DOCUMENT RESUME

ED 164 356 SE 026 362

AUTHOR Arbetter, Carolyn Clinton; And Others
TITLE USMES Student Study: Effects of USMES on Student Problem-Solving Abilities and Attitudes.
INSTITUTION Education Development Center, Inc., Newton, Mass.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE 78
GRANT NSF-SED-6901071
NOTE 99p.; For related documents, see SE 026 361-363; Contains occasional light and small type
AVAILABLE FROM Education Development Center, 55 Chapel Street, Newton, Massachusetts 02160 ($5.65)

EDRS PRICE MF-$0.83 HC-$4.67 Plus Postage.
DESCRIPTORS Curriculum; Educational Research; *Elementary Education; Elementary School Mathematics; Elementary School Science; *Integrated Curriculum; *Mathematics Education; *Program Evaluation; Science Course Improvement Project; *Science Education

ABSTRACT

This study is one of three designed to evaluate the effectiveness of the Unified Science and Mathematics for Elementary Schools (USMES) program. The student study examined the program's impact on students' development of problem-solving abilities and on their attitudes. Conditions affecting the program's impact were identified, including the amount and type of USMES experience students received, their schools, grade levels, and scholastic achievement levels. Instruments were developed and administered to a small number of carefully studied schools. The USMES program did not appear to have a statistically significant effect on the problem-solving ability of the students. However, the students did seem to enjoy doing USMES work. (BB)

UNIFIED SCIENCES AND MATHEMATICS FOR ELEMENTARY SCHOOLS:
Mathematics and the Natural, Social, and Communication Sciences in Real Problem Solving

This report was prepared with the support of the National Science Foundation, Grant No. SED-6901071. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Education Development Center
55 Chapel Street
Newton, Massachusetts 02160

Executive Summary of the USMES Student Study*

The USMES Student Study examined a number of effects of the USMES program on students. The specific issues selected for study were suggested by a substantial body of past research and by the USMES staff's desire to provide developers and users of USMES units with useful new information on the program's impact under a variety of conditions, as well as with prototype diagnostic instruments.

Methods

To measure student problem-solving skills and strategies the USMES staff developed a consumer product research problem (called the Pencil Problem) that would enable them to look at a few key indicators:

1. Does the student consider the possible importance of many factors, like product performance, durability, and cost?
2. Does the student systematically perform experiments and collect useful information?
3. When advising others about the product, does the student present relevant information in effective ways?

*Copies of the final report of the Student Study are available from the USMES Office at Education Development Center.

The Pencil Problem asks students, in groups of three, to examine six brands of pencils and recommend which brand their principal should buy. Students are then questioned about their approaches toward solving the problem and observed while they try the pencils. A similar, though shorter, process is repeated using erasers. Students' verbal and written responses are recorded and later coded.

This consumer product research setting seemed to offer a realistic small-scale problem as well as a useful source of observations about the students. The potential richness of observations suggested that the instrument might be suitable for diagnostic use by USMES classroom teachers.

To measure student attitudes, the USMES staff developed a three-page multiple-choice Questionnaire on Attitudes Toward Real Problem Solving (QARPS) to assess three broad types of student attitudes:

1. Attitudes toward working on real problems and producing effective solutions

2. Attitudes toward the group interaction involved in real problem solving

3. Attitudes toward a number of specific problem-solving activities.

The QARPS also included a section only for students with USMES experience, covering their attitudes toward USMES as a school program.

The student sample came from twenty-four classes in four elementary schools that differ in size and composition of student bodies, and in the nature of their surrounding communities (see Table below). All USMES classes above the first grade were included, as well as a number of comparison classes that were not doing USMES at the time. Classes chosen for comparison were those judged by school principals as most equivalent to the USMES classes in grade level, student ability level, and teacher's experience.

The QARPS was administered to all students in all sample classes; the Pencil Problem was administered to a random sample of students in each class.

The USMES classes included in the study had used a wide variety of published USMES units as well as several locally-developed units.

Three members of the USMES research staff visited each school during April and May 1977 to administer the QARPS and the Pencil Problem. A procedure was devised to ensure that, in administering the Pencil Problem to a group of students they would not know what class the students had come from. This helped rule out biased observation or conversational tone.

Results and Conclusions

The mean scores of the USMES students were higher than those of the comparison students on virtually every major outcome. However, these differences were only statistically significant on some of the outcomes. The effects of USMES were generally small, compared to the effects of School, Grade, and Achievement Level.

CHARACTERISTICS OF SCHOOLS
(All figures are approximate)

	A	B	C	D
Number of Students in School	300	800	700	1,600
Geographic Region	Mid-central	Southwest	North-central	Southeast
Community Population	50,000	1.5 million	25,000	650,000
Community Type	Small city	Large city	Suburb	Large city
Surrounding Neighborhood	Expensive houses & university	Inexpensive houses & shopping center	Moderate to expensive houses & apartments	Moderate houses
Racial/Ethnic Composition	99% white 1% Asian, Black	49% Spanish surnamed 49% white 2% Black, Asian, Native American	99% white 1% Black & Spanish surnamed	98% Black 2% white, Spanish surnamed, other

Because of the small and uneven effects of different amounts of USMES experience, a linear regression analysis showed only moderate statistical significance for Amount of USMES on four of the six major outcomes and poor significance levels on the other two outcomes. One of the statistically significant regression results was negative; the other three were positive.

The effects of Amount of USMES were most pronounced in a few settings, such as School D. In that school, the comparison group scored the lowest of all four comparison groups on the three Pencil Problem outcomes. Yet, the USMES group which had received a substantial amount of USMES experience in School D scored the highest of all four schools' "substantial USMES" groups on two of the three outcomes and second-highest on the other. Similar effects of Amount of USMES were found in the 6th grade (however, since many of the 6th graders in the sample came from School D, the 6th grade and School D results may be related).

The motivational effects of USMES did not seem to be so localized. The responses to questions directed at students with USMES experience confirmed once again a repeated finding: the great majority of USMES students enjoy USMES work, find it important, and would like to do more. This finding appeared to be true in all the schools and at all grade levels.

Several factors seemed to account for the difference between motivational and problem-solving effects of USMES. Evidently, "little" amounts of USMES (USMES experience in the past only, or less than 20 hours of USMES in the current school year with no past experience) produced inconsistent results. Substantial amounts of USMES experience--including amounts greater than the 20 hours which defined "substantial" in the Student Study--may be necessary for consistent, significant problem-solving results.

The specific activities done in an USMES class may also be crucial. The USMES School Study* found evidence that opportunities for skill development were often closely associated with the occurrence of serious investigations of data, extended discussions among students, and other recommended USMES activities. Furthermore, in the Student Study the "substantial" Amount of USMES group appeared to have engaged in significantly higher levels of many recommended activities than the "little" Amount of USMES group, and the "substantial" group produced the most consistently positive results.

In addition, the surprisingly large effects of School (which included some community differences in the Student Study), Grade, and Achievement Level suggested that all of these factors made important contributions to the impact of USMES experience.

Finally, the influence of the school administration appeared to be crucial. In School B, USMES had its most positive effect on attitudes, and in School D USMES had its most positive effect on problem-solving skills. In each school, these were the outcomes that the principal of the school desired most to promote (see the USMES School Study report). The particular style of USMES chosen in each school emphasized the desired outcomes and was successful in improving it.

The results of this study suggested that the choice of USMES units (and the local development of units) should be based on a careful consideration of the outcomes that would be most valuable for a particular school and class. Furthermore, it appeared that individual students may experience a wide range of effects of USMES experience. While virtually all students seem to enjoy USMES work, there seems to be considerable variation from student to student in just what is learned. For consistent, significant results on problem-solving skills, the Student Study and the USMES School Study suggest that "regular" USMES units of some duration--with recommended USMES teaching methods --should generally be preferred over abbreviated or haphazardly-created units.

The QARPS and the Pencil Problem seem to be useful instruments for diagnostic purposes. The QARPS is easily administered, easily analyzed, and shows satisfactory discrimination. The Pencil Problem probes real-problem-solving abilities more deeply than "pencil and paper" tests, yet can be administered and analyzed in practical ways.

A more detailed summary of the Student Study's findings may be found on pp. 23-37 of the Student Study report.

---

*Available from the USMES project office at Education Development Center.

ACKNOWLEDGMENTS

This study was designed and administered by the USMES research staff:

Carolyn Clinton Arbetter, Coordinator for Research Studies
Daniel S. Cooper, Senior Research Associate
George H. Stalker, Research Associate
Felicia DeMay Weitzel, Research Associate
Nancy Gardner Weiner, Administrative Assistant for Research Studies
Phyllis A. Gentile, Secretary/Data Processing Assistant

Each made substantial contributions, too numerous to mention, at the same time they were conducting studies on USMES resource teams and schools. Daniel Cooper was the principal author of the report and had major responsibility for the design of the study and the analysis and interpretation of the data. Carolyn Arbetter and Daniel Cooper designed the instruments and with Felicia Weitzel administered them at the four schools. George Stalker designed the data entry and processing techniques. Nancy Weiner ably handled travel arrangements, budgets, and liaison with the schools. Phyllis Gentile entered a huge amount of data into computer files accurately and efficiently. Joan Labby drafted much of the Executive Summary.

Earl Lomon, USMES Project Director, and Janet Whitla, EDC Associate Director, provided very useful advice and suggestions throughout the course of the study. The overall design was based in part on guidelines established by the Evaluation Panel of the USMES Planning Committee. Chaired by Irving Morrissett, the panel includes Jack Borsting, Camilla Fano, Abraham Flexer, and Fred Johnson. This group, along with Wayne Welch, reviewed the preliminary analyses and findings and helped set priorities for the limited amount of time available for analysis, interpretation, and reporting of results. Our special thanks to all for their encouragement and insights.

Martha Allegro quickly and carefully typed most of the final report; Jeanne McDonald also skillfully typed several sections. John Saalfield designed the cover, and Paula Lakeberg did the illustrations for the technical reports. Throughout the past years, other USMES staff members have facilitated our work in a variety of ways and we appreciate their help.

Our special thanks to the Feldberg Computer Center of Brandeis University for financial and technical assistance when a belated need arose for further analysis of some of the data.

Finally, and most especially, we thank the principals, teachers, and students of the four participating schools for their unfailing and cheerful cooperation. Although they are not named in this report to preserve their anonymity, they have been valued colleagues throughout the study.

Carolyn Clinton Arbetter

TABLE OF CONTENTS

INTRODUCTION

The USMES Project conducted three studies in 1976-1977 to learn more about USMES students, schools, and resource teams. The Student Study examined the USMES program's impact on students' development of problem-solving abilities and on their attitudes. A major goal of the Student Study was to identify conditions affecting the program's impact, including the amount and type of USMES experience students received, their schools, grade levels, and scholastic achievement levels.

The general strategy of the Student Study was to develop instruments for assessing key problem-solving abilities and attitudes, then to administer the instruments in a small number of carefully studied schools. This strategy allowed us to study a wide variety of students and USMES classes, at the same time that we collected fairly extensive observations about the particular schools, communities, teachers, and class activities. It also enabled us to design instruments that met another important criterion--that they be suitable for diagnostic use by teachers and administrators.

This report covers the major goals, methods, and findings of the Student Study. Only findings on a few key outcomes and causal factors are covered; because of serious time constraints, we have not attempted to reach final conclusions but rather to summarize our impressions about the major findings. The report is intended for use by funders, school people, and others interested in the impact of USMES on students. It is accompanied by brief technical reports on the instruments used and an appendix showing detailed findings.

BACKGROUND

The specific issues selected by the USMES staff for attention in the Student Study reflected a number of influences. A major influence was the substantial body of past research on USMES. One summary of that research listed the following findings:*

- "In interviews with outside evaluators, teachers overwhelmingly reported that USMES was fostering development of practical problem-solving abilities and social skills in their students. Furthermore, the administrators, teachers, and students who were interviewed by Boston University personnel in 1975 unanimously agreed that students enjoyed USMES and looked forward to using it.
- "Although inconclusive results were obtained from the Playground Problem** and the Picnic Problem**, apparently because of limitations in the instruments themselves, the results from the Notebook Problem** showed dramatic gains in some specific problem-solving skills for USMES students in contrast with much less positive results for comparison students. Further development of instruments to assess problem-solving abilities is needed, and efforts in this direction are being made.
- "Assessments of student achievement in reading, mathematical computation, and several other subject areas have found that USMES students were progressing at least as rapidly as comparison students. In virtually every study where differences appeared between USMES and comparison groups, the trends favored the USMES group. Most of these differences were small, but they are encouraging and warrant further investigation.
- "Data gathered with classroom observation scales has shown that USMES students spend more time in self-directed, active, and creative behaviors than students in comparison classes. USMES students also generate more ideas in both large and small group discussions.

*See Chapter V of the fourth edition of the USMES Guide, EDC, 1976.
**Instruments developed at Boston University between 1971 and 1975 to assess problem-solving skills for evaluations of USMES.

- "All of the available evidence on affective outcomes (from interviews and questionnaires) suggests that students find USMES challenging and inviting. Moreover, the program may have significant positive effects on their appreciation for science."

Taken together the findings of past research suggested that further research on student attitudes and problem-solving abilities might be especially valuable.*

Another major influence on the choice of issues was our desire to provide useful information to people involved in the development of USMES units. This group includes not only members of the USMES central staff, their funders, and their advisors, but also a large number of school teachers and administrators; trained USMES teachers and administrators have always been the primary source of ideas for USMES published units, and they continue to develop additional units for local use. For the benefit of all these people involved in USMES development, it also seemed important to learn more about USMES's impact on student attitudes and problem-solving abilities, under a variety of conditions.

In addition to providing developers and users of USMES with new information, we wanted to provide them with diagnostic instruments on problem-solving abilities and attitudes, at least in prototype form. However, practical constraints meant that we could only expect to develop one instrument on each outcome. Hence, our research issues had to be measurable with a single attitude instrument and a single problem-solving instrument, both adaptable for use by others.

While the choice of issues to be examined in the Student Study involved balancing many concerns, the decision to design the study around existing classes and to rely almost entirely on data collected near the end of the school year was a less complicated one: we had no practical alternative. Time and budget constraints required us to study classes which were already committed to their 1976-1977 instructional programs before we contacted them. Furthermore, the time needed for instrument development prevented us from collecting research data with the instruments until the spring of 1977.

However, we believed that the limitations imposed by such a research design would not prevent us from gathering useful information, because we believed we could overcome many of the limitations of the design. For example, we believed that we could locate groups of students who were essentially comparable except for their use (or non-use) of USMES. Furthermore, we believed that we could collect information other than pre-test scores, (e.g., recent achievement test scores) which would bear upon the comparability of students receiving different amounts of USMES. Finally, we knew

*A reanalysis by the USMES research staff of the Boston University study of USMES and "basic skills" has shown that further research in this area is also needed. See Appendix I.

that we could restrict our interpretations of Student Study data in accordance with the limitations imposed on the study. In general, we took a position which we believe is similar to that of Donald Campbell and Julian Stanley when they wrote:

> "From the standpoint of the final interpretation of an experiment..., every experiment is imperfect.... (The experimenter) should seek out those artificial and natural laboratories which provide the best opportunities for control. But beyond that he should go ahead with the experiment and interpretation, fully aware of the points on which the results are equivocal."*

Our decision to conduct the Student Study in only four schools, and the selection of the particular schools in the study, was dictated by a number of requirements. We wanted to collect a relatively large amount of information about the types of USMES experience received by each USMES class in the study, through site visits, teacher reports, and interviews; but to do so, we had to restrict the number of schools. We also wanted to work with schools where we would find students who had prior USMES experience. A wide range of student abilities, school types, and surrounding communities was also needed, to help identify conditions affecting the impact of USMES on students. Furthermore, logistical considerations demanded that we work with schools having a relatively large number of USMES classes and schools in which we could conduct a related USMES research study. Finally, because we would be asking schools for considerable assistance and because time was short, we had to work with schools which were at least moderately familiar to the USMES central office. We were fortunate that the schools we decided to study were all willing to work with us and were, in fact, extraordinarily helpful collaborators throughout the research.

*Campbell, D.T. & Stanley, J.C. *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally & Co., 1966.

VARIABLES STUDIED

Outcome Variables

Since problem-solving skills and strategies are a central concern of USMES, a major focus of the Student Study was on these outcomes. There are serious problems in assessing students' problem-solving abilities, however. One is the absence of commercial instruments for measuring the kinds of problem-solving abilities which USMES tries to foster. Another is the length of time it takes to solve a real problem. USMES classes usually work for a number of weeks on a single problem and it is difficult to imagine a group of students working meaningfully through all the stages of real problem solving in less than several hours. To assess a student's performance in all the stages of real problem solving would require so much time that only a few students could be assessed. Yet our desire to study students with a wide variety of USMES experiences (and comparable students with no USMES experience) meant that we had to assess a relatively large number of students.

To resolve this dilemma, we decided to focus on only a few key indicators of students' problem-solving abilities. Our decision was based on the belief that important skills and strategies used during at least <u>some</u> stages of real problem solving could be systematically observed in a reasonable amount of time. While such observations would not constitute a complete measure of real problem solving competence, they would provide useful information about the level of students' problem-solving development. This view had been supported by the USMES Planning Committee's Evaluation Panel, which stated:

> "The Panel held...hope for the feasibility of developing measures of 'indicators.' Indicators were understood to mean abilities or characteristics that presumably would accompany the growth of problem-solving ability. Such indicators can be looked upon as desirable in themselves as well as indicators of problem-solving ability."*

The problem-solving indicators on which the Student Study focused included the following:

1. While working on a consumer product research problem, does the student consider the possible importance of many factors, like the product's performance, durability, and cost, or does the student overlook many of the factors?

**Guidelines for Evaluation of USMES Outcomes*, EDC, 1975.

2. In the same setting, does the student try to systematically collect information that is relevant and useful?

3. When advising others about the product investigated, does the student present relevant information in effective ways?

These problem-solving indicators were selected because they seemed to include outcomes that were of central concern to USMES and were reasonably feasible to assess. Furthermore, the consumer product research setting seemed to offer a realistic, if small-scale, problem for students as well as a rich source of observations about students. The richness of potential observations also made them suitable for diagnostic use by USMES classroom teachers.

Problem-solving skills and strategies were not the only outcomes of concern in the Student Study. The other major focus of the Student Study was student attitudes, because some of the most important USMES goals involve relatively subtle changes in attitudes. For example, USMES experience is expected to foster a sense of confidence in students about their capacity for finding good solutions to real problems and for implementing their solutions. Furthermore, USMES experience is expected to promote a positive attitude toward working with other students in real problem solving.

To decide which attitudes were most important for the Student Study, the USMES staff reviewed a number of instruments developed for previous research on USMES and on other curricula. In addition, both the USMES staff and a Boston University USMES evaluation team had completed earlier reviews of several commercially available attitude instruments. None of these reviews located an instrument which seemed to cover an appropriate set of attitudes. However, some questionnaire items used in previous research on the USMES program suggested topics that might be incorporated in a new questionnaire.

On the basis of these reviews, the USMES staff decided to develop a student attitude questionnaire for the Student Study. It was to assess three broad types of student attitudes:

1. attitudes toward working on real problems and producing effective solutions

2. attitudes toward the group interaction involved in real problem solving

3. attitudes toward a number of specific problem-solving activities

In addition, the questionnaire was to include a section only for students with USMES experience, covering their attitudes toward USMES as a school program they had taken part in.

Independent Variables

A central concern in the Student Study was to examine the impact of varying amounts of USMES experience. Consequently, we viewed "amount of USMES" (both the extent of a student's USMES experience this year, if any, and whether or not the student had prior USMES experience) as a key independent variable. After examining reports of actual USMES use this year, we decided to divide "amount of USMES" into three broad levels:

1. "none": no USMES experience, past or present
2. "little": either USMES experience in the past alone, or up to 20 hours of USMES this year (but none in the past)
3. "substantial": either USMES experience in the past as well as the present, or over 20 hours of USMES this year (with or without past USMES)

Other independent variables given high priority for study included school (i.e., the four different schools in the study) and grade level. These two independent variables seemed important both for identifying conditions affecting the impact of USMES and for statistical "control" of the findings.

For the same reasons two additional independent variables were identified for study, but they were given lower priority because of difficulties involved in measuring them. These two variables were achievement test scores and "type of USMES."

Achievement test data were difficult to analyze, because the four schools of the study used a number of different achievement tests, administered at different times of the year, and scored in different ways (e.g., local percentiles versus national grade-equivalents). For our purposes, we could not justify asking students and schools to submit to a special round of achievement testing just for the USMES Student Study; hence, we opted to use only achievement data the schools were routinely collecting and to analyze it cautiously. Ultimately, we only attempted to collect achievement data for the relatively small number of students tested on problem-solving indicators.

The independent variable, "type of USMES," was intended to be a panel rating (by USMES staff members) of USMES activities in the USMES classes of the study along several dimensions. Over the course of the past year, we experimented with a number of possible rating dimensions, but none seemed simultaneously easy to code and quantitatively related to even the most obvious student outcomes. Consequently, we eventually decided not to attempt a quantitative analysis of "type of USMES," but rather to use the information about USMES class activities as an informal aid in interpreting findings on other variables.

INSTRUMENTS AND OTHER DATA SOURCES

The main instrument for measuring problem-solving indicators was the Pencil Problem*, which was developed for the Student Study. The Pencil Problem asks students, in groups of three, to look over six brands of pencils and to recommend which brand their principal should buy for students to use. The students know that their principal really will buy $10.00 worth of pencils based on their recommendations. Students are asked several questions about their approaches toward solving the problem and the reasons for their choice, and they are systematically observed while they try the pencils. Students are also asked to look over five brands of erasers, under similar conditions, and to tell how they would choose an eraser if they had time. Their responses to the questions are tape-recorded, and they are also asked to make written recommendations to their principal about their choice of pencils. The verbal and written responses are coded with checklists that cover the essential phrases of virtually all common responses. Scores on a number of problem-solving indicators, such as considering most of the important information about each alternative, can be derived from the codes. A more complete description of the Pencil Problem is given in the accompanying *Technical Report on the Pencil Problem*.

Students attitudes were assessed with the Questionnaire on Attitudes toward Real Problem Solving (QARPS), which was also developed for the Student Study. The QARPS is a three-page, multiple-choice questionnaire with two sections. The first section, with 20 items, is for both USMES students and students with no USMES experience. The second section (on a separate page) is only for USMES students; it contains nine attitude items about the USMES program and a checklist about experiences the students may have had while doing USMES work. A more complete description of the QARPS is contained in the accompanying *Technical Report on the Questionnaire on Attitudes Toward Real Problem Solving*.

Information on students' USMES experience was obtained from several sources. For students in USMES classes this year, weekly teacher reports (or interviews with teachers who were unable to send reports) provided information about the amount of time students spent on USMES and the nature of their activities during USMES time. The report and interview forms were developed on the basis of several years' experience with USMES teacher logs of class activities. The questions on the forms were designed to provide information on the critical aspects of the students' USMES experience. Samples of the teacher report and interview forms are included in Appendix J. The amount of USMES experience a student had this year was calculated (in hours) on the basis of each class's total USMES work for the year. In addition, each school provided information about whether a student

*The Pencil Problem also includes a short "eraser problem" as well as a much longer "pencil problem."

had done USMES work in past years. Thus, the research data included the amount of USMES experience this year (if any) for every class in our sample and also whether or not each student in the sample had done USMES in the past.

Achievement test data came from different tests in different schools. Two schools used the *Iowa Tests of Basic Skills*, one used the *SRA Assessment Survey Achievement Series, Multilevel Edition*, and one used the *Metropolitan Achievement Tests*. The only common reporting format for achievement scores, among the four schools, was national grade-equivalents. We recognized that grade-equivalents are hazardous data in general, and especially when they come from different tests. However, we decided to use the available grade-equivalent data, and to interpret them cautiously, rather than to ignore achievement levels entirely.

SAMPLE

The sample of students in the Student Study came from four very different elementary schools. Some key characteristics of the four schools are indicated in Table 1. As indicated in the table, the schools differ not only in the size and composition of their student bodies, but also in the nature of the communities in which they are located.

All USMES classes in the four schools, above the first grade, were included in the Student Study. (Kindergarten and first grade classes were not included, because our instruments were not appropriate for very young students.) In addition, a number of comparison classes which were not doing USMES this year were included. Since we were primarily interested in USMES classes which appeared from teacher reports to have done significant amounts (at least 20 hours) of USMES this year, we tried to include comparison classes which were most equivalent to the relatively active USMES classes. To select the comparison classes, we provided each principal with a list of the active USMES teachers and asked the principals to identify, for each of these classes, a comparison class which had not done any USMES this year and was similar in terms of grade level, ability level of students, and experience of teacher. Occasionally, there were no comparison classes available at appropriate grade levels, and in those cases, a class at the closest grade level was selected. The distribution of USMES and comparison classes in the Student Study, by school and grade level, is shown in Table 2.

As shown in the table, the practical limitations of availability of classes resulted in (1) differences in the grade-level distributions from school to school and (2) within some schools a mild imbalance of grade levels between the USMES and comparison groups. Furthermore, our evidence about the comparability of USMES and comparison schools is limited to the principals' judgments (and some students' achievement test scores). These restrictions limit our ability to rule out some challenges to our findings and to generalize from the findings. Nonetheless, we believe that the set of classes in the Student Study represented an informative diversity of USMES experiences and the best available sources of comparison data.

The QARPS was administered to each of the classes recorded in Table 2. The Pencil Problem, however, was administered to only a sample of the students in each class, because the Pencil Problem could be administered to only three students at a time. (The QARPS was administered to a whole class at a time.) The selection of a Pencil Problem sample from each class went through five stages:

TABLE 1

CHARACTERISTICS OF SCHOOLS

(All figures are approximate)

	A	B	C	D
Number of Students in School	300	800	700	1,600
Geographic Region	Mid-central	Southwest	North-central	Southeast
Community Population	50,000	1.5 million	25,000	650,000
Community Type	Small city	Large city	Suburb	Large city
Surrounding Neighborhood	Expensive houses and university	Inexpensive houses and shopping center	Moderate to expensive houses and apartments	Moderate houses
Racial/Ethnic Composition of School	99% white 1% Asian, Black	49% Spanish surnamed 49% white 2% Black, Asian, Native American	99% white 1% Black and Spanish surnamed	98% Black 2% white, Spanish surnamed, other

TABLE 2

DISTRIBUTION OF USMES AND COMPARISON CLASSES IN STUDENT STUDY

Grade Level	SCHOOL A		SCHOOL B		SCHOOL C		SCHOOL D		TOTAL	
	USMES Classes	Comparison Classes	USMES Classes	Comparison Classes	USMES Classes	Comparison Classes	USMES Classes	Comparison Classes	USMES Classes	Comparison Classes
2		2								2
2-3	2								2	
3			2	1					2	1
3-4					2	2	1		3	2
4			2	1			1	1	3	2
4-6	4	1							4	1
5			1	1	1	2	3	1	5	4
5-6				1						1
6			2				3	1	5	1
TOTAL	6	3	7	4	3	4	8	3	24	14

1. Students who had participated in the field-testing of the Pencil Problem in the fall of 1976 were excluded.*

2. The remaining students were divided into two sections: those who had participated in USMES before the 1976-1977 school year and those who had not. (There were students with past USMES experience in many comparison classes, as well as many USMES classes.) The schools provided information about past USMES experience for students in each class of the sample.

3. If a class contained more than one grade level (e.g., a combination of second and third-grade students), three students were chosen at random (using a calculator program to generate random numbers) from each grade level, in each class section (past/no past USMES experience). Each group of three students so chosen was administered the Pencil Problem as a group. If there were less than three students available in any category (e.g., "third grade, no past USMES experience"), then that category was dropped for that class; dropping of such categories was necessary to ensure that the Pencil Problem was always administered to homogeneous groups of three students.

4. If a class contained only one grade level, two groups of three students each were chosen at random from each section of the class. Each group of three students was administered the Pencil Problem together. There was one exception to this rule: in USMES classes which appeared to have been relatively inactive (i.e., whose teachers had reported doing less than 20 hours of USMES this year), we selected only one group of three students from each section.** Again, if there were too few students available to form a complete group of three students in any category, we avoided forming groups with less than three students (e.g., if only five students were available in a category from which we intended to form two groups, only one group was formed instead).

---

*In the pretesting, only USMES classes (or classes expected to do USMES) had been involved. Since the pre-test experience could have significantly affected students' performance on the spring Pencil Problem, and since almost all of the affected students would have been in USMES classes (disregarding the few classes expected to do USMES which actually did no USMES this year), to have included these students would have meant introducing a potentially serious bias into the spring data. Excluding these students reduced the available number of students in most USMES classes by up to six students. On the average, less than 15% of the students in USMES classes and less than 5% of the students in comparison classes were excluded. The excluded students had been chosen by their teachers as "typical" of their classes. Their exclusion may have caused a mild reduction in the representativeness of the sample, from USMES classes especially, but this loss seemed less hazardous than the potential bias which would have resulted from including them.

**Since a very few classes which appeared to have been relatively inactive were later found to have been more active than it had appeared, these classes are underrepresented in overall USMES results. We do not believe this problem seriously distorts our results since it occurred infrequently.

5. We selected, at random, "standby" students within each category as replacements for students who might be absent. When possible, we selected 50% as many standby students as the number of students originally selected in a category. (This number proved sufficient in all cases.)

The final distribution of Pencil Problem groups (each group containing three students) is shown in Table 3. Not surprisingly, the distribution shown in this table resembles the distribution of classes shown in Table 2. Hence, the constraints on generalizing from such a distribution which were discussed above apply to Pencil Problem results as well as QARPS results.

TABLE 3

DISTRIBUTION OF PENCIL PROBLEM GROUPS*

Grade Levels	SCHOOL A		SCHOOL B		SCHOOL C		SCHOOL D		TOTAL	
	USMES Classes	Comparison Classes	USMES Classes	Comparison Classes	USMES Classes	Comparison Classes	USMES Classes	Comparison Classes	USMES Classes	Comparison Classes
2		8								8
2-3	5								5	
3			4	2					4	2
3-4					5	6	3		8	6
4			4	2			2	3	6	5
4-6	11	5							11	5
5			3	3	3	6	7	4	13	13
5-6				2						2
6			6				8	4		4
TOTAL	16	13	17	9	8	12	20	11	61	45

*Each group contains three students.

THE USMES PROGRAM IN THE STUDENT STUDY SCHOOLS

Publications of the USMES project have described its program in the following ways:

> "Unified Sciences and Mathematics for Elementary Schools: Mathematics and the Natural, Social, and Communications Sciences in Real Problem Solving (USMES) was formed in response to the recommendations of the 1967 Cambridge Conference on the Correlation of Science and Mathematics in the Schools. Since its inception in 1970, USMES has been funded by the National Science Foundation to develop and carry out field trials of interdisciplinary units centered on long-range investigations of real and practical problems (or "challenges") taken from the local school or community environment."*
>
> "Children working on USMES units tackle problems in getting to school; problems in the classroom, the lunchroom, and the playground; problems in consuming, producing, and advertising goods; problems in finding the best way to learn things, to make decisions, and to communicate certain kinds of information. They work on real problems like a dangerous crosswalk near the school or classroom furniture that doesn't fit the students in the class....
>
> "USMES engages students in all aspects of the problem-solving process: definition of the problem; observation; measurement; collection of data; analysis of data using graphs, charts, statistics; discussion; formulation and trial of solutions; development and clarification of values; decision making; working with others in small groups; and communicating findings to others. In addition, students become more inquisitive, more critical in their thinking, and more self-reliant."**

USMES classes in the four Student Study schools used a wide variety of USMES published units, and several locally-developed units, as shown in Table 4.

The total amount of time devoted to each USMES unit ranged from 4.5 hours to over 100 hours. Most units took between approximately 10 and 50 hours. Usually (but not always) all the students in an USMES class were involved in USMES activities.

*_USMES Guide_, EDC, 1976.
**USMES brochure, EDC, 1976.

TABLE 4

USMES UNITS USED IN STUDENT STUDY SCHOOLS

Published USMES Units	Number of Classes Which Used Unit
Advertising	1
Bicycle Transportation	1
Consumer Research	1
Describing People	2
Designing for Human Proportions	1
Design Lab Design	1
Dice Design	1
Growing Plants	5
Manufacturing	1
Mass Communications	1
Orientation	3
Planning a Special Occasion	2
Protecting Property	1
School Zoo	1
Using Free Time	2
Weather Predictions	2

Locally Developed Units	Number of Classes Which Used Unit
Keeping Restrooms Clean	1
Compiling a Cookbook	1
Fire Safety	1
Improving School	3
Metric Awareness	1
Purchasing a School Mat	1
School Spirit	1
Television Survey	1
Traffic in School	1

DATA COLLECTION PROCEDURES

Three members of the USMES research staff visited each of the four schools, in turn, during April and May 1977. They administered the QARPS to each class in the sample, reading the items aloud if students had any difficulty in reading the items for themselves. The same staff members administered the Pencil Problem to the selected groups of three students, using a procedure which ensured that administrators never knew, at the time of a Pencil Problem administration, from which class the three students had come. That procedure was to have each administrator's next group of Pencil Problem students be selected (more or less at random out of all the remaining groups) and brought to the administration area by one of the other administrators. We took great pains to abide by this procedure, in order to prevent administrators from being biased in their observations or in their conversational tone.

DATA ANALYSIS METHODS

Data Processing

Data were entered into computer files and formatted with specially-developed computer programs which performed a number of verification procedures to ensure accuracy of the data entry. In addition, a number of entries were spot-checked, and basic cross-distributions were analyzed to further verify the accuracy of data entry. Scaling, recoding, and statistical calculations were performed with the computer programs of the *Statistical Package for Social Sciences* (SPSS), Version 6.

Coding of Outcome Variables

For purposes of analysis, students' responses on the Pencil Problem and QARPS were coded as follows:

1. Pencil Problem: Total Number of Factors Considered. For this outcome variable, students received one point for each (of nine possible) critical factors that they mentioned in discussing the quality and cost of pencils and erasers. The possible factors were (1) how well a pencil writes, (2) how well a pencil can be sharpened, (3) how well a pencil erases, (4) the cost of a pencil, (5) the kind of lead in a pencil, (6) the size of a pencil, (7) the strength of a pencil, (8) how well an eraser erases, and (9) the cost of an eraser. The first seven factors (those concerning pencils) might have been mentioned at any of three places in the Pencil Problem--in response to "How will you decide on a pencil?" or "Which pencil do you think (your principal) should buy and why?" or on the "Advice to Principal" sheet. However, students received no more than one point for each factor, no matter how many times they mentioned it.

2. Pencil Problem: Number of Investigations Carried Out. On this variable, students received one point for each pencil they wrote with, one point for each pencil they sharpened, one point for each pencil they erased with, and one point for every other kind of investigation (like comparing the lengths of pencils) they carried out.

3. Pencil Problem: Response to "How Would You Convince Other Students That the Pencil You Chose was the Best One?" On this variable, students received one point for each pencil factor they said they would "show other kids" or "tell other kids about." In addition, they received a point if

they said they would "let other kids try the pencils out" (or the equivalent) and another point if they said they would ask other students for their opinions about pencils.

4. QARPS: Attitudes Toward Working on Real Problems and Producing Effective Solutions. All the QARPS variables, including this one, involved summing students' responses over several QARPS items. The possible responses on each item had been coded as follows (with a few exceptions noted below):

   - "Agree a lot": +2
   - "Agree a little": +1
   - "Disagree a little": -1
   - "Disagree a lot": -2

   The eight QARPS items on which the Attitudes Toward Working on Real Problems variable was based included:

   - "If kids show that a change is needed, then grownups will go along."
   - "I figure out what I'm going to do before I start something."
   - "Kids usually can't come up with good ways to solve problems around school." (Coded in reverse, e.g., "Disagree a lot" coded +2.)
   - "Usually there is just one way to solve a real life problem." (Coded in reverse.)
   - "I like to make things better around school."
   - "I think I am good at solving real life problems."
   - "As soon as I think of one way to solve a problem, I don't think about other ways." (Coded in reverse.)
   - "I like working on problems that have more than one answer."

5. QARPS: Attitudes Toward Group Interaction.
   This variable was based on the following seven QARPS items:

   - "I like to work in small groups with other students."
   - "I like to help decide what we do in class."
   - "I like to work by myself." (Coded in reverse.)
   - "If I think other kids won't like my idea, then I keep quiet about it." (Coded in reverse.)

- "I like to ask questions."
- "I like to talk to other students about my ideas."
- "I like to listen to other students talk about their ideas."

6. QARPS: Attitudes Toward Specific Problem-Solving Activities. This variable was based on the following five QARPS items:

- "I like to write about my ideas."
- "I like to measure things when it helps solve real life problems."
- "I like to make charts and graphs that show things I found out."
- "I like to use numbers to solve real life problems."
- "I like to do surveys to find out what other people think."

While these six variables were admittedly crude composites of students' responses, they allowed us to analyze a manageably small, relatively coherent set of outcomes. Time did not permit an analysis of the technical qualities of these composite variables, but we believed that they were logically related to the six outcomes on which the Student Study was focused.

Stages of Analysis

Prior to the formal analysis of results, the Pencil Problem data were "aggregated" into mean results for each administration group. The administration groups were the groups of three students to whom the Pencil Problems had been administered at once, and aggregation had the effect of producing one set of average scores for each administration group rather than one set of scores for each individual student.

The aggregation was necessary to meet a technical requirement of the statistical significance tests used in our analysis. The significance tests we used* assume that each observation on which the test is based is "independent" of other observations. However, it was clear to administrators of the Pencil Problem that the three students in each administration group were often influenced by the responses of the other students in their group. There seemed to be no basis for treating individual students' Pencil Problem responses as independent observations, under the circumstances. To have done so would have inflated the number of supposedly independent observations, since there were three times as many individuals as groups. The inflated number of observations would have produced artificially favorable significance levels, since the significance tests are strongly affected by the number of observations. Aggregating the Pencil Problem data into administration group mean scores prevented this problem.

*These were primarily F-tests for analyses of variance and for multiple regressions.

For similar reasons, we experimented with a different kind of aggregation of both QARPS and Pencil Problem data. For these experiments, we divided each class into two sections: (1) students with USMES experience before this year and (2) students with no USMES experience before this year. Then we aggregated the data into mean scores for each class section. Our purpose was to examine the possibility that the scores of students with comparable USMES experience, both this year and in the past, should be treated as "not independent." In these experiments the aggregated class section data produced results similar to those produced by the original data. Therefore, we have not included class section results in this report, except for two samples in Appendices G and H. The complete results for class sections are available from the USMES office at Education Development Center.*

Our main analysis of the Student Study data took place in two stages: (1) examination of each outcome variable's apparent relationships with each independent variable**, considering just one independent variable at a time; and (2) analysis of each outcome variable's relationships with each independent variable** after all the other independent variables were taken into effect. The second stage of analysis made allowances for incidental relationships among the independent variables, like a disproportionate number of non-USMES students coming from certain grade levels.

---

*We also conducted a separate set of analyses using a modified coding of past USMES experience. This coding only applied to students who had indicated (on the cover of the QARPS) a different amount of past USMES experience from the amount given in school records. Ultimately, we opted for using the school records in cases of conflict, to be consistent with our sampling strategy. However, results of this modified coding are also available from the USMES office.

**Achievement data were used only in the second stage of analysis and only for Pencil Problem outcomes. In these analyses, composite achievement levels (combining mathematics and verbal scores), expressed in months above or below grade level (adjusted for the date of test administration), were used. As noted above, such data must be interpreted cautiously.

RESULTS

In the first stage of analysis, we examined the mean scores of each outcome variable by amount of USMES (see Table 5).*

TABLE 5

MEAN SCORES ON EACH OUTCOME VARIABLE BY AMOUNT OF USMES

Outcome Variable	Amount of USMES		
	None	Little	Substantial
PENCIL PROBLEM			
Total Number of Factors Considered	3.01	3.52	3.35
Number of Investigations Carried Out	6.64	7.09	6.90
Responses to "How Would You Convince Other Students That the Pencil You Chose was the Best One?"	1.03	1.19	1.15
QARPS			
Attitudes Toward Working on Real Problems and Producing Effective Solutions	4.11	3.71	5.14
Attitudes Toward Group Interaction	4.02	4.41	5.10
Attitudes Toward Specific Problem-Solving Activities	2.03	3.01	2.44

Table 5 shows that the mean scores for Substantial amount of USMES exceeded those for None on all Six outcomes. Surprisingly, the mean scores for Little USMES are the highest of all on four outcomes and the lowest on one.

However, mean scores alone do not show the whole pattern of results. For instance, some differences between mean scores are so small as to be statistically insignificant, considering the margins of error involved. Table 6 shows

*Appendices A-F give more detailed results for the QARPS.

the "one standard error confidence intervals" around the mean scores shown in Table 5. A confidence interval may be interpreted as follows: allowing for chance errors in sampling and measurement, we can feel confident that about two thirds of the mean scores we might obtain in similar circumstances would fall within this interval. Also, about 95% of these mean scores would fall within an interval twice as large (a "two standard error" interval).

TABLE 6

ONE STANDARD ERROR CONFIDENCE INTERVALS AROUND MEAN SCORES

Outcome Variables	Amount of USMES		
	None	Little	Substantial
PENCIL PROBLEM			
Total Number of Factors Considered	2.81 to 3.21	3.25 to 3.79	3.20 to 3.50
Number of Investigations Carried Out	6.01 to 7.27	6.27 to 7.91	6.43 to 7.36
Responses to "How Would You Convince Other Students That the Pencil You Chose was the Best One?"	0.85 to 1.21	1.05 to 1.33	1.06 to 1.24
QARPS			
Attitudes Toward Working on Real Problems and Producing Effective Solutions	3.80 to 4.43	3.41 to 4.02	4.94 to 5.34
Attitudes Toward Group Interaction	3.72 to 4.31	4.11 to 4.71	4.90 to 5.30
Attitudes Toward Specific Problem-Solving Activities	1.74 to 2.32	2.73 to 3.28	2.26 to 2.62

Table 6 indicates that on all six outcomes there is some "overlap" among the confidence intervals for None, Little, and Substantial. Figures 1 through 6 show these overlaps in graphical form.

Table 6 and Figures 1 through 6 show that the one standard error confidence interval for Substantial Amount of USMES actually overlaps that for None on all the Pencil Problem outcomes and on Attitudes Toward Specific Problem-Solving Activities. The confidence interval for Little overlaps either the

Confidence Intervals and Means for None (N), Little (L), and Substantial (S) Amounts of USMES

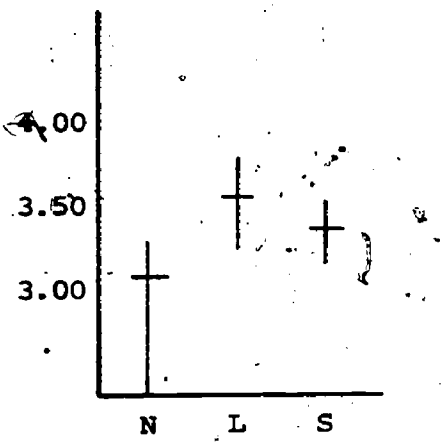


Figure 1. Total Number of Factors Considered (Pencil Problem).

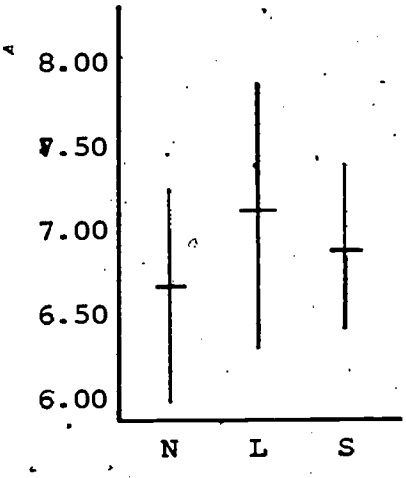


Figure 2. Number of Investigations Carried Out (Pencil Problem).

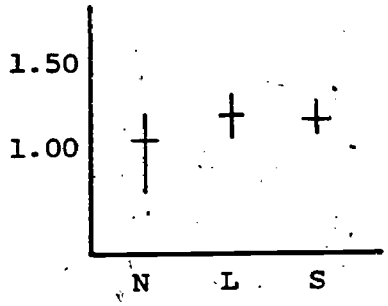


Figure 3. Responses to "How Would You Convince Other Students That the Pencil You Chose was the Best One?" (Pencil Problem)

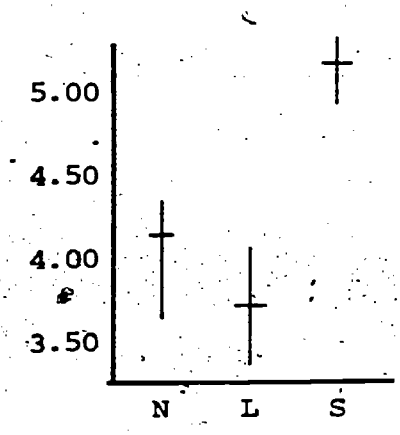


Figure 4. Attitudes Toward Working on Real Problems and Producing Effective Solutions (QARPS).

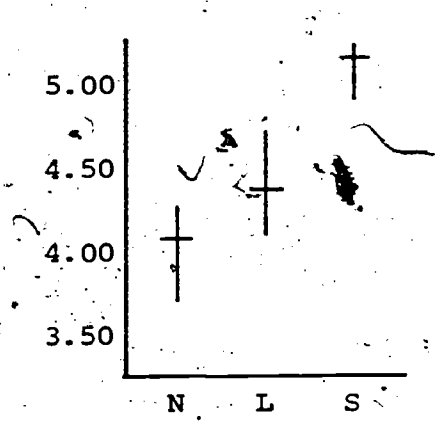


Figure 5. Attitudes Toward Group Interaction (QARPS).

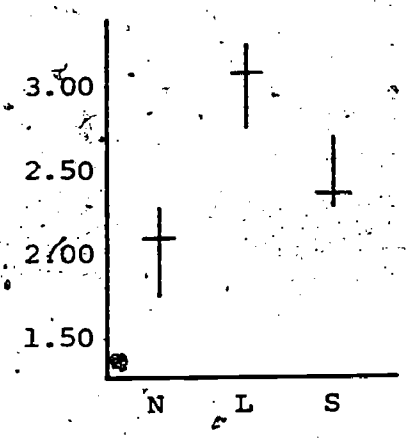


Figure 6. Attitudes Toward Specific Problem-Solving Activities (QARPS).

None or the Substantial interval on all outcomes except Attitudes Toward Specific Problem-Solving Activities.*

Table 7 shows that there were some pronounced differences among the results for the four different schools. It is interesting to note that each school ranked first on at least one outcome; school D ranked first on three outcomes.

TABLE 7

MEAN SCORES ON EACH OUTCOME VARIABLE BY SCHOOL

Outcome Variable	School A	School B	School C	School D
PENCIL PROBLEM				
Total Number of Factors Considered	3.40(2)	2.94(4)	3.37(3)	3.45(1)
Number of Investigations Carried Out	7.36(2)	5.59(4)	7.85(1)	6.85(3)
Responses to "How Would You Convince Other Students that the Pencil You Chose was the Best One?"	1.00(3)	1.49(1)	1.00(3)	1.02(2)
QARPS				
Attitudes Toward Working on Real Problems and Producing Effective Solutions	6.16(1)	4.47(2)	3.22(4)	4.42(3)
Attitudes Toward Group Interaction	4.87(2)	4.85(3)	2.95(4)	5.46(1)
Attitudes Toward Specific Problem-Solving Activities	1.87(3)	2.67(2)	0.74(4)	3.69(1)

Note. Numbers in parentheses show relative ranks of schools.

*An analysis of variance confirmed what these figures suggest: that the superiority of Substantial over None is only statistically significant at the .05 level on two outcomes: Attitudes Toward Working on Real Problems and Producing Effective Solutions, and Attitudes Toward Group Interaction. Similarly, the differences between Little Amount of USMES and the other two Amounts of USMES are only statistically significant on the outcome Attitudes Toward Specific Problem-Solving Activities. (At the .33 significance level, both the Substantial and the Little group also scored significantly better on another outcome--Total Number of Factors Considered.)

Table 8 indicates that there were also noteworthy differences by grade. However, the pattern of results by grade is more consistent than the pattern by school: the upper grades produced the highest mean scores on all outcomes, and the sixth grade ranked number one on five of the six outcomes.

TABLE 8

MEAN SCORES ON EACH OUTCOME VARIABLE BY GRADE

Outcome Variable	Grade 2	3	4	5	6
PENCIL PROBLEM					
Total Number of Factors Considered	3.13(3)	2.95(4)	3.51(2)	3.13(3)	3.60(1)
Number of Investigations Carried Out	5.67(4)	4.88(5)	6.46(3)	7.66(2)	7.95(1)
Responses to "How Would You Convince Other Students That the Pencil You Chose was the Best One?"	1.07(4)	1.05(5)	1.22(1)	1.09(3)	1.16(2)
QARPS					
Attitudes Toward Working on Real Problems and Producing Effective Solutions	5.96(2)	3.40(5)	4.20(3)	4.19(4)	6.11(1)
Attitudes Toward Group Interaction	4.00(4)	3.20(5)	4.33(3)	4.69(2)	6.37(1)
Attitudes Toward Specific Problem-Solving Activities	2.15(4)	2.22(3)	2.56(2)	2.06(5)	3.18(1)

Note. Numbers in parentheses show relative ranks of grade levels.

The pattern of results by school and by grade suggests a need to treat the earlier results by Amount of USMES with some caution. As suggested by Table 2 (page 12), a disproportionate number of the classes in School D and in the sixth grade were USMES classes. Therefore, the apparently favorable conditions in School D and in the sixth grade may be somewhat confused with USMES effects.

A possible way to resolve this confusion is to examine the Amount of USMES results school-by-school and grade-by-grade. Tables 9 and 10 show the mean scores on each outcome variable by Amount of USMES within each school and grade, respectively.

TABLE 9

MEAN SCORES ON EACH OUTCOME VARIABLE BY SCHOOL AND AMOUNT OF USMES

Outcome Variable	School	Amount of USMES		
		None	Little	Substantial
PENCIL PROBLEM	A	3.21	3.61	3.42
	B	2.75	2.67	3.12
Total Number of Factors	C	3.50	3.72	3.00
Considered	D	2.67	3.76	3.65
	A	8.13	6.44	7.31
Number of Investigations	B	6.21	3.75	5.76
Carried Out	C	6.83	9.78	7.17
	D	5.29	7.24	7.33
Responses to "How Would You	A	0.83	1.44	0.91
Convince Other Students That	B	1.58	1.00	1.57
the Pencil You Chose was the	C	1.11	1.00	0.92
Best One?"	D	0.57	1.24	1.12
QARPS	A	6.78	5.08	6.41
	B	3.64	2.97	5.26
Attitudes Toward Working on Real	C	2.82	3.03	3.50
Problems and Producing Effec-	D	4.76	3.54	4.93
tive Solutions				
	A	4.52	4.90	4.92
Attitudes Toward Group	B	4.14	3.60	5.51
Interaction	C	1.94	2.58	3.59
	D	5.32	5.01	5.84
	A	0.41	3.17	1.71
Attitudes Toward Specific	B	2.03	3.07	2.97
Problem-Solving Activities	C	0.39	0.97	0.83
	D	4.06	3.59	3.62

TABLE 10

MEAN SCORES ON EACH OUTCOME VARIABLE BY GRADE AND AMOUNT OF USMES

Outcome Variable	Grade	Amount of USMES		
		None	Little	Substantial
PENCIL PROBLEM	2	2.87	3.56	3.17
	3	2.42	2.93	3.40
Total Number of Factors	4	3.43	3.94	3.36
Considered	5	3.13	3.44	3.04
	6	2.67	3.78	3.76
	2	5.87	5.33	5.67
	3	3.67	5.40	5.33
Number of Investigations Carried	4	7.14	8.06	5.36
Out	5	8.50	8.78	6.86
	6	4.56	6.33	8.96
	2	1.00	1.44	0.67
Responses to "How Would You Con-	3	1.08	1.40	0.67
vince Other Students That	4	1.10	0.89	1.44
the Pencil You Chose was the	5	1.20	1.22	0.98
Best One?"	6	0.33	1.11	1.33
QARPS	2	7.00	4.97	7.25
	3	3.33	2.36	3.86
Attitudes Toward Working on Real	4	3.60	3.84	4.65
Problems and Producing Effec-	5	4.03	3.59	4.54
tive Solutions	6	5.33	3.85	6.80
	2	4.00	4.21	3.25
	3	2.53	3.18	3.61
Attitudes Toward Group Interaction	4	4.32	4.16	4.39
	5	4.23	4.31	5.06
	6	5.33	5.82	6.68
	2	-0.50	3.79	2.13
	3	1.49	3.45	2.17
Attitudes Toward Specific Problem-	4	2.97	2.60	2.32
Solving Activities	5	1.86	2.34	2.01
	6	2.44	3.77	3.16

In general, School B shows below-average benefits from USMES on problem-solving outcomes but above-average benefits on attitude outcomes. School D shows the opposite pattern of results. Grade 6 shows above-average benefits on all outcomes, especially problem-solving outcomes. However, the patterns shown in Tables 9 and 10 are not uniform because of the relatively small number of students in each entry. Furthermore, these results are hard to summarize because there are so many of them.

A different approach to separating the effects of Amount of USMES from those of school and grade is regression analysis. Tables 11 through 14 show a summary of regression results for each outcome variable and independent variable.

TABLE 11

REGRESSION ANALYSIS RESULTS FOR AMOUNT OF USMES

Outcome Variables	B*	S.E. (B)*	Incr. R^2 (%)	R^2 Other (%)	Signif.*
PENCIL PROBLEM					
Total Number of Factors Considered	0.07	0.16	0.2	11.4	.66
Number of Investigations Carried Out	0.61	0.43	1.8	26.9	.15
Responses to "How Would You Convince. . ."	0.07	0.11	0.4	8.9	.54
QARPS					
Attitudes Toward Working on Real Problems. . .	0.28	0.18	0.2	6.2	.13
Attitudes Toward Group Interaction	0.33	0.18	0.3	6.8	.06
Attitudes Toward Specific..Activities	0.23	0.17	0.1	6.1	.17

*Note. Column headings are defined in text below.

TABLE 12

REGRESSION ANALYSIS RESULTS FOR SCHOOL

Outcome Variables	B*	S.E. (B)*	Incr. R^2 (%)	R^2 Other (%)	Signif.*
PENCIL PROBLEM					
Total Number of Factors Considered	-	-	4.8	6.9	.21
Number of Investigations Carried Out	-	-	5.4	23.3	.13
Responses to "How Would You Convince. . ."	-	-	8.4	0.9	.08
QARPS					
Attitudes Toward Working on Real Problems	-	-	4.5	1.9	.00
Attitudes Toward Group Interaction	-	-	2.7	4.4	.00
Attitudes Toward Specific..Activities	-	-	5.9	0.3	.00

*Note. Column headings are defined in text below.

TABLE 13

REGRESSION ANALYSIS RESULTS FOR GRADE

Outcome Variables	B*	S.E. (B)*	Incr. R^2 (%)	R^2 Other (%)	Signif.*
PENCIL PROBLEM					
Total Number of Factors Considered	0.04	0.13	0.1	11.6	.77
Number of Investigations Carried Out	1.18	0.36	9.7	19.0	.00
Responses to "How Would You Convince. . ."	0.00	0.09	0.0	9.3	.99
QARPS:					
Attitudes Toward Working on Real Problems. . .	0.62	0.14	1.9	4.6	.00
Attitudes Toward Group Interaction	0.70	0.13	2.5	4.7	.00
Attitudes Toward Specific Activities	-0.12	0.13	0.1	6.2	.66

*Note. Column headings are defined in text below.

TABLE 14

REGRESSION ANALYSIS RESULTS BY ACHIEVEMENT LEVEL

Outcome Variables	B*	S.E. (B)*	Incr. R^2 (%)	R^2 Other (%)	Signif.*
PENCIL PROBLEM					
Total Number of Factors Considered	0.04	0.02	5.4	6.3	.03
Number of Investigations Carried Out	0.12	0.04	6.5	22.1	.01
Responses to "How Would You Convince. . ."	0.00	0.01	0.1	9.1	.72

*Note. Column headings are defined in text below.
Achievement data were only available for the Pencil Problem sample.

In Tables 11 through 14 the column headed "B" gives "unstandardized regression coefficients" (conventionally called "B's" in statistics). A B indicates the expected difference in an outcome variable for each difference of one unit in the corresponding independent variable. For instance, the B of 0.04 on the top line of numbers in Table 13 means that the expected difference in Total Number of Factors Considered corresponding to a difference of one grade level, in this sample, is +0.04 Factors.

For Amount of USMES, the units are quantities of USMES experience over one or more school years. These units were coded as follows:

0 for no USMES experience, past or present;

1 for "little" USMES experience--either some amount less than 20 hours this year or else past USMES experience only;

2 for "substantial" USMES experience--either experience in the past as well as this year or else over 20 hours of experience this year.

Thus, for example, the B of 0.07 on the top line of numbers in Table 11 means that students with "little" USMES experience would be expected to consider 0.07 more factors in the Pencil Problem than students with no USMES experience, and likewise that students with "substantial" USMES experience would be expected to consider 0.07 more factors than students with "little" USMES experience.

For the independent variable Achievement Level, the units are months above or below grade level in composite mathematics and verbal achievement. Thus, the B of 0.04 on the top line of numbers in Table 14 means that the expected difference in Total Number of Factors Considered corresponding to one month of grade-equivalent achievement is 0.04. However, the figures for Achievement Level should be read with special caution, as indicated on pages 9 and 22.

There are no entries in the column headed "B" in Table 12, because there are no "units" to express the different categories of the independent variable School. B's would not be appropriate statistics in Table 12.

The column headed "S.E. (B)" in Tables 11 through 14 gives "Standard Errors" of the corresponding B's. These standard errors are analogous to the standard errors of mean scores that were used to create Table 6. The standard errors of the B's may be interpreted as follows: allowing for chance errors, we can feel confident that approximately 68% of the B's computed in similar circumstances would fall within one standard error of the reported B's and that approximately 95% of similarly computed B's would fall within two standard errors of the reported B's. Thus, for example, about 68% of the similarly computed B's for Total Number of Factors Considered and Amount of USMES (top line of Table 11) should fall within plus or minus 0.16 (one standard error) of 0.07 (the reported B), and about 95% should fall within 0.32 (two standard errors) of 0.07.

The column headed "R^2 Other (%)" and the column headed "Incr. R^2" in Tables 11 through 14 involve "squared multiple correlation coefficients" or R^2's. These are statistical measures of the relationship between an outcome variable and one or more independent variables. "Incr. R^2", or the increment in R^2, may be interpreted as the percent of the total variation in an outcome variable which is "accounted for," or explained, by a corresponding independent variable after the effects of all the other independent variables are taken into account. It gives a conservative estimate of the effect of an independent variable because it removes from consideration not only the effects of other independent variables but also any joint effects of two or more independent variables. For instance, the joint effects of Amount of USMES and of Grade (which are impossible to separate because there were a disproportionate number of USMES classes in the Grade 6 sample) are removed from consideration when the increment in R^2 is computed for either Amount of USMES or for Grade.

Hence, the increment in R^2 provides a relatively conservative measure for both Amount of USMES and Grade.

The numbers in the column headed "R^2 Other" represent the effects of all other independent variables (and any inseparable joint effects) for each increment in R^2. The sum of an increment in R^2 and its corresponding "R^2 Other" represents the percent of an outcome variable's total variation which is accounted for by all of the independent variables.*

As an example of the R^2 measures, consider the top line of Table 14, showing results for the outcome variable Total Number of Factors Considered and for the independent variable Achievement Level. In this case, the "R^2 Other" of 6.3% means that all the independent variables except Achievement Level (i.e., Amount of USMES, School, and Grade) accounted for 6.3% of the variation in Total Number of Factors Considered. The increment in R^2 of 5.4% means that Achievement Level accounted for 5.4% of the variation in Total Number of Factors Considered after all the other independent variables were accounted for. In all, 11.7% (6.3% + 5.4%) of the variation in Total Number of Factors Considered was accounted for by all the independent variables. The remaining 88.3% of the variation was not accounted for; presumably it was due to chance effects and to unmeasured factors.

While the increment in R^2 can be useful for comparing the influence of different independent variables, its practical significance depends on the situation. For example, it may be interesting to note that School accounts for more of the variation in Attitudes Toward Working on Real Problems and Producing Effective Solutions than does either Grade or Amount of USMES; however, for teachers and administrators, School and Grade may be unchangeable "givens." Amount of USMES may be the only variable which represents an opportunity for improving these attitudes. In this case, the magnitude of the increment in R^2 may not be as important as the fact that the increment--and the B--are simply positive for Amount of USMES.

The column headed "Signif." in Tables 11 through 14 gives the significance levels for corresponding increments in R^2 and standard errors of B (the significance level applies to both). The significance level represents the probability that the reported results could have occurred by chance alone, and a low significance level is therefore desirable. How low a significance level "should be" depends on the risks which an educator may be ready to accept. For scientific research, it is common to look for significance levels of .05 or lower, i.e., no more than a 1-in-20 chance that a reported "finding" was really due to luck. However, a school principal who is willing to accept a risk of up to one chance in three, say, because he or she believes that the potential benefits outweigh the risks, should presumably be willing to accept a significance level of up to .33.

A few general observations can be made about the results reported in Tables 11 through 14. First, it is important to remember that the lower significance levels on QARPS outcomes are partly a result of the much greater sample size for QARPS analyses (about 1000 individual students) than for

*Note that Achievement Level is one of the independent variables for the Pencil Problem outcomes but not for the QARPS outcomes.

Pencil Problem Analyses (about 100 administration groups of three students each); large sample sizes tend to give low significance levels. An important pattern appears throughout the "Incr. R^2" and "Signif." columns of Table 11: each of the independent variables has strong relationships with only some of the dependent variables. Amount of USMES has its strongest relationships with Number of Investigations Carried Out (in terms of increment in R^2) and with Attitudes Toward Group Interaction (in terms of significance level); it has modest relationships with the other two QARPS outcomes and almost negligible relationships with the other Pencil Problem outcomes. School has very low significance levels on the QARPS outcomes but more modest significance levels on the Pencil Problem outcomes (this may be in part an artifact of the large QARPS sample sizes, since School's increments in R^2 are only a little larger for Pencil Problem outcomes than for QARPS outcomes). Grade has relatively strong relationships with Number of Investigations Carried Out, Attitudes Toward Working on Real Problems and Producing Effective Solutions, and Attitudes Toward Group Interaction, but practically negligible relationships with the other outcomes. Achievement Level has relatively strong relationships with Total Number of Factors Considered and with Number of Investigations Carried Out (despite the fairly large amounts of missing data on achievement) but almost no relationship with responses to "How Would You Convince Other Students That the Pencil You Chose was the Best One?"

In reviewing Tables 11 through 14 it may be appropriate to recall that, for reasons discussed above, the regression analysis may somewhat understate the possibly complex effects of different amounts of USMES experience. Any non-linear effects (e.g., situations where Little USMES produced positive results of greater magnitude than Substantial USMES) would not improve the increment in R2 or the significance level for Amount of USMES. Furthermore, some effects of USMES may have been obscured by the imbalance in the sample of classes (e.g., in Grade 6 and School D). Therefore, it may be best to balance the impressions gained from Tables 11 through 14 with those gained from the earlier tables when drawing conclusions about the effects of USMES.

Finally, the results for the QARPS questions directed to USMES students are shown in Table 15.

The results of Table 15 parallel those of several similar, earlier studies (see, for example, page 38 of The USMES Guide, 4th ed.). Both the overall very positive attitude toward doing USMES work and the divided opinion about whether USMES is "hard work" have been observed in all the earlier studies, although the exact format of the questions--as well as the times and locations of the studies--have varied.

TABLE 15

PERCENTAGE OF USMES STUDENTS GIVING EACH RESPONSE ON QARPS ITEMS DIRECTED TO STUDENTS WITH USMES EXPERIENCE

QARPS Item	Response			
	"Agree a lot"	"Agree a little"	"Disagree a little"	"Disagree a lot"
21. I think USMES work is fun.	74.2	19.0	3.1	3.7
22. I think USMES work is boring.	5.0	8.9	16.9	69.2
23. Doing USMES is hard work.	16.0	34.2	22.0	27.8
24. I don't know why we do some things in USMES.	16.5	28.5	26.0	29.1
25. Doing USMES makes me think.	59.4	28.6	7.4	4.6
26. I think USMES work is confusing.	9.8	20.1	26.7	43.4
27. I think USMES work is important.	68.2	21.0	7.3	3.6
28. In USMES it's hard to decide what to do next.	22.0	36.4	24.7	16.9
29. I would like to do more USMES.	70.2	16.7	6.4	6.7

CONCLUSIONS

Effects of USMES on Students

The mean scores of the Substantial Amount of USMES group were higher than those of the None group on every outcome, and the mean scores of the Little Amount of USMES group were higher than those of the None group on every outcome but one*. However the Substantial Amount of USMES group did not always score higher than the Little Amount of USMES group. (See Table 5, p. 23, and Figures 1-6, p. 25.) Also, these effects were only statistically significant on some of the outcomes (see footnote on page 26).

The magnitude of these USMES effects were generally small, compared to the effects of School, Grade, and Achievement Level. On each outcome there was at least one other independent variable with a markedly larger influence than Amount of USMES. Also, the effects of Little Usmes were often erratic: on four outcomes the Little Amount of USMES group had a mean score above both the Substantial and the None group, while on another outcome the Little Amount of USMES group had a mean score below both the other groups. Partly for these reasons, a linear regression analysis showed only moderate significance levels (.06 to .17) on four outcomes for Amount of USMES and poor significance levels (over .50) on the other two outcomes. One of the statistically significant regression coeffecients was negative (see Table 11, p.30).

The effects of Amount of USMES were most pronounced in a few settings, such as School D. The None group in School D scored the lowest of all four None groups on all the Pencil Problem outcomes, yet the Substantial Amount of USMES group in School D scored the highest of all four Substantial groups on two of these outcomes and second-highest on the third (see Table 7, p. 26). Similar effects of Amount of USMES were found in the 6th grade (see Table 8, p. 27); however, since many of the 6th graders in the sample came from School D, the 6th grade and School D results may be related.

The motivational effects of USMES did not seem to be so localized. The responses to questions directed at students with USMES experience confirmed once again a repeated finding: the great majority of USMES students enjoy USMES work, find it important, and would like to do more (see Table 15, p. 35). This finding appeared to be true in all the schools and at all grade levels (see Appendices D and E, pp. 41-42).

Factors Influencing the Effectiveness of USMES

Why weren't the motivational effects of USMES translated into more consistent results on the main Pencil Problem and QARPS outcomes? We can only speculate on this issue, but it appears that several factors were important.

Little amounts of USMES (USMES experience in the past only, or less than 20 hours of USMES in the current year with no past experience) evidently produced inconsistent results. Although earlier research (see Chapter V of The USMES Guide, 4th ed.) has shown that time spent on USMES is unlikely to impede "basic" skills development, neither is a brief exposure to USMES certain to improve problem-solving skills significantly. Substantial amounts of USMES experience--including amounts greater than the 20 hours which defined "substantial" in this study--may be necessary for consistent, significant results.

---

*Moreover, of the 20 QARPS items completed by all students, there were five items that produced statistically significant results for Amount of USMES; on all five of these items, the Substantial Amount of USMES group produced the most favorable responses, and on four of the five the Little Amount of USMES group produced the next-most-favorable responses (see Appendix A, p. 38).

The specific activities done in an USMES class may also be crucial. The USMES School Study* found evidence that opportunities for skill development were often closely associated with the occurrence of serious investigations of data, extended discussions among students, and other recommended USMES activities. Furthermore, in the present study the Substantial Amount of USMES group appeared to have engaged in significantly higher levels of many recommended activities than the Little Amount of USMES group (see Appendix F, p. 43), and the Substantial group produced the most consistently positive results.

In addition, the surprisingly large effects of School (which included some community differences in this study), Grade, and Achievement Level suggest that all of these factors make important contributions to the impact of USMES experience. Certainly they should not be overlooked in future research on USMES and related programs.

Finally, the influence of the school administration should not be ignored. In School B, USMES had its most positive effect on attitudes, and in School D, USMES had its most positive effect on problem-solving skills. In each school, these were the outcomes that the principal of the school desired most to promote (see the USMES School Study report). The particular style of USMES chosen in each school emphasized the desired outcomes and was successful in improving it.

Some Implications

The results of this study suggest that the choice of USMES units (and the local development of units) should be based on a careful consideration of the outcomes that would be most valuable for a particular school and class. Furthermore, teachers should be aware that individual students may experience a wide range of effects of USMES experience. While virtually all students seem to enjoy USMES work, there seems to be considerable variation from student to student in just what is learned.** For consistent, significant results, this study and the USMES School Study suggest that "regular" USMES units of some duration--with recommended USMES teaching methods--should generally be preferred over abbreviated or haphazardly-created units.

In reviewing the results of this study, it seems clear that additional research is needed on the impact of extended (especially multi-year) USMES and on real problem solving at upper grade levels. The trends in this study suggest that these varieties of USMES may prove especially fertile for the development of effective problem-solving abilities and attitudes. Research on a local or regional level may prove just as valuable as national research on these issues.

For diagnostic and other kinds of research, we believe that the QARPS and the Pencil Problem may be useful instruments. The QARPS is easily administered, easily analyzed, and shows satisfactory discrimination. The Pencil Problem probes real-problem-solving abilities more deeply than "pencil and paper" tests yet can be administered and analyzed in practical ways.

Additional discussions of the findings reported here are needed. We hope that readers will join us in those discussions.

---

*Available from the USMES project office at Education Development Center.

**A comparison of Appendix H (p. 45) with Table 11 (p. 30) shows that the average effects of USMES on groups of students may be more predictable than effects on individual students.

APPENDIX A

Percentage of Students Responding "Agree a lot" or "Agree a little" on QARPS Items 1-20, by Amount of USMES

QARPS Item	Amount of USMES			Significance*
	None	Little	Substantial	
1. If kids show that a change is needed, then grownups will go along.	66.7	67.1	62.8	
2. I like to work in small groups with other students.	84.5	86.5	84.1	
3. I like to write about my ideas.	59.9	70.3	62.0	B
4. I figure out what I'm going to do before I start something.	88.8	88.7	88.5	
5. Kids usually can't come up with good ways to solve problems around school.	53.4	52.7	39.6	A
6. I like to help decide what we do in class.	82.3	87.8	88.6	
7. Usually there is just one way to solve a real life problem.	58.6	58.6	48.2	A
8. I like to measure things when it helps solve real life problems.	67.7	75.0	73.4	
9. I like to work by myself.	60.3	59.5	56.7	
10. I like to make things better around school.	92.2	89.6	90.9	
11. If I think other kids won't like my idea, then I keep quiet about it.	76.2	70.7	58.9	A
12. I like to ask questions.	69.1	73.9	73.4	
13. I like to make charts and graphs that show things I found out.	64.5	64.7	60.0	
14. I think I am good at solving real life problems.	54.1	56.1	62.5	
15. I like to talk to other students about my ideas.	77.3	77.7	80.5	
16. I like to use numbers to solve real life problems.	47.8	56.1	51.4	
17. As soon as I think of one way to solve a problem, I don't think about other ways.	47.6	59.3	39.1	A
18. I like to listen to other students talk about their ideas.	89.2	86.0	90.4	
19. I like to do surveys to find out what other people think.	78.0	76.9	82.7	
20. I like working on problems that have more than one answer.	74.0	70.0	74.3	

*Key. A: $p < .01$; B: $.01 \leq p \leq .05$; blank: $.05 < p$ (chi-squared test).

APPENDIX B

Percentage of Students Responding "Agree a lot" or "Agree a little" on QARPS Items 1-20, by School

QARPS Item	School A	B	C	D	Significance*
1. If kids show that a change is needed, then grownups will go along.	56.7	68.0	51.6	74.8	A
2. I like to work in small groups with other students.	86.0	82.2	78.8	89.5	A
3. I like to write about my ideas.	54.9	66.5	49.2	74.5	A
4. I figure out what I'm going to do before I start something.	87.4	90.7	86.4	88.9	
5. Kids usually can't come up with good ways to solve problems around school.	24.8	50.0	34.8	62.8	A
6. I like to help decide what we do in class.	89.3	86.5	85.3	86.7	
7. Usually there is just one way to solve a real life problem.	32.6	58.4	51.6	62.7	A
8. I like to measure things when it helps solve real life problems.	70.9	75.7	61.0	77.1	A
9. I like to work by myself.	59.5	55.0	65.0	56.0	
10. I like to make things better around school.	90.2	92.5	83.6	94.3	A
11. If I think other kids won't like my idea, then I keep quiet about it.	60.7	65.5	70.1	66.2	
12. I like to ask questions.	75.7	70.5	67.2	75.3	
13. I like to make charts and graphs that show things I found out.	53.7	66.1	46.7	73.1	A
14. I think I am good at solving real life problems.	59.6	54.8	51.6	67.0	A
15. I like to talk to other students about my ideas.	79.1	79.3	68.1	85.4	A
16. I like to use numbers to solve real life problems.	47.0	53.7	45.1	56.6	B
17. As soon as I think of one way to solve a problem, I don't think about other ways.	30.8	43.9	47.5	58.9	A
18. I like to listen to other students talk about their ideas.	92.5	89.7	81.4	90.8	A
19. I like to do surveys to find out what other people think.	85.1	72.2	78.1	85.5	A
20. I like working on problems that have more than one answer.	75.3	76.8	59.3	76.8	A

*Key. A: $p < .01$; B: $.01 \leq p \leq .05$; blank: $.05 < p$ (chi-squared test).

APPENDIX C

Percentage of Students Responding "Agree a lot" or "Agree a little" on QARPS Items 1-20, by Grade

QARPS Item	Grade 2	3	4	5	6	Significance*
1. If kids show that a change is needed, then grownups will go along.	70.9	56.1	68.5	62.7	68.5	B
2. I like to work in small groups with other students.	80.0	79.9	80.7	87.0	90.5	A
3. I like to write about my ideas.	54.5	65.9	64.3	58.9	68.3	
4. I figure out what I'm going to do before I start something.	90.9	92.1	86.7	87.0	89.7	
5. Kids usually can't come up with good ways to solve problems around school.	38.2	53.0	47.2	44.8	42.5	
6. I like to help decide what we do in class.	87.3	86.6	84.3	87.0	89.7	
7. Usually there is just one way to solve a real life problem.	41.8	64.6	57.8	53.0	41.3	A
8. I like to measure things when it helps solve real life problems.	70.4	72.4	70.3	73.0	74.8	
9. I like to work by myself.	67.3	70.7	63.1	53.5	47.3	A
10. I like to make things better around school.	92.7	86.0	91.9	90.6	93.7	
11. If I think other kids won't like my idea, then I keep quiet about it.	74.5	81.1	68.4	64.9	49.8	A
12. I like to ask questions.	69.1	73.2	72.9	70.1	75.2	
13. I like to make charts and graphs that show things I found out.	52.7	61.6	64.0	59.9	65.2	
14. I think I am good at solving real life problems.	49.1	54.9	53.8	62.0	66.4	B
15. I like to talk to other students about my ideas.	72.2	76.2	77.8	78.0	85.6	
16. I like to use numbers to solve real life problems.	56.4	50.0	51.2	49.7	54.1	
17. As soon as I think of one way to solve a problem, I don't think about other ways	40.0	54.3	48.0	50.2	31.8	A
18. I like to listen to other students talk about their ideas.	98.2	86.0	86.2	87.7	94.1	A
19. I like to do surveys to find out what other people think.	83.6	72.6	79.8	78.7	88.3	A
20. I like working on problems that have more than one answer.	78.2	75.0	72.8	69.6	76.2	

*Key. A: $p < .01$; B: $.01 \leq p \leq .05$; blank: $.05 < p$ (chi-squared test).

APPENDIX D

Percentage of USMES Students Responding "Agree a lot" or "Agree a little" on QARPS Items 21-29, by School

QARPS Item	School A	B	C	D	Significance*
21. I think USMES work is fun.	90.8	90.9	91.0	97.2	B
22. I think USMES work is boring.	15.2	11.0	14.0	15.0	
23. Doing USMES is hard work.	53.0	42.3	46.5	55.1	
24. I don't know why we do some things in USMES.	52.4	36.8	48.5	44.0	B
25. Doing USMES makes me think.	86.1	88.3	84.8	90.2	
26. I think USMES work is confusing.	33.7	31.5	22.0	29.6	
27. I think USMES work is important.	84.7	95.1	75.8	93.1	A
28. In USMES it's hard to decide what to do next.	61.7	54.0	54.0	60.9	
29. I would like to do more USMES.	79.9	87.7	89.0	90.2	

*Key. A: $p < .01$; B: $.01 \leq p \leq .05$; blank: $.05 < p$ (chi-squared test)

APPENDIX E

Percentage of USMES Students Responding "Agree a lot" or "Agree a little" on QARPS Items 21-29, by Grade

QARPS Item	Grade 2	3	4	5	6	Significance*
21. I think USMES work is fun.	100.0	86.1	91.6	93.2	97.4	A
22. I think USMES work is boring.	0.0	16.8	17.4	17.9	6.8	A
23. Doing USMES is hard work.	50.0	49.5	49.4	47.1	54.2	
24. I don't know why we do some things in USMES.	45.0	59.4	51.9	45.1	31.9	A
25. Doing USMES makes me think.	85.7	86.1	87.6	85.9	91.5	
26. I think USMES work is confusing.	40.0	41.2	31.4	32.5	19.2	A
27. I think USMES work is important.	94.7	94.1	89.7	82.0	92.6	A
28. In USMES it's hard to decide what to do next.	68.4	70.3	61.9	56.6	49.7	A
29. I would like to do more USMES.	90.0	82.2	88.6	88.9	89.5	

*Key. A: $p < .01$; B: $.01 \leq p \leq .05$; blank: $.05 < p$ (chi-squared test)

APPENDIX F

Percentage of USMES Students Checking Each Response on OARPS Item 30, by Amount of USMES (Little vs. Substantial)

Response	Amount of USMES		Significance*
	Little	Substantial	
30. When we do USMES, I usually do these things:			
measure things	39.6	35.8	
do surveys of people	32.0	47.1	A
build things	35.1	41.6	
ask questions	42.2	62.0	A
write about things we find out	41.3	52.2	A
work in small groups	60.0	72.4	A
make graphs or charts	20.4	35.1	A
do math	25.3	28.2	
help decide what we do	47.6	63.3	A
talk about my ideas	42.7	63.3	A
work by myself	23.1	22.2	

*Key. A: $p < .01$; B: $.01 \leq p \leq .05$; blank: $.05 < p$ (chi-squared test)

APPENDIX G

Sample of Class Section Results: Mean Scores on Each QARPS Outcome by Amount of USMES*

Outcome Variable	Amount of USMES		
	None	Little	Substantial
QARPS			
Attitudes Toward Working on Real Problems and Producing Effective Solutions	4.45	2.91	5.49
Attitudes Toward Group Interaction	3.92	3.87	5.15
Attitudes Toward Specific Problem-Solving Activities	1.52	2.58	2.49

*Comparable to bottom half of Table 5.

APPENDIX H

Sample of Class Section Results: Regression Analysis Results for Amount of USMES*

Outcome Variable	B**	S.E. (B)**	Incr. R^2 (%)**	R^2 Other (%)**	Signif.**
QARPS					
Attitudes Toward Working on Real Problems and Producing Effective Solutions	.58	.33	2.7	21.5	.08
Attitudes Toward Group Interaction	.44	.30	1.9	23.7	.15
Attitudes Toward Specific Problem-Solving Activities	.09	.32	0.9	11.5	.63

*Comparable to bottom half of Table 11.

**Column headings are defined in text after Tables 11-14.

APPENDIX I

Why More Research on USMES and Basic Skills is Needed.

We believe there were technical flaws of a very serious nature in both the 1973-74 data and the 1974-75 data used by the Boston University evaluation team*. The USMES staff conducted a careful, computerized analysis of the original data, during the fall of 1975 and the spring of 1976. Among the flaws we found were the following:

1. In the data collected during school year 1973-74:

   - Eight of the 12 studied classes contained missing or erroneous data, for which the investigators made no reported adjustment.

   - Eleven of the 12 classes had invalid or incorrect student-identification codes for at least some of the data, while seven of the classes actually had invalid or incorrect codes for the majority of their data; again, no adjustment for these flaws was reported.

2. In the data collected during school year 1974-75:

   - Approximately one-half of the data cards used in computing the reported results had either invalid identification codes or invalid data (e.g., letters where there should have been numbers, numbers in the wrong areas, etc.); apparently the data cards were automatically generated from machine-readable scoring sheets, without an adequate screening for "dirty" data (like stray pencil marks on the sheets, sheets out of order, etc.).

   - Twenty-eight of the 60 classes contained data that was biased and invalid, because the classes were not properly matched; in fact, there were not even equivalent numbers of observations at each grade level--the USMES data contained more observations than the comparison data at some grade levels and less at other grade levels--even though the scaling of the achievement scores ensured that scores would vary with grade levels.

Because of these flaws in the data collected by the Boston University team, we believe that the impact of USMES experience on students' achievement scores has not been adequately studied.

In general, our own analyses support the conclusions of the Boston University team that, "Clearly, USMES students do not fall behind their control counterparts in their performance on tests of basic skills," and that "though it is not statistically significant (at the .05 level), there is a noteworthy

*M. Shann, An Evaluation of Unified Sciences and Mathematics for Elementary Schools During the 1973-74 School Year, Boston University, August 1975; and M. Shann, N. Reali, H. Bender, T. Aiello, and L. Hench, Student Effects of an Interdisciplinary Curriculum for Real Problem Solving: The 1974-75 USMES Evaluation, Boston University, December 1975.

trend for the growth rates of the control classes to fall behind those for the USMES classes at the upper grade levels."* However, we believe that more accurate, detailed evidence on these conclusions is needed by teachers and administrators.

*Shann, et al., The 1974-75 USMES Evaluation.

APPENDIX J:

Sample Teacher Weekly Summary Form (including First Report on a Unit and Last Report on a Unit) and Teacher Interview Form

SUPPLEMENTARY **FIRST** REPORT ON A UNIT

(Please enclose with weekly summary for first week.)

1. How many students are in the class? _____

2. How did the challenge arise? (Was the problem already being discussed by the students? Did you initiate discussion of the problem? Did the problem arise in some other way?)

3. How do you and the students describe the challenge? (What is the "wording" of the challenge?)

TEACHER'S WEEKLY SUMMARY

Your name: _____

Unit: _____

Week covered by this report: _____

School: _____ Grade level: _____

1. Please comment on student planning this week. (Do you think they too much or too little time on planning?)

2. Were any new groups formed this week? (If so, please describe the groups, their tasks, and the process by which they were formed.)

3. What obstacles did students encounter in this week's work? How did they deal with them?

4. Did you conduct any skills sessions this week?
   What topics were covered?
   What made you decide that the session(s) were needed?

5. Did you have to provide any redirection for the students this week?
   What sort of intervention did you make?
   What made you decide that intervention was needed:

6. Were there any activities you hoped or expected to arise, which did not?

7. What else do we need to know about what happened this week?

8. We would like to have copies of any student work that you can send us.

PLEASE BE SURE YOU HAVE COMPLETED THE FOLLOWING SUMMARIES FOR ALL USMES SESSIONS THIS WEEK.

Monday

duration of session: _____ no. of students actively involved: _____

MAJOR ACTIVITIES: _____ whole class
_____ small group

Briefly describe activities:

DISCUSSIONS: _____ whole class _____ student-run _____ at beginning of session
_____ small group _____ teacher-run _____ in middle of session
_____ at end of session

Briefly describe discussion(s):

Tuesday

duration of session: _____ no. of students actively involved: _____

MAJOR ACTIVITIES: _____ whole class
_____ small group

Briefly describe activities:

DISCUSSIONS: _____ whole class _____ student-run _____ at beginning of session
_____ small group _____ teacher-run _____ in middle of session
_____ at end of session

Briefly describe discussion(s):

Wednesday

duration of session: _____ no. of students actively involved: _____

MAJOR ACTIVITIES: _____ whole group
_____ small group

Briefly describe activities:

DISCUSSIONS: _____ whole class _____ student-run _____ at beginning of session
_____ small group _____ teacher-run _____ in middle of session
_____ at end of session

Briefly describe discussion(s):

Thursday

duration of session: _____ no. of students actively involved: ____

MAJOR ACTIVITIES: _____ whole class
_____ small group

Briefly describe activities:

DISCUSSIONS: _____ whole class _____ student-run _____ at beginning of session
_____ small group _____ teacher-run _____ in middle of session
_____ at end of session

Briefly describe discussion(s):

Friday

duration of session: _____ no. of students actively involved: _____

MAJOR ACTIVITIES: _____ whole class
_____ small group

Briefly describe activities:

DISCUSSIONS: _____ whole class _____ student-run _____ at beginning of session
_____ small group _____ teacher-run _____ in middle of session
_____ at end of session

Briefly describe discussion(s):

SUPPLEMENTARY **LAST** REPORT ON A UNIT

(Please enclose with weekly summary for last week.)

1. Please comment on the ending of unit activities.
   Was the problem finally resolved or set aside?
   If it was resolved, what was the resolution?

2. What made the class decide that work on the unit should end?

3. Please comment on how effectively students were able to work together over the course of the unit.

USMES STUDENT STUDY
Teacher Interview

NAME OF TEACHER_____ SCHOOL_____

USMES UNIT_____ CITY_____

INTERVIEWER_____ DATE_____

1. DATES OF UNIT_____ NO. SESSIONS PER WEEK_____
   AVERAGE LENGTH OF TIME PER SESSION:_____
2. NUMBER OF CHILDREN ACTIVELY INVOLVED PER SESSION:_____
3. CHALLENGE as understood by teacher and students:("wording" of challenge?)

4. Briefly describe the activities that took place:

5. Was the problem finally resolved, or set aside?
   What was resolution:

   OR

   What made class decide to set it aside:

6. Were any small groups formed during work on the unit?
   How many children per group:

   AND

   Tasks of each group:

7. Did the major activities occur when the class was together as a whole or in small groups?

8. Did most discussions occur when class was together as a whole or in small groups?

   Fraction teacher-run:

   AND

   Fraction student-run:

9. Did you conduct any skill sessions during this unit?

   Topics covered:

   AND

   Why sessions were needed:

10. Did you have to provide any redirection for students at any point?

    Kind of intervention:

    AND

    Why intervention needed:

11. Please comment on student planning throughout the unit. Do you think they spent too much or too little time on planning?

12. Please comment on how effectively students were able to work together over the course of this unit. (What problems and successes did they encounter?)

USMES Technical Report

The Pencil Problem: An Instrument to Measure Real Problem-Solving Abilities

UNIFIED SCIENCES AND MATHEMATICS FOR ELEMENTARY SCHOOLS:
Mathematics and the Natural, Social, and Communication Sciences
in Real Problem Solving

This report was prepared with the support of the National Science Foundation, Grant No. SED-6901071. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Education Development Center
55 Chapel Street
Newton, Massachusetts 02160

DEVELOPMENT OF THE PENCIL PROBLEM: BACKGROUND

Since no commercially available instruments which would adequately measure real problem-solving skills had ever been located and since past efforts to develop such measures had met with considerable difficulties, a major goal of the 1976-1977 USMES research program was to try to develop such an instrument. An important requirement for the instrument was that it also be suitable for diagnostic use by teachers.

In 1971 Professor Bernard Shapiro from Boston University began developing instruments to measure student problem-solving abilities. The first such instrument, used in 1971-1972, was the Notebook Problem. In this test individual students are presented with three different types of notebooks. Each student is asked to select one notebook as best for his or her class and to give reasons for the selection. Cost information is given to the students; paper, pencils, and rulers are available if students wish to use them. No time limit is imposed, but the testing has always been completed in one class period. Two outcomes are measured:

- Does the student name any reasons for his or her choice that could be objectively measured, like size, cost, or weight of the notebooks? (This variable is called "Reasons for Selection.")

- Are any of the student's reasons based on actual tests? If not, are any based on suggested tests? (This variable is called "Highest Level of Warrant.")

These two variables focus on two specific problem-solving skills--identifying quantifiable differences among alternative solutions and designing tests to measure those differences.

In 1971-1972, samples of two to six students selected at random from a selection of thirty-one USMES and twenty-two non-USMES classes (matches for grade level and school) were tested with the Notebook Problem. Dr. Shapiro's report* on this testing concludes that "in

*Bernard J. Shapiro, *The Notebook Problem: Report on Observation of Problem-Solving Activity in USMES and Control Classrooms.* EDC, 1973.

terms of the two dependent variables studies, the USMES experience had, irrespective of units or teachers involved, a marked and positive effect on the students' problem-solving behavior." There were acknowledged shortcomings in the design, administration, and analysis of this study, but as noted in the report, the "redeeming aspect to these difficulties lies in the clarity and consistency of the actual results." (For a summary of the results, see Chapter V of *USMES Guide*, fourth edition, EDC, 1976.)

As part of a study of economic learning conducted in 1975-1976, Arthur Ellis and Allen Glenn of the University of Minnesota administered a modified Notebook Problem (using written responses) to seventy-eight fifth and sixth graders. The students were assigned at random among four classes receiving four different types of instruction (USMES Manufacturing unit; economic games and simulations; economic workbook; and geography--as a control). The USMES students scored highest on both Notebook Problem outcomes. Although the number of classes involved was limited, the results were again encouraging.*

For the school year 1972-1973, Dr. Shapiro and his assistant Susan Rogers developed a new instrument called the Playground Problem. Groups of five students are selected at random from sample classes and are asked to plan a new playground for their school. They are provided with a catalogue of playground equipment, cost data, measuring instruments, and a hypothetical budget of $2,000. The scoring protocol yields twelve group scores: four "behavioral" scores (motivation to accept problem, commitment to task, efficiency in allocating jobs, and nature of group leadership), four cognitive scores (variable identification, measurement, calculation, and recording), and four scores on the playground plan produced (on the scale, labels, landmarks, and area designations).

For the 1974-1975 evaluation, Professor Mary Shann and her staff developed a third instrument in this genre, the Picnic Problem. Groups of five students, selected at random from sample classes, develop plans for a class picnic. The students are supplied with a photograph showing various foods they may select and a map showing the location of a school and three parks from which they may choose a picnic site. Measuring instruments, cost data, and a hypothetical budget of twenty-five dollars are also provided. The scoring protocol includes the four behavioral scores and four cognitive scores derived for the Playground Problem.

The Playground Problem was administered in 1972-1973, 1973-1974, and 1974-1975. The Picnic Problem was administered in 1974-1975. Unfortunately, Dr. Shapiro has never reported on the 1972-1973 results and the information gained from the later testing is limited. Analysis carried out on the 1973-1974 Playground Problem data yielded no significant difference among the groups of classes for any of the outcome

*Arthur K. Ellis and Allen D. Glenn, *Effects of Real and Contrived Problem Solving on Economic Learning* (in preparation).

variables. In the view of the Boston University evaluators, the best explanations for these results probably "lie with the unsatisfactory testing procedures and with the instrument itself."*

The analysis of the 1974-1975 Playground Problem and the Picnic Problem data also yielded no significant differences between the USMES and non-USMES groups. The report on these analyses again questions the appropriateness of the instruments. The evaluators concluded "the more objective instruments to measure problem solving skills are still too unsophisticated to give an accurate reading."**

The Playground and Picnic Problems presented students with larger, more complex problems than the Notebook Problem and ones thought to be of more interest to students. In these ways, they resembled the real problems of USMES units more closely than the simple Notebook Problem. However, they were not administered as real problems while the Notebook Problem was; it could be more adequately solved in the small amount of testing time available and the problem and its variables were more tangible and less hypothetical. The scoring protocol for the later problems was also more ambitious and more removed from the data itself.

For these reasons, our best alternative for developing an appropriate instrument seemed to be continued development of an "Assessment of Problem-Solving Approaches" (or APSA), which had evolved from the Notebook Problem. The APSA, as originally created in the fall of 1975, presented students with a more carefully scripted notebook selection problem. When revised for field-testing in the fall of 1976, the APSA was still basically a consumer research problem, but students were asked this time to decide which of several pencils they thought their school should buy for students to use. Students, especially younger ones, were found to be more interested in pencils and pencils presented more options for choice. Students were presented with samples of the pencils and encouraged to "look over" the pencils and to test them in whatever ways they saw fit. They were told that the school would buy $10.00 worth of pencils, and they were shown the price of each pencil. (To maintain the realism of the problem situation, after the testing was completed we did supply each school with $10.00 worth of pencils selected on the basis of the students' choices.) Students were also asked a number of questions about how they decided on a pencil, how they would convince others that their choice was the best, and related issues. Their responses were tape-recorded, and the administrator (a member of the USMES research staff) took notes on their investigations. Typically, the testing sessions took about fifteen minutes.

---

*M.H. Shann, *An Evaluation of Unified Science and Mathematics for Elementary Schools (USMES) During the 1973-74 School Year*. Boston University, 1975.

**M.H. Shann, N.C. Reali, H. Bender, T. Aiello, L. Hench, *Student Effects of an Interdisciplinary Curriculum for Real Problem Solving: The 1974-75 USMES Evaluation*. Boston University, 1975.

The fall field-testing involved 125 students in the four schools of the 1976-1977 USMES Student Study. The students came from classes, grades 1-6, scheduled to do USMES this school year. They were selected by teachers as "representative" of students in their classes.

While there was a complete script, we deliberately altered a few parts of it during the field-testing to experiment with possible improvements. We recognized that these alterations would produce data which were not fully standardized, but the experiments made important contributions to the development of the instrument. They showed, for example, that students were less reticent when tested three at a time. This strategy thus produced more information than testing students individually, while it also allowed for testing a larger sample of students. Furthermore, we found that it was easy for the administrator to note differences in students' strategies, such as the number of pencils they wrote with or the reasons they gave for their choices, even when several students were working on the APSA at once. Student responses also suggested a number of refinements in the wording of questions and in the materials which the students saw and used during APSA sessions.

Students' responses to the APSA questions were transcribed and studied primarily to suggest further revisions in the instrument. Overall, however, we found that students were eager to work on the problem of choosing pencils. They also expressed confidence in their ability to make good recommendations about pencils, and they were able to make their choices within a reasonable amount of time. (Most students took 3-5 minutes.)

THE PENCIL PROBLEM: PRESENT VERSION

The spring version of the instrument, which has come to be known as the Pencil Problem, consisted of the attached script and a number of other materials. Six brands of pencils were used--two 15¢ pencils with #2 leads, two 15¢ pencils with #2½ leads, one 8¢ pencil with a #2 lead, and one 8¢ pencil with a #2½ lead. All were national brands, painted yellow with pink erasers. The pencils were arranged in a display box which showed the price of each pencil; six of each brand were included, three of which were sharpened and three unsharpened. Scrap paper, rulers, and hand-held sharpeners were available to students. In addition, since most students in the fall seemed to overlook the $10 limitation on the purchase of pencils, we included a small sign with the figure "$10" and a picture of a pencil among the materials in front of students during their investigations. (See attached illustration of Pencil Problem materials.)

As indicated in the script, the last part of the instrument included a question about erasers. For this question, the sample pencils were removed from the table and a box containing five types of erasers and a "$10" sign were presented to students. (An illustration is

attached.) The five types were a 35¢ soft pink eraser, a 35¢ hard red eraser, a 35¢ gum eraser, a 25¢ ink and pencil eraser, and a 25¢ white plastic eraser. The purpose of this part of the instrument was to provide a crude indicator of the reliability of the instrument across different consumer research problems.

Three students at a time worked on the problem. Each student was asked to respond verbally to the following questions, over the course of the Pencil Problem session:

1. "Do you think kids can give (principal's name) good advice about which pencil to buy?"

2. "How will you decide on a pencil?"

3. "Which pencil do you think (principal's name) should buy, and why?"

4. "How would you try to convince other students that the pencil you chose was the best one?"

5. "What would you do to decide which eraser (he/she) should buy?"

The students' responses were tape-recorded and they were also asked to indicate their choice of pencils and their reasons in writing. Their written responses were given on an "Advice to Principal" sheet (attached). Student behavior was observed as they worked with the pencils and the observations recorded on an observer's sheet (also attached).

The Pencil Problem was administered in April and May 1977 to 336 students in the four schools of the 1976-1977 USMES Student Study. Each student's verbal and written responses were coded on coding sheets (attached) designed to cover every common type of response. Observations of student investigations were also coded. For the verbal responses, two different coders listened to the tape recording for each student and then independently coded the student's responses. Coded responses were entered into computer files.

The ways in which the coded responses were scored and analyzed are described in the report on the Student Study. Briefly, three outcomes were examined:

- Total Number of Factors Considered
  (How many of nine critical factors do students mention, at any of several points in the Pencil Problem session, when discussing the quality and cost of pencils and erasers?)

- Number of Investigations Carried Out
  (How many pencils do students write with, sharpen, and erase with? How many other investigations do they carry out?)

- Response to "How Would You Convince Other Students That the Pencil You Chose was the Best One?" (How many factors would they tell or show other students? Would they have other students try them out? Would they ask other students' opinions?)

This is not the only possible method of scoring Pencil Problem responses, however. In fact, other methods would probably be more appropriate for diagnostic uses of the instrument. To date there has not been an adequate opportunity to explore other scoring methods, and users of the Pencil Problem will need to determine a method which is appropriate for their purposes.

PENCIL PROBLEM SCRIPT

Start tape.

HELLO. MY NAME IS (name). WHAT ARE YOUR NAMES?

Write full names and ages on observer's sheet.

AND HOW OLD ARE YOU?

I AM INTERESTED IN HOW STUDENTS SOLVE PROBLEMS. HERE IS THE PROBLEM WE WILL BE TALKING ABOUT.

Point to $10.00 sign.

YOUR PRINCIPAL, (name), IS GOING TO SPEND $10.00 TO BUY PENCILS FOR YOU AND OTHER STUDENTS TO USE. THE PROBLEM IS WHICH PENCIL TO BUY. I'M ASKING LOTS OF KIDS TO LOOK OVER SOME PENCILS AND THEN SAY WHICH ONE THEY THINK (principal's name) SHOULD BUY AND WHY. THE PENCIL HE/SHE CHOOSES WILL DEPEND ON THE ADVICE YOU GIVE HIM/HER.

BEFORE YOU TRY TO DECIDE ON A PENCIL, LET ME ASK YOU THIS. DO YOU THINK KIDS CAN GIVE (principal's name) GOOD ADVICE ABOUT WHICH PENCIL TO BUY? I'D LIKE TO HEAR WHAT EACH OF YOU THINK, SO I'LL ASK EACH OF YOU.

Ask each student.

- DO YOU THINK KIDS CAN GIVE (principal's name) GOOD ADVICE ABOUT WHICH PENCIL TO BUY, (name)?. . . WHY DO YOU THINK SO?

Ask each student.

- O.K., NOW, HOW WILL YOU DECIDE ON A PENCIL, (name)?

NOW TAKE SOME TIME TO LOOK OVER THE PENCILS. THINK ABOUT WHAT (principal's name) WILL WANT TO KNOW. THEN DECIDE WHICH PENCIL YOU THINK (principal's name) SHOULD SPEND THE $10.00 ON, AND WHY?

Point to pencils.

HERE ARE THE SIX KINDS HE/SHE IS THINKING ABOUT.

Point to prices.

HERE IS HOW MUCH EACH KIND COSTS.

YOU CAN USE ANY OF THE THINGS ON THE TABLE. YOU CAN WORK SEPARATELY OR TOGETHER--WHICHEVER YOU WANT--BUT YOU DON'T HAVE TO AGREE ON THE SAME PENCIL. IT'S O.K. TO TALK WHILE YOU'RE WORKING WITH THE PENCILS. SO GO AHEAD AND TAKE A FEW MINUTES TO DECIDE.

(Record students' actions on observer's form. It's best if you seem to be reading, writing, etc. Wait until the students seem to have finished examining the pencils before proceeding with the questions.)

● HAVE YOU FINISHED? O.K., NOW I WANT TO ASK EACH OF YOU WHICH PENCIL YOU THINK (principal's name) SHOULD BUY AND WHY.

Ask each student. WHICH PENCIL DO YOU THINK (principal's name) SHOULD BUY, (name)? . . . WHY DID YOU DECIDE ON THAT ONE?

Give "Advice to Principal" sheet to each student. ● NOW, WE HAVE TO LET (principal's name) KNOW WHICH PENCIL YOU THINK HE/SHE SHOULD BUY AND WHY, SO WHY DON'T YOU WRITE DOWN YOUR IDEAS ON THESE SHEETS.

Point to top part. CHECK OFF WHICH PENCIL YOU CHOSE UP HERE.

Point to bottom part. THEN WRITE WHY YOU CHOSE THAT ONE DOWN HERE. YOU DON'T HAVE TO WRITE SENTENCES: YOU CAN JUST LIST YOUR IDEAS, AND YOU DON'T HAVE TO FILL UP THE WHOLE PAGE. WRITE YOUR NAME AT THE BOTTOM.

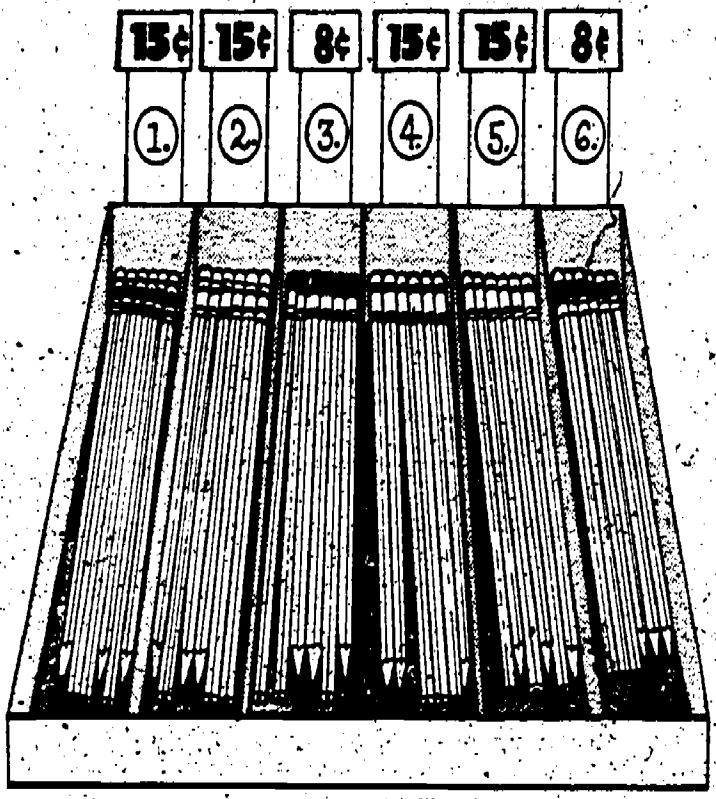
Collect "Advice to Principal" sheet.

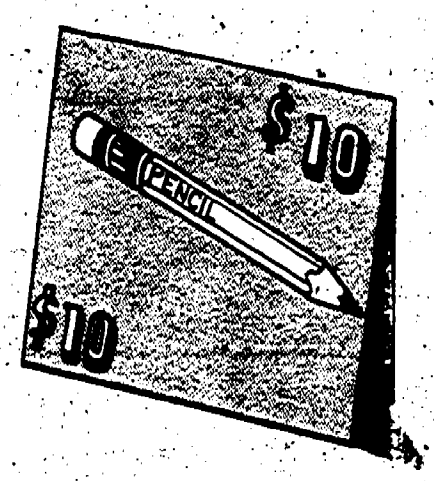
Ask each student. ● NOW, LET ME ASK YOU THIS. HOW WOULD YOU TRY TO CONVINCE OTHER STUDENTS THAT THE PENCIL YOU CHOSE WAS THE BEST ONE, (name)?

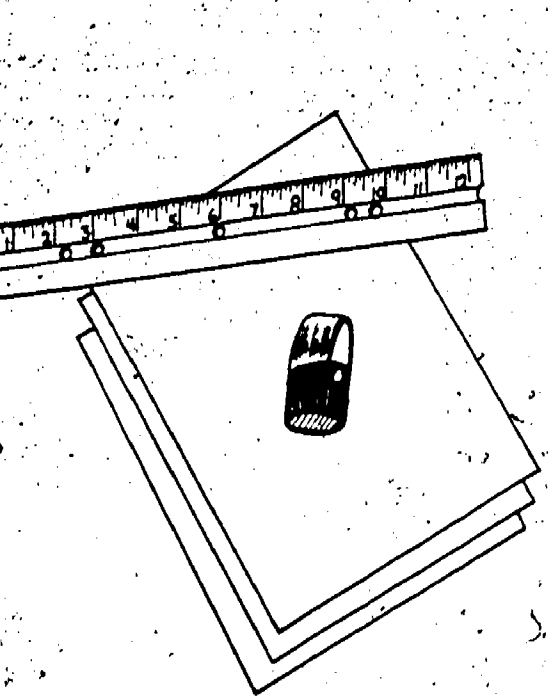
Display erasers and $10.00 sign. ● NOW, ONE LAST THING. HERE'S A DIFFERENT PROBLEM. (Principal's name) IS ALSO THINKING ABOUT BUYING $10.00 WORTH OF ERASERS FOR STUDENTS TO USE. THESE ARE THE KINDS THAT HE/SHE IS THINKING ABOUT. WE DON'T HAVE TIME NOW FOR YOU TO TRY TO CHOSE ONE, BUT IF YOU HAD AS MUCH TIME AS YOU WANTED,

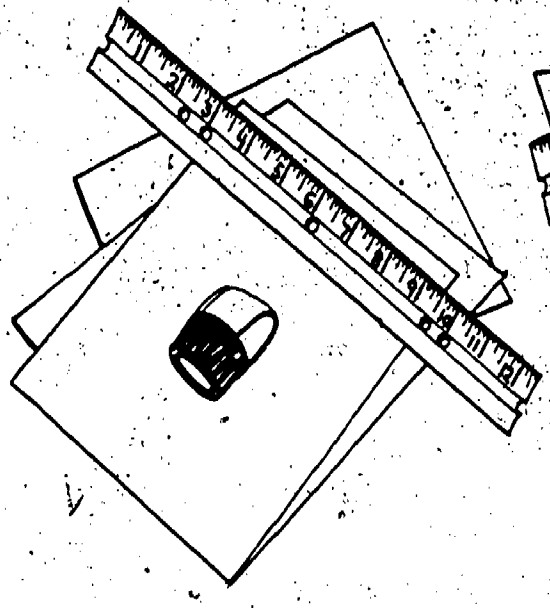
Ask each student. WHAT WOULD YOU DO TO DECIDE WHICH ERASER HE/SHE SHOULD BUY, (name)?

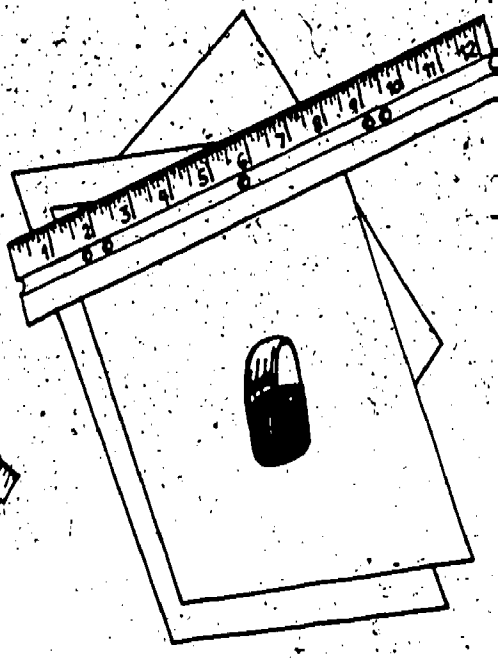
● O.K., YOUR IDEAS HAVE BEEN A BIG HELP. YOU SHOULD BE HEARING ABOUT THE PENCILS IN A COUPLE OF DAYS. THANK YOU VERY MUCH.

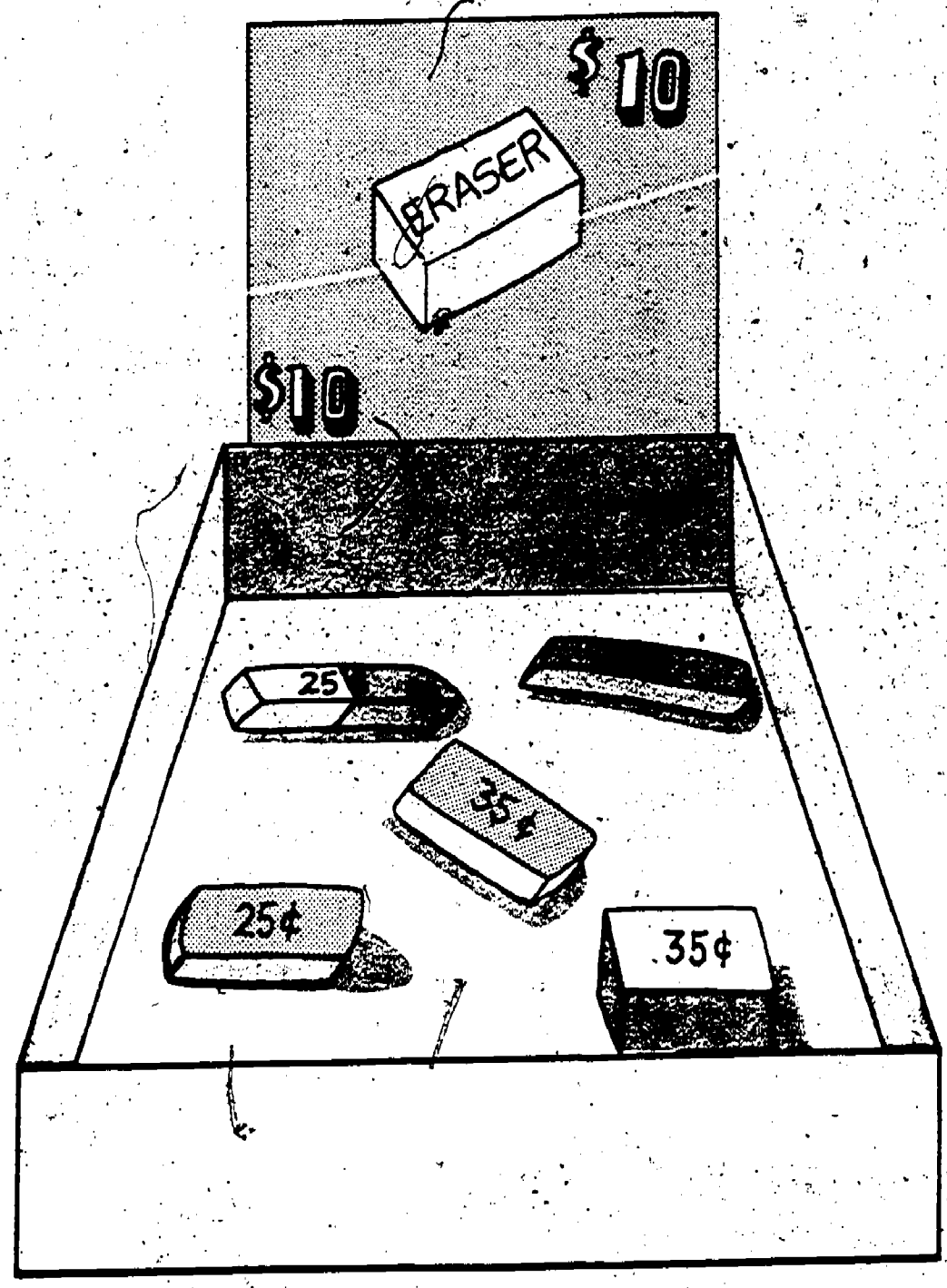

15¢
15¢
8¢
15¢
15¢
8¢
1.
2.
3.
4.
5.
6.


$10
PENCIL
$10






PENCIL


$10
ERASER
$10
25
35¢
25¢
35¢

ADVICE TO PRINCIPAL

Which Pencil To Buy:

- ☐ 1. Dixon Ticonderoga #2 5/10
- ☐ 2. Dixon Ticonderoga #2
- ☐ 3. Green-Rite #F
- ☐ 4. Eberhard Faber MONGOL #F
- ☐ 5. Eberhard Faber MONGOL #2
- ☐ 6. Eberhard Faber #2

Why Buy That Pencil:

Your Name

PENCIL PROBLEM OBSERVER'S SHEET SCHOOL ______________ DATE __________ TIME ________

TAPE # ______

NAME: ____________________

AGE: ____________________

Apprx. No. written with ______

Apprx. No. erased with ______

Apprx. No. sharpened ______

What he/she wrote:

___name(s)
___lines
___pictures
___numbers
___other writing (specify)

____

____

Other "investigations"

___calculating
___dropping
___flexing pencils
___measuring (what & how)

____

____

___other (specify)

____

____

NAME: ____________________

AGE: ____________________

Apprx. No. written with ______

Apprx. No. erased with ______

Apprx. No. sharpened ______

What he/she wrote:

___name(s)
___lines
___pictures
___numbers
___other writing (specify)

____

____

Other "investigations"

___calculating
___dropping
___flexing pencils
___measuring (what & how)

____

____

___other (specify)

____

____

NAME: ____________________

AGE: ____________________

Apprx. No. written with ______

Apprx. No. erased with ______

Apprx. No. sharpened ______

What he/she wrote:

___name(s)
___lines
___pictures
___numbers
___other writing (specify)

____

____

Other "investigations"

___calculating
___dropping
___flexing pencils
___measuring (what & how)

____

____

___other (specify)

____

____

CODER _____ TAPE _____ STUDENT'S NAME _____

RESPONSES TO: DO YOU THINK KIDS CAN GIVE (principal) GOOD ADVICE ABOUT WHICH PENCIL TO BUY?

1

1. Yes
2. No
3. It depends/maybe

"YES" REASONS

4. Kids use/write/work with pencils
5. Kids need pencils
6. Kids want/like pencils

7. Kids know about pencils
8. Kids have more/better ideas/are smart
9. Kids can find out/think of...

10. Kids can tell/show principal
11. Principal needs/would like advice
12. Principal might get wrong pencil

13. Kids want to give advice

14. Some pencils break/feel bad/are too dark/light/thick/thin/etc.

2

"NO" REASONS

1. Principal is buying so he/she should decide
2. Principal would know about price
3. Kids would choose different pencil than principal
4. Kids would choose wrong pencil
5. Kids would spend too much
6. Kids argue too much

"IT DEPENDS/MAYBE" REASONS

7. Depends on kids' age
8. Depends on kids' experience
9. Depends on how smart they are
10. Depends on the pencil

OTHER REASONS

1. Different kids like different pencils
2. If principal gives them to us, we won't have to buy them
3. I have some of the pencils

4. I don't know/not sure, etc.
5. (Unintelligible)

STUDENT'S NAME _____

RESPONSES TO: HOW WILL YOU DECIDE ON A PENCIL?

4

FIRST PHRASE

1. I'd choose/pick/take/decide on the one that...
2. (No verb) The one that/If it/How it...
3. It depends on...
4. I'd see/find/look for one that/which...
5. I'd see/find out if it...
6. I'd see/find out how it...
7. I'd see/find out/look at what it...
8. I like...

9. I'd write with them/use them
10. I'd sharpen them
11. I'd erase with them
12. I'd measure them
13. I'd bend them
14. I'd hold them/feel them
15. I'd ask other kids...
16. I'd try them out/test them
17. I'd look them over

5

SECOND PHRASE

1. Writes/works
2. Writes/works well/good/best/nice
3. Writes dark
4. Writes light
5. Writes medium/not too light or dark
6. Writes smooth/doesn't scratch
7. Writes easy
8. Writes neat
9. Writes small/thin

10. Sharpens
11. Sharpens well/good/best
12. Sharpens without breaking lead
13. Sharpens without chewing up wood

14. Erases
15. Erases well/good/best/has a good eraser
16. Has a big eraser
17. Erases clean
18. Erases without crumbs
19. Erases without tearing paper

20. Costs
21. Costs less/cheapest/lowest
22. Costs more/expensive
23. Costs 8¢
24. Costs 15¢

MORE SECOND PHRASES

1. Has a good lead
2. Has a #2 lead
3. Has a strong lead/won't break

4. Is good/right size
5. Is long/big
6. Is thick/fat
7. Is thin/skinny

8. Pencil is strong/won't break/won't bend
9. Has good wood
10. Pencil lasts long

11. Feels good/is comfortable
12. Is pretty/looks good/has good color
13. Is sharp
14. Has good name
15. Is good quality
16. Is best

OTHER RESPONSES

17. (Names a specific pencil)
18. I don't know/not sure, etc.

19. (Unintelligible)

Counter #'s: From_____to_____ STUDENT'S NAME _____

Machine #:_____ . RESPONSES TO: WHICH PENCIL DO YOU THINK (principal) SHOULD BUY, AND WHY?

7

PENCIL CHOSEN

1. Dixon Ticonderoga #2 5/10 (medium)
2. Dixon Ticonderoga #2 (soft)
3. Green-Rite #F
4. Eberhard Faber MONGOL #F
5. Eberhard Faber MONGOL #2
6. Eberhard Faber #2

WRITING

7. Writes/works well/good/better/best/nice
8. Draws well/good/better/best/nice
9. Writes neat
10. Writes easy
11. Writes smooth
12. Doesn't scratch/dig into paper
13. Writes small/thin/skinny
14. Writes thick/fat
15. Writes light/not dark
16. Writes dark/not light
17. Writes heavy
18. Writes medium/not too light or dark
19. Doesn't smear

SHARPENING

20. Sharpens well/good/better
21. Gets a sharp point
22. Lead doesn't break
23. Lead doesn't fall out
24. Wood doesn't crumble/get chewed up

8

ERASING

1. Erases well/good/clean/good eraser
2. Eraser is strong
3. Eraser is big
4. Eraser is put in well
5. Eraser doesn't come out
6. Eraser doesn't wear down
7. Eraser doesn't leave crumbs
8. Eraser doesn't rip paper/make holes

COST

9. Is low/cheap/inexpensive/less
10. Is high/expensive
11. Is 8¢
12. Is 15¢
13. Is a good deal

LEAD

14. Has good/best lead
15. Has a #2 lead
16. Has strong lead/lead doesn't break
17. Has hard lead/not too soft
18. Has soft lead/not too hard
19. Lead lasts long/won't wear down-fast

9

SIZE

1. Is good/right size
2. Is long/big
3. Is short/small
4. Is thick/fat
5. Is thin/skinny

OTHER DESCRIPTIONS ABOUT PENCIL

6. Pencil is strong/won't break/won't bend
7. Pencil lasts long/won't be used up quickly
8. Pencil has good wood
9. Feels good/is comfortable
10. Is pretty/looks good/has good color
11. Is sharp
12. Has good name
13. Is good quality/good/great
14. (Names bad things about other pencils)

OTHER RESPONSES

15. I'm used to it/have used it a lot
16. I like it
17. I don't know/not sure, etc.
18. (Unintelligible)

STUDENT'S NAME _____

RESPONSES TO: HOW WOULD YOU TRY TO CONVINCE OTHER STUDENTS THAT THE PENCIL YOU CHOSE WAS THE BEST ONE?

10

FIRST PHRASE

1. I'd tell them/I'd say...
2. I'd show them...
3. I'd compare...
4. No first phrase

5. I'd let them...
6. I'd have them/ask them to/tell them to...

7. I'd ask them...
8. I'd see...

*9. If they didn't like it...

11

SECOND PHRASE

1. (about how it writes)
2. (about how it sharpens)
3. (about how it erases/about eraser)
4. (about its cost)
5. (about its lead)
6. (about its size)
7. (about its feel)
8. (about pencil breaking)
9. (about how long it lasts)
10. (about how good it is)
11. (about different pencils--nothing specific
12. (about _____)

13. Write/use/draw with it
14. Erase with it
15. Try it/test it
16. Look at it
17. Vote
18. Compare
19. Decide

20. If/how they like it/what they think of it
21. What they like/think

12

*1. I'd try to convince them
*2. I'd ask which they liked
*3. I'd compare
*4. I wouldn't mind

OTHER RESPONSES

5. I don't know/not sure, etc.
6. (Unintelligible)

STUDENT'S NAME _____

RESPONSES TO: WHAT WOULD YOU DO TO DECIDE WHICH ERASER HE/SHE SHOULD BUY?

13

FIRST PHRASE

1. I'd choose/pick/take/decide on the one that...
2. (No verb) The one that/If it/How it...
3. It depends on...
4. I'd see/find/look for one that/which...
5. I'd see/find out if it
6. I'd see/find out how
7. I'd see/find out/look at what it...
8. I like...

9. I'd erase with them/use them
10. I'd write and erase with them

11. I'd measure them
12. I'd bend them
13. I'd hold them/feel them
14. I'd ask other kids...
15. I'd try them out/test them
16. I'd look them over

14

SECOND PHRASE

1. Erases/works
2. Erases/works well/better/best
3. Doesn't smear/erases clean
4. Doesn't break/crumble/fall apart
5. Doesn't tear paper/make holes
6. Lasts long

7. Is big/long
8. Is strong
9. Is hard
10. Doesn't bend too much

11. Costs
12. Costs less/cheapest/lowest
13. Costs more/expensive
14. Costs 25¢
15. Costs 35¢

15

MORE SECOND PHRASES

1. Feels good/is comfortable
2. Is pretty/looks good/has good color
3. Has good name
4. Is good quality
5. Is best

OTHER RESPONSES

6. (Names a specific eraser)
7. (Names a specific eraser and gives reasons)
8. (Names bad things about other erasers)
9. I don't know/not sure, etc.
10. (Unintelligible)

STUDENT'S NAME _____

-RESPONSES TO: ADVICE TO PRINCIPAL SHEET

PENCIL CHOSEN 16

1. Dixon Ticonderoga #2 5/10 (medium)
2. Dixon Ticonderoga #2 (soft)
3. Green-Rite #F
4. Eberhard Faber MONGOL #F
5. Eberhard Faber MONGOL #2
6. Eberhard Faber #2

WRITING

7. Writes/works well/good/better/best/nice
8. Draws well/good/better/best/nice
9. Writes neat
10. Writes easy
11. Writes smooth
12. Doesn't scratch/dig into paper
13. Writes small/thin/skinny
14. Writes thick/fat
15. Writes light/not dark
16. Writes dark/not light
17. Writes heavy
18. Writes medium/not too light or dark
19. Doesn't smear

SHARPENING

20. Sharpens well/good/better
21. Gets a sharp point
22. Lead doesn't break
23. Lead doesn't fall out
24. Wood doesn't crumble/get chewed up

ERASING 17

1. Erases well/good/clean/good eraser
2. Eraser is strong
3. Eraser is big
4. Eraser is put in well
5. Eraser doesn't come out
6. Eraser doesn't wear down
7. Eraser doesn't leave crumbs
8. Eraser doesn't rip paper/make holes

COST

9. Is low/cheap/inexpensive/less
10. Is high/expensive
11. Is 8¢
12. Is 15¢
13. Is a good deal

LEAD

14. Has good/best lead
15. Has a #2 lead
16. Has a strong lead/lead doesn't break
17. Has hard lead/not too soft
18. Has soft lead/not too hard
19. Lead lasts long/won't wear down fast

SIZE 18

1. Is good/right size
2. Is long/big
3. Is short/small
4. Is thick/fat
5. Is thin/skinny

OTHER DESCRIPTIONS ABOUT PENCILS

6. Pencil is strong/won't break/won't bend
7. Pencil lasts long/won't be used up quickly
8. Pencil has good wood
9. Feels good/is comfortable
10. Is pretty/looks good/has good color
11. Is sharp
12. Has good name
13. Is good quality/good/great
14. (Names bad things about other pencils)

OTHER RESPONSES

15. I'm used to it/have used it a lot
16. I like it
17. I don't know/not sure, etc.
18. (Unintelligible)

ADMINISTRATOR'S OBSERVATIONS

Student's name: _____

WHAT STUDENT WROTE 19

1. Names
2. Lines
3. Pictures
4. Numbers
5. Scribbles, squiggles
6. Words
7. Letters
8. Shadings
9. Other writing

INVESTIGATIONS 20

1. Measured lengths of pencils with ruler
2. Compared lengths of pencils by sight
3. Compared sizes of erasers
4. "Weighed" pencils by feel
5. Observed time to sharpen
6. Picked up and looked at/ examined details of construction
7. Pressed point to test lead
8. Breaks lead
9. Smeared writing
10. Handled erasers
11. Calculating
12. Dropping
13. Flexing pencils

21. Age _____

22. Written w/ _____

23. Erased w/ _____

24. Sharpened _____

USMES Technical Report

The Questionnaire on Attitudes Toward Real Problem Solving

UNIFIED SCIENCES AND MATHEMATICS FOR ELEMENTARY SCHOOLS:
Mathematics and the Natural, Social, and Communication Sciences in Real Problem Solving

This report was prepared with the support of the National Science Foundation, Grant No. SED-6901071. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Education Development Center
55 Chapel Street
Newton, Massachusetts 02160

BACKGROUND

In deciding which student attitudes were most important for the 1976-1977 research, the USMES staff reviewed a number of instruments used in previous research. Existing published instruments had been reviewed earlier, both by the USMES staff and by the Boston University USMES evaluation staff, but no published instrument had been found which covered the particular attitudes most relevant to the USMES type of real problem solving. However, some of the items used by the USMES staff in the 1972-1973 district implementation experiment in Lansing, Michigan and by the Boston University group in their 1974-1975 evaluation program suggested topics which might be of interest.

Eventually, the USMES staff agreed that four types of attitudes should be studied in the 1976-1977 research: (1) attitudes towards working on real problems and producing effective solutions; (2) attitudes toward group interaction; (3) attitudes towards specific problem-solving activities (like working with numbers, surveying, and writing); and (4) attitudes toward USMES (for students with USMES experience). These categories of questions were thought to cover the main attitudinal goals of the USMES program and also the key attitudes which might affect the students' motivation to get fully involved in USMES activities.

The USMES research staff, in consultation with others in the USMES staff and with other evaluation specialists at EDC, developed a three-page questionnaire covering the above topics, during the spring and summer of 1976. A draft version of the questionnaire which was called the Questionnaire on Attitudes towards Real Problem Solving (QARPS) was sent for review to experienced USMES teachers and administrators in Iowa, Texas, and Minnesota. Their suggestions were incorporated in the fall version of the QARPS.

THE FALL QARPS

The fall version of the QARPS consisted of two parts: (1) twenty questions for both USMES and non-USMES students on attitudes toward working on real problems, group work, and specific real problem solving activities; and (2) eleven questions for experienced USMES students

on attitudes toward USMES. The front page of the questionnaire contained space for identification information and directions for students. A sample of the cover letter which accompanied the fall QARPS is attached.

In the fall, thirty-three classes (849 students) in grades 1 through 6 of the four schools in the Student Study received the QARPS. These classes included most of the classes doing USMES this year in these schools, although very few classes had begun USMES work at the time of the fall QARPS administration, in mid-October, and several other classes at the same grade levels in the schools. Virtually all students in each class filled out the QARPS.

Inspection of the fall results showed that some questions were unclear, especially to younger students. These were reworded in the spring version. Also there appeared to be some unnecessary redundancy and some omitted topics. To make up for this some of the fall questions were replaced with new ones in the spring QARPS.

THE SPRING QARPS

The spring version of the QARPS (attached) also contained twenty questions on attitudes toward real problem solving. Eleven of these questions contained minor changes of wording, and two were substitutions for questions used in the fall. The spring version also contained nine questions on attitudes toward USMES, and a checklist on USMES activities; this section was intended for USMES students only. One of these questions was a minor rewording of a fall question, and three were new to the spring version. No cover letter was prepared for the spring version, since it was administered entirely by the USMES research staff.

In the spring, thirty-seven classes (1,011 students) in grades 2 through 6 filled out the QARPS. The sample of classes was chosen to assure coverage of key variables in the Student Study.

RELIABILITY AND VALIDITY

We have not been able to conduct technical studies of the instrument, such as formal reliability studies or systematic development of norms. However, we have examined several kinds of evidence about the qualities of the instrument. None of these types of evidence are as strong as we wish by themselves, but they are reasonably persuasive when considered together. They include the observed consistency of responses on many items from fall to spring and the fairly satisfying inter-correlations of responses to different questions, roughly as we had predicted.

The positive reactions received from teachers and others about the appropriateness of QARPS questions provides some reassurance about the content validity of the instrument. Responses to some of the QARPS questions also show expected trends by age and school.

SCORING AND ANALYSIS

For the Student Study analysis, the twenty questions on real problem solving were combined into three scales corresponding to the three types of attitudes the instrument was designed to measure:

Attitudes Toward Work on Real Problems and Producing Effective Solutions

(8 items)

- "If kids show that a change is needed, then grown-ups will go along."
- "I figure out what I'm going to do before I start something."
- "Kids usually can't come up with good ways to solve problems around school." (Coded in reverse, e.g., "Disagree a lot" coded highest.)
- "Usually there is just one way to solve a real life problem." (Coded in reverse.)
- "I like to make things better around school."
- "I think I am good at solving real life problems."
- "As soon as I think of one way to solve a problem, I don't think about other ways." (Coded in reverse.)
- "I like working on problems that have more than one answer."

Attitudes Toward Group Interaction

(7 items)

- "I like to work in small groups with other students."
- "I like to help decide what we do in class."
- "I like to work by myself." (Coded in reverse.)
- "If I think other kids won't like my idea, then I keep quiet about it." (Coded in reverse.)
- "I like to ask questions."
- "I like to talk to other students about my ideas."
- "I like to listen to other students talk about their ideas."

Attitudes Toward Specific Problem-Solving Activities
(5 items)

- "I like to write about my ideas."
- "I like to measure things when it helps solve real life problems."
- "I like to make charts and graphs that show things I found out."
- "I like to use numbers to solve real life problems."
- "I like to do surveys to find out what other people think."

These three scales do not represent the only way to analyze QARPS results. In fact, for informal, diagnostic purposes it may be more appropriate to look at responses item by item for individual students and, using means, for the class as a whole. To date there has not been time to explore other scoring methods and QARPS users will want to investigate other procedures for themselves.

Responses to the ten questions about the USMES program (on the back page of the QARPS) may be examined item by item to reveal how individual students or the class as a whole is experiencing USMES. Responses to past administrations of similar items are reported on in Chapter V of the fourth edition of the *USMES Guide* and may be used for comparison purposes.

COVER LETTER FOR FALL QARPS

TO: Teachers and others who plan to use the Questionnaire on Attitudes toward Real Problem Solving (QARPS)

FROM: The USMES research staff

The purpose of the QARPS is to assess some attitudes that may be related to students' performance in solving "real problems". It is intended for use both by teachers who want to examine the progress of their own classes and by researchers or curriculum supervisors who are studying the impact of programs to teach real problem solving. We expect that students in grades 4-8 will be able to read and fill out the QARPS form on their own; students in grade 3 or below may be able to fill out their own responses if their teacher reads each item out loud.

The first twenty items (pages 1 and 2) should be appropriate for both USMES students and students who have done no USMES work. The last section of the QARPS (page 3) applies to USMES students only and should be ignored when other students are using the questionnaire.

Students should understand that the word "problem" as used in this questionnaire refers to practical, open-ended problems--like trying to improve some situation in the school or community--not word problems or workbook exercises. For this reason, we suggest that you read the "Directions for students" on the cover of the QARPS out loud before the students begin to fill out the questionnaire.

Depending on your purposes in using the QARPS, you may want to have the students provide more or less background information than is indicated on the cover sheet. The information that is requested will be most useful for USMES's research purposes, and we hope you will include it in any data that you can send to us (see below). We have not asked for students' names, since it often seems best on this kind of questionnaire to let students remain anonymous. However, names or identification numbers can be added in a blank part of the cover sheet, if you prefer.

Comments and suggestions about this "final draft" are welcome. In addition, we would like to see what results you obtain so we can assess the usefulness of the questionnaire. We will be happy to compute tallies and percentages for you (as long as our funding permits). Please send comments, completed questionnaires, or results you tally yourself, to Dan Cooper or Carolyn Arbetter at the USMES office at EDC.

QARPS

SPRING VERSION

Questionnaire on Attitudes toward Real Problem Solving

Your name: _____

Your teacher's name: _____

Did you do USMES last year? _____

DIRECTIONS FOR STUDENTS:

We would like to know what you think about some things around school. Your answers will be used to try to make some school work more interesting.

Inside there are some questions that we would like you to answer. This is not a test, and there are no "right" answers. We just want to know how you feel.

Each question asks if you agree or disagree with something other students have said. Here is an example for you to try. Please make a mark through the circle that goes with your choice:

I like to write stories.

- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

It may be hard to make up your mind about some questions. Just pick the answer which is closest to how you feel. Don't spend too long on any one question.

Several questions are about "solving problems." They don't mean math problems or problems in work books. They mean real life problems, like trying to make something you do in school safer or more fun.

If you have any questions, you should ask them now.

477

unified sciences and mathematics for elementary schools/edc/55 chapel st./newton/mass./02160

1. If kids show that a change is needed, then grownups will go along.
   - ○ Agree a lot
   - ○ Agree a little
   - ○ Disagree a little
   - ○ Disagree a lot

2. I like to work in small groups with other students.
   - ⊙ Agree a lot
   - ○ Agree a little
   - ○ Disagree a little
   - ○ Disagree a lot

3. I like to write about my ideas.
   - ○ Agree a lot
   - ○ Agree a little
   - ○ Disagree a little
   - ○ Disagree a lot

4. I figure out what I'm going to do before I start something.
   - ○ Agree a lot
   - ○ Agree a little
   - ○ Disagree a little
   - ○ Disagree a lot

5. Kids usually can't come up with good ways to solve problems around school.
   - ○ Agree a lot
   - ○ Agree a little
   - ○ Disagree a little
   - ○ Disagree a lot

6. I like to help decide what we do in class.
   - ○ Agree a lot
   - ○ Agree a little
   - ○ Disagree a little
   - ○ Disagree a lot

7. Usually there is just one way to solve a real life problem.
   - ⊙ Agree a lot
   - ○ Agree a little
   - ○ Disagree a little
   - ○ Disagree a lot

8. I like to measure things when it helps solve real life problems.
   - ○ Agree a lot
   - ○ Agree a little
   - ○ Disagree a little
   - ○ Disagree a lot

9. I like to work by myself.
   - ○ Agree a lot
   - ○ Agree a little
   - ○ Disagree a little
   - ○ Disagree a lot

10. I like to make things better around school.
    - ○ Agree a lot
    - ⊙ Agree a little
    - ○ Disagree a little
    - ○ Disagree a lot

11. If I think other kids won't like my idea, then I keep quiet about it.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

12. I like to ask questions.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

13. I like to make charts and graphs that show things I found out.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

14. I think I am good at solving real life problems.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

15. I like to talk to other students about my ideas.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

16. I like to use numbers to solve real life problems.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

17. As soon as I think of one way to solve a problem, I don't think about other ways.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

18. I like to listen to other students talk about their ideas.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

19. I like to do surveys to find out what other people think.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

20. I like working on problems that have more than one answer.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

21. I think USMES work is fun.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

22. I think USMES work is boring.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

23. Doing USMES is hard work.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

24. I don't know why we do some things in USMES.
- ○ Agree a lot
- ⊙ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

25. Doing USMES makes me think.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

26. I think USMES work is confusing.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

27. I think USMES work is important.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

28. In USMES it's hard to decide what to do next.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

29. I would like to do more USMES.
- ○ Agree a lot
- ○ Agree a little
- ○ Disagree a little
- ○ Disagree a lot

30. When we do USMES, I usually do these things: (Put a check next to the ones you usually do.)

____ measure things
____ do surveys of people
____ build things
____ ask questions
____ write about things we find out
____ work in small groups
____ make graphs or charts
____ do math
____ help decide what we do
____ talk about my ideas
____ work by myself

edc

Thank you.