

Section 1: Introduction

The content knowledge of future teachers and their competency in using instructional strategies (pedagogical skills) was assessed with a task and scoring rubric designed specifically to understand pre-service teachers' abilities to use and teach mathematics. The Performance Assessment in Mathematics (PAM) was developed and piloted as part of a joint project between members of the Department of Mathematics and the Department of Education and Educational Psychology at Western Connecticut State University along with local mathematics specialists and was supported by funding through the Connecticut State Department of Education (Refer to Appendix A for a copy of the accepted proposal: Outcome Based Assessment in Elementary and Secondary Teacher Preparation). The purpose of the instrument is to assess pre-service teachers' abilities to use appropriate language related to mathematics, to apply appropriate instructional methods and materials, and to diagnose mathematical problems of students while imparting accurate mathematical concepts using the standards of the National Council of Teachers of Mathematics (NCTM) and the Connecticut Competencies for Teachers (CCT).

The NCTM standards address five content areas that should be included in every mathematics lesson (NCTM, 2001): communication, connections/transfer, problem solving, representations, and reasoning/proof. To assess teachers' abilities to use these standards, the instrument presents released items from either the Connecticut Mastery Test (CMT) or the Connecticut Academic Performance Test (CAPT) and directs respondents to: (a) formulate ideal answers for the selected mathematical problem (Content Knowledge), (b) identify national and state standards, as well as content and skills related to the problem (Content Knowledge), (c) write what should be said to students, instructing them about how to accomplish the mathematical task (Pedagogical Content Knowledge), and (d) address responses from students in

order for the teacher to diagnose any difficulties in mathematics as well as to redirect the students in understanding the mathematical concept (Student Learning and Assessment). Refer to Appendix B for the Directions for the final version of the PAM task along with student responses to the elementary and secondary versions of the task. The most recent rubric is located in Appendix C. The task and rubric used to pilot test the PAM in October Are located on Appendix D and E. Student responses are also included in this Appendix E. Appendix F includes the task, rubric, and student responses related to the summer 2002 pilot test of the PAM.

To assess responses to the PAM, a scoring rubric was developed and refined at each stage of pilot testing. For the final version, each of the four sections of the PAM can be examined indicating whether or not the performance standard is Fully Met (3), Partially Met (2), Not Adequately Met (1), or has No Response (0). Appendix C provides a copy of the final version of the rubric.

Both the task format and rubric were developed by a group of seven faculty members from WestConn's Mathematics Department in the School of Arts and Sciences and the School of Professional Studies' Department of Education and Educational Psychology. The teaming of faculty from both schools resulted in group-ownership over the products, resulting in the establishment of this performance task and rubric as teaching tools in mathematics and education courses. A second team included two WestConn education faculty members as well as mathematics specialists from three local school districts. This latter group was called upon to provide input into the development process for the task and rubric, objectively review the task and rubric, aid in the selection of mathematics problems for the task, develop exemplary responses to the task, and use the rubric to score responses of pre-service teachers.

Validity and reliability for any instrument are always being pursued through an evolving process. Content validity related to this project begins by assessing the component parts of the task and relating its content to established standards in education. An alignment table in Appendix G parallels the expected exemplary performances of the four parts of the PAM with categories from the following standards: the National Council for the Accreditation of Teacher Education (NCATE) (2002); the National Council of Teachers of Mathematics (NCTM) Content Standards (2002); the NCTM Professional Standards for Teaching Mathematics (1991); the Connecticut Common Core of Teaching (1999), the Connecticut Competencies for Teachers (CCT) III (2002); the Beginning Educator Support and Training Program (BEST) 2001-2002 Rubrics (2001-2002); the BEST Program Portfolio Resources: Guiding Questions (2001-2002); and the Interstate New Teacher Assessment and Support Consortium (INTASC) Standards (2002).

This performance task was developed over a period of one and a half years. During this time, three pilots studies were initiated using each of the three iterations of the task. Input for refining the task was obtained from the task and rubric developers, local mathematics specialists, university faculty members, the students who completed the task, Connecticut State Department of Education grant monitors, and from a group of independent grant evaluators. These many comments served to improve the validity of the instrument.

Plans for obtaining data to contribute to criterion-related validity will include correlations of four types of scores. Performance ratings on the PAM will be correlated with students' grades in their mathematics preparation courses as well as in a pre-student teaching course involving methods and materials for the classroom, and with an observational rating of a mathematics lesson during student teaching.

The construct under investigation is the ability to effectively teach mathematics. An initial examination of construct validity will begin by administering the PAM to individuals at different periods in their preparation for teaching in order to record changes in the measured concept over time, within and between groups. Students will be assessed when they enroll in their first mathematics courses, while still completing their general education requirements, when they are completing their professional semester of educational methods courses and field experiences, during student teaching, and during their beginning year of teaching.

As a result of a pilot study, inter-rater reliability indices for each section of the PAM were calculated between mathematics specialists in local schools, members of WestConn’s mathematics faculty, and WestConn’s education faculty. These data will continue to be examined with future administrations of the PAM assessment tool.

It is planned that the construct related to this performance assessment tool will continue to be integrated into mathematics and education courses at WestConn and that the instrument will be administered to all teacher certification candidates during the professional semester, immediately prior to student teaching. In the event of an evaluation indicating that the standards are “Not Adequately Met,” an individual will be required to participate in a review workshop and to retake the assessment.

The List of Proposed Deliverables. The original deliverables related to this project are indicated below. All deliverables, as originally stated, have been completed. The location of each response within this document is represented in Table 1.

Table 1

Proposed Deliverables and Responses.

Proposed Deliverables	Response to Deliverable
1. The investigators will work with the university administration to support and encourage cooperation between faculty in the School of Arts and Sciences and the School of Professional Studies.	Section 1: Introduction, <i>Members of the Workgroup and Their Institutional Roles</i> ,
2. The investigators will align NCATE, NCTM, INTASC, and CCT standards with Western Connecticut State University Classes. Designing this alignment will provide the necessary direction that team members will use to complete the project in a meaningful way.	Appendix G
3. The investigators will identify performance outcomes based on NCTM standards in elementary and secondary education. At least one aspect of the task will be related to each NCTM process standard (communication, connections, problem-solving, reasoning and proof, and representations) to assess the students' performance.	Appendix B
4. The investigators will develop a rubric to be used in the analysis of the performance tasks for elementary and secondary mathematics education.	Appendix C
5. The investigators will create a template to help translate the task to other disciplines.	Appendix L

Members of the Workgroup and Their Institutional Roles. This project was completed by two teams. Team 1 consisted of university personnel from the School of Arts and Sciences' Mathematics Department (n=3) and The School of professional Studies' Department of Education and Educational Psychology (n=4). This was the primary workgroup, which met on a regular basis, beginning with weekly meetings (spring 2002 and summer 2002) and continuing

with biweekly meetings (fall 2002 and spring 2003). A consultant from the Connecticut State Department of Education attended these meetings, as her schedule allowed. Conceptualizing the task and rubric were the primary goals of this team. All members worked together sharing perspectives related to curriculum, instruction in mathematics, child development, educational psychology, theories of adult learners, and psychometrics.

The five members of Team 2 were consulted to review various versions of the task and rubric, to assist in the selection of mathematical problems for the task, to score the rubric, and to provide general feedback about the characteristics of a teacher of mathematics. Their advice was sought via the Internet, through the regular mail, at designated meetings, and through personal conversations at the university or in local schools. The professional role of each team member is provided in Table 2.

Table 2
The Professional Roles of Members from Team 1 and Team 2 in the Development of the PAM.

School/ Department/ Team	Group Member	Institutional Role
Professional Studies: Department of Education and Educational Psychology		
Team 1	Dr. J. Burke	Professor of Instructional Technology for Undergraduate Students in the Teacher Certification Program
	Dr. T. Cordy	Department Chair
	Dr. M. Daria	Undergraduate Program Coordinator

School/ Department/ Team	Group Member	Institutional Role
Team 2	Dr. M. Delcourt	Graduate Program Coordinator, Interim Director of the Ed.D. Program in Instructional Leadership
	Dr. T. Canada	Assoc. Professor of Education, Educational Psychology for Undergraduate Students in the Teacher Certification Program and for Graduate Students in the Counseling Program
	Dr. L. Stambler	Coordinator of Student Teaching
Arts and Sciences: Mathematics Department		
Team 1	Dr. D. Burns	Assoc. Professor of Mathematics, Education Advisor, Coordinator of Secondary Education Programs in Arts and Sciences
	Dr. R. Kutz	Department Chair
	Dr. P. Maida	Assoc. Professor of Mathematics, Coordinator of the Math Lab
PK-12 school faculty		
Team 2	Ms. R. Finney	Department Chair of Mathematics at Ridgefield High School
	Ms. A. Fishman	Department Chair of Mathematics at Brookfield High School
	Ms. R. Santos	Elementary School Mathematics Specialist, Danbury Public Schools

Specific Outcome Events from the Spring of 2002 Through the Spring of 2003

Initial Planning Period March: April 2002

In this phase we determined the nature and scope of the performance task. Several meetings of the core group took place in which we discussed the program and how a performance task might fit with our current curricular sequence. A significant moment occurred when Dr. Kutz shared an assignment that he has been using in his Elementary Methods course for the last 10 years. The members of the workgroup were immediately drawn to this assignment that had a distinct performance character. We made plans over the next few meetings to try this “task” with a group of students to determine the type and quality of the information we could gather. At the end of May, we administered the assignment to a group of Juniors in the Elementary Education Certification Program. Refer to Appendix F for a copy of this initial task. This provided us with a collection of responses to use as fuel for our summer discussions. Copies of the task were provided to our extended workgroup partners for review. We also encouraged them to provide us with any information they thought to be relevant for revising the task and scoring the responses.

Specific Development of the Task and Scoring Rubric: May – August 2002

The summer months were devoted to the review of the performance data gathered in the spring and to the generation of a set of performance criteria. One of our first realizations was that the task did not provide a rigid enough structure for gathering data that were comparable and consistent. At the same time, we realized that we needed to determine the qualities of data that we were hoping to illicit. By the end of the summer we able to complete a performance table that gave four levels of performance in six target areas. We also had a better idea of how the task

ought to be framed to support the data we hoped to collect. By the end of August, the energy of the workgroup had been devoted to the elementary performance task. The thinking was that the template would expedite the creation of the secondary performance task, since the only needed change would be the actual problem involved with the task.

Revision and Piloting: September – December 2002

As the school year started again we had high hopes of piloting our task in September. This would hopefully give us enough time to score and revise the task while developing the secondary task. The writing of the new task took longer than expected because it turned out to be lengthier and more complicated than anticipated. We were unable to share the task with our extended workgroup partners until October. After more revision, we finally piloted this revision in November of 2002. Scoring with our performance table was somewhat problematic and several meetings were used to discuss the scoring and exactly how the table should be used. Also during October and November, we began to assemble the secondary task. After some discussion, we confirmed our original belief that the six areas included in the task were as important at the secondary level as they were at the elementary level. Work began on a project that was parallel to and similar to the elementary task. As we neared the December 13 progress report meeting, we were confident that the underlying concepts related to our elementary task were on schedule and not as confident about the secondary task. Subsequently, we requested and were granted a deadline extension.

At this time, we also wanted to streamline the task and to use standard problems for the task. Thus, we abandoned the idea of fitting the task into a lesson format, as this was too cumbersome, and we adopted items from the CMT and CAPT for use in our performance task. These decisions helped us to revise the task to its present format.

Revision and Final Preparation: January – March 2003

The December 13 meeting was a turning point for many on our group. We came away from the day convinced that the central concept of our task was sound, though it had become clouded and confusing to explain. In late December, we met to discuss some structural improvements that essentially laid out what was to become the current task. At the same time, we were convinced that the two tasks (elementary and secondary) were more similar than different. In January, we met in a day-long session with Georgette Nemr to work through the details of this revision as well as to revise the rubric. The rest of the spring was spent revising this task and the rubric.

In order to examine the content of the final task, all members of Team 1 and Team 2 provided exemplary responses to at least one of the tasks. In other words, we each completed the assignment we were asking our certification candidates to complete. Team members with strengths at a specific grade level (elementary or secondary) completed the task at that level, while others completed both the secondary and elementary tasks. These responses will serve as initial benchmarks in the category of “Fully Met” until satisfactory responses are obtained from the teacher candidates.

A review of the rubric was obtained from each team member as he or she applied the rubric to student data that was already collected in the summer of 2002. While assessing an old version of the task with a revised rubric was not an ideal match in content, this exercise provided insight into the usefulness of the rubric as well as an opportunity to collect initial estimates of inter-rater reliability. At the end of March we were ready with two fully developed performance tasks and a scoring rubric. We needed only final approval to proceed with the pilot.

Additional Piloting and Writing of the Final Report: April – May 2003

After the task was administered to a sample of elementary certification students and a sample of secondary certification students, all members of the two Teams scored the completed assignments. The reliability coefficients for these data were recorded and the final report was written (Documents related to meetings and relevant correspondences can be found in Appendix H).