DOCUMENT RESUME

ED 382 685 TM 023 226

AUTHOR	Telese, James A.; Kulm, Gerald
TITLE	Performance-based Assessment of At-risk Students in Mathematics: The Effects of Context and Setting.
SPONS AGENCY	Department of Education, Washington, DC.
PUB DATE	Apr 95
CONTRACT	RE117E20049
NOTE	71p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 18-22, 1995).
PUB TYPE	Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Analysis of Variance; *Context Effect; *Educational Assessment; Educational Environment; Generalizability Theory; Grade 8; *High Risk Students; Interrater Reliability; Junior High Schools; *Junior High School Students; *Mathematics Achievement; Problem Solving; Scoring; *Test Construction
IDENTIFIERS	*Performance Based Evaluation; Texas Assessment of Academic Skills

ABSTRACT

A team of university and public school mathematics educators designed performance-based mathematics assessment tasks designed to align with the Texas Assessment of Academic Skills for 93 students who had been identified as at-risk in mathematics. Scenarios were developed based on four contexts: (1) familiar activity; (2) social issue; (3) hands-on; and (4) technology. Each context was administered in three settings: individually, aided by a proctor, and small group. The data analysis consisted of two repeated measures analyses of variance with context and setting, and content and setting as the main factors. The repeated measures were operations, concepts, or problem-solving scores. The results indicated that context was not significant, but content was significant. Setting was significant in both analyses. Generalizability studies (G-studies) were conducted to measure dependability of raters and students. The G-studies indicated that the six raters were dependable when assigning scores. The problem-solving domain was the most dependable knowledge domain rated and the concept domain was least dependable. An appendix provides scoring rubrics and sample questions. Thirteen tables and four figures. (Contains 48 references.) (SLD)

Performance-based Assessment of At-risk Students in Mathematics: The Effects of Context and Setting

James A. Telese
Penn State University at Harrisburg
777 West Harrisburg Pike
Middletown, PA 17057

717-948-6725
Jat6@psu.edu

Gerald Kulm
Montana State University
STEP Project
Linfield Hall
Bozeman, MT 59717

406-994-5655

Paper presented at the annual meeting of the American Education Research Association, San Francisco, CA, April, 1995.

This work was funded by the U. S. Department of Education under grant number RE117E20049. Any opinions and findings, or conclusions and recommendations are those of the authors and do not necessarily rflect the views of the U. S. Department of Education.

Abstract

A team of university and public school mathematics educators designed performance-based mathematics assessment tasks for students who have been identified as at-risk in mathematics. Scenarios were developed based on four contexts: familiar activity, social issue, hands-on, and technology. Each context was administered in three settings: individually, aided by a proctor, and small group. The data analysis consisted of two repeated measures ANOVAs with context and setting, and content and setting as the main factors. The repeated measures were operations, concepts, or problem-solving scores. The results indicated that context was not significant, but content was significant. Setting was significant in both analyses. Generaliziability studies were conducted to measure dependability of raters and students. The G-studies indicated that the raters were dependable when assigning scores. The problem-solving domain was the most dependable knowledge domain rated and the concept domain was least dependable.

Traditional standardized tests of mathematics achievement have been linked to the Industrial Age's testing paradigm reflecting an assembly line notion of education where students were the raw materials for processing and teachers transmitted mathematical information as bits to be memorized through drill and practice (Romberg, 1992). In turn, mathematics education of the Industrial Age reflected the minimalist notion that the masses should have the ability to follow clear and simple directions, to read newspapers, and perform basic mathematical calculations (Romberg, 1992). A broader standardization of curricula, instruction and testing evolved assimilating immigrants into a common culture, but still focusing on a Eurocentric, middle class, predominantly male view of mathematics and its uses (NCTM, 1993). Consequently, a stereotypical mathematics classroom has come to be characterized by teacher-centered lectures, robbing the students of creativity, exploration, and curiosity (Schmittau, 1991; Romberg, 1992).

The stereotypical mathematics classroom fails miserably for most students today (Romberg, 1992), and especially for the at-risk student population (Hilliard, 1990; Secada, 1992). Historically, stake-holders have used the results from traditional standardized tests for student classification, which has led to the alienation of some students from a challenging mathematics curriculum. As a result of this practice students are sorted by "social class, race, ethnicity, language background, gender, and other demographic characteristics" (Secada, 1992, p. 26) and face a curriculum that presents mathematics through low-level rote drill and practice exercises rather than foster higher-order thinking and problem-solving skills

(Hilliard, 1990; Oakes, Ormseth, Bell, & Camp, 1990; Secada, 1992). The stereotypical mathematics classroom's curriculum and associated testing practices have underestimated what disadvantaged students are capable of doing, depriving some students of meaningful or motivating contexts for learning (Means & Knapp, 1991).

The National Council of Teachers of Mathematics (1993) contended that assessments tasks need to engage students' interest and involvement in order for the assessment to reveal mathematical competencies. The purpose of this exploratory study is to examine at-risk students responses to performance-based assessment tasks that vary in content, context, and setting, thus adding to and expanding the knowledge base of performance assessments. A collaboration was established between public school teachers and university researchers in an attempt to answer the following questions:

1. What is the degree of influence that context (i.e. a familiar activity, a social issue, hands-on construction, and video presentation) exerts on eliciting the Texas Assessment of Academic Skills (TAAS) test's mathematical knowledge domains of problem solving, operations, and concepts from at-risk students?

2. How does the administration setting of performance assessment (i.e., aided, small group, or individually in a large group) impact at-risk students' mathematical performance?

3. Which setting is most effective in promoting at-risk students' mathematical performance in relation to the TAAS mathematical content categories?

4. How dependable are raters of a situated performance-based assessment?

5. Which TAAS mathematical knowledge domain, operations, concepts, or problem-solving can be dependably rated?

Theoretical Framework

The Importance of Culturally Relevant Content for Learning and Assessment

Researchers (e.g. Aronowitz & Giroux, 1993; Cobb, 1994; Jacob & Jordan, 1993a) have suggested that contexts arise from the students' past experiences, are collected within their cultural frameworks, and contribute to shape educational outcomes. A cultural framework represents a social world that constitutes a relationship between persons acting on their ever changing world (Lave, 1993). In the present study, contexts are used to attach meaning and relevancy to authentic classroom activity systems that integrate the subject, the object, and the materials into a unified whole (Engeström, 1993). Many cognitive researchers and theorists (Bransford & Vye, 1989; Cobb, 1986; Davis & Maher, 1990; Greeno, 1991; Kulm, 1990; Wittrock, 1991) have characterized meaningful learning as reflective, constructive, and self-regulated. Activity systems aide in establishing a classroom culture where students are engaged in meaningful learning activities. Through the use of contexts during instruction or assessment, students have the opportunity to make connections between their experiences and their mathematical capabilties (Saljo & Wyndhamn, 1993). For example, Shavelson, Webb, & Lehman (1986) concluded that as a result of experiencing content material in relevant contexts, problem-solving abilities and deeper understanding of concepts were developed. The stereotypical mathematics classroom fails to present meaningful learning activities. The "students often acquire algorithms and decontextualized definitions that lie inert. . ." (Lave, 1993, pg. 8). Consequently, mathematical knowledge remains dormant when classroom activities do not capitalize on students' experiences through meaningful and relevant contexts.

Mathematical learning activities are often presented in either an imposed or authentic problem format. Although completing several imposed problems could be considered an activity system, an imposed problem is a pseudo-problem such as the dull exercises found in a textbook (Borba, 1990). On the one hand, students are asked to solve pseudo-problems, which are not problems for them personally. But some students solve these just to get a good grade while others remain unmotivated to do the exercises regardless of the resulting grade. On the other hand, as an activity system, solving authentic problems involve using "real mathematics, realistic situations, questions or issues that might actually occur in a real-life situation, and realistic tools and resources" (Lesh & Lamon, 1992, p. 18). For example, in a study of education majors and nursing students Ross, McCormick, and Krisak (1986) determined that problems situated in contexts that were congruent to the students' majors increased their achievement levels in mathematics. The tasks had an authentic nature to them, representing actions that the students would do in the professions of education or nursing. Students who encounter an authentic problem are more likely to make attempts at solving the problem. However, consideration should be given to the notion that not all problems will be interesting to all students since problems are culturally bounded, what is interesting for someone depends, in part, on the cultural traditions of that person (Borba, 1990). Hence, through dialogue, the classroom teacher must become aware of and gain an understanding of their students' interests, beliefs, or their collective cultural traditions.

Toward Situated Performance-based Assessment

A learning theory, situated cognition, developed by Brown, Collins, and Duguid, (1989) has as a foundational tenet that situations coproduce knowledge through activity since obtaining knowledge, including abstract technical concepts, is a process wherby meanings are inherited from the context of the activity and situation in which knowledge is developed. Similarly, Cobb (1994) contended that knowledge acquired on the basis of context requires the learning mind to integrate with one's world whereby students are acting and interacting with their social world in conjunction with conceptual learning activities. For a situated assessment to incorporate aspects of the students' "social world," the activities should have an authentic nature. Some authentic tasks are presented as one of several isolated questions that represent "realistic" situations. In contrast, a situated assessment elaborates on the realistic aspect of authentic tasks. As a problem-solving activity, the questions may be woven together into a life-like scenario that encourages possible intellectual activities performed by students in daily life.

Diverse learners bring to school a variety of interests, aptitudes, and experiences. The administration of a situated performance assessment should capitalize on various learning styles (i.e., tactile, social, visual, auditory). Usually, students encounter a highly structured and anxiety producing administration associated with traditional testing. A high stress setting is poorly suited to accommodate different learning styles, especially for those who have been culturally socialized and conditioned to prefer a learning style that is informal, collaborative, tactical, or affective (Gay, 1988).

The recognition that knowledge has a social and situated basis (Brown, Collins, & Duguid, 1989; Lave, 1993; Rogoff & Lave, 1984), has four implications for instruction and assessment. First, problems should be designed to encourage student demonstration of in-depth knowledge of targeted principles. Second, in relation to instruction, the development of knowledge and skills should be built on what students know and can do. Third, the variety of contexts allows students to exercise newly acquired problem-solving abilities, and to promote generalizability from a single situation-specific context to broader domains (Gitomer, 1993). Fourth, assessment tasks and their administration settings should be created to collectively involve the cultural traditions or experiences of the students.

Advantages of Performance Assessment

Baker, O'Neil, and Linn, (1993) have defined performance-based assessment as a type of testing that permits demonstration of understanding and skill in applied, procedural, or open-ended formats.The proponents of performance-based assessment see three advantages for its use. First, the students are required to demonstrate knowledge in situations that are more life-like and more complex, and which transcend the simplistic multiple-choice questions. Second, the assessments are more familiar to the students in that the tasks are similar to instructional activities (Frechtling, 1991) whose contextual cues elicit appropriate skills and dispositions (Marzano et al, 1989). Third, the results are considered more valid since performance assessments can be designed to be closely linked to instructional goals and provide a clearer, more accurate, and deeper understanding of what students know and can do (Ziomek & Maxey, 1993), which may have a greater impact towards the improvement of instructional programs. Performance assessments become indistinguishable from the goals that encourage students to create, perform, produce, or do something that aid in ascertaining their knowledge (Baker, O'Neil, & Linn, 1993) Thus, for classroom teachers, results from performance assessments have the promise of being more useful indicators of students' capabilities (Frechtling, 1991).

Critics of authentic assessment flinch at the idea of using performance assessments for accountability purposes (e. g., Frechtling, 1991). Since performance assessments are considered classroom practice, it is questionable to translate them into tools for accountability Frechtling (1991). The reasons cited include: i) performance assessments are too time consuming and too costly to develop; ii) scoring is frequently complex, examining processes and products; iii) the use of the classroom teacher as a critical factor in the scoring process would introduce more subjectivity into the scoring procedures; and iv) stakeholders of accountability testing find that it becomes more difficult to use alternative assessment results for assessing districts/schools and individual students.

Validity and Reliability

Although there is a trend to expand the definition of validity especially with regard to performance assessment, critics believe that the terms reliability and validity are not appropriate for testing under real-life conditions (Messick, 1989; Linn, 1995). Previously, Messick (1989) has argued that content-related validity is an adequate and sufficient consideration. More recently, the evaluation criteria for validity has included both evidential and consequential considerations (Messick, 1994). As a unifying force, construct validity can not be ignored since performance-based assessment has the potential for aligning instruction with assessment although it has limits related to consequences and interpretations of assessment results (Linn, 1995). The consequential aspect of validity is important in that the performance assessment process must take care not to alienate students by addressing the issue of fairness for diverse student populations (Telese, & Kulm, in press).

The reliability or the generalizability of performance assessment results, relative to the scoring process, is considered to be a flaw of some of the new performance-based assessments (Linn, 1995). Three criticisms of scoring performance assessments that are considered to hinder obtaining high reliability are i) rating is highly subjective, ii) raters can not be consistent, and iii) scores vary across tasks. It is generally accepted that the use of multiple raters can improve reliability, but when individuals are rated by two different raters a bias may result with one rater being more or less lenient than another (Houston, Raymond, & Sevc, 1991). Researchers (e.g., Baker 1992; Dunbar, Kortez and Hoover, 1991) have suggested that error due to raters can be held to a minimum when students have the same tasks, and there is careful training of raters on the interpretation of the scoring rubric's criteria. The items may be a larger problem than raters. Shavleson, Baxter, and Pine (1992) found a large error due to the Person x Task interaction. Another limitation of the generalizability of results from an on-demand assessment is related to duration. There may be times when an item may take more than one class period. For high generalizability coefficients to result, Linn (1995) contended that "10 tasks is beyond the realm of reasonable possibilities . . ." (p. 10).

The educational testing community must bare in mind the purpose for the various assessment systems. There is not one assessment system suited for all purposes. Advocates for the use of performance assessment are attempting to meet the challenges and criticisms cited, which should not limit their use. Performance assessments by their nature can have a greater impact in the classroom through the alignment of instructional practices with assessment. Performance assessments, in particular situated performance assessments, can be tools to gauge the effectiveness of both classroom instruction and student capabilities. Real-life scenarios offer a greater potential to establish equity in assessment practices since they can be designed to be congruent to students' experiences and cultural traditions, and permit students to demonstrate their knowledge by doing something.

DESIGN AND METHODOLOGY

Situated Task Development

The tasks' framework reflected an attempt to align a performance assessment with the eighth grade Texas Assessment of Academic Skills (TAAS) test's five mathematical content areas: (a) statistics and probability, (b) algebra, patterns, relations, and functions, (c) number concepts, (d) measurement concepts, and (e) geometric concepts. For each content area on the TAAS test, students are evaluated on mathematical operations, concepts, and problem solving. Teachers of at-risk students were invited to participate in every phase including the development and administration of the tasks, and the design of the scoring rubrics. The teachers, as co-developers of the tasks, offered suggestions for improving the wording of items, whether or not to adapt a context to suit cultural backgrounds or special student interests, and to make adjustments for mathematics or other subject matter knowledge. Once the adjustments were made, a final draft of the tasks was written and pilot tested. The pilot test was helpful in identifing any potential problems that might have been caused by deficiencies in the students' reading comprehension, writing ability, or by unfamiliarity with the tasks' contexts, and to ascertain the degree of relevancy of the tasks to students of differing ethnic and cultural backgrounds.

Based on the teachers' observations, daily dialouge with their students, and other research, the tasks (situations) were designed around four general contexts (see Table 1): a familiar activity, a social issue, hands-on construction, and video presentation. In order to determine whether it was a particular context or content knowledge that elicited the performance, each of the five content areas were represented within each context by including at least two questions from each content area. As a result, ten target questions were associated with each scenario.

Table 1
Assessment Scenarios

General Context	Focused Content	Situated Setting
Familiar Activity	Numbers and numeration	A pizza party
Social issue of current interest	Statistics and graphing	Foods, diets, and heart attacks
Hands-on construction	Geometry and measurement	Building a kite
Technology and visual information	Functions and patterns	Video story of an eagle rescue (Jasper Woodbury)

A familiar activity was selected so that it would be relevant to students through some possible experience from their daily lives. The social issue context was chosen as an activity that may be performed in a daily setting, such as reading a brochure. The issue of a healthy diet was decided upon because of a recent effort placed on good health via the media and is related directly to the students. The choice of a hands-on construction context was based on research literature which suggested that at-risk students find learning interesting and enjoyable when using hands-on materials (Cole & Griffin, 1987; Greeno, 1991). Kite building was the hands-on construction task agreed upon by the teachers. Their previous experiences with kite building activities used in the classroom led them to conclude that kites are both engaging and useful for eliciting mathematical knowledge from their students. The technology-video context was chosen simply because students enjoy watching television at home and at school. The questions within each scenario were connected, flowing from one question to the next, rather than the isolated types which are often found on traditional standardized tests. Consequently, four situated, real-life situations, based on the contexts, were established: planning a pizza party, reading information about fat and cholesterol from a brochure, building a kite, and watching *Rescue at Boone's Meadow* a video from the *Jasper Woodbury* series produced by Vanderbilt University.

Pre-assessment Tasks

In order to ensure that all students had a similar level of familiarity with the situations, a pre-assessment procedure was incorporated. The pre-assessments were designed to introduce the students to various features of the situated context. The pre-assessment activity for the pizza scenario involved having the students participate in a pizza trivia game that provided background information about pizza's history and nutritional information. For the diet situation, students were introduced to a brochure that listed consumer information from a health organization about fat and cholesterol of familiar food items and other health facts related to diet. The third pre-assessment activity involved the hands-on experience of building a kite. Students built a model of a kite to have available during the performance assessment from plastic straws, string, and butcher paper. Also, the students were provided a history of kites, and basic information on aerodynamics. The fourth pre-assessment activity used the Jasper Woodbury video (Learning Technology Center, 1992), *Rescue at Boone's Meadow*, as an engagement tool and to provide background information about ultra light aircraft.

Evaluator Training

Scoring rubrics based on the TAAS test's mathematical knowledge domains of concepts, operations, and problem solving were written incorporating aspects from the Vermont's mathematics portfolio assessment project (Vermont State Board of Education, 1990) and recommended performances from the *Curriculum and Evaluation Standards for School Mathematics* (NCTM, 1989). The teachers' comments and reviews of the scoring rubric were integrated into the final rubric (see Appendix). The rubrics were designed so that each targeted question could be rated for each of the three TAAS knowledge domains regardless of the content area (see Table 2). The scale was chosen because the TAAS test assigns a one, two, three, or four to indicate degree of mastery of a content area and to make future equating of the assessments possible.

Table 2
Third Training Trial- Raters' Assigned Scores For Kite's Target Question #6

	Operations	TAAS Domains Concepts	Problem Solving
	Students	Students	Students
	1 2 3 4 5 6	1 2 3 4 5 6	1 2 3 4 5 6
Raters			
1	4 4 3 0 3 1	4 4 3 0 2 1	4 3 2 0 2 1
2	3 4 2 0 3 1	4 4 3 0 3 1	4 4 3 0 2 1
3	4 3 1 0 3 1	4 3 1 0 3 1	4 2 1 0 3 1
4	4 3 3 0 4 2	4 3 3 0 3 2	4 3 3 0 3 2
5	4 4 2 0 4 1	4 3 2 0 2 1	4 3 2 0 2 1
6	3 4 4 0 2 1	4 4 4 0 3 2	4 4 4 0 2 2
Percent Agreement	95.6	95.6	94.4

A training session in the rubric's use was conducted by scoring the field tested tasks. The raters in the study were the four participating teachers and the two university researchers. A second practice session was held to furhter standardize interpretations of the rubric's levels. The kite scenario was randomly selected for the second training session. From the total number of papers, six student responses were selected. The procedure for the first round of scoring required that each rater score the same two target questions. Once each rater had scored the two questions, the scores were tabulated in order to determine any discrepancies. Agreement was obtained when the scores matched or differed by one. Similarly, a second pair of different targeted questions were rated. This procedure was followed one more time. A percent agreement was calculated as a measure of consistency by totaling the number of possibilities (90) and dividing the number of agreements for each student rated. With each succeeding round, the raters' percent agreement became more consistent approaching 96%. Inter-rater reliability was not an issue at this point. However, generalizability theory was employed in the research design to determine rater reliability and the reliability of the sceanarios to elicit competencies from the students will be discussed in the following section.

Generalizability Study.

A two facet mixed design G-study was employed for each G-study. The random facets were the raters and persons (students). Since five content areas were examined, the fixed facet was considered to be the TAAS content area. Six separate analyses were conducted using the raters and students as the object of measurement for each of the TAAS knowledge domains of concepts, operations, and problem solving (see Table 3).

Table 3

Generalizability Research Design Matrix

Facets		Knowledge Domain G-Studies		
Random	Fixed	Concepts	Operation	Probl. Sol
Raters	TAAS Content Area			
Students	TAAS Content Area			

The data were analyzed using the GENOVA program version 2.2 (Crick & Brennan, 1984). A subsample of 20 students' papers was randomly selected. Each of the students' papers were rated by two different randomly assigned raters. Although there were ten target questions, the scores for each of the five content areas were averaged over the two related questions.

Analysis of Variance Research Population and Cell Assignment

Teachers. Three eighth grade public school teachers from urban districts in Texas volunteered to participate. One teacher from school A is a white male. His school has a large percentage of Hispanic students. A second teacher from school B is a white female; her school's student population has a predominately white ethnicity. The third teacher from school C is an African American female whose school has a large African American student population.

Students. The research student population consisted of 139 eighth grade students of which 93 were classified according to state guidelines as being at-risk. One criteria was a TAAS test mathematics section score below the 30th percentile. The number of participating students at school A was 34, school B had 32 and school C had 27 (see Table 4).

Table 4
Number of Students in Each Cell

	School B	
Hands-on/Aided 9	Social Issue/Individual 14	Technology/Group 9
	School A	
Familiar activity/Aided	Social Issue/Group	Hands-on/Individual
7	5	6
Technology/Individual	Familiar activity/Group	Social Issue/Aided

4	8	4
	School C	
Familiar Activity/Individual 10	Technology/Aided 10	Hands-on/Group 7

Context-setting combinations. The teachers' classes were randomly assigned to one of twelve context-setting combinations which were randomly selected from the total number of possible combinations. The three administration settings were: (a) individually without assistance, (a) individually with assistance, and (c) small groups consisting of two to five students. Teachers at schools B and C decided, for practical reasons, to use three different class periods. The six classes involved were each randomly assigned, to limit the effect of time of day, to one of the twelve context-setting combinations. Since there were 12 assessment cells, consisting of three context-setting combinations, for three teachers, the teacher at School A recruited a colleague to participate who was oriented to the study and administration procedures by the participating teacher at school A. Each of the two classes in this school was randomly assigned to an assessment cell (see Table 5). This allowed the three remaining context-setting combinations to complete the matrix of twelve context-setting assessment combinations.

Administration of the Situated Performance-based Assessment

The final tasks were administered three weeks prior to TAAS testing. The project leaders traveled to each school to assist in the administration of the tasks, which was intended to model the administration of a large-scale assessment such as the TAAS. Two class periods were necessary, one for the pre-assessment, and a second class period on the following day for the actual assessment. Both at-risk students and non at-risk students in each classroom participated.

The procedures and the tasks were matched relative to each situated context so that the most effective and useful information could be obtained from the presentation. Similar procedures were followed at each site regarding the performance assessment's administration settings. The students assigned to the individual setting were told that they could not ask for any assistance. These students sat at individual desks as they would in a traditional type testing situation. The students who were assigned to the small group setting were told that they could discuss questions with each other, but each student had to do their own work. Students who were assigned the aided setting were told that they could ask for assistance from the proctor. The proctor recorded the students' questions and the response given to the student as a prompt or a micro-teaching comment. The teacher made available any tools that they might need such as compasses, protractors, rulers, and calculators. The students were permitted 45 minutes for working on the assessment since this time period was the shortest of the three schools. The first day consisted of presenting the pre-assessment activities. After announcements at school A, the classes was divided into three groups, one for each of the predetermined context and setting orders. On the following day students were administered the assessment tasks.

RESULTS

Analysis of Variance Procedure

The use of a MANOVA was considered for the research design. However, the balanced nature of a MANOVA and associated cell assignments required the teachers to devote more time than they anticipated for data collection. The teachers were not willing to give-up one to two weeks from their class schedules, which included preparing for the TAAS test. Consequently, after consulting with a statistical expert, an analysis of variance was used to analyze the effects of context, content and the setting.

To ensure a differentiation between the content and the context, the analysis consisted of two separate repeated measures analysis of variance procedures. The repeated measures were the scores from the TAAS knowledge domains of concepts, operations, and problem-solving. The first analysis used context and setting as the main factors, and the second analysis used the TAAS content and administrative setting as the main factors. The number of levels of Context was four levels, the setting had three levels, and there were five levels for the TAAS content. The repeated measures in each of the analyses were the mathematical knowledge domains of problem solving, operations, and concepts.

Means for The Knowledge Domains. The TAAS content area scores were obtained by summing the scores related to each of the five content areas. If there were two scores for a question related to a content area on an individual, then the scores were summed and averaged. If one score was missing, then the provided score was recorded, and when both scores were missing, the score was read as missing for that item. There were a total of 15 content area mean scores, five for each of the TAAS knowledge domains. This procedure was followed for all the questions per student to obtain values for each of the TAAS content areas and knowledge domains. Tables 5, 6, and 7 provide the mean content area scores for each setting and context.

Table 5
Mean Scores for Individual Setting over TAAS Content Area and Knowledge Domains

	Situations			
Content Area	**Familiar Activity**	**Hands-on**	**Technology**	**Social Issue**
Statistics & Probability	Procedures 2.05 (20, 1.28)	2.40 (10, 0.84)	2.00 (1, -)	2.66 (28, 0.97)
	Concepts 2.15 (20, 1.14)	2.40 (10, 0.84)	3.00 (1, -)	2.66 (28, 0.97)
	Prob. Solving 2.15 (20, 1.22)	2.10 (10, 0.88)	2.00 (1, -)	2.72 (28, 0.88)

Table 5 (cont.)

Content Area	Familiar Activity	Hands-on	Technology	Social Issue
Algebra, Patterns, Relations & Functions	Procedures 2.27 (11, 1.10)	1.38 (8, 0.74)	2.29 (7, 1.25)	2.42 (26, 1.14)
	Concepts 2.15 (20, 1.14)	1.25 (8, 0.46)	2.29 (7, 0.95)	2.38 (26, 1.10)
	Prob. Solving 2.15 (20, 1.23)	1.13 (8, 0.35)	2.14 (7, 1.07)	2.35 (26, 1.06)
Number Concepts	Procedures 2.00 (9, 1.22)	1.10 (10, 0.32)	2.17 (6, 1.33)	2.42 (24, 1.13)
	Concepts 2.11 (9, 1.27)	1.30 (10, 0.48)	2.50 (6, 1.22)	2.33 (24, 1.17)
	Prob. Solving 2.00 (9, 1.22)	1.20 (10, 0.42)	2.17 (6, 1.17)	2.33 (24, 1.17)
Measurement Concepts	Procedures 1.00 (1, -)	1.29 (7, 0.76)	2.00 (4, 1.15)	2.19 (16, 0.98)
	Concepts 1.00 (1, -)	1.29 (7, 0.76)	1.75 (4, 0.96)	2.19 (16, 0.98)
	Prob. Solving 1.00 (1, -)	1.29 (7, 0.76)	1.75 (4, 0.96)	2.19 (16, 1.11)

Geometric Concepts	Procedures 0.00	2.00 (10, 1.05)	2.00 (3, 1.00)	1.55 (11, 1.04)
	Concepts 0.00	2.00 (10, 1.05)	2.00 (3, 1.00)	1.64 (11, 1.03)
	Prob. Solving 0.00	2.00 (10, 1.05)	2.00 (3, 1.00)	1.64 (11, 1.03)

Note: The ordered pair, (n, s.d.), represents number of responses and the standard deviation.

Table 6

Mean Scores for Aided Setting over TAAS Content Area and Knowledge Domains

	Situations			
Content Area	**Familiar Activity**	**Hands-on**	**Technology**	**Social Issue**
Statistics & Probability	Procedures 2.52 (14, 0.94)	2.12 (17, 1.05)	1.50 (2, 0.71)	2.75 (8, 0.71)
	Concepts 2.64 (14, 0.86)	2.12 (17, 1.17)	1.50 (2, 0.71)	2.87 (8, 0.64)
	Prob. Solving 2.65 (14, 0.86)	2.12 (17, 1.17)	1.50 (2, 0.71)	2.75 (8, 0.71)
Algebra, Patterns, Relations & Functions	Procedures 2.75 (12, 1.22)	2.88 (16, 1.20)	2.75 (4, 1.26)	2.25 (8, 0.89)
	Concepts 2.75 (12, 1.14)	2.75 (16, 1.13)	2.75 (4, 1.26)	2.25 (8, 0.89)
	Prob. Solving 2.75 (12, 1.14)	2.56 (16, 0.96)	2.75 (4, 1.26)	2.25 (8, 0.71))

Table 6 (cont.)

Content Area	Familiar Activity	Hands-on	Technology	Social Issue
Number Concepts	Procedures 2.17 (12, 1.19)	1.83 (18, 0.92)	2.67 (6, 0.51)	2.20 (5, 0.84)
	Concepts 2.42 (12, 1.08)	1.89 (18, 0.83)	2.67 (6, 0.51)	2.20 (5, 0.84)
	Prob. Solving 2.17 (12, 1.03)	1.78 (18, 0.81)	2.67 (6, 0.51)	2.00 (5, 0.71)
Measurement Concepts	Procedures 1.73 (11, 1.10)	2.56 (19, 1.12)	2.20 (10, 1.03)	3.33 (3, 0.58)
	Concepts 2.27 (11, 0.90)	2.63 (19, 1.16)	2.00 (10, 1.05)	3.33 (3, 0.58)
	Prob. Solving 2.09 (11, 1.04)	2.63 (19, 1.16)	2.00 (10, 0.99)	2.67 (3, 1.15)
Geometric Concepts	Procedures 2.40 (10, 1.07)	3.22 (22, 1.05)	1.33 (3, 0.57)	2.67 (3, 1.53)
	Concepts 2.60 (10, 0.84)	3.09 (22, 1.01)	1.00 (3, 0.00)	2.67 (3, 1.53)
	Prob. Solving 2.50 (10, 0.85)	3.09 (22, 1.02)	1.00 (3, 0.00)	2.33 (3, 1.53)

Note: The ordered pair, (n, s.d.), represents number of responses and the standard deviation.

Table 7
Mean Scores for Group Setting over TAAS Content and Knowledge Domains

	Situations			
Content Area	Familiar Activity	Hands-on	Technology	Social Issue
Statistics & Probability	Procedures 2.80 (10, 1.23)	1.00 (1, -)	2.13 (16, 1.20)	3.00 (2, 0.00)
	Concepts 3.00 (10, 1.05)	1.00 (1, -)	2.25 (16, 1.13)	3.00 (2, 0.00)
	Prob. Solving 2.80 (10, 1.03)	1.00 (1, -)	2.25 (16, 1.13)	3.00 (2, 0.00)
Algebra, Patterns, Relations & Functions	Procedures 2.70 (10, 1.26)	2.00 (1, -)	3.17 (18, 0.92)	2.22 (9, 0.83)
	Concepts 2.90 (10, 1.20)	2.00 (1, -)	3.11 (18, 0.96)	2.22 (9, 0.97)
	Prob. Solving 2.80 (10, 1.23)	2.00 (1, -)	3.11 (18, 0.96)	2.11 (9, 0.93)
Number Concepts	Procedures 3.22 (9, 0.83)	1.00 (3, 0.00)	2.78 (14, 0.98)	3.00 (2, 0.00)
	Concepts 3.33 (9, 0.86)	1.33 (3, 0.58)	2.71 (14, 0.83)	3.00 (2, 0.00)
	Prob. Solving 3.33 (9, 0.86)	1.00 (3, 0.00)	2.64 (14, 0.84)	3.00 (2, 0.00)

Table 7 (cont.)

	Situations				
Content Area	Familiar Activity	Hands-on	Technology	Social Issue
Measurement Concepts	Procedures 2.50 (6, 0.55)	1.50 (2, 0.71)	2.07 (14, 0.83)	0.00
	Concepts 2.67 (6, 1.03)	1.50 (2, 0.71)	2.21 (14, 0.80)	0.00
	Prob. Solving 2.50 (6, 1.04)	1.50 (2, 0.71)	2.14 (14, 0.77)	0.00
Geometric Concepts	Procedures 1.50 (6, 0.84)	1.33 (3, 0.58)	2.53 (19, 1.07)	0.00
	Concepts 1.50 (6, 0.84)	1.67 (3, 0.58)	2.53 (19, 1.07)	0.00
	Prob. Solving 1.33 (6, 0.52)	1.33 (3, 0.58)	2.47 (19, 1.07)	0.00

Note. The ordered pair, (n, s.d.), represents number of responses and the standard deviation.

Analysis of Variance

ANOVA Procedure. The number of observations used in the analyses was 2,790. The mainframe SAS program version 6.08 was employed to analyze the data. The analysis was complex and cumbersome. The within-subjects portion of the full model was extremely large due to the large number of data points used by the program for the interaction term Items x TAAS Knowledge Domain x Student nested in Setting x Context. As a reslut, the sum of squares was not calculated for the full model by the program (see Table 8).

Since the sum of squares could not be calculated, the analysis was conducted from two perspectives, one perspective used the full model without the student interactions terms. The second perspective included an analysis for each of the within-subjects factors to obtain separate error terms. A separate summary was produced for each category of the Within-subjects components.

Table 8
Analysis of Variance Results of TAAS Knowledge Domains with Context and Setting as Between Subjects and TAAS Knowledge Domains and Items as Within Subject Factors

Between Subjects

Source of Variation	df	Sum of Squares	Mean Squares	F	Pr > F
Setting (S)	2	44.30	22.15	3.30 i	0.0422
Context (C)	3	33.01	11.00	1.64 i	0.1871
SC	6	65.40	10.90	1.62 i	0.1530
Student(SC) i	77	516.76	6.71		

Within Subjects

Source of Variation	df	Sum of Squares	Mean Squares	F	Pr > F
Domain (D)	2	1.12	0.56	3.29 j	0.0399
S*D	4	0.12	0.03	0.18 j	0.9484
C*D	6	2.13	0.36	2.12 j	0.0540
S*C*D	12	1.32	0.11	0.65 j	0.7965
Student*D (S*C) j	154	25.87	0.17		
Item (I)	9	63.32	7.04	10.51 k	0.0001
I*S	18	98.69	5.48	8.18 k	0.0001
I*C	27	187.12	6.93	10.34 k	0.0001
I*S*C	54	98.61	1.83	2.73 k	0.0001
I*D	18	2.41	0.13	0.19 k	0.9999
I*S*D	36	2.30	0.06	0.09 k	1.0000
I*C*D	54	4.13	0.08	0.12 k	1.0000
I*S*C*D	108	6.27	0.06	0.09 k	1.0000
Residual k	1186	789.08	0.67		

Note. Domain is the abbreviated form for the TAAS knowledge domain areas of procedures, concepts, and problem solving. The letters i, j, and k denote the error term used to calculate the F-values.

Since the within-subjects portion of the model could not be calculated by the mainframe computer, the variances were apportioned in order of importance. This allowed the sum of squares for the interaction Item by TAAS knowledge domain to be calculated after the main effects. The purpose of the partitioning was to tease out the effect of the student interaction terms on the sum of squares for the model so that a determination could be made concerning which student interaction sum of squares should be used as an error term. The three student interaction sum of squares were (a) Item x Student nested in Setting x Context, (b) Student x TAAS Knowledge Domain nested in Setting x Context, and (c) Items x TAAS Knowledge Domain x Student nested in Setting x Context. The sum of squares for (c) was not produced by the program.

ANOVA Results for Context. The between subjects analysis indicated that the main effect, Setting, was significant; the students differed in their performance levels within the various settings. The main effect context was not significant at the 0.05 level indicating that the contexts did not differ in eliciting mathematical competencies. The interaction Setting x Context was not significant which indicated that the students performed similarly regardless of the context-setting combinations.

The within-subjects factor, TAAS Knowledge Domains, was significant with an F-value of 3.29 and $p < 0.05$. This finding indicated that students varied in their abilities relative to the TAAS knowledge domains of operations, concepts, and problem solving skills. The Context x TAAS knowledge domain interaction approached significance, indicating that the students' TAAS knowledge domains appeared to be affected by the context of the assessment (see Table 8).

The item component for within-subjects analysis was significant which indicated that the students' performance varied depending on the item. The two-way interactions, Item x Setting, and Item x Context, were significant, and the three-way interaction, Item x Setting x Context was significant (see Table 8). These results indicated that performance on individual items was influenced by the setting and/or context in which the items were experienced. Apparently, the items produced different effects depending on the setting and context in which they were solved.

Three-way Interaction Effects with Context

The within-subjects three-way interaction, Item x Setting x Context, provided insight into individual item performance by the students in each setting and context. Results attributed to this interaction indicated that the contexts of the situated performance assessment had fairly similar effects for each of the TAAS knowledge domains. The interactions between the setting and context with problem solving scores for each setting, individual, aided, and small group are presented in figures 1, 2, and 3 respectively. Since problem-solving involves both conceptual and procedural (operations) knowledge, the raters tended to assign similar scores for each of the TAAS knowledge domains. For example, when a target question was rated a two for operations, it was also rated a two for concepts and a two for problem-solving. This rating pattern was unexpected. The problem-solving scores were selected for illustration because they are representative of the scores for the TAAS knowledge domains.

Generally, the hands-on context appeared to produce lower problem-solving mean scores (see Figures 1 to 3). The mean problem-solving scores were lower in the individual setting than in the aided or group settings. The aided setting appeared to raise the mean scores (see Figure 2). The familiar activity and technology contexts appeared to produce higher problem-solving scores in individual and group settings (see Figures 1 & 3).

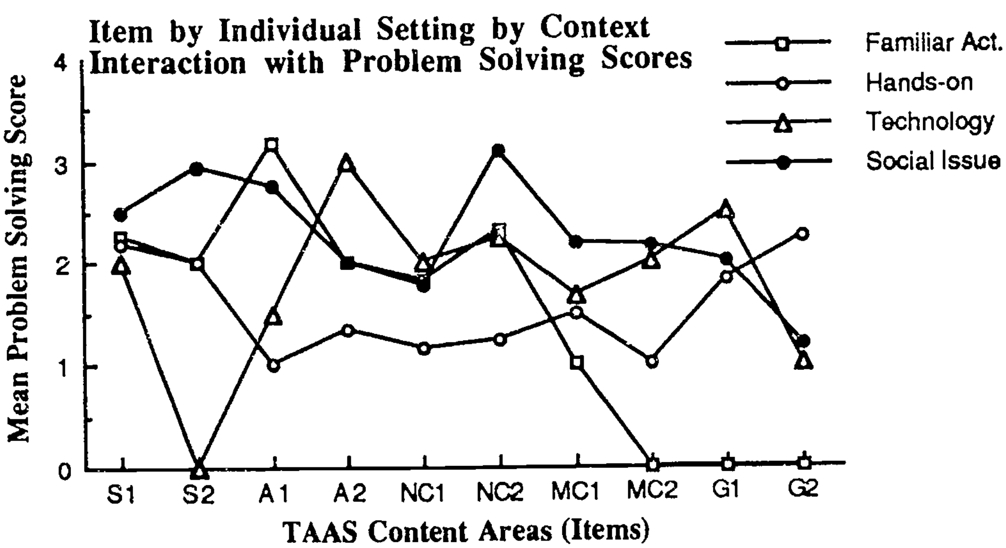


Figure 1. Three-way Interaction Item by Individual Setting by Context for Problem Solving Scores

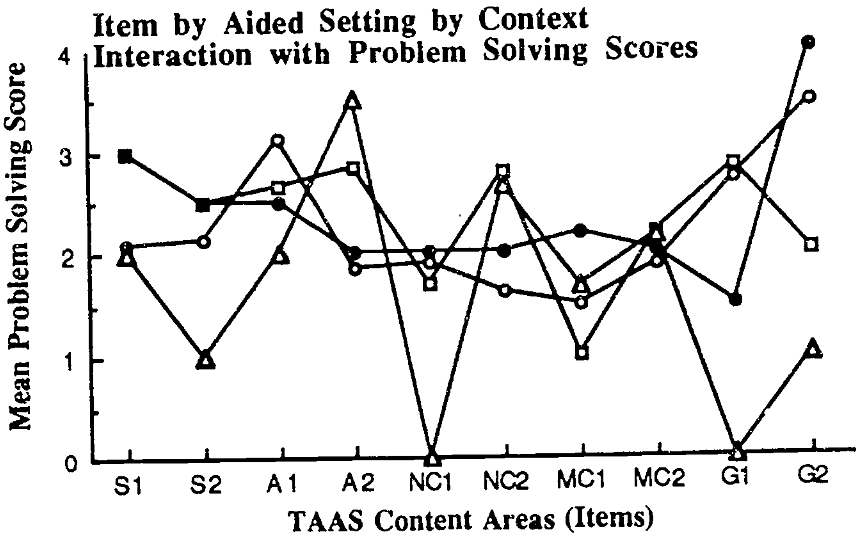


Figure 2. Three-way Interaction Item by Aided Setting by Context for Problem Solving Scores

Item S2 from the technology context appears to be an outlier; students did not answer this question. Likewise, time constraints might have hindered students in completing items MC2, G1, and G2 of the familiar activity context. The group setting appeared to produce a wider range of mean problem-solving scores while the aided and individual settings appeared to have less of a separation between the scores.

Generally, the contexts of familiar activity, social issue, and technology appeared to generate higher mean scores depending on the knowledge domain, and the hands-on context produced the lowest scores regardless of setting. The group setting appeared to produce higher problem-solving scores for the familiar activity scenario than the other situated contexts (see Figure 3).

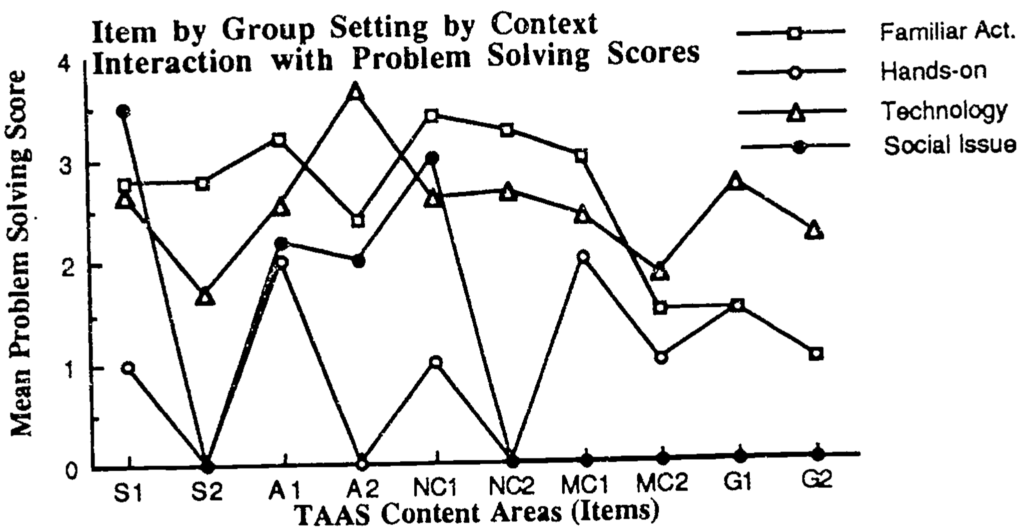


Figure 3. Three-way Interaction Item by Group Setting by Context for Problem Solving Scores

These three-way interactions may be influential for two reasons. First, the lower mean problem solving scores for the hands-on context may have influenced the interaction. Second, the technology outlier items NC1, and G1 were not answered, perhaps due to time constraints.

The group setting appeared to elicit a greater variety of mean scores for the hands-on scenario (see Figure 3). This variation may be due to the students' lack of classroom experience working in groups. In contrast, the social issue and the technology scenarios seemed to have a large role in eliciting problem solving abilities from the students as indicated by the higher mean scores. It appeared that the hands-on, social issue, and the technology scenario were influential in eliciting some mathematical understanding. The geometry- related items were near the end of the familiar activity scenario, which might explain the lack of response for items MC2, G1, and G2.

<u>ANOVA Results for Content</u>. The second analysis of variance procedure used the TAAS content and administravtive setting as the main effects. Table 9 presents the results of the repeated measures analysis of variance tests of hypotheses for between-subjects effects.

Table 9
Analysis of Variance for Setting and Content as Main Factors

Source	df	Sum of Squares	Mean Square	F Value	Pr > F
Setting (S)	2	61. 39	30.70	5.12	0.0095
Error	50	299.98	5.99		
TAAS Content(TC)	4	7.00	1.74	3.23	0.0136
TC * S	8	10.06	1.26	2.33	0.0207
Error(TC)	200	108.01	0.54		
TAAS Knowledge Domain (D)	2	0.15	0.08	0.83	0.4376
S * D	4	0.44	0.11	1.20	0.3153
Error(D)	100	9.07	0.09		
TC * D	8	16.58	2.07	2.10	0.0350
TC * D * S	16	49.12	3.07	3.11	0.0001
Error(TC*D)	400	395.36	0.99		

The TAAS knowledge domains factor was not significant with an F-value of 0.83, $p < 0.05$ indicating that the results of the analysis of variance for the content areas were similar regardless of the knowledge domains. The administration's setting, however, was significant which indicated that students performed differently in each administrative setting (see Table 10). The content area was also significant which indicates that the students performed better on some content areas. The three way interaction, Content area by Knowledge domain by Setting, was significant, which indicated that the students performed better on some content areas within certain settings (see Table 9).

Figure 4 illustrates the interaction between the administrative setting and content for problem-solving. The mean problem-solving scores for the content areas algebra and number concepts were the greatest in the group setting. In the aided setting, the mean scores were enhanced for the content areas algebra, geometry, and measurement concepts while the content area, statistics and probability, had nearly equal means (see Figure 4). Also, the students in the aided and individual settings performed equally well on the items related to number concepts. The items in the aided setting related to geometry produced greater mean scores than in the other two settings.

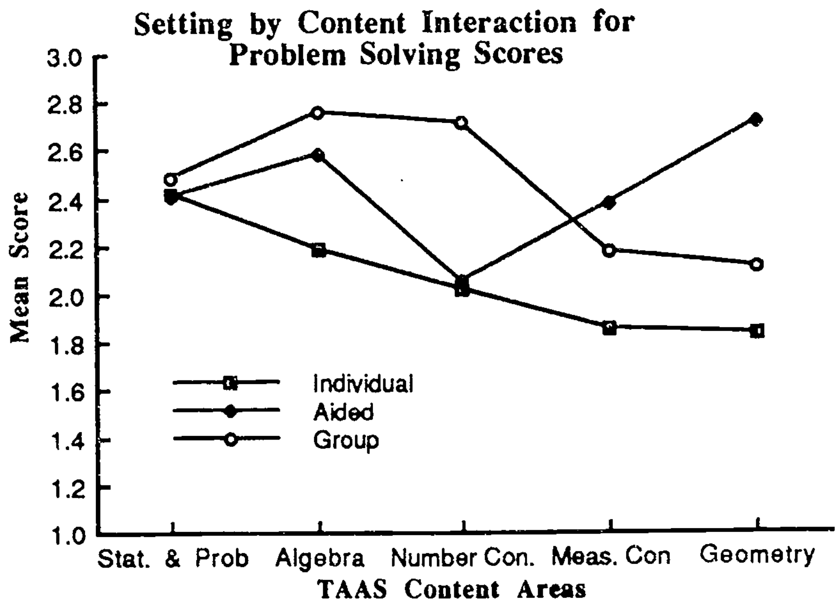


Figure 4. Setting by Content Interaction for Problem Solving Scores

Generalizability

Rater Consistency. Table 10 presents the results for the operations scores with raters as the object of measurement. The small total variance percentage of 4.4% for the Raters x Persons interaction and 3.4% for the raters indicated that the evaluator training was effective in producing consistent socores over individuals' mathematical operation knowledge. The G-coefficient of 0.65 indicated that the raters were consistent in assigning operations scores to the students' responses. The interaction Persons x Content contributed 47.6% of the total variance, which indicated that the students had different mathematical operation skills.

Table 10
Rater GENOVA Estimates of Variance Components for the TAAS Knowledge Domain-Operations Scores

Source of Variation	Sums of Squares	df	Mean Squares	Estimated Variance Components	Percentage of Total Variance
Persons (P)	4.832	19	4.832	0.452	31.3
Raters (R)	0.911	1	0.911	0.006	3.4
Content (C)	10.007	4	2.502	0.023*	1.6
PR	6.014	19	0.317	0.063	4.4
RC	0.283	4	0.071	0.0	0.0
PC	120.543	76	1.586	0.687	47.6
RPC	16.168	76	0.213	0.213	14.7

Note. Percentages do not equal 100% due to rounding.
G-coefficient = 0.65
*Quadratic Form

Table 11 presents the results for the concepts scores with raters as the object of measurement. The Persons x Content interaction was 47.5% of the total variance. This indicated that the students held widely different views concerning the degree of difficulty of the mathematical concepts within the content areas. The interaction Raters x Persons contributed 8.9% of total variance which indicated that the raters disagreed when assigning concept scores. The low G-coefficient of 0.37 indicated that the assignment of concept scores was not consistent or reliable.

Table 11
Rater GENOVA Estimates of Variance Components for the TAAS Knowledge Domain-Concepts Scores

Source of Variation	Sums of Squares	df	Mean Squares	Estimated Variance Components	Percentage of Total Variance
Persons (P)	92.593	19	4.874	0.421	28.9
Raters (R)	1.051	1	1.051	0.004	0.3
Content (C)	4.550	4	1.138	0.0*	0.0
PR	12.624	19	0.664	0.132	8.9
RC	0.580	4	0.145	0.0	0.0
PC	120.950	76	1.591	0.693	47.5
RPC	15.620	76	0.206	0.206	14.2

Note. Percentages do not equal 100% due to rounding.
G-coefficient = 0.37
*Quadratic Form

Table 12 illustrates the variance percentages for the problem solving scores. The G-coefficient of 1.0 was a result of the raters who had zero variance. This indicated that the raters were very consistent in interpreting problem solving scores. The greatest percentage of the total variance was due to the Persons x Content interaction with 46.6%. This indicated that the students viewed the content has having different degrees of difficulty. The next highest percentage of total variance, 28%, was due to the persons facet. This indicated that the students had different levels of problem solving abilities.

Table 12
Rater GENOVA Estimates of Variance Components for the TAAS Knowledge Domain-Problem Solving Scores

Source of Variation	Sums of Squares	df	Mean Squares	Estimated Variance Components	Percentage of Total Variance
Persons (P)	83.205	19	4.379	0.388	28.0
Raters (R)	0.045	1	0.045	0.0	0.0
Content (C)	11.843	4	2.961	0.369*	2.7
PR	9.605	19	0.506	0.101	7.3
RC	0.668	4	0.167	0.0	0.0
PC	112.758	76	1.484	0.630	46.6
RPC	16.933	76	0.223	0.223	16.1

Note. Percentages do not equal 100% due to rounding.
G-coefficient = 1.00
*Quadratic Form

Student Consistency. A summary of the results for the students as the object of measurement has been collasped into Table 13 The table contains the percentage of total variance and the G-coefficients for each of the TAAS knowledge domains. For each knowledge domain, the Persons x Content interaction contributed between 30 to 47% of the variance. Although the students had less variance when performing operations, the students viewed the assessment's difficulty in relation to their ability levels and had varying performance levels dependening on the content areas. Similarly, the next largerst portion of variance, about 28% was contributed by the students. This indicated that regardless of the knowledge domain, the students had various levels of proficiency.

Table 13
Summary of The Persons G-Studies Percent of Total Variance

	Percent of Total Variance for Mathematical Knowledge Domains		
Source of Variation	Operations	Concepts	Problem Solving
Persons (P)	30.8%	28.9%	28.1%
Raters (R)	0.3	0.3	0.0
Content (C)	1.5*	0.0*	2.6*
PR	5.7	9.1	7.4
RC	0.0	0.0	0.0
PC	37.1	47.6	45.1
PRC	24.5	14.1	16.8
G-coefficient	0.92	0.86	0.88

Note: Percentages do not total 100% due to rounding.
*Quadratic Form

The G-coefficient of 0.92 for the operations scores indicated that the assessment was very consistent in eliciting operations skills from the students. The analysis for the concepts scores produced a G-coefficient of 0.86 which indicated that the assessment was reliable in elicting concept scores. The problem solving score analysis produced similar results. The G-coefficient was 0.88, which indicated that the assessment was reliable in eliciting problem-solving responses from the students.

Discussion

Response to the Research Questions

The between-subjects analysis of the main effects for context (situations) and administration setting indicated that the contexts did not differ across contexts. Overall, the students demonstrated similar performances regardless of which context they experienced. The two-way within-subjects interaction TAAS knowledge domain by context approached significance. Yet, the contexts may have had some effect on eliciting the TAAS knowledge domains from the students indicated by the significant within-subjects two-way interaction item by context. This was possibly the result of whether the students had insufficient time to complete the tasks or their lack of interest. After a discussion with the teachers, it was concluded that the unanswered questions were due to the limited time frame since the teachers were reluctant to keep the students from attending their other classes. Ideally, the students should have been permitted to respond to the assessment with an unlimited time.

The within-subjects analysis of the contexts, familiar activity, technology and the social issue scenario, produced the greater mean scores when compared to the hands-on context mean scores. Although the context main effect indicated that the contexts did not differ from each other, there is some evidence suggested by both the within subjects two-way interaction of item by context, and the three-way interaction item by setting by context, that the context had some effect on the students' performance on certain items for a particular context. The influence of the interactive effects of context in this study lends support to the situated learning theory of Brown, Collins, and Duguid (1989) who suggested that situated contexts aid in accessing important mathematical structures and perceiving associated cues. Lave (1993) suggested that contexts are part of one's social world. The greatest means were produced by the at-risk students within the social contexts rather than in the hands-on scenario. This observation concurs with Young and Kulikowich's (1993) finding that instruction connected to students' past experiences and knowledge is influential in sparking certain abilities during the learning process. The contexts were designed to activate the students' attributes and abilities although some students had insufficient time to complete the tasks. Those students who attempted a larger portion of the tasks were were apparently guided cognitively toward their mental models of the content areas which were perceived from contextual clues and information in support of Means and Knapp's (1991) finding. The pizza scenarioan the diet scenario appeared to be more effective than the kite-hands-on scenario. Although the kite building was used in previous lessons before the research project began, it was incongruent to the at-risk students' experience. Apparently, kite building is not a usual activity for them. A context's design, as an aid in producing a more descriptive picture of students' mathematical capabilities, should have relevance and represent meaningful activities, which foster in the students a willingness to demonstrate their mathematical knowledge.

The analysis of administrative setting main effect indicated a similar pattern of results, regardless of the TAAS knowledge domain. Generally, students who experienced the aided setting had lower mean scores than those who were administer the tasks in the group setting. The students who administered the assessment in the individual setting had the lowest mean scores in comparison to the aided or group administrative settings. The students tended to have greater consistency of scores in the aided setting. The higher group mean scores lend support to findings of Phelps and Damon (1989) who found achievement gains in students' mathematical understanding from working in groups. The higher group mean scores also support the findings of De Avila (1986) that Mexican-American students showed gains in mathematics learning. A study by Gilbert and Gay (1985) also determinded that African-American students respond better when working in cooperative groups.

With regard to the content areas, the most effective administrative setting for eliciting number concepts, geometry concepts and algebra, patterns, and function concepts was the group setting. The aided setting in comparison to the individual setting was effective in eliciting algebra, relations and function concepts, and geometry concepts. However, the content area of statistics and probability had the greater mean scores in the individual setting than either group or aided settings, regardless of knowledge domains.

The setting by context interaction was not significant. The students produced similar performances regardless of context in each setting. The within-subjects three-way interaction of items, setting and context indicated that the students performed better on certain items than others within the different contexts and settings. The students had higher scores in the familiar activity and social issue contexts for the content areas of number concepts, statistics and probability, and algebra, relations and functions. The hands-on context's inherent geometric nature is reflected by the larger mean scores for the geometry content area than in the other content areas. Hence, the group and aided settings were influential in raising the mean scores regardless of context or content, and the aided setting helped to improve the geometry content mean scores regardless of context.

In summary, the mean values were useful indicators of general student performances. The means for the hands-on context in the individual setting appeared to be the lowest in comparison to the mean scores for the other three contexts. The social issue and the technology contexts appeared to produce the highest mean scores for the content areas of (a) statistics and probability, (b) algebra, patterns, relations and functions, and (c) number concepts. This trend may be due to the nature of the social issue and technology contexts which had those content areas embedded in them. Although the familiar activity produced means nearly equal to the other contexts, the content areas of measurement concepts and geometric concepts had means of zero. Because of time constraints, this result was possibly due to these items which were encountered near the end of the assessment restricting student performance.

The generalizability study indicated that the raters were dependable in assigning scores. This finding is contray to the finding by Houston, Raymond, and Sevc (1991) that a bias may result when one individual is rated by two raters. The problem solving G-coefficient has two possible explanations. First, the students were consistent and had similar socres, thereby restricting the range of scores. The students may have had similarly weak problem-solving skills which tended to produce consistently low scores. Second, the raters had very little variance when assigning problem solving scores. This explanation is related to the first; the students had very similar problem-solving skills. Hence, the raters were able to assign similar scores. Also, the raters may have had congruent interpretations of the problem-solving criteria. Since at-risk students are often placed in classrooms were problem-solving skills are not emphasized (Oakes, Ormseth, Bell, & Camp, 1990; Secada, 1992), the first explanation is most likely to have occured.

The G-studies indicated that the problem-solving domain was the most accurately rated domain whereas the concept domain posed the most difficulty for the raters. An improved training program to clarify the meaning and recognition of mathematical concepts elicited from students would possibly raise the scoring reliabilty of mathematical concepts.

CONCLUSIONS

This study showed that both the group and aided settings were influential in revealing at-risk students' mathematical competencies. The generally higher mean scores associated with the group administrative setting is consistent with the literature concerning at-risk students' enhanced achievement as a result of working in groups. The aided adminstrative setting permitted students to clarify misunderstandings whether in reference to the wording of a question or content. The opportunity to ask for clarifications possibly reduced the assessment environment to a less stressful one, which may have contributed to the higher mean scores than in the individual setting.

An important finding of this study was that the at-risk students demonstrated similar performances regardless of the contexts for each TAAS knowledge domain. The assessment contexts seemed to have similar roles in revealing at-risk students' mathematical competencies. However, some differences in the situated activities were apparent in influencing students' abilities to demonstrate their mathematical understanding. For example, the situated activity of planning a pizza party (familiar context) made particular mathematical content areas such as number concepts, and fractions more accessible to the students than the other contexts. The social issue context (diet) had an inflence on the students' ability to make mental connections to their schema related to statistical information. Apparently, the students' ability to retrieve certain mathematical information was fostered by certain contexts, as a result of having mathematical knowledge associated with the particular contexts.

These results on context-based assessment have implications for the classroom. Since the assessment contexts produced similar results in revealing student abilities in mathematics, the classroom teacher could develop a context-based activity related to their students' interests and cultural backgrounds in order to foster mathematical learning. Rather than simply realizing that they have an interest or enjoyment of the context itself, it is important to know in which contexts students have mathematical knowledge.

During the design process of a contextualized performance-based assessment, decisions sometimes center on whether to create contexts which have inherent mathematical content in order to expose that content or to design an assessment which crosses several content areas. Based upon the within-subjects results, it is recommended that the design of a context (scenario) elicits specific mathematical content rather than attempting to cross several mathematical content areas. Finally, another consideration for the design of a situated performance-based assessment involves the careful development of conceptually-based questions. It is recommended that instructional wording of the concept-based questions, in comparison to a procedural based question, should be written clear enough so that the student can respond with a demonstration of conceptual understanding rather than a procedural skill. Also, the rubric's design, for assigning concept scores, should be very specific. There should be little room for rater misunderstandings as to what constitutes a conceptual response. This could be accomplished through careful wording of the rubric's descriptors for assigning concepts scores.

Limitations

A limitation of the research design that possibly had an effect on the results was placing emphasis of an assessment order in one school in which the students did not attempt to put forth their best efforts. However, the contexts of the assessment did stimulate a few of the students to respond. This did provide an indication of the effectiveness of a contextualized basis for assessment and teaching. Another limitation was the small number of participants in the social issue/group assessment order which produced several unanswered questions.

The raters had more difficulty assigning the concepts scores than the other two knowledge domains. The difficulty arose from possible misunderstanding of the rubric's description concerning concepts a trend which was not evident during rater training. Another possible explanation is that the tasks did not lend themselves to elicit conceptual understanding in an effective manner. Time became a crucial factor during the assessment process, making it difficult for some students to have an opportunity to complete as many of the items as anticipated within the 45-minute class periods. This constraint could have been corrected by permitting the students to continue working or by returning on the third day to allow students to complete unanswered questions.

Implications for Further Study

There is a growing body of literature (e.g., Brown, Collins, & Duguid, 1989; Lave, 1993; Young & Kulikowich, 1993) which suggests that interest in an activity sparks certain abilities. The findings of this study lend support to this literature base because each of the contexts were designed to be both meaningful and relevant to the students with a strong attempt to stimulate student interest in the contexts. However, the degree of interest in the contexts should be examined to gain careful insights into the effect student interest exerts on performance levels. The study provided indications that a certain context might influence performance levels for particular content area. Further research is needed to examine how the results would change when the same content question is placed into each context. Also, the study provides support to Cobb (1994) who contends that constructivism and sociocultural factors tell half of a good story, "the sociocultural perspective gives rise to theories of the conditions for the possibility of learning . . . whereas theories developed from the constructivist perspective focus on both what students learn and the processes by which they do so" (p. 18).

Situated learning theory (Brown, Collins, & Duguid, 1989) suggests that contexts provide cues for students when they are accessing mathematical knowledge from a context. Whether in the form of an assessment or class activity, a delineation of what embedded mathematical content knowledge at-risk students have within a context is needed. The results may not be unique to at-risk students. As a follow-up study, the comparison between at-risk and non at-risk students is considered. Students regardless of backgrounds may benefit from situated performance-based assessment or learning. Because time became a crucial element which seemed to hinder student performance, a similar study should be conducted which would allow the students more time. A research design to include a larger number of students for each cell may prove enlightening in an effort to minimize the effect of small cell sizes. Further research is necessary to gain insight into the complex role context or situated activities have in prompting mathematical abilities. Contexts have the potential to hinder or enhance understanding and exert a range of influences when attempts are made to transfer mathematical abilities. Indications from the study reveal that mathematics tasks should connect to students' social environment when they are expected to use their mathematical knowledge in different situations.

References

Aronowitz, S., & Giroux, H. A. (1993). Education under siege (2nd. ed.). Westport, CT: Bergin & Garvey.

Baker, E. L. (1992). The role of domain specifications in improving the technical quality of performance assessment. (Tech. Report). Los Angeles: UCLA Center for Research on Evaluation Standards and Student Testing.

Baker, E., O'Neil, H., & Linn, R. (1993). Policy and validity prospects for performance-based assessment. American Psychologist, 48 (12), 1210-1218.

Borba, M. (1990). Ethnomathematics and education. For The Learning of Mathematics, 10(1), 39-43.

Bransford, J., & Vye, N. (1989). A perspective on cognitive research and its implications in instruction. In L. B. Resnick & L. E. Klopfer (Eds.), Toward the thinking curriculum: Current cognitive research (pp. 173-205). Alexandria, VA: Association for Supervision and Curriculum Development.

Brown, J., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. Educational Researcher, 18(1), 32-41.

Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. Educational Researcher, 23(7), 13-20.

Cobb, P. (1986). Contexts, goals, beliefs, and learning mathematics. For The Learning of Mathematics, 6(2), 2-9.

Cole, M. & Griffin, P. (1987). Improving science and mathematics education for minorities and women: Contextual factors in education. Madison, WI: Wisconsin Center for Education Research, School of Education, University of Wisconsin-Madison.

Crick, J., & Brennan, R. (1983). Manual for GENOVA: A generalized analysis of variance system (ACT Technical Bulletin, No. 43). Iowa City, IA: ACT.

De Avila, E. (1986). Bilingualism, cognitive function, and language minority group membership. In R. Cocking & J. Mestre (Eds.), Linguistic and cultural influences on learning mathematics (pp. 101-123). Hillsdale, NJ: Lawrence Erlbaum Associates.

Davis, R. B., & Maher, C. A. (1990). Constructivist view on the teaching of mathematics (Monograph No. 4). Journal for Research in Mathematics Education. Reston, VA: NCTM.

Dunbar, S., Koretz, D., & Hoover, H. (1991). Quality control in the development and use of performance assessments. Applied Measurement in Education, 4(4), 289-303.

Engeström, Y. (1993). Developmental studies of work as a test bench of activity theory: The case of primary care medical practice. In S. Chalkin & J. Lave (Eds.), Understanding practice: Perspectives on activity and context (pp. 64-103). New York: Cambridge University Press.

Frechtling, J. (1991). Performance assessment: Moonstruck or the real thing. Educational Measurement: Issues and Practices, 10(4), 23-25.

Gay, G. (1988). Designing relevant curricula for diverse learners. Education and Urban Society, 20(4), 327-340.

Gilbert, S., & Gay G. (1985). Improving the success in school of poor black children. Phi Delta Kappan, 67(3), pp. 133-137.

Gitomer, D. (1993). Performance assessment and educational measurement. In R. E. Bennett & W.C. Ward (Eds.), Construction versus choice in cognitive measurement: Issues in constructed response performance testing, and portfolio assessment (pp. 1-28). Hillsdale, NJ: Lawrence Erlbaum Associates.

Greeno, J. (1991). Number sense as situated knowing in a conceptual domain. Journal for Research in Mathematics Education, 22(3), 170-218.

Hilliard, A. (1990). Misunderstanding and testing intelligence. In J. I. Goodlad, & P. Keating (Eds.), Access to knowledge: An agenda for our nation's schools (pp. 145-158). New York: College Entrance Examination Board.

Houston, W. M., Raymond, M. R., & Svec, J. C. (1991). Adjustments for rater effects in performance assessment. Applied Psychological Measurement, 15(4), 409-421.

Jacob, E, & Jordan, C. (1993). Understanding minority education: Framing the issues. In E. Jacob & C. Jordan (Eds.), Minority education: Anthropological perspectives (pp. 3-15). Norwood, NJ: Ablex Publishing Corp.

Kulm, G. (1990). New directions for mathematics assessment In G. Kulm (Ed.), Assessing higher-order thinking skills in mathematics (pp. 71-78). Washington, DC: American Association for the Advancement of Science.

Lave, J. (1993). The practice of learning. In S. Chalkin & J. Lave (Eds.), Understanding practice: Perspectives on activity and context (pp. 3-30). New York: Cambridge University Press.

Learning Technology Center. (1992). Rescue at Boone's meadow. Nashville, TN: Learning and Technology Center, Vanderbilt University.

Lesh, R., & Lamon, S. (1992). Assessing authentic mathematical performance. In R. Lesh & S. Lamon (Eds.), Assessment of authentic performance in school mathematics (pp. 17-62).

Washington, DC: American Association for the Advancement of Science.

Linn, R. (1995). Performance assessment: Policy promises and technical measurement standards. Educational Researcher, 23(9), 4-14.

Marzano, R., Brandt, R., Hughes, C. S., Jones, B. F., Presseisen, B. Z., Rankin, S. C., & Suhor, C. (1989). Dimensions of thinking. Alexandria, VA: Association for Supervision and Curriculum Development.

Means, B., & Knapp, M. (1991). Models for teaching advanced skills to educationally disadvantaged students. In B. Means & M. Knapp (Eds.), Teaching advanced skills to educationally disadvantaged students (pp. 1-20). Washington, DC: Department of Education, Office of Planning, Development, and Evaluation.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. Educational Researcher, 23(2), 13-23.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 13-103). New York: Macmillan.

National Council of Teachers of Mathematics. (1993). Assessment standards for school mathematics: Working draft. Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (1989). Curriculum and evaluation standards for school Mathematics. Reston, VA: NCTM.

Oakes, J., Ormseth, T., Bell, R., & Camp, P. (1990). Multiplying inequalities: The effects of race, social class, and tracking on opportunities to learn mathematics and science. Santa Monica, CA: The Rand Corporation.

Phelps, E., & Damon, W. (1989). Problem solving with equals: Peer collaboration as a context for learning mathematics and spatial concepts. Journal of Educational Psychology, 81(4), 639-646.

Rogoff, B., & Lave, J. (Eds.) (1984). Everyday cognition: Its development in social context. Cambridge, MA: Harvard University Press.

Romberg, T. (1992). Assessing mathematics competence and achievement. In A. R. Tom (Ed.), Toward a new science of educational testing and assessment. (pp. 23-52). Albany, NY: State University of New York.

Ross, S., McCormick, D., & Krisak, N. (1986). Adapting the thematic context of mathematical problems to student interests: Individualized versus group-based strategies. Journal of Educational Research, 79(4), 245-252.

Saljo, R., & Wyndhamn, J. (1993). Solving everyday problems in the formal setting: An empirical study of the school as a context for thought. In S. Chalkin & J. Lave (Eds.), Understanding practice: Perspectives on activity and context (pp. 327-343). New York: Cambridge University Press.

Secada, W. G. (1992). Race, ethnicity, social class, language, and achievement in mathematics. In D. A. Grouws (Ed.), Handbook of research on mathematics teaching and learning (pp. 623-660). New York: Macmillan Publishing Company.

Schmittau, J. (1991). Mathematics education in the 1990s: can it afford to ignore its historical and philosophical foundations? Educational Theory, 41(2), 121-133.

Shavelson, R, J., Baxter, Gail, P., & Pine, J. (1992). Performance assessments: Political rhetoric and measurement reality. Educational Researcher, 21(4), 22-27.

Shavelson, R. J., Webb, N. M., & Lehman, P. (1986). The role of symbol systems in problem solving: A review of literature. CSE Report No. 269. Los Angeles, CA: CRESST.

Telese, J., & Kulm, G. (in press). Mathematics assessment scenarios for at-risk students. The Urban Review.

Vermont State Board of Education. (1990). Looking beyond "The answer": The report of Vermont's mathematics portfolio assessment. Mont Pelier, VT: Vermont State Board of Education.

Wittrock, M. (1991). Testing and recent research in cognition. In M. C. Wittrock, & E. L. Baker (Eds.), Testing and cognition (pp. 5-16). Englewood Cliffs, NJ: Prentice Hall.

Young, M., & Kulikowich, J. (1992, April). Anchored instruction and anchored assessment: An ecological approach to measuring situated learning. Paper presented at the Annual Meeting of the American Education Research Association, San Francisco, CA

Ziomek, R., & Maxey, J. (1993). To nationally test or not to nationally test: That is the question! Measurement and Evaluation in Counseling and Development, 26(1), 64-68.

Appendix

Scoring Rubrics.

Operations

0 - No response

1 - Incorrect or very limited use of operations, more than one major error or omissions.

2 - Some correct use of number operations but a major error or with several minor errors

3 - Appropriate use of number operations with possible slips or omissions, but without significant errors.

4 - Extended use of number operations without errors in calculations; appropriate use of models or representations.

Concepts

0 - Lack of evidence to determine knowledge, or no attempt made.

1 - Wide gaps in concept understanding, major errors made based on lack of conceptual knowledge.

2 - Some evidence of conceptual understanding, but difficulty in using models, diagrams, and symbols for representing concepts or translating from one mode to another mode. Some evidence of the concept's properties.

3 - Good evidence of conceptual knowledge. No major misconceptions; responses contain accurate use of models, diagrams, and symbols with evidence of translation form one mode to the other. Recognition of the meaning and interpretation of concepts. Some evidence of using concepts to verify or explain procedures.

4 - Clear understanding of concepts and associated procedures. Effective use of models, diagrams, and symbols with broad translation from one mode to another. Recognition of the meaning and interpretation of concepts to explain or verify procedures or conclusions.

Problem Solving

0 - No response

1 - Unworkable approach, incorrect or no use of mathematical representations, poor use of estimation, evidence for lack of understanding.

2 - Appropriate approach, estimation used, implemented a strategy, possibly reasoned decision making, solution with observations.

3 - Workable approach, used estimation effectively, mathematical representation used appropriately, reasoned decision making inferred, judged reasonableness of solution.

4 - Efficient/sophisticated approach, estimation used effectively, extensive use of mathematical representations, explicit reasoned decision making. Solutions with connections, synthesis, or abstraction.

Social Issue (diet)
Questions

Note: the brochure is not illustrated.

Instructions for the student. Use the information from the brochure to answer the questions. You may use any tool like a calculator or a ruler. Answer each question as completely as you can. Your thinking is important, so write complete sentences when you are describing your thinking.

1. From the Love your Heart Brochure, use the fat and cholesterol graph, describe the relationship between the amount of cholesterol in the blood and the percent of calories from fat.

2. The average San Franciscan consumes 90 kg of meat in one year. About how many grams of meat does that person eat in one day?

3. What is the <u>percent</u> of the calories from fat in a hot dog?

4. Use the mystery foods list, and information from the brochure. Do you think the list is made up of all animal sources or plant products? Write the reasons for your choices.

Mystery Foods List

Mystery Food	Fat content	Cholesterol Content
Food A	5 g	45 mg
Food B	26 g	72 mg
Food C	16 g	74 mg

5. What statement can you make about the amount of cholesterol in most plant foods? Justify your answer.

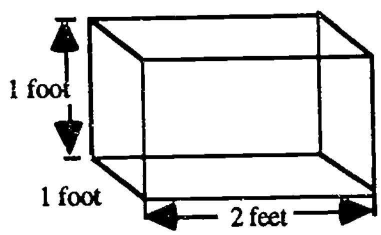
6. A friend of yours said, "Bacon is typical of animal sources of food because it has about 7 grams of fat per serving size." Assume that the meat and dairy products listed in the brochure are representative of all animal sources. How would you go about arguing for, or against the statement?

7. Using the foods in the brochure, write lunch menus for someone who lives in Guatemala City, and San Francisco. Explain your reasoning for designing the menus.

8. Your friend usually eats two eggs, bacon, buttered toast every morning, and you eat a bowl of oatmeal, a piece of fruit, and toast with jelly. Why is your breakfast healthier?

9. Design a display stand that will hold at least 50 oranges that average 4 inches in diameter.

10. Suppose that you own an orange orchard. Your containers for shipment to HEB are 1 foot high, 1 foot wide and 2 feet long and holds 54 oranges. How many of your containers will fit into a truck that is 10 feet long, 6 feet wide, and 12 feet high? Estimate the number of trucks that you would need if your orchard produced 50,000 oranges.



PIZZA PARTY

Instructions For the Student. Answer each question as completely as you can. Your thinking is important, so write complete sentences when you are describing your thinking. You may use a calculator, ruler, or any other tool.

Let's order pizza!

Pizza Palace's Prices

	Large	Medium	Per Slice
1 Topping	$5.99	$4.99	$1.75
2 Toppings	$7.50	$6.25	$2.25
3 Toppings	$9.99	$7.50	$3.25

Napoli's Pizza Palor Prices

	Large	Medium	Per Slice
1 Topping	$6.49	$4.89	$1.00
2 Toppings	$8.75	$6.99	$2.75
3 Toppings	$8.99	$6.50	$3.00

You would like to invite five of your friends to eat with you. The restaurant will slice the pizza into any number of slices you want, up to 12 slices.

1. Decide how many pizzas you want, and how many slices your pizza is to be cut. Why did you choose the number of pizzas and slices that you did?

2. How much will it cost altogether (including tax), if the tax rate is 7.5%?

3. Cheese comes with every pizza. The other toppings are mushroom, pepperoni, and sausage. How many different pizzas can you order using the topping choices?

4. If everyone is to get a fair share, how many pieces will each person get to eat?

5. Which would be a better deal, to buy the pizza by slices or by the whole pizza? Why?

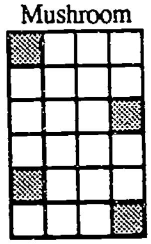
6. What fraction of the pizza does each person get?

7. How many pieces does it take to make half the pizza?

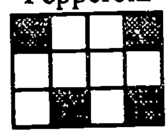
8. If you eat one-fourth of the pizza, then how many pieces did you eat?

9. **Use the pictures below.**
Suppose you ate mushroom pizza and your friend ate pepperoni pizza The shaded portion represents the pieces of a pizza that were eaten.

Mushroom



Pepperoni



a) What fraction of the mushroom pizza did you eat?

b) What fraction of the pepperoni pizza was eaten by your friend?

c) If both of you had eaten from one pizza, then how much is left?

d) Later, more of your friends came by. You offered them pizza from the mushroom pizza, how many friends can you serve if they each get 2 slices?

Let's collect the money.

Use the information from questions 1 and 2 on page 1 to help you complete questions 10, 11, and 12 on this page.

10. About how much should each member of the party pay? (Don't forget the tip for the wait person or delivery person)

Let's have something to drink!

11. A container of soda holds 2 liters and costs $1.20. Our drinking glasses hold 250 ml. How many glasses of the drink can we get from the container?

12. How many glasses can each person have in order to use all of the soda?

Here's the pizza!

13. How many 4 inch square slices can be served from a rectangular pizza with a length of 18 inches and a width of 12 inches?

14. What size would you make a rectangular pizza so that it will have the same area as a round pizza that is 14 inches in diameter?

Rescue At Boone's Meadow

You may use the information given to you such as the map and data sheet when you need them to answer the questions.

1. Use the given map. A flight plan is needed for rescuing the eagle. With Cumberland as a starting point, draw a picture of one possible flying route to and from Boone's Meadow on the map, include distances and a scale.

2. Use the information and map from the video.
A. List the possible routes you might use to rescue the eagle.

B. Describe how you would rescue the eagle, include the route that you think is best, the time it would take, and who would fly the ultralight. Show your calculations, and justify your answer.

3. Use the formula: P = W + F + C
Pay load = Weight of pilot + Fuel tank's weight when full + Cargo

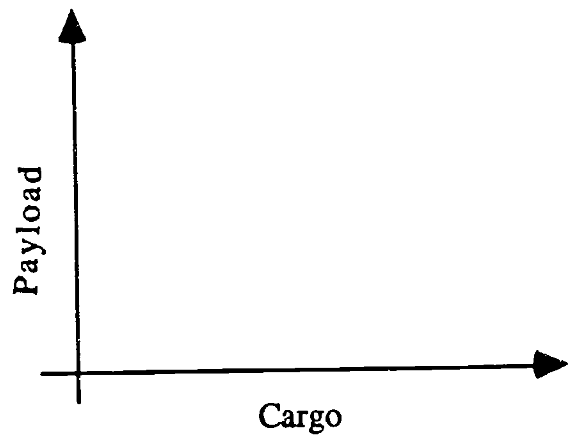
A. What is the pay load if you are the pilot who weighs 120 pounds, your tank full of fuel weighs 42 pounds, and you are to carry 50 pounds of food and water to a camper?

B. An ultralight's pay load is 250 pounds. The pilot weighs 175 pounds, and the cargo the pilot is to carry weighs 50 pounds. A gallon of fuel weighs 6 pounds per gallon. How much fuel can the pilot carry so that a take off is possible? Use P = W + F + C

4a. Complete the chart.

Payload	Cargo
210 lb.	10 lb.
215	
220	20
	25
230	
240	

4b. Graph the data from the chart.



4c. How did you find the payload when the cargo's weight is 25 pounds?

4d. Use the formula: $P = 170 + 6n$ where P is the pay load, and n is the number of gallons in the tank. Complete the chart for the condition, $170 < P < 200$ and then graph the information from the chart. What can you conclude from the graph?

P	n

5. Draw a design for a fuel tank and figure out how much gas it would hold. Why did you choose the shape that you did?
(1 cubic foot holds 5 gallons)

Kite Assessment

Instructions for the student. Answer each question as completely as you can. Your thinking is important, so write complete sentences when you are describing your thinking. You may use any tool(s) that you think will help you to answer questions like a calculator, a compass, a straight-edge, or a protractor.

Part 1 The Kite

1. Tell why the design you made on the kite is symmetric.

2. How many triangles does the kite form?

3. Draw a picture of all the different triangles that you can find in your kite.

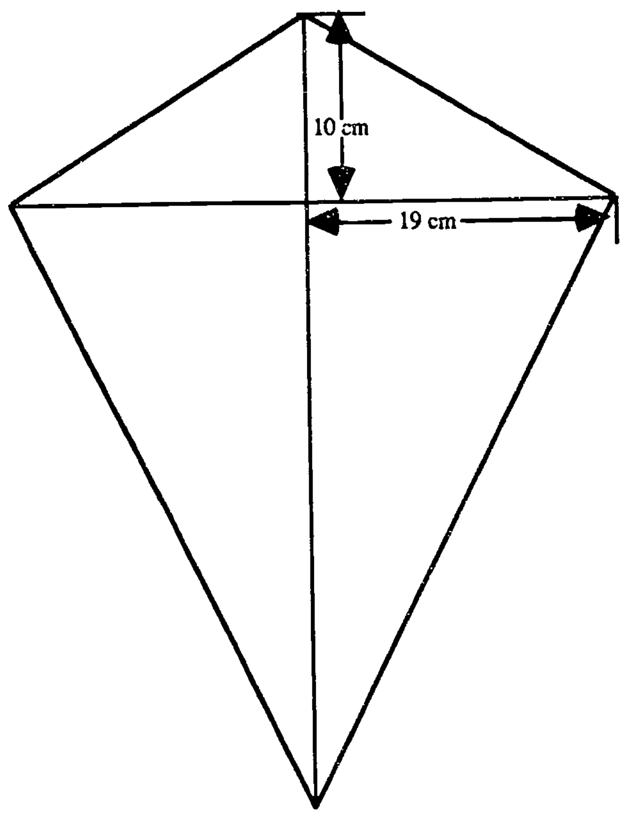
4. The spine is the longer, vertical brace and the spar is the shorter, horizontal brace. Measure the angles formed by the intersection of the spine and spar. What are the angles called?

5. If you wanted to build a bigger kite with the same proportion as the kite that you built, what would be possible lengths for the spine and spar?

6. Estimate the area of the design that you used to decorate your kite. How did you get the estimate?

7. Use the diagram and information to find the Kite's perimeter. What is the perimeter? Show your work.

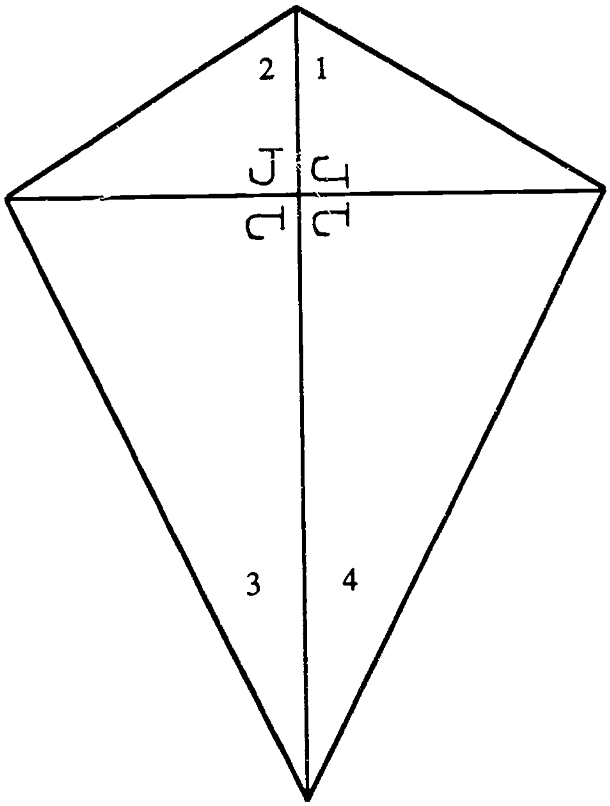
Information
The spar's length is 38 cm.
The spine's length is 45 cm.



8. Juanita decided to decorate her kite with a "J" as the pictures below shows.

i) From picture A, choose the transformation, on the right, that she used to move the "J," and place the letter in the blank on the left. You may use a transformation more than once.

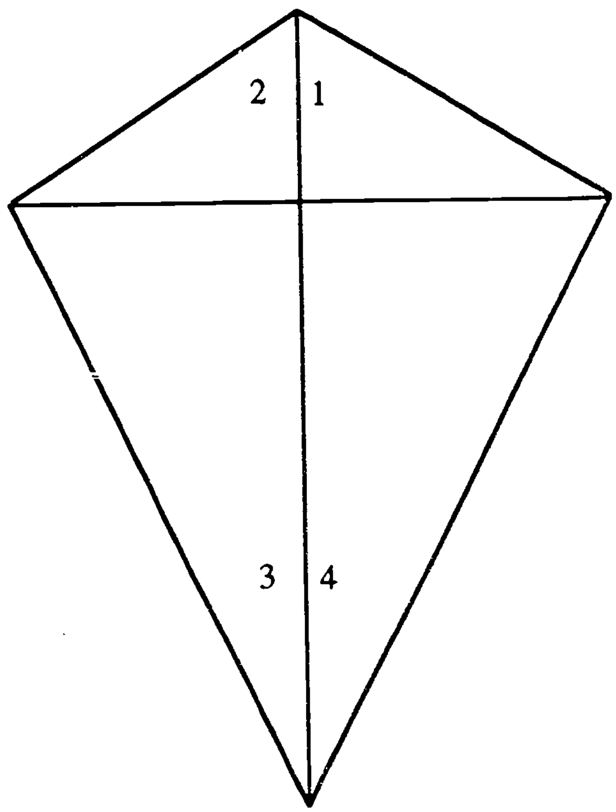
Picture A



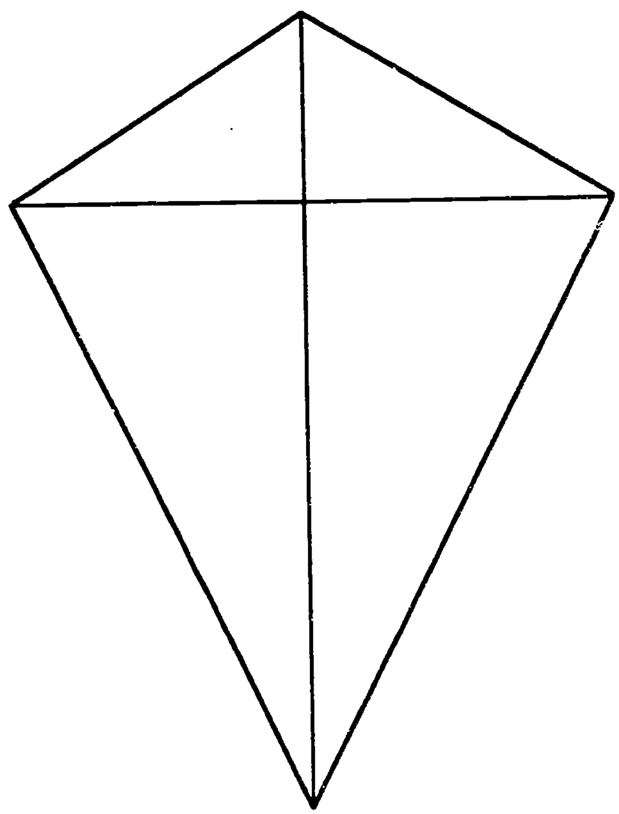
Movement	Transformation
______From quadrant 1 to 2	A. Rotation
______From quadrant 4 to 3	B. Translation
______From quadrant 3 to 4	C. Reflection
______From quadrant 4 to 1	

8ii) In picture B, place your first initial in the quadrants using the same transformations as Juanita used.

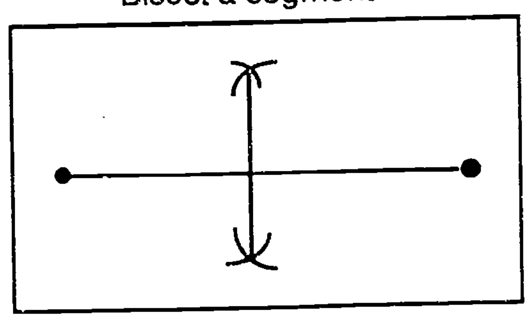
Picture B



9. Using only the compass, and straight edge find the midpoints of the sides of the kite, then join the midpoints. What does the new picture resemble (look like)?



Bisect a segment



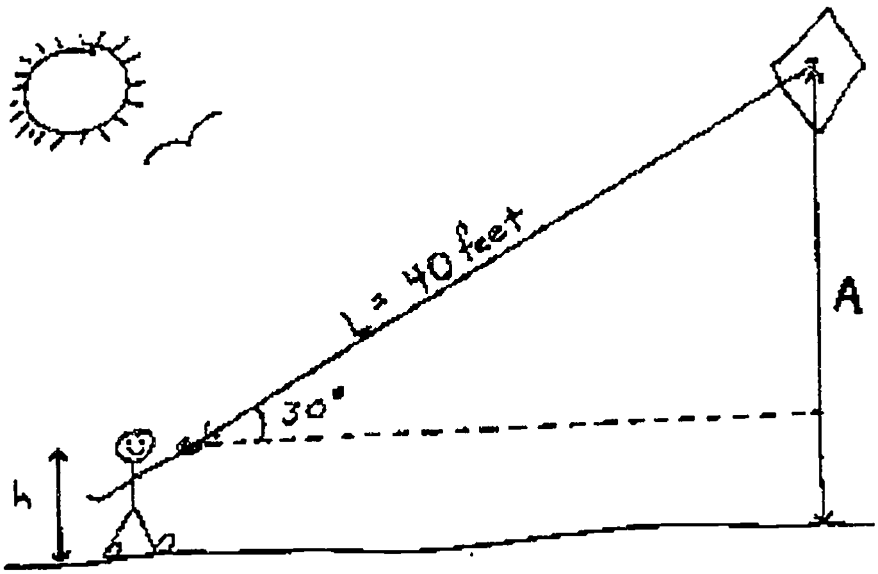
Part 2 Flying the kite.

The altitude of a kite is found by measuring (1) the angle from the ground to the kite using a tool called a clinometer, and (2) the length (L) of the string when the angle is 30 degrees.

Use the equation below to find the **Altitude (A)** of the kite:

$$A = \frac{L}{2} + h$$

where **L** = the length of the string, and **h** = your height. If you are not sure of your height, make a guess.



Use $A = \frac{L}{2} + h$ for questions 10 and 11.

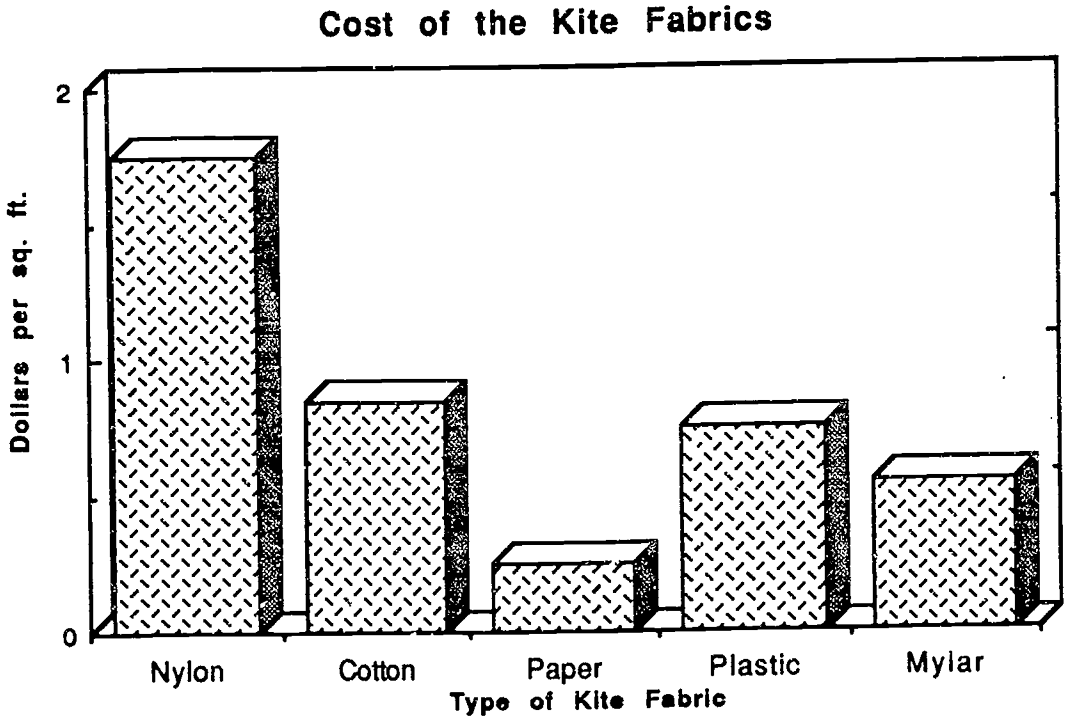
10. Determine the altitude of the kite in the above picture.

11. If your height (h) is 5 feet, how much string would you need so that the kite can reach a height of 105 feet when the angle is 30 degrees?

Frequency Table. Total Number of Kites Damaged and Entered in Contest

Type of Fabric	**Damaged**	**Entered**
Nylon	2	6
Cotton	3	8
Paper	12	48
Plastic	3	9
Mylar	4	6

Bar
Graph.



12. Use the above frequency chart and bar graph. What material would you use to make your kite if you were entering a kite fighting contest? Why would you use that fabric?

Frequency Table. Total Number of Kites Damaged and Entered in Contest

Type of Fabric	Damaged	Entered
Nylon	2	6
Cotton	3	8
Paper	12	48
Plastic	3	9
Mylar	4	6

13. Use the above frequency table. Which fabric has the highest probability of becoming damaged in a future contest? How did you find which fabric has the highest probability?

14. Suppose that you have $10.00 to spend on fabric to make a kite. Which fabric would you buy and how many square feet would the kite have?