

January 30, 2007

## **2006/2007 PRAC Assessment Grant Summary of Progress**

### **“Secondary Education Benchmark Development”**

**Submitted by Paula Magee and Jenny Conner**

#### **Introduction**

The PRAC grant is being used to support the work on the first stage of a two-stage project involving the development (Stage 1) and validation (Stage 2) of a secondary education assessment protocol and instrument (video clip and questions, referred to from now on as “the benchmark”). The benchmark examines pre-service teachers’ (participants) abilities to meet established Teacher Education learning outcomes at a large urban university in the Midwest. Its task involves watching and analyzing a video clip of a teaching episode. This type of activity is referred to as the video case method (Beck, King, & Marshall, 2002). Stage 1 of the project focuses on the development and refinement of the instrument, protocol, and the scoring rubric. During Stage 2, the validity of the instrument will be determined. Ultimately, the results of this benchmark will help assist teacher education faculty in identifying strengths and weaknesses of the secondary education program.

#### **Developing and Refining the Instrument, Protocol, and Rubric**

The first stage of this project (the stage supported by the PRAC grant) is almost completed. During this stage, the benchmark was field-tested (FT) three times – once in September 2005 (FT 1), once in April 2006 (FT 2), and once in December 2006 (FT 3). Funding from PRAC supported the second and third field tests.

#### **Field-Test 1**

January 30, 2007

Sources of data collected included the field-tested benchmark student products and post-benchmark questionnaires.

A Language Arts lesson was presented to ten pre-service secondary education student participants via video clip. Participants analyzed this lesson through a series of prompts (the benchmark) prepared by the researchers.

To ensure standardization in the delivery of the instrument participants completed the benchmark in a computer lab at a pre-determined time. Each participant had access to a computer with Microsoft Word and Windows Media Player. Each was supplied with a copy of the video taped lesson (on CD), headphones, and a hard copy of the directions, which included the prompts. Participants had the opportunity to view the video clip as many times as necessary. Responses were typed into a Microsoft Word document and all participants had one and half hours to complete the prompts. Upon completing the benchmark, participants filled out a questionnaire that asked them to reflect on various aspects of the experience. These responses were also typed into a Microsoft Word document.

The benchmark results and questionnaires were analyzed to determine the practicality of the protocol procedures as well as the depth of student responses related to pedagogical skills and knowledge. Conceptual categories were created and themes from the data, through an analysis where “induction and deduction are in constant dialogue” (Erikson, 1986), were identified. In other words, researchers looked for patterns that “emerged” from the data. Results facilitated changes in the instrument and protocol prior to FT 2 (see “Results” section).

### Field-Test 2

Sources of data collected included the field-tested English Language Arts benchmark student products from five participants (using the revised instrument), post-benchmark

January 30, 2007

questionnaires, and focus group data. The computer lab protocol followed during FT 1 was also followed during FT 2, except that participants were given two hours to complete FT 2.

In a focus group setting, the five FT 2 participants were asked to discuss the extent to which the task provided them with the opportunity to express their depth of knowledge. Additionally the participants were asked to comment on the logistics of the benchmark experience.

Data from the FT 2 were analyzed in the same manner as FT 1 data.

### Field-Test 3

Prior to this third field testing session, a video clip to be used for the social studies benchmark was identified. Sources of data collected for this field test included five products from the English Language Arts benchmark and five products from the social studies benchmark.

## **Results**

Stage 1 of this project focused on developing and revising the benchmark instrument. The original instrument was field-tested, and factors that seemed to negatively affect the quality of student responses were identified. Results are based on benchmark results from all fifteen participants involved in FT 1 and FT 2, questionnaire data collected from these same participants immediately after completing the assessment, and the focus group data. Analysis of this data allowed researchers to identify the following as key factors that affected practicality (logistical issues) and effectiveness (depth of pedagogical knowledge) of the benchmark.

### 1. Writing Prompts and Task Directions

The data suggest that the way in which writing prompts are developed, as well as the general guidelines given for responding to the prompts, significantly affects students' tendencies to

January 30, 2007

provide more than mere generalizations about the lesson they have analyzed. For FT 1, prompts were somewhat leading and attempted to focus students on specific teaching skills (i.e., “Describe the impact of the type of questions that the teacher asked on the lesson.”). Directions for overall task completion, however, were very vague (“Type your responses to the questions below.”) This resulted in responses with very little depth that made sweeping generalizations about the sixteen minute teaching episode on the video clip without any specificity to support the assertions. Very few specific examples from the clip were used to illustrate the assertions, and the analyses tended to lack any evidence that students had more than a commonsense notion of good practice in the classroom. During FT 2, when the writing prompts were left more open-ended (“Describe and analyze a one to three-minute segment of the video clip.”), but the task directions included an indication of expectations for the depth of responses (“Your work will be evaluated for its ability to analyze and reflect on the video clip from a pedagogical perspective.”), the quality of the responses improved significantly.

## 2. How and when students are introduced to the assessment

For FT 1, students were simply told (verbally) approximately one week prior to completing the assessment that the task would involve viewing a video clip of a teaching episode and then answering questions about what they saw. There was no explanation of the purpose of the task and students weren’t given enough information about how their work would be evaluated to prepare for the task. In addition to producing work that lacked depth and suggested a very superficial knowledge base, questionnaire data suggested that participants felt that their minimal knowledge and understanding about the nature of the task adversely affected their abilities to complete it in a way that adequately reflected their true knowledge and skills. For FT 2, the

January 30, 2007

protocol for completing the assessment as well as a detailed written explanation for the purpose of the assessment and how results would be evaluated were given to the students several weeks prior to the task. This same document was e-mailed to the students a few days prior to the day of the assessment, to remind them of its contents and impress upon them the importance of the information it provided. The document also invited students to bring any materials to the assessment that they wanted (i.e., books and notes) to support their completion of the task. This revised introduction to the task seemed to affect not only the quality of the products, but also resulted in participants feeling more prepared for and confident about the task. Questionnaire and focus group data suggested that students at least perceived that they knew what was expected of them while analyzing the video clip. And although most reported not using the materials they brought with them to help them complete the task, it seems that simply being invited to bring whatever textual support they wanted helped them understand the purposes and expectations of the assessment task.

### 3. The set-up and protocol of the task itself

According to questionnaire and focus group data, the task set-up and protocol seemed to support the quality of the results in that they allowed each student to work through the task at their own discretion, helped students feel comfortable, and supported the goals of the task. Students were allowed to view all or portions of the video clip as many times as they wanted, and they could do so privately, without other students or the assessment facilitator knowing how often or which portions of the clip were being reviewed. Additionally, they could easily go back and forth between the video clip and the word processing document on which they were writing their responses to the prompts. In addition to student comfort, this seemed to help students understand

January 30, 2007

that the goal of the task was not to provide a general sense of the teaching episode, but to study and analyze portions of the clip.

The social studies and English Language Arts benchmarks completed by students during the third field test are currently being used to develop and refine a rubric. That work will be completed by June 2007.

### **National Conference**

We will be presenting our progress thus far on the benchmark at the American Educational Research Association conference in Chicago this April. Money from the PRAC grant will be used to support these efforts.

### **REFERENCES**

Beck, R. J., King, A., & Marshall, S. K. (2002). Effects of videocase construction on preservice teachers' observations of teaching. *The Journal of Experimental Education*, 70(4), 345-361.

Erikson, F. (1986). Qualitative methods in research on teaching. In (Wittrock, Ed.) Handbook on Research on Teaching. 3<sup>rd</sup> edition. New York: Macmillan.