

# Testing the Problem-Solving Skills of Students in an NCTM-oriented Curriculum

Carmen M. Latterell

---

An interesting question concerns how well NCTM-oriented students do on standardized mathematics tests. Another important question that has received less attention is: Are standardized tests truly measuring the skills that NCTM-oriented students have? Would other tests reveal skills that differentiate NCTM-oriented students from traditional students? Moreover, what are these skills? This paper contributes to the answers to these questions, and finds that students in one NCTM-oriented curriculum displayed such qualities as engagement, eagerness, communication, flexibility, and curiosity to a much higher degree than traditional students did. The implication given is that not only should we use standardized tests, but we should revise them and/or supplement them to measure the qualities that are not currently being measured.

The National Council of Teachers of Mathematics (NCTM) has been promoting standards for teaching secondary mathematics (NCTM, 2000; NRC, 1989). At the same time, many curriculum projects have been developed for the purpose of providing NCTM-oriented curricula for teachers (e.g., the NSF-funded curricula). However, the implementation of these NCTM-oriented curricula has not been without problems (Curcio, 1999). In fact, the term “math wars” has been used to describe the controversies (Senk & Thompson, 2003b; Schoen, Fey, Hirsch, & Coxford, 1999). Some of these controversies surround the issue of standardized testing (Hancock & Kilpatrick, 1993). In fact, some parents have indicated concern that if their children are in NCTM-orientated curricula, they will be at a disadvantage on standardized tests (Senk & Thompson, 2003a). Of course, standardized tests might change in time, but it is an interesting question whether NCTM-oriented students actually are at a disadvantage on standardized tests. Perhaps it is an interesting question because standardized testing should change in time. In other words, the question of interest is not simply “Are students in NCTM-oriented curricula at a disadvantage in standardized testing?” but rather “Are standardized tests the best measure of everything that students in NCTM-oriented curricula can do?” And then if the answer is that other tests can reveal abilities that standardized tests do not, what are these other abilities?

---

*Carmen M. Latterell is an assistant professor of mathematics at the University of Minnesota Duluth. Her research interests include the testing of mathematics. This interest includes the types of questions seen in this article as well as such issues as placement testing into undergraduate mathematics courses.*

Standardized achievement tests are viewed as “blunt” measuring instruments by some (Kilpatrick, 2003, p. 479). Certainly issues are more complex when standardized tests attempt to measure students' problem-solving ability. Since problem solving is regarded as a process (NRC, 1989; NCTM 2000) and not a product, standardized tests must be well constructed to measure problem solving adequately. While researchers say that standardized tests cannot measure problem solving, it is important to remember that testing is “value laden and socially constructed” (Gipps, 1999, p. 370). Are standardized tests measuring the abilities that we value for students in NCTM-oriented curricula?

Now that NCTM-oriented curricula have been put into action for several years, researchers are able to examine these issues with research studies. This study contributes to the literature by examining some of the issues involved in the testing of students in a particular NCTM-oriented curriculum, the Core Plus Mathematics Project Curriculum (CPMP). The overarching research question is how to better assess students who are in a NCTM-oriented curriculum to measure as fully as possible their problem-solving abilities.

To answer this, some comparisons between NCTM-oriented students and traditional students are made, as well as some comparisons between NCTM-oriented students' results on different types of tests. The reader is cautioned that the intent of this study is not to compare NCTM-oriented students to traditional students, per se. Many studies (Senk & Thompson, 2003a) have already been done (this paper will review some of them as applies to CPMP in a later section), giving evidence that students in NCTM-oriented

curricula can perform as well as students in traditional curricula on standardized tests.

The current study attempts to examine the subtleties of testing students in a NCTM-oriented curriculum. For example, although these students might do as well as others on standardized tests, are there other tests that are a better measure of these students' problem solving abilities? Or even, are there other tests that will measure abilities that standardized tests do not and that traditional students may not have? And, perhaps most important, what is the nature of these abilities that are not (if in fact they are not) being measured by standardized testing?

The research questions are:

1. How do the scores of students in CPMP compare to students in a traditional curriculum on a standardized problem-solving test?
2. How do the scores of students in CPMP compare to students in a traditional curriculum on a parallel constructed-response problem-solving test?
3. Are there differences that qualitative data can illustrate between the manners in which pairs of students in CPMP and pairs of students in a traditional curriculum solve non-routine constructed-response problem-solving items?

#### Method

The basic method was to give a standardized problem-solving test to students in CPMP and to comparable students in a traditional curriculum. Effort was made to have similar ability students in both sets of curricula before entering the curriculum, using Iowa Test of Educational Development (ITED; Feldt, Forsyth, Ansley, & Alnot, 1993) data. In addition, a test parallel to the standardized test but with constructed-response items was also given to both sets of students. Finally, both sets of students were given a performance assessment. This section describes the curricula, the students, a survey of the classroom environments, and the tests.

#### *The CPMP Curriculum*

As stated earlier, the NCTM-oriented curriculum used CPMP, the Core Plus Mathematics Curriculum. The author had access to students in the state of Iowa, and a large number of schools in Iowa had implemented CPMP for several years. Therefore, CPMP was chosen as the curriculum to represent NCTM-oriented curricula. Yet, no claim is made that CPMP is better or worse than other NCTM-oriented curricula, or that the results, therefore, would be the same under other curricula. Regardless, the researcher

judged it would be better not to have more than one NCTM-oriented curriculum and thus possibly confound the results.

CPMP has developed student and teacher materials for a three-year high school mathematics curriculum for all students and a fourth-year course for college bound students. The main theme of CPMP is mathematics as sense making. Students investigate problems set in real-life contexts within an integrated curriculum that includes algebra and functions, geometry and trigonometry, statistics and probability, and discrete mathematics. The curriculum for each year is seven units with a capstone section which is “a thematic two-week, project-oriented activity that enables students to pull together and apply the important mathematical concepts and methods developed in the entire course” (Schoen & Ziebarth, 1998b, p. 153). Mathematical modeling is emphasized throughout the curriculum. Graphing calculators are used. Additional characteristics of CPMP are that the curriculum is designed to be accessible to all students, to engage the students in active learning, and to provide multidimensional assessment (Hirsch & Coxford, 1997). Assessment opportunities are embedded within the curriculum and include students' answers to questions in class, groupwork, student journals, quizzes, in-class and take-home end-of-unit assessments, cumulative written assessments, and extended projects (Hirsch, Coxford, Fey, & Schoen, 1995).

Inclusion of topics in the CPMP curriculum is based on the merits of the topics themselves; that is, the topics must be important in their own right (Schoen et al., 1999). The instructional sequence follows a four-step process labeled as launch, explore, share and summarize, and apply. The “launch” sets the context for what is to follow and consists of a class discussion of a problem. The “explore” is usually a group or pair activity in which students investigate the problems and questions. “Share and summarize” brings the class back together to discuss key concepts and methods. “Apply” is time in which individual students practice what they learned (Hirsch et al., 1995).

CPMP authors have conducted numerous and extensive research into the effectiveness of CPMP. At the end of Course 1, CPMP students' average score on the *Ability to Do Quantitative Thinking* (a subtest of the nationally standardized Iowa Tests of Educational Development) was significantly higher ( $p < .05$ ) than algebra students in traditional curricula (Schoen & Hirsch, 2003). At the end of Course 3, CPMP students performed significantly better ( $p < .05$ ) on concept and

application tasks but significantly poorer on algebraic manipulation tasks when compared with Algebra II students in traditional curricula (Huntely, Rasmussen, Villarubi, Sangtong, & Fey, 2001). Using SAT 1, mathematics scores of CPMP III students versus Algebra II students showed no significant differences (Schoen, Cebulla, and Winsor, 2001). When American College Testing (ACT) Assessment Mathematics Test means were used, the Algebra II students had significantly higher scores (Schoen, Cebulla, and Winsor, 2001). Using placement tests constructed with items from a test bank from the Mathematical Association of America, CPMP Course 4 (N=164) versus Precalculus (N=177) students showed no significant differences on algebraic symbol manipulation skill but a significant difference in favor of CPMP on concepts and methods needed for the study of calculus (Schoen & Hirsch, 2003). Another study examined proof competence, as well as perceived need for proof using CPMP III students versus traditional Algebra II students; no significant differences were found (Kahan, 1999).

This research suggests CPMP is a curriculum that is at least as effective as a traditional curriculum except in some by-hand manipulation skills when these skills are needed outside of application context and graphing calculators are not allowed. In many areas (such as problem solving in context and with graphing calculators), CPMP students outperform traditional students. The design of the CPMP curriculum suggests a curriculum that is indeed in alignment with NCTM recommendations. The research indicates CPMP is effective in promoting the achievement of student reasoning, communication, problem solving, and representation. In traditional skills (such as by-hand manipulation skills), the CPMP curriculum might be less effective than traditional curricula. In less traditional skills (such as problem solving in multiple representations), the CPMP curriculum might be more effective than traditional curricula (Schoen et al., 1999; Schoen & Ziebarth, 1998a).

### *The Traditional Curriculum*

There is, of course, no such curriculum as “traditional curriculum” and as NCTM’s standards become increasingly respected, even the most traditional of curriculum begins to look like NCTM-oriented curriculum. The researcher fully acknowledges that there is no true dichotomy between NCTM-oriented curricula and traditional curricula any longer, but rather a continuum.

For the purposes of this study, I chose to define the traditional curriculum to not be one of the NSF-funded, nor identify itself in any manner as reform. The traditional curriculum should lack an emphasis on groupwork and graphing calculators, while emphasizing hand symbolic manipulations as the only meaning of algebra. This approach moved the researcher to the traditional end of the continuum, but once there, could traditional curriculum be described in and of itself?

The traditional curricula in this study have an emphasis on separate units of mathematical content (in our case, algebra), which includes an emphasis on procedures (although conceptual understanding is present as well). The teacher most often serves as the “teller” of information (i.e., the students were not engaged in discovery work). Students most often work individually. Testing is usually easily accomplished through short-answer or even multiple-choice items. Computation is important. In our case, since these were traditional algebra classes, the curriculum included solving (by hand) algebraic equations with solution processes not dependent on technology.

### *Students*

A sample of five Iowa schools using CPMP agreed to participate in this study. As previously mentioned, within the state of Iowa, there were many schools that had implemented CPMP. The researcher had access to a list of all Iowa schools that had CPMP in place for several years. A subset of these schools was identified and contacted, so that the subset were schools spread throughout Iowa. Five of these schools agreed to participate. This resulted in 230 students.

The researcher also had access to existing ITED data. This data was gathered for the schools that had already agreed to participate. Another sample of Iowa schools was gathered according to two characteristics: The size and location of the schools had to match one of the five already in the sample, and the mathematics curricula had to be of a traditional nature. The existing ITED data was used to narrow this sample to schools whose average school score was the same as those already in the sample. This resulted in a much smaller set of schools, but out of this set, five did agree to participate. This resulted in 320 students.

### *The Classroom Environment Form*

At this point, the teachers from each classroom completed a Classroom Environment Form (written by the researcher) to provide data on such interests as the textbook, availability of graphing calculators,

groupwork practices, and assessment practices (see Table 1). The purpose of gathering this data was two-fold.

One purpose was to supply background and context for the discussion of the results. Although the purpose of this study was not to compare CPMP students to traditional students, it is through that comparison that the researcher is able to distinguish differences in testing results that may be created as a consequence of NCTM-oriented curricula. However, if the intended curriculum (whether that is CPMP or traditional) is not the enacted curriculum, then the worth of this comparison is questionable. (The researcher acknowledges that there is also the achieved curriculum.)

The Classroom Environment Form was also intended to supply possible explanations for results. An example might illustrate this. CPMP students work in groups as part of the curriculum. If it turns out that CPMP students work better under testing practices that include groups, perhaps that is reasonable. However, just because CPMP authors call for students to learn in groups, do the students actually spend time in groups? Further, perhaps the traditional students also learn in groups, and thus groupwork is not really a difference

between the two curricula. Clearly the Classroom Environment Form will not necessarily give a full disclosure of a classroom. It will only give what the teacher chooses to say.

### *The Standardized Test*

A standardized problem-solving test, the subtest from the Iowa Test of Educational Development (ITED) titled *Test Q: Ability to Do Quantitative Thinking*, was used. The stated purpose of the ITED is to “provide objective, norm-referenced information about high school students’ development in the skills that are the long-term goals of secondary education—skills that constitute a major part of the foundation for continued learning” (Feldt et al., 1993, p. 4). Test Q at level 15 (grade nine) consists of 40 multiple-choice items with five response options. Students are given 40 minutes to complete the test.

The questions, based on realistic situations, are “practical problems that require basic arithmetic and measurement, estimation, data interpretation, and logical thinking” (Feldt, et al., 1993, p. 13). In addition, some of the questions test more abstract concepts. “The primary objective of the test is to

Table 1  
*Comparison of the Two Types of Classrooms*

	NCTM-oriented	Traditional
Textbook	CPMP	<i>Algebra I Explorations and Applications</i> , McDougal Littell, 1997 (3 classrooms)  <i>Algebra I Applications and Connections</i> , Macmillan/McGraw Hill, 1992 (2 classrooms)
Graphing calculators	constantly available	One teacher answered constantly available. The remaining teachers stated that they were available sometimes during classes and on tests, but less frequently during homework.
Mean percent estimate for the amount of class time that students are in groups	86%	22%
Groupwork outside of class	Four teachers encouraged students to work in groups outside of class.	Three teachers encouraged their students to work in groups outside of class.
Multiple-choice tests used	Never used	2 teachers sometimes used them.
Students receive partial credit for work and explanations.	All teachers said “always”.	All teachers said “frequently”.
Students receive credit for writing out reasons on graded assessments.	All teachers said “always”.	All teachers said “frequently”.
Students are encouraged to give reasons and not just an answer.	All teachers said “always”.	All teachers said “frequently”.

measure students' ability to use appropriate mathematical reasoning, not to test computational facility under pressure" (Feldt et al., 1993, p. 13).

### *The Constructed-Response Test*

ITED has two parallel forms. One of the forms was given intact as the previous test. To create the second test, the remaining form was modified in the following manner: The researcher converted each item to a constructed-response item. For most items, this simply meant eliminating the choices and keeping the stem. For items in which the choices completed a sentence, the item was changed to form a question. For a small number of items, the item was completely reworded, but the purpose of the item and the context remained the same. The researcher scored the constructed-response tests using a scoring key in which each problem is worked including numerical calculations and a verbal explanation.

### *The Performance Assessment*

Due to time constraints, a subset of the sample consisting of two classrooms of CPMP students and two classrooms of traditional students solved problems in pairs using non-routine items while the researcher observed. This resulted in approximately 100 students. The following two problems, each written out on separate pieces of paper, were given to pairs of students.

1. How many keystrokes are needed to put page numbers on a paper of length 124 pages?
2. Three friends, returning from the movie Friday the 13th Part 65, stopped to eat at a restaurant. After dinner, they paid their bill and noticed a bowl of mints at the front counter. Sean took  $\frac{1}{3}$  of the mints, but returned four because he had a momentary pang of guilt. Faizah then took  $\frac{1}{4}$  of what was left but returned three for similar reasons. Eugene then took half of the remainder but threw two that looked like they had been slobbered on back into the bowl. (He felt no pangs of guilt—he just didn't want slobbered-on mints.) The bowl had only 17 mints left when the raid was over. How many mints were originally in the bowl? (Herr & Johnson, 1994, p. 303)

There are various solution processes for each of the questions. For example, on the first problem, the student could count. Another method would be to reason as follows. The page numbers 1 through 9 are one digit each. Thus, nine keystrokes are needed. The page numbers 10 through 99 are two digits each; thus  $2 \times 90 = 180$  keystrokes are needed. The page numbers

100 through 124 are three digits each; thus  $3 \times 25 = 75$  keystrokes are needed. The sum of 9, 180 and 75 is 264, which is the answer.

For the second problem, one could use algebra, producing a rather complicated, linear equation to be solved. An easier approach is to begin at the end. There were 17 mints left in the bowl. Just before that, Eugene threw two back in, so there were 15. Before that, Eugene took half of the mints and left 15 in the bowl. There were 30 mints in the bowl before Eugene began. Before that, Faizah put three mints back, leaving 30, so that was 27. Faizah took  $\frac{1}{4}$ , leaving 27. So, there must have been 36 mints in the bowl before Faizah. Right before Faizah, Sean returned four to the bowl. So, there must have been 32. Sean took  $\frac{1}{3}$  of the mints, leaving 32, so there must have been 48. This brings us to the beginning of the problem, in which there were 48 mints.

The CPMP students had been working in pairs throughout the school year (this assessment was conducted in May). The same pairs were used for this assessment. The classroom teachers paired the traditional students. The students were asked to work on these problems and think out loud with each other on how to solve the problems. In addition, the students were asked to write out their solution process after they were happy with their solution. It was emphasized to the students that the researcher was not interested in a numeric answer, but interested in the process that they used when solving the problem. The students were observed with the researcher taking notes, but not intervening, and videotaped while solving these problems. The field notes, annotated transcripts, and student work were then analyzed for the quality of the problem-solving strategies and processes, the overall success of the students in solving the problems, the quality of the cooperative work, and other emergent themes in the process of problem solving.

### Results with Discussion

Given that the students were members of intact classes, it was determined that the individual student should not be the unit of analysis in analyzing the results from the standardized test and the constructed response test. So, schools were used as the unit of analysis. In comparing CPMP to traditional schools on the standardized test, the traditional schools had a mean proportion correct score of .52 (.06 standard deviation) and the CPMP schools had a mean proportion correct score of .52 (.05 standard deviation). Clearly there is no significant difference between whether a school was CPMP or traditional.<sup>1</sup>

The standardized test in the format of constructed responses fared about the same. The mean proportion correct score for CPMP schools was .65 (.13 standard deviation) and traditional schools had a mean proportion correct score of .67 (.14 standard deviation). Again, there is no significant difference between whether a school was CPMP or traditional.

On the performance assessment, considering only the correctness of a solution, there again was no difference between the CPMP students and the traditional students. Most of the students, whether from CPMP or traditional, were able to solve the typewriter problem. A slightly higher number of CPMP students than traditional were able to solve the mints problem. This may be due to the fact that although the problem was much easier solved just working backwards, traditional students attempted to use algebra to solve the mints problem. However, if one counts only a numeric answer, on both problems, there was close to equivalence between the two groups of students.

A performance assessment allows one to look at other issues than just correctness of an answer. Five themes emerged from the analysis of the performance data (see Table 2).

For ease of reporting, the discussion of the themes will be given in Polya's framework for viewing the problem-solving process (Polya, 1945/1973). Polya outlined these stages as: (1) getting to know the problem, (2) forming a solution plan, (3) carrying out the solution plan, and (4) looking back.

The first two themes occurred in Polya's getting to know the problem stage. CPMP students were more engaged in the problems than the traditional students were and CPMP students were more eager to work in pairs than the traditional students were.

CPMP students immediately became engaged in the problems. This was evidence in the words that CPMP students used. "Do we get to work on these?" was actually a question a CPMP student shouted with excitement toward the researcher as soon as the researcher handed her the problem sheets. "I bet I can solve this" and "we are good at problems" were two other comments that the researcher interpreted as positive.

In addition to expressing interests and enthusiasm for solving problems, the CPMP students talked with each other about what the problems meant. For example, on the typewriter problem, the majority of CPMP pairs immediately began to act out what it might mean to type numbers on pages of a book. One CPMP student said to her partner, "See, I know, it is that... well, 35, that would be a 2. See?", meaning that the number 35 has 2 digits in it.

In analyzing the annotated transcripts for students in the traditional curriculum, the researcher could not find a single positive comment toward problem solving. In addition, one traditional student told the researcher, "We are really bad at math" which promoted another to say, "He [the teacher] hates us." Several pairs of traditional students stated that they did not want to work the typewriter problem or that they could not work the typewriter problem because they did not understand what the problem was asking. Several traditional pairs read the problem and then stated, "I don't get it." However, on the mints problem, many of the traditional students stated that they probably could solve it, as it required algebra and they were in an algebra course.

As students were getting to know the problems, the cooperation aspect was the second theme that emerged.

Table 2  
*Five Themes of Student Problem Solving Activity*

Theme	Description
Students' engagement in problems	CPMP students were more engaged in the problems than the traditional students. CPMP students engaged in the problem-solving process beyond a numeric answer. CPMP students asked to discuss the problems after they were done solving them.
Students' enthusiasm for working in pairs	CPMP students were more eager and able to work in pairs than traditional students.
Students' use of symbol manipulation for their solutions	The traditional students considered and actually used algebraic techniques.
Students' ability to communicate mathematically	CPMP students were able to write about their mathematical processes.
Students' flexibility in their solutions	CPMP students looked for more than one solution path.

In terms of the cooperation aspect, the CPMP students began cooperating as they were getting to know the problem. The traditional students didn't cooperate until possibly the looking back stage when they compared answers. Traditional students became familiar with the problems individually, and mostly inside their heads (not readily observable behavior). Since, CPMP students cooperated while getting to know the problem, the researcher could observe some of the ways that they interacted with the problem. For example, sometimes they tried to act out the situation. Other times they described the situations to each other and tried to clarify by talking to each other what the situation really meant. So, "getting to know the problem" is trying to understand the situation, and CPMP students did this cooperatively, while traditional students did not.

To summarize, while the CPMP students discussed the problem, the traditional students worked separately. The only time the traditional students engaged together as a pair was when they compared answers with each other. This point was made obvious with many of the traditional pairs literally telling each other to work on the problem and "we will check our answers at the end." A minority of traditional students was unwilling to compare answers, however, saying, "Solve it yourself."

The first two themes (CPMP students were more engaged in the problems than the traditional students were and CPMP students were more enthusiastic to work in pairs than the traditional students were.) continued during Polya's (3) carry out the plan stage. The CPMP students engaged in conversation with each other about the problems, and actually worked on the problems together. An excerpt from the annotated transcripts might illuminate this. For the sake of this example, the students will be labeled simply Student A and Student B. The students are working on the typewriter problem.

*Student A:* How many two-digit numbers are there?

*Student B:* I don't know. 11, 12, 13, ... [mumbles] ... let's ...

*Student A:* It will take too long to count them. 10, too.

*Student B:* Oh, yes. There are 10 through 100. 99.

There are 10 through 99.

*Student A:* So there are 89.

*Student B:* Yes, 89.

At this point, the students continued in the problem and started to examine the number of 3-digit numbers that existed. The reader might be worried about the 89 (knowing that the correct number is 90), but the reader

is asked to be patient, as we will return to this same pair later in this section.

In addition to the first two themes continuing, another theme (Students' ability to communicate mathematically) emerged. The CPMP students were able to write about their mathematical processes, demonstrating an ability to communicate mathematically. Only one-fourth of the traditional students showed anything other than a numeric answer. All of the CPMP students showed something other than their answer, with most showing the steps and thought processes of their solution. For example, on the typewriter problem, many of the CPMP students wrote out the number of 1-digit numbers, 2-digit numbers, and 3-digit numbers, labeling each respectively. On the mints problem, several CPMP students wrote the words "work backwards" or just "backwards" on their sheet. Other CPMP students showed the backward progression of numbers, but did not use the word "backwards." On the mints problem, the traditional students who did write something down (again this was about one-fourth of them) wrote out the algebraic equation that they had developed. Although some traditional students did work backwards and solve the problem, they simply showed the answer in a box. When the researcher asked them to show their solution process, they said that there was nothing to show.

A theme (Students' use of symbol manipulation for their solutions) emerged during (2) Polya's forming a plan stage. The traditional students were much more likely than the CPMP students to think of using symbol manipulation. In fact, the traditional students at times seemed to skip the (1) getting to know the problem stage and enter right into (2) forming a plan, with that plan being to use algebraic symbols. The traditional students felt that they could work the mints problem, because they thought there would be manipulations of symbols for doing so. The traditional students, unlike the CPMP students, were very quick on the mints problem to use symbol manipulation.

Polya's final stage, (4) looking back, revealed the final theme that emerged from the analysis of students problem solving activity: student's flexibility to have different solutions. The CPMP students looked for more than one solution path and they stayed engaged in the process beyond a numeric answer. The fact that a numeric answer did not signify "done" to the CPMP students seemed important to the researcher. The CPMP students seemed to value finding a variety of solution paths, whether they lead to a numeric answer or not. Yet, the numeric answer was viewed to be

valuable as well. In particular, the numeric answer was viewed as a method of checking the various solution paths.

So, the researcher was actually amazed to find that after the CPMP students had written down a solution strategy, the vast majority of them asked each other if there was another way to solve the problem, and then continued to work on the problem. Although some wrote down more than one solution process, most did not. Although the alternative solution paths did not show up in writing, most CPMP students talked about alternative solution paths. There was no evidence any of the traditional students searched for more than one solution. This process of finding a second solution path on the part of the CPMP students seemed to serve two purposes. It appeared to be a check on the numeric answer, but it also appeared to be an aid for helping each other if one of the pair did not understand the previous process. In this manner, the numeric answer was important, but not the end goal. CPMP students repeatedly asked each other if they were in agreement on a process.

An example returns us to the students who thought that there were 89 two-digit numbers. This pair actually decided to count the digits by starting with 1 and ending with 124. (Many traditional students counted the digits, as well, but this was their only solution process.) These students then discovered that their first answer did not match their second answer. The students then tried to go back over their solution to find out where there might be a mistake.

Finally, when the CPMP students stated that they were done, every pair asked the researcher to discuss the problems with them. None of the traditional students wished to see the solution processes for either question even when the researcher offered to work the problems. The CPMP students seemed to view part of the problem-solving process as explaining to each other and the researcher what they were thinking. Thus, the problem continued even when the CPMP students had exhausted their solution paths. One pair of CPMP students asked the researcher if she knew whether class time would be spent on the problems on the following day. The student stated that she wanted to show the class her solution and ask class members what their solutions were. When the researcher told her that actually she really did not know if more class time would be spent on these problems, the student ran to the chalkboard and began to act out what she would tell the class if she was given the time. Her partner watched patiently throughout this process, and even put in comments at appropriate places. Then the

partner turned to the researcher and asked, "Should I go to the chalkboard, too?" This enthusiasm for the problem-solving process was present in all the CPMP students to a more or lesser degree, and appeared to be virtually absent in the traditional students.

While the paper-and-pencil tests show CPMP students and traditional students being equivalent in their problem-solving abilities, the performance assessment paints a picture of CPMP students excelling at problem-solving characteristics and traditional students lagging considerably behind. Obviously, the researcher is not suggesting that the results (especially the performance assessment results, which were with 100 students and 4 classrooms, while the paper-and-pencil results had considerably more students and from 10 schools) generalize to all CPMP students and all traditional students. However, the reader is reminded that the goal of this study is really not to compare these two types of students, but rather to comment on what aspects of students' "problem solving ability" may not be seen in testing. Clearly, the performance test revealed aspects of problem solving that were present in at least this sample of CPMP students that the other tests did not reveal with the CPMP students. None of the tests revealed these aspects of problem solving in the traditional students. It could be that the traditional students lacked these aspects of problem solving. Or, it is possible that the traditional students had these aspects of problem solving, but were unable to demonstrate them under any of these testing situations. This last case, however, seems unlikely to the researcher.

This study suggests that it is time to examine how we are testing. If indeed we test what we value, do we value the problem-solving skills that are tested by the paper-and-pencil tests in this study, for example? Or do we value the problem-solving skills that became apparent in the performance test? Do we value engaging in the problem-solving process beyond a numeric solution? Do we value solving a problem more than one way? These abilities may indeed be present while taking paper-and-pencil tests, but they are difficult to measure. Of course, it is possible to give a paper-and-pencil test with constructed-response items, and tell students to find more than one solution process. It is, however, an interesting observation that without being told to find more than one solution process, CPMP students were inclined to do so. An implication to this study, then, is to call for an adjustment to how researchers and even classroom teachers evaluate NCTM-oriented programs. Standardized testing is not the best method when



seeking to document types of students' mathematical features (e.g., engagement) seen in this study.

However, the implications can take researchers in the other direction, as well. Rather than a call to stop using standardized tests, perhaps this study calls for a revision of standardized tests. As stated in an earlier section, there have been numerous studies on the CPMP curriculum. Many of these use standardized testing materials, perhaps because this area concerns the public. Recall the mention of parental concerns in the introduction to this paper. However, it might be that it is time for testing and measurement experts to create alternatives to the standardized tests, so that researchers could attempt to evaluate CPMP and other NCTM-oriented curricula from a new perspective. Of course, researchers can write their own testing instruments, but there are advantages (including validity, reliability, and the standardization of the testing itself) to using a standardized test. An implication to this study, then, is to call for a massive revision of standardized tests.

Although it may be an obvious point, it probably should not go without saying that there is no question the CPMP students had more experience with aspects of problem solving than the traditional students. The CPMP students were more used to working together, writing about mathematics, discussing mathematics, not necessarily valuing a numeric answer, and looking for more than one solution. The CPMP students had had more opportunity to work with a variety of problems. If the traditional students had these experiences, would they too have these skills? This may be true, although it might be argued that if traditional students had these experiences they would not be referred to as traditional students. The traditional students were quick to think of algebraic techniques, and indeed the traditional students had more experience with algebra. Perhaps this returns us to a previous point: what it is that we value in problem solving? This is not to suggest that we have an either/or situation, as there is nothing to prevent us from including more traditional in reform, or including more reform processes in traditional.

### Summary

The problem-solving ability measured on the standardized tests did not show differences between CPMP and traditional schools. In spite of the absence of statistically significant differences, there were qualitative differences between CPMP students and traditional students observed in the performance assessment. For example, CPMP students appeared to

be more engaged in the problem-solving process. They looked for alternative solutions processes; they worked together as a group; and they showed more steps on the written work.

With this rather large sample, the CPMP students were not at a disadvantage on standardized problem-solving tests, in the sense that they scored as well as the traditional students. However, the alternative assessment did reveal some aspects of the CPMP students that the standardized tests did not. Regardless of one's interest in the qualities (such as students' engagement, flexibility etc. discussed in emergent themes) that the standardized test did not measure, it appears to be the case that one can succeed on a standardized test while being in the CPMP curriculum. And this study leaves us with the implication that perhaps it is time to reconceptualize standardized testing, and what, as researchers and classroom teachers, we really want to test.

### REFERENCES

- Curcio, F. R. (1999). Dispelling myths about reform in school mathematics. *Mathematics Teaching in the Middle School*, 4, 282-284.
- Feldt, L. S., Forsyth, R. A., Ansley, T. N., & Alnot, S. D. (1993). *Iowa Tests of Educational Development (Forms K & L)*. Chicago: The Riverside Publishing Company.
- Gipps, C. (1999). Socio-cultural aspects of assessment. In A. Iran-Nejad, & P. D. Pearson (Eds.), *Review of Research in Education*, (pp. 355-392). Washington, DC: The American Educational Research Association.
- Hancock, L., & Kilpatrick, J. (1993). Effects of mandated testing on instruction. In Mathematical Sciences Education Board, *Measuring what counts: A conceptual guide for mathematics assessment* (pp. 149-174). Washington, DC: National Academy Press.
- Herr, T., & Johnson, K. (1994). *Problem solving strategies: Crossing the river with dogs and other mathematical adventures*. Emeryville, CA: Key Curriculum Press.
- Hirsch, C. R., & Coxford, A. F. (1997). Mathematics for all: Perspectives and promising practices. *School Science and Mathematics*, 97, 232-241.
- Hirsch, C. R., Coxford, A. F., Fey, J. T., & Schoen, H. L. (1995). Teaching sensible mathematics in sense-making ways with the CPMP. *Mathematics Teacher*, 88, 694-700.
- Huntely, M.A., Rasmussen, C. L., Villarubi, R. S., Sangtong, J., & Fey, J. T. (2001). Effects of standards-based mathematics education: A study of the Core-Plus Mathematics Project algebra and functions strand. *Journal for Research in Mathematics Education*, 31, 328-361.
- Kahan, J. A. (1999). *Relationships among mathematical proof, high-school students and reform curriculum*. Unpublished doctoral dissertation. University of Maryland.
- Kilpatrick, J. (2003). What works? In S. Senk, & D. Thompson (Eds.), *Standards-based School Mathematics Curricula* (pp. 471-488). Mahwah, NJ: Erlbaum.

- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- National Research Council (1989). *Everybody counts: A report to the nation on the future of mathematics education*. Washington, DC: National Academy Press.
- Polya, G. (1945/1973). *How to solve it*. Princeton, NJ: Princeton University Press. (Original work published 1945).
- Schoen, H. L., Cebulla, K. J., & Winsor, M. S. (2001). *Preparation of students in a standards-oriented mathematics curriculum for college entrance tests, placements tests, and beginning mathematics courses*. Paper presented at the Annual Meeting of the American Educational Research Association. Seattle, WA.
- Schoen, H. L., Fey, J. T., Hirsch, C. R., & Coxford, A. F. (1999). Issues and options in the math wars. *Phi Delta Kappan*, 80(6), 444-453
- Schoen, H. L., & Hirsch, C. R. (2003). Responding to calls for change in high school mathematics: Implications for collegiate mathematics. *American Mathematical Monthly*, 110, 109-123.
- Schoen, H. L., & Ziebarth, S. W. (1998a). *Assessments of students' mathematical performance*. A Core-Plus Mathematics Project field test progress report. Unpublished manuscript, University of Iowa.
- Schoen, H. L., & Ziebarth, S. W. (1998b). High school mathematics curriculum reform: Rationale, research, and recent developments. In P. S. Hlebowitsh & W. G. Wraga (Eds.), *Annual review of research for school leaders* (pp. 141-191). New York: Macmillan.
- Senk, S., & Thompson, D. (2003a). *Standards-based school mathematics curricula*. Mahwah, NJ: Erlbaum.
- Senk, S., & Thompson, D. (2003b). School mathematics curricula: Recommendations and issues. In S. Senk & D. Thompson (Eds.), *Standards-based School Mathematics Curricula* (pp. 3-27). Mahwah, NJ: Erlbaum.
- Swafford, J. (2003). Reaction to the high school curriculum projects research. In S. Senk & D. Thompson (Eds.), *Standards-based School Mathematics Curricula* (pp. 457-468). Mahwah, NJ: Erlbaum.

---

<sup>1</sup> The use of school as the unit of analysis resulted in an enormous loss of power. The question may be asked as to whether the lack of significance observed here is due to this lack of power. Two factors suggest that this is probably not the case. First, the values of the means analyzed here are very similar across groups. Second, the analysis using the student as the unit of analysis showed a very similar pattern of nonsignificant results.