

## Section

### 3

# Task Development

As the overview history in the first section of this report indicates, we have been piloting midpoint assessment tasks since Fall 2001 and exit portfolios since fall 2002. Our pilots have been done with CCSU students at the point in the program when the assessment will be given. From the start, we have been concerned about both validity and reliability of the assessments, as we knew that they would be important to student progress through the program.

## Validity

### Midpoint Planning Task

To insure the validity of assessments, we spent considerable energy and time creating alignment charts showing the standards CCSU candidates are expected to meet upon program completion. Those charts formed the basis for all future development work. The first task we developed was the midpoint planning task. Working in a large CCSU faculty group that included representatives from each content area, we reviewed the standards alignment charts and began extensive discussions of what skills, knowledge, and dispositions were crucial for students entering student teaching. When we arrived at consensus about what skills, knowledge, and dispositions were essential across all content areas, we began to think about how these might be assessed. We were able to sort a set of related knowledge and skills that centered on planning and analysis of both content and student work. (We concluded that dispositions really needed to be assessed in a different way and set those aside.)

All our work has undoubtedly been influenced by the CCSU faculty members who had worked in the BEST program. Therefore, it should be no surprise that we concluded that we could create an assessment that used real student work to simulate the planning tasks faced by student teachers. As we discussed the demands of student teaching, we concluded that we wanted the task to be completed independently and in a timed setting, as timely completion of plans was often a problem for less effective student teachers. As a group, we hammered out the basic sections of a task and developed initial questions. The development of task questions and directions required careful attention to language as not every term was used in the same way in every content area. Each content area indicated the NCATE folio outcomes that the task aligned with. These are noted on the task sheet.

To insure that the task would validly assess the desired skills and knowledge, it was important that every prospective CCSU student teacher had had adequate opportunity to learn the K-12 content used in the prompts. Content faculty in each program identified k-12 content that every teacher candidate had had multiple opportunities to learn. In elementary, for example, we identified math problem solving strategies as content that was used in different ways in Math 113, 213, and 412. Science identified process skills that were relevant and known by students in every certification area of science. History selected a task that any history or social studies student would have the background to address. At that point, we collaboratively developed a format for the content area prompts. Each content area then developed a first prompt and had it reviewed by the entire group.

The first pilot was small, using selected portions of the groups in the pre-student teaching semester in fall 2001. Each content area tried to select pilot participants who represented the range of performance in the group. We used the student performances on the fall 2001 pilot to help us refine the task and develop the initial draft of the rubric. That group's feedback led us to provide each student with hard copy of all prompt materials and a one page copy of task questions. In spring 2002, we eliminated one section of the task (reflection) and made significant language changes in the task based on the responses students gave in the initial pilot. We did no formal interviews or surveys but each content area faculty proctor asked students who completed the pilot assessment for their feedback.

The same cycle of revision based on student performance data has occurred throughout the piloting process. In the fall 2002 pilot, we used whole groups and used the scoring process to further refine the rubric and make some minor language changes in the task. In spring 2003, we used whole groups and made the task and rubric available to students before testing. The spring 2003 group completed the task on WebCT rather than simply in Word. This pilot has led us to plan a slightly longer period – the web-based format seemed to slow students down a bit. We think 2 hours and 30 minutes may need to be available rather than 2 hours. In May 2003, we also made some wording changes in the rubric that we believe will lead to better alignment of expectations and scores across content areas on items 1 and 7.

We have informally examined the relationship between performance on the midpoint assessment and performance in student teaching. With the exception of some failures in student teaching that were clearly related to dispositions, which the task does not purport to assess, it appears that students who do poorly on the assessment do poorly in student teaching, especially in the area of planning. A formal study of these data is planned for summer and fall 2003. The redesigned student teaching evaluation should yield more reliable student teaching data and facilitate meaningful comparisons of performance.

#### Exit Portfolio

The exit portfolio task was developed by the cross-disciplinary team in May 2002. The portfolio task was initially piloted in fall 2002 in elementary education with every elementary student teacher (50 students). In January 2003, we used those pilot portfolios to develop a rubric and to refine the task. The spring 2003 pilot with every elementary student teacher (75 students) and science student teachers (4 students) enabled us to further refine the task and rubric.

## **Reliability**

### Midpoint Planning Task

We have worked on scoring consistency for the midpoint planning assessment in three ways. First, we decided that every response needed to be scored by two raters in the content area and that all responses should be scored “blind.” Second, we did our initial pilot scoring and a good bit of fall 2002 scoring in “camp” where we all were able to look at what we were scoring and compare our ratings. This process helped us refine our rubrics and identify difficult spots in scoring. In fall 2002, we designed and piloted formal scorer training. The agenda for that scorer training is included in this report. Spring 2003 responses were scored by scorers trained in fall 2002.

We have not yet done a formal study of interrater reliability. We have consistently used collaborative scoring, scorer conferences, and analysis of student performance patterns to look for scoring issues. Since the student performance data suggested that there were differences in the expectations held in different content areas for items 1 and 7 on the rubric, we also worked in May 2003 to clarify the rubric language to insure that expectations were consistent across content areas. We have asked scorers for spring 2003 to submit scores independently and expect to formally assess interrater reliability for spring 2003 during summer 2003.

### Exit Portfolio

At this point, we have scored exit portfolios collaboratively with an emphasis on task refinement and rubric development. We expect to employ the same reliability processes for the portfolio as we have for the midpoint assessment.

## **Scoring Protocols**

The development of the both midpoint and exit rubrics was driven by student performances and described in previous sections. The development of the rubrics and the refinement of the tasks have been intertwined. Plans for scorer training are included in the previous section.