4
<b>Professionalism (PRF)</b>	Professional and Ethical Practice 1, 2 Reflection and Continuous Learning 3, 4 Leadership and Collaboration 5, 6	Principle 7, 9, 10	Proposition 2, 3, 4, 5	1.6, 1.7, 1.8, 3.3, 4.4

Furthermore, all of the candidate proficiencies as outlined above in Eastern Connecticut State University's conceptual framework are aligned with the Connecticut Common Core of Teaching, INTASC principles, and NCATE standards. (See the alignment chart on the next page.) Because the proposed assessment task is closely aligned with ECSU's Education Unit Conceptual Framework and because the conceptual framework is aligned with the Connecticut Common Core of Teaching, INTASC principles, and NCATE Standards, the assessment workgroup strongly believes that the task for teacher candidates developed through the project is valid.

The validity of this assessment task is further insured because the workgroup carefully planned that the candidates would have the opportunity to learn these concepts and skills in their content and methods courses prior to their student teaching. In particular, teacher candidates take methods courses in the Education Department prior to the semester of their student teaching. Because this assessment task will be administered during methods courses in late October, the candidates have sufficient opportunity to learn the needed pedagogical skills and concepts in September and early October. In particular, candidates will have an opportunity to design a lesson that meets the requirements of this assessment when they take the *Learning and Teaching* course. Thus, this assessment task takes into consideration the candidate's level of preparation (i.e., we do not expect them to demonstrate knowledge of or skill in something they have not had the opportunity to learn).

Clearly, the proposed assessment task and rubric are vastly improved in terms of clarity compared to the pilot task and rubric. Every workgroup member, as well as outside experts (Kathy Lake from Alverno College, Georgette Nemr, Title II Director, Connecticut State Department of Education, and Irv Richardson from TBA Consulting) spent a substantial amount of time reading and analyzing every word of the task and rubric. In addition, several of the teacher candidates from the pilot study were interviewed after they completed the assessment task. Feedback from the teacher candidates helped pinpoint confusing task prompts and unclear expectations. Furthermore, every team member, outside expert, and interviewed teacher candidates thought that the assessment task was an important step in preparing the candidates for successful experiences in student teaching. Some reflections and reviews from various subgroups and reviewers, as well as candidates who participated in the pilot, appear in Appendix A.

**Reliability:** Reliability is measured in terms of the consistency in scoring the assessment task. The reflections about the pilot assessment task and rubric are provided in Appendix A. There was extensive discussion among workgroup members on nearly every word used in the assessment task and rubric during the course of this project. From these discussions scorers got a pretty good understanding of what the assessment task demanded from the candidates. Thus, it was somewhat surprising that even with this extensive discussion, the scoring results in the pilot study were not highly consistent. For example, the math pilot scores in two indicators "Analysis of Student Work" and "Feed back to Students" from all the three scorers did not show the consistency as expected. The scores on these two indicators from the math workgroup follow.

## Math Pilot Scores

### Analysis of Student Work

	Scorer A	Scorer B	Scorer C	Consensus
Candidate 1	3, 2, 3, 2	2, 3, 3, 3	2, 2, 2, 3	2, 2, 3, 3
Candidate 2	3, 1, 1, 1	2, 2, 2, 2	1, 2, 1, 1	2, 2, 1, 1
Candidate 3	2, 1, 3, 1	2, 2, 1, 1	1, 1, 1, 1	2, 1, 1, 1
Candidate 4	2, 1, 3, 2	1, 2, 2, 2	1, 1, 2, 1	1, 1, 2, 2
Candidate 5	2, 2, 3, 2	2, 2, 2, 2	2, 1, 1, 1	2, 2, 2, 2
Candidate 6	2, 2, 2, 2	1, 1, 1, 1	1, 1, 1, 1	1, 1, 1, 1
Candidate 7	3, 1, 1, 2	2, 2, 2, 1	2, 1, 2, 1	2, 1, 2, 1

### Feedback to Students

	Scorer A	Scorer B	Scorer C	Consensus
Candidate 1	2, 2, 2, 2	3, 3, 2, 3	1, 1, 2, 2	2, 2, 2, 2
Candidate 2	1, 1, 1, 1	2, 1, 1, 1	1, 1, 0, 0	1, 1, 1, 1
Candidate 3	1, 1, 2, 1	1, 1, 1, 1	1, 0, 0, 0	1, 1, 1, 1
Candidate 4	2, 1, 2, 1	2, 2, 2, 1	1, 1, 2, 1	2, 1, 2, 1
Candidate 5	1, 1, 3, 2	2, 2, 1, 2	1, 1, 1, 1	1, 1, 1, 2
Candidate 6	1, 2, 2, 2	1, 2, 1, 2	1, 0, 0, 1	1, 1, 1, 2
Candidate 7	2, 1, 1, 1	1, 1, 1, 1	1, 0, 1, 0	1, 1, 1, 1

Because of these scoring discrepancies, the mathematics subgroup held extensive discussions on what the problems were. Similar discussions were held in other subgroups as well. These discussions lead to revision of the wording of assessment task prompts and further development of the rubric. (Reflections on discussions related to scoring problems can be found in Appendix A.) Due to the improved clarity of the assessment task prompts and the close revision and more detailed rubric, the workgroup believes that consistency in scoring should no longer be a substantial problem. Unfortunately because no candidate was available in January when the task and rubric were revised, the improved task and rubric could not be administered for a second pilot. The group, therefore, attempted to use the revised task and rubric to score old pilot responses. Even though this is not a fair scoring, the consistency between the scorers did improve on indicators that remained same. We expect that the consistency among the scorers will improve substantially when we implement the revised task and rubric in Fall Semester 2003.

**Scoring Protocols:** The above descriptions and the documentations provided in Appendix A illustrate that the development of the rubric to evaluate the assessment task was an extensive endeavor. The current revision has utilized candidate performances at various levels. As

described above, the revised task and rubric have been enhanced both in terms of validity and reliability.

In Fall Semester 2003, the revised task and rubric will be fully implemented. The workgroup members who participated in this project will score these candidates' work. Because every workgroup member has committed to this project through fall 2003, the training of scorers to establish the consistency will not arise.

In the future, we plan to extend this competency-based assessment to other areas of certification. Hence, as new people are enlisted to expand ECSU's competency-based assessment of teacher candidates, they will need to be trained before scoring candidates' work. The workgroup members in the Education Department will be responsible for training future evaluators. Both the pilot task and rubric and the newly revised task and rubric in conjunction with candidates pilot responses will be used to train future evaluators. In addition, plans for training sessions will be updated as new data and materials become available. The Dean of Education and Professional Studies is very supportive of this idea.